



*Jan Sansoni*<sup>1</sup>

*Nick Marosszeky*<sup>2</sup>

*Yun-Hee Jeon*<sup>3</sup>

*Lynn Chenoweth*<sup>4</sup>

*Graeme Hawthorne*<sup>5</sup>

*Madeleine King*<sup>6</sup>

*Marc Budge*<sup>7</sup>

*Siggi Zapart*<sup>8</sup>

*Emily Sansoni*<sup>9</sup>

*Kate Senior*<sup>10</sup>

*Patsy Kenny*<sup>11</sup>

*Lee-Fay Low*<sup>12</sup>

**Suggested citation**

Sansoni J, Marosszeky N, Jeon Y-H, Chenoweth L, Hawthorne G, King M, Budge M, Zapart S, Sansoni E, Senior K, Kenny P, Low L (2007) ***Final Report: Dementia Outcomes Measurement Suite Project***. Centre for Health Service Development, University of Wollongong.

## Author Affiliations

1. Director, Australian Health Outcomes Collaboration and Principal Research Fellow, Centre for Health Service Development, University of Wollongong
2. Research Fellow, Centre for Health Service Development, University of Wollongong
3. Fellow, Australian Primary Health Care Research Institute, Australian National University
4. Professor, Health and Ageing Research Unit, University of Technology Sydney and South Eastern Sydney and Illawarra Area Health Service
5. Senior Lecturer, Department of Psychiatry, University of Melbourne
6. Associate Professor, Centre for Health Economics Research and Evaluation, University of Technology Sydney
7. Head, Department of Geriatric Medicine, Australian National University Medical School
8. Research Officer, Centre for Health Economics Research and Evaluation, University of Technology Sydney
9. Research Assistant, Centre for Health Service Development, University of Wollongong
10. Division Head, Education and Training, Menzies School of Health Research
11. Senior Research Officer, Centre for Health Economics Research and Evaluation, University of Technology Sydney
12. Postdoctoral Research Fellow, Dementia Collaborative Research Centre for Assessment and Better Care Outcomes, University of New South Wales

## Acknowledgements

The project team would like to thank all members of the National Expert Panel for their assistance and in particular the contributions of Dianne Wikstrom, Prof. Henry Brodaty and Dr. Rod McKay to the deliberations of the Expert Measurement Group. The project team would also like to thank Dr Marion Haas, Deputy Director, Centre for Health Economics Research and Evaluation for her feedback; and Elizabeth Cuthbert, Darcy Morris and Erin Gleeson from the Centre for Health Service Development for their administrative, editing and research support.

## Peer Review Statement

This report has been subject to peer review by an international expert in dementia, in order to assess its scientific merit and its contribution to the literature and clinical practice. In doing so, this report was evaluated in terms of its clinical utility, practicality and validity. The international expert was independently consulted by the Australian Government Department of Health and Ageing to conduct the review, as part of the Department's quality and evaluation processes. The review made a number of suggestions which were used to improve the final version of this report. In conclusion, the international expert found that "this report is thorough, up to date, well written, and gives a comprehensive overview of the most useful and well validated, assessment tools for the field of dementia diagnosis, cognitive and functional assessment and care."

## Table of Contents

<b>1</b>	<b>EXECUTIVE SUMMARY</b>	<b>1</b>
1.1	Project Aim and Description-----	1
1.2	Recommendations Concerning Clinical Terminology and Diagnostic Classification-----	2
1.3	Methods of Instrument Review-----	3
1.4	The Recommended Measures-----	6
1.4.1	Dementia Staging and Descriptive Measures	6
1.4.2	Health Related Quality of Life and Health Status Measures	8
1.4.3	Instruments for the Assessment of Cognitive Status	10
1.4.4	Multi-attribute Utility Measures	11
1.4.5	Measures of Social Isolation and Participation	13
1.4.6	Measures of the Associated Symptoms of Dementia	16
1.4.7	Measures of Function	19
1.4.8	Measures of Patient and Carer Satisfaction	21
1.5	Measurement Issues-----	27
1.5.1	Recommendations Concerning Cognitive Impairment and the Capacity to Self Rate	27
1.5.2	Recommendations Concerning Proxy Assessment	27
1.5.3	Recommendations Concerning Assessment with Culturally and Linguistically Diverse (CALD) Populations	28
1.5.4	Recommendations Concerning Assessment with Aboriginal and Torres Strait Islander Populations	28
1.6	Implementation Issues-----	29
1.7	Conclusion-----	30
<b>2</b>	<b>INTRODUCTION</b>	<b>34</b>
2.1	Revised Project Timelines and Reporting Requirements-----	34
2.2	An Outline of the Second Report-----	34
2.3	The Final Report-----	35
2.4	Meetings of the Dementia Outcomes Measurement Suite National Expert Panel and the Expert Measurement Group-----	35
2.5	An Overview of the Literature Search and Instrument Review Processes-----	36
<b>3</b>	<b>THE STANDARDIZATION OF CLINICAL TERMINOLOGY</b>	<b>40</b>
3.1	Background Issues-----	40
3.2	Detailed Examination of Health Classification Systems-----	41
3.2.1	Definitions and Diagnostic Criteria for Dementia	41
3.2.2	The International Classification of Functioning, Disability and Health (ICF)	50
3.2.3	Outcome Measurement of Dementia	51
3.2.4	Differential Diagnosis of Dementia	52
3.2.5	Severity of Dementia	54

3.2.6	Behavioural and Psychological Symptoms of Dementia (BPSD)	55
3.3	Recommendations	57
<b>4</b>	<b>DEMENTIA STAGING AND DESCRIPTIVE INSTRUMENTS</b>	<b>62</b>
4.1	Justification for Selection of Dementia Staging and Descriptive Measures for Review	62
4.2	Dementia Staging and Descriptive Measures	62
4.3	Search Strategies	63
4.4	Selecting the Measures for Comprehensive Review	63
4.5	Selecting Contender Instruments for Review	64
4.6	The Process of Reviewing the Best Five Measures	70
4.7	Strengths and Weaknesses of the Selected Instruments	70
4.7.1	Global Deterioration Scale (GDS)	70
4.7.2	Clinical Dementia Rating Scale (CDR)	71
4.7.3	Dementia Severity Rating Scale (DSRS)	73
4.7.4	Blessed Dementia Scale (BDS)	74
4.7.5	Sandoz Clinical Assessment – Geriatric (SCAG)	74
4.8	Summary of Instrument Scores and the Comparative Ranking of Instruments	75
4.9	Recommendations Concerning Dementia Staging and Descriptive Instruments	75
<b>5</b>	<b>HEALTH RELATED QUALITY OF LIFE INSTRUMENTS AND DEMENTIA</b>	<b>81</b>
5.1	Quality of Life in Dementia	81
5.2	Generic Health Status and Health Related Quality of Life Measures	82
5.3	Dementia Specific Health Related Quality of Life Measures: Initial Literature and Impact Search	87
5.4	Results of Detailed Review and Rating	88
5.4.1	Quality of Life in Alzheimer’s Disease (QOL-AD)	94
5.4.2	DEMQOL	96
5.4.3	Quality of Life in Late-Stage Dementia (QUALID)	97
5.4.4	Dementia Quality of Life Instrument (DQOL)	98
5.4.5	Alzheimer Disease Related Quality of Life (ADRQOL)	99
5.4.6	Cornell Brown Scale for Quality of Life in Dementia (CBS)	100
5.5	Patient Versus Proxy (Carer) Report of HRQOL	102
5.6	Recommendations	102
<b>6</b>	<b>INSTRUMENTS FOR THE ASSESSMENT OF COGNITIVE STATUS</b>	<b>108</b>
6.1	Cognition in Dementia	108
6.2	Measuring Cognitive Status in Dementia	108
6.3	Measurement Instruments	108
6.3.1	In-depth Clinical Neuropsychological Tests	109
6.3.2	Mental Status Tests (Including Simple Screening Tests)	109
6.3.3	Combination of Tests	109

6.4	Reviewed Instruments-----	110
6.4.1	Modified Mini Mental State Exam (3MS) (Teng and Chui, 1987)	112
6.4.2	Alzheimer's Disease Assessment Scale – Cognition (ADAS-Cog)	113
6.4.3	General Practitioner Cognition Scale (GPCOG)	116
6.4.4	Rowland Universal Dementia Assessment Scale (RUDAS)	117
6.4.5	Minimum Data Set – Cognition (MDS-COG)	118
6.4.6	Kimberley Indigenous Cognitive Assessment (KICA-Cog)	119
6.4.7	Other Approaches to Cognitive Assessment	119
6.5	Recommendations-----	120
<b>7</b>	<b>ECONOMIC EVALUATION IN DEMENTIA CARE AND THE INCORPORATION OF THE PATIENT PERSPECTIVE</b>	<b>129</b>
7.1	Economic Evaluation in Dementia-----	129
7.1.1	A Review of Cost-Utility Analysis (CUA) Dementia Studies	130
7.2	The Axioms of Utility Measurement-----	132
7.2.1	Measuring Utilities Using MAU-Instruments	133
7.3	Utility Instrument Review-----	135
7.4	Comparison of Instruments-----	138
7.5	Instrument Responsiveness-----	147
7.6	Conclusions and Recommendations-----	150
7.6.1	Summary Comments	151
7.6.2	Research Recommendations	155
<b>8</b>	<b>MEASURES OF SOCIAL ISOLATION AND ITS ASSESSMENT IN OLDER ADULTS</b>	<b>166</b>
8.1	Background-----	166
8.2	Method and Review Criteria-----	169
8.2.1	The Review Criteria	170
8.3	Review of the Instruments-----	171
8.3.1	DUKE-UNC Functional Social Support Questionnaire	171
8.3.2	Friendship Scale	173
8.3.3	De Jong Gierveld Loneliness Scale	174
8.3.4	Medical Outcomes Study Social Support Survey	179
8.3.5	Norbeck Social Support Questionnaire	181
8.3.6	Sarason Social Support Questionnaire	184
8.3.7	UCLA Loneliness Scale	187
8.3.8	Three-item Loneliness Scale	191
8.4	Discussion and Recommendations-----	193
8.4.1	Recommendations	197
<b>9</b>	<b>MEASURES OF THE ASSOCIATED SYMPTOMS OF DEMENTIA</b>	<b>208</b>
9.1	Introduction-----	208
9.1.1	Initial Search Strategies	208

9.2	Behavioural and Psychological Symptoms of Dementia (BPSD)-----	208
9.2.1	Decision Making Strategies	210
9.2.2	Neuropsychiatric Inventory (NPI)	216
9.2.3	Behavioural Pathology in Alzheimer's Rating Scale (BEHAVE-AD)	219
9.2.4	Dementia Behaviour Disturbance Scale (DBDS)	220
9.2.5	Neurobehavioural Rating Scale (NRS)	220
9.2.6	Consortium to Establish a Registry for Alzheimer's Disease – Behavior Rating Scale for Dementia (CERAD-BRSD)	221
9.2.7	Recommendations Concerning BPSD Instruments	223
9.3	Differential Diagnosis: Delirium -----	224
9.3.1	Decision Making Strategies	226
9.3.2	Confusion Assessment Method (CAM)	231
9.3.3	Delirium Rating Scale-Revised-98 (DRS-R-98)	232
9.3.4	Recommendations Concerning Delirium Instruments	234
9.4	Individual Symptom Measures for Associated Symptoms -----	235
9.4.1	Introduction	235
9.4.2	Aggression	235
9.4.3	Agitation	245
9.4.4	Conclusions Concerning Measures of Agitation	251
9.4.5	Anxiety	252
9.4.6	Apathy	255
9.4.7	Depression	259
9.5	Other Omnibus Measures: HoNOS 65+ -----	263
9.6	Conclusions and Recommendations -----	264
<b>10</b>	<b>MEASURES OF FUNCTION FOR DEMENTIA</b>	<b>284</b>
10.1	Introduction -----	284
10.1.1	Importance of the Measurement of Function for People with Dementia	284
10.1.2	A Simple Working Definition of Function	284
10.1.3	A Possible Analysis Framework for Functional Assessment Instruments	285
10.1.4	Challenges for Functional Assessment Instruments	285
10.1.5	Summary of the Measurement Literature for the Assessment of Function for People with Dementia	286
10.1.6	Challenges for Generic Functional Assessment Instruments when used with People with Dementia	287
10.1.7	Challenges for Dementia Specific Functional Assessment Instruments	287
10.1.8	Recent Research Highlights	288
10.2	Selection of Instruments -----	288
10.2.1	Literature Search	288
10.2.2	Short-listed Instruments	290
10.2.3	Reviewed Instruments	295
10.2.4	Instrument Rankings – Summary Rating	300

10.2.5	Recommendations	300
<b>11</b>	<b>MEASURES OF PATIENT AND CARER SATISFACTION</b>	<b>308</b>
11.1	Patient Satisfaction-----	308
11.1.1	Defining Patient Satisfaction	308
11.1.2	Method	310
11.1.3	The Review Criteria	313
11.1.4	Review of Items and the Instruments	313
11.1.5	Discussion	322
11.1.6	Recommendations	330
11.2	Carer Satisfaction with Services: A Review-----	332
11.2.1	Introduction	332
11.2.2	Method	335
11.2.3	The Review Criteria	336
11.2.4	Review of Items and the Instruments	337
11.2.5	Discussion and Recommendations	345
11.2.6	Recommendations	348
11.3	Other Informal Care Outcome Measures-----	348
11.3.1	Carers' Experience	349
11.3.2	Carer Health and Well-being	350
11.3.3	Carer Satisfaction with Services	350
11.3.4	Relationships Among Measures	350
11.3.5	Conclusion	351
<b>12</b>	<b>MEASUREMENT AND IMPLEMENTATION ISSUES</b>	<b>364</b>
12.1	Introduction-----	364
12.2	Cognitive Impairment and Self-Report-----	364
12.2.1	Cognitive Impairment and the Capacity to Self Rate	364
12.2.2	Methods to Facilitate Self-Completion	365
12.2.3	Recommendations	366
12.3	Proxy Measurement-----	366
12.3.1	Definition of Proxy Measurement	366
12.3.2	The Importance of Direct Measurement	367
12.3.3	Highlights of Recent Research	368
12.3.4	Advantages of Proxy Measurement	368
12.3.5	Disadvantages of Proxy Measurement	369
12.3.6	Characteristics Affecting Scores for Patients and Proxies	370
12.3.7	Suitable Domains of Proxy Measurement	371
12.3.8	Proxy / Informant Instruments	372
12.3.9	Recommendations when using Proxy Measures	373
12.3.10	Areas for Further Research	373
12.4	Dementia Measurement Issues with Culturally and Linguistically Diverse (CALD) Populations	374

12.4.1	Introduction	374
12.4.2	Assessment Issues	375
12.4.3	Current Research Developments	377
12.4.4	Guidelines for Assessment of Non-English Speaking People with Dementia	378
12.4.5	Valid and Reliable Dementia Outcome Measures for CALD Populations	379
12.4.6	Recommendations for the Assessment of CALD Populations	392
12.5	Dementia Assessment Issues for Aboriginal and Torres Strait Islander People-----	393
12.5.1	Introduction	393
12.5.2	Assessment Issues	393
12.5.3	General Difficulties with Dementia Measures in Indigenous Settings	394
12.5.4	Issues with the Tools for Assessment	394
12.5.5	Discussion	395
12.6	Implementation Issues-----	396
12.6.1	Introduction	396
12.6.2	Should Measures be Mandated or Recommended?	396
12.6.3	The Application of the Instruments in Different Settings and for Different Stages of Dementia	397
12.6.4	Further Issues Concerning the Application of the Instruments	402
12.6.5	Training Issues	408
12.6.6	A Dissemination Strategy	408
12.6.7	Identified Research Gaps	409
<b>13</b>	<b>CONCLUSIONS AND RECOMMENDATIONS</b>	<b>418</b>
13.1	Clinical Terminology and Classifications Systems-----	418
13.2	Recommended Assessment Instruments for Dementia-----	419
13.2.1	Dementia Staging and Descriptive Instruments	419
13.2.2	Health Related Quality of Life and Health Status Instruments	419
13.2.3	Instruments for the Assessment of Cognitive Status	419
13.2.4	Multi-attribute Utility Measures	420
13.2.5	Measures of Perceived Social Isolation and Social Support	421
13.2.6	Measures of the Associated Symptoms of Dementia	422
13.2.7	Measures of Function	423
13.2.8	Measures of Patient and Carer Satisfaction	425
13.3	Measurement Issues-----	426
13.3.1	Recommendations Concerning Cognitive Impairment and the Capacity to Self Rate	426
13.3.2	Recommendations Concerning Proxy Assessment	426
13.3.3	Recommendations Concerning Assessment with Culturally and Linguistically Diverse (CALD) Populations	427
13.3.4	Recommendations Concerning Assessment with Aboriginal and Torres Strait Islander Populations	427
13.4	Implementation Issues-----	428
13.4.1	Identified Research Gaps	429

13.5 Conclusion-----	431
<b>14 REPORT REFERENCE LIST</b>	<b>433</b>
<b>GLOSSARY OF TERMS</b>	<b>508</b>

## List of Tables

Table 1	Table of Criteria and Weights for Instrument Ranking.....	5
Table 2	Summary of Comparative Ratings for Dementia Staging and Descriptive Instruments .....	7
Table 3	Summary of Comparative Ratings for Dementia-Specific HRQOL Instruments.....	9
Table 4	Summary of Ratings for Cognitive Instruments .....	10
Table 5	Summary of Ratings for MAU Instruments.....	12
Table 6	Summary Assessing Social Isolation Instruments Against the Study Criteria .....	14
Table 7	Summary of Ratings for BPSD Global Instruments .....	17
Table 8	Summary of Ratings for Delirium Instruments .....	18
Table 9	Summary of Ratings for the Generic Measurement of Function Instruments.....	19
Table 10	Summary of Ratings for the Dementia Specific Measurement of Function Instruments.....	21
Table 11	Summary Assessing Patient Satisfaction Instruments Against the Study Criteria .....	23
Table 12	Summary Assessing Carer Satisfaction Instruments Against the Study Criteria .....	26
Table 13	Table of Criteria and Weights for Instrument Ranking.....	38
Table 14	Definitions and Diagnostic Features of Dementia .....	43
Table 15	Comparisons of the APA and WHO Classifications of Dementia .....	45
Table 16	Examples of Clinical/Practice Guidelines for Diagnosis of Dementia (with levels of evidence classification) .....	48
Table 17	Differential Diagnosis of Dementia, ICD-10-AM (2002).....	53
Table 18	The Severity of Dementia .....	55
Table 19	First-level Assessment of Global Dementia Measures.....	65
Table 20	Summary Sheet - Selected Dementia Staging and Descriptive Instruments .....	67
Table 21	Summary Sheet - Evaluation of Selected Dementia Staging and Descriptive Instruments .....	68
Table 22	Summary of Ratings for Dementia Staging and Descriptive Instruments .....	75
Table 23	Summary of the Six Short-listed Dementia-specific HRQOL Instruments (Part 1) .....	89
Table 24	Summary of the Six Short-listed Dementia-specific HRQOL Instruments (Part 2) .....	91
Table 25	Summary of Comparative Ratings for Six Short-listed Dementia-Specific HRQOL Instruments .....	101
Table 26	Summary of Ratings for Cognitive Instruments .....	120
Table 27	Health Care Costs, OECD Countries 1980-2000, Percentage of Gross Domestic Product ...	129
Table 28	Content of Descriptive Systems of MAU-Instruments .....	140
Table 29	Responsiveness of Selected MAU-Instruments to Various Health Conditions (a).....	149
Table 30	Summary Assessing the Utility Instruments Against the Study Criteria .....	154
Table 31	Definitions of the Social Functioning – Social Isolation Continuum.....	167
Table 32	Summary Assessing Social Isolation Instruments Against the Study Criteria.....	195
Table 33	Decision Summary of the BPSD Global Leading Contenders.....	212
Table 34	Summary of Ratings for BPSD Global Instruments.....	223
Table 35	DSM-IV-TR and ICD-10 Diagnostic Criteria of Delirium .....	225

Table 36	Decision Summary of the Delirium Leading Contenders .....	227
Table 37	Summary of Ratings for Delirium Instruments .....	234
Table 38	Decision Summary of the BPSD Global Leading Contenders .....	238
Table 39	Summary of Ratings for Aggression Instruments.....	245
Table 40	Decision Summary Table for Agitation Instruments .....	246
Table 41	Summary of Ratings for Measures of Agitation.....	252
Table 42	Short-listed Anxiety Instruments.....	254
Table 43	Summary of Ratings for Anxiety Instruments.....	255
Table 44	Short-listed Apathy Instruments .....	257
Table 45	Summary of Ratings for Apathy Instruments .....	259
Table 46	Short-listed Depression Instruments .....	260
Table 47	Summary of Ratings for Depression Instruments .....	262
Table 48	HoNOS 65+ Scales and Factors .....	263
Table 49	ADL Instruments.....	291
Table 50	IADL Instruments.....	292
Table 51	Combination Instruments .....	293
Table 52	Short-listed Measurement of Function Instruments .....	294
Table 53	Summary of Ratings for the Generic Measurement of Function Instruments.....	297
Table 54	Summary of Ratings for the Measurement of Dementia Specific Function Instruments .....	300
Table 55	Purpose of Instruments Reviewed .....	323
Table 56	Content Validity (Coverage).....	324
Table 57	Scoring of the Instruments .....	326
Table 58	Validity Evidence .....	327
Table 59	Reliability and Responsiveness Evidence.....	328
Table 60	Additional Criteria .....	329
Table 61	Summary Assessing Patient Satisfaction Instruments Against the Study Criteria.....	331
Table 62	Summary Assessing Carer Satisfaction Instruments Against the Study Criteria.....	346
Table 63	Main Findings from the Literature Review of Neumann, et al. (2000).....	371
Table 64	List of Proxy / Informant Instruments .....	372
Table 65	Recommended DOMS Instruments - Analysing Items for Acculturation and other issues.....	380

## List of Figures

Figure 1	Dementia and its Outcomes in the Structure of the ICF (Source: AIHW, 2007).....	52
Figure 2	Differences in Obtained Utility on the EQ-5D, by UK and US Scoring Algorithms .....	136
Figure 3	Data Distribution Issues for the EQ-5D .....	142
Figure 4	Characteristics Affecting Patient Scores and Proxy Ratings as Outlined by Snow, et al. (2005a) .....	370
Figure 5	List of Issues when Using and Interpreting Assessment Tools for CALD Populations.....	376
Figure 6	A Matrix Model for the Recommended Instruments.....	407



# 1 Executive Summary

## 1.1 Project Aim and Description

The purpose of this project is to develop a set of recommended measures/tools for routine use in the assessment, diagnosis, screening and outcomes monitoring of dementia conditions and the evaluation of treatments that are applicable for the Australian health care context. By developing a set of recommended measures it is hoped to standardise the assessment and evaluation procedures used in this field to enhance comparability of findings across research and practice settings. Put simply, we are trying to create a tool-kit of measures for clinicians and researchers to use with people with dementia, in order to assist with communication across the field. A related aim is to make recommendations concerning the clarification and standardization of the clinical terminology applicable in this field. To enhance comparisons between studies it is important that standardized approaches to diagnosis and patient classification be undertaken.

Although this project covers instruments that are useful for all stages of assessment (screening, prognosis and evaluation) this project has a particular focus on the assessment of outcomes. With respect to outcome evaluation in the context of dementia, where deterioration is part of the expected progress of the condition, it should be noted that positive outcomes of interventions may be expressed in terms of the maintenance of function or a reduction in the rate of decline rather than in terms of cure. Whilst psychometric features such as reliability and validity are relevant to instruments used at any stage of assessment, instruments that are used for outcome evaluation and monitoring must be sufficiently sensitive and responsive to detect changes in the person's/group's condition over time.

There are some limitations to the project's scope. With regard to measures used to assess cognitive impairment, a scoping exercise was undertaken by Prof Chenoweth concerning the cognitive measures. A decision was made by the DOMS-EMG that the project should focus on the instruments/tools that are available for use in routine care and this would exclude many of the more detailed neuropsychological instruments or instruments that require specialist training for their administration and interpretation. Feedback on this issue has also been obtained from other clinicians associated with the project and the DOMS-NEP. It was thought that a follow up project could undertake a more detailed assessment of the neuropsychological instruments to determine recommendations for this specialty.

Other issues outside this project's scope include comprehensive geriatric assessments for care or treatment planning like the 75+ health assessment, the interRAI or the EASY-Care. A recent review of these measures was conducted by the Lincoln Centre for Ageing and Community Care Research in 2004 (Lincoln Centre for Ageing and Community Care Research, 2004). Goal attainment scaling, recently advocated by Rockwood (2007), to individualise outcome measurement for people with dementia, has also not been examined in this project.

It should also be noted that the review of terminology in Section 3 indicates that recognition of mild cognitive impairment (MCI) is important and clinicians need to be vigilant about its further development to dementia, however there is insufficient evidence as yet to embrace MCI as a new diagnosis. At the first meeting of the National Expert Panel (NEP), the members agreed that given MCI is not fully established as a proper diagnosis and as the DOMS project focuses on the clinical phase of diagnosis it is best not to be included in this project. It is also noted that a related project by Cherbuin, et al. (2006) has a specific focus on reviewing dementia screening instruments to facilitate early detection of dementia and MCI.

The scope of this project has also been confined to an examination of carer satisfaction with health services and thus a detailed review of measures to assess carer burden, carer appraisal and carer wellbeing are not included in the scope of this report. However, it is acknowledged that carer

satisfaction may well be influenced by carer burden and carer wellbeing. Section 11.3 briefly outlines some of the interrelationships between these constructs. It is recommended that a review of instruments used to assess these domains could form a follow-up project to this report.

An assessment of the issues concerning safety/ risk assessment is outside the scope of this project. It is recommended that a further project be undertaken to examine risk assessment issues (e.g. elder abuse, aggression, self harm etc) for people with dementia (refer Section 12.6).

This project has been advised by two expert groups – the National Expert Panel (DOMS-NEP) and the Expert Measurement Group (DOMS-EMG). The National Expert Panel contains representatives from key dementia groups across Australia. The Expert Measurement Group consists of members of the project team with acknowledged expertise in the area of psychological measurement. The terms of reference and the membership of these groups can be found in Appendices 1 and 2 of this report.

## **1.2 Recommendations Concerning Clinical Terminology and Diagnostic Classification**

Section 3 of this report provides a detailed discussion of these issues. The recommendations below have been based on the review of literature, clinical feedback and these recommendations have also been reviewed by the National Expert Panel.

It is recommended that:

- The ICD-10-AM is used to inform the diagnostic classifications for dementia and its subtype given this system is already in place in collecting national data in Australia.
- The ICD-10-AM and ICD-10 are used for diagnostic criteria for dementia and Alzheimer's disease (AD). Following consultation it seemed appropriate to recommend the ICD-10 instead of DSM-IV. Clinicians do not necessarily follow either of the classifications as they often rely on their clinical judgement. Given that majority of the health related information is collected based on the ICD-10 and the ICD-10-AM it is more efficient for clinicians to use one system rather than two (i.e. DSM-IV diagnostic criteria and ICD-10 for coding exercise).
- For research, the DSM-IV is preferred as it is more inclusive of mild to moderate dementia and most epidemiological studies use the DSM-IV because of ease of comparison with prior studies. However this is not mandatory, providing the study states the type of the classification used, as there is no evidence available to say the DSM-IV is superior to the ICD-10.
- In terms of differential diagnosis (DD) and diagnoses of frontotemporal dementia (FTD) and dementia with Lewy bodies (LBD), additional criteria are used: the National Institute of Neurologic Disorders and Stroke and the Association Internationale pour la Recherche et l'Enseignement en Neurosciences (NINDS-AIREN) (Roman, et al. 1993) for DD of Vascular dementia from Alzheimer's type; the Lund-Manchester criteria for FTD (1994) and the consensus criteria for LBD (McKeith, et al. 2005).
- Mild cognitive impairment (MCI) is not to be included in this project as a diagnostic entity, however screening measures for those who are suspected of cognitive impairment need to be considered.
- For assessing the severity of dementia, the CDR scale has been used for two main reasons: the AIHW recommends this and, in addition to three stages of dementia, the CDR allows room to record abnormal cognitive function without necessarily labelling it as MCI. It is well validated and widely recognised. Similarly the GDS has also been widely used to assess the severity of dementia. A detailed review of these instruments is provided in Section 4 and in Appendix 5.
- The ICF may be used as a conceptual framework for classification of measurement scales. However, given its early developmental status as a classification system in Australia, hence its unfamiliarity among clinicians and researchers, and lack of evidence relating to validity and reliability of the classification, it is deemed beyond the scope of the DOMS project to provide a definite recommendation on this subject.

- Behavioural and psychological symptoms of dementia (BPSD) are an integral part of dementia outcome measures. The guidelines provided by the International Psychogeriatric Association (IPA) are to be used for the definitions. Whilst the AIHW recommends Caldwell and Bird's guideline for the severity of BPSD, it has been suggested that a more widely recognised measure is selected for this project. Readers are referred to the discussion in Section 9 of this report.

### **1.3 Methods of Instrument Review**

An initial overall literature search was undertaken (MEDLINE, PsycINFO) on twenty key terms (e.g. dementia, cognition, memory, function, Quality of Life etc). The major texts in the field were examined which included psychometric texts containing instrument reviews (e.g. McDowell, 2006; Bowling, 2001, 2005) as well as those containing instrument reviews applicable for dementia and assessment of the elderly (e.g. Burns, 2004; Kane and Kane, 2000; Lezak, 2004; McKeith, 1999). This process identified a list of instrument names and then searches were undertaken on all measures identified.

A database was then developed which provided comparative data for instruments for each domain / category (Associated Symptoms, Cognitive, Comprehensive, Dementia Staging and Description, Function, Health Related Quality of Life, Miscellaneous, Neuropsychological, Patient Satisfaction, Social and Utility Measures). This database included 844 named instruments.

An impact sheet was then developed for consideration by the review teams and the DOMS-EMG. This considered MEDLINE, text and web impacts, presence in instrument databases (e.g. PROQOLID) and its use in clinical practice. This process usually identified the leading twelve or so instruments for consideration in each category.

Additional selection criteria were then applied to reduce this to the leading 5-6 instruments in each domain / category. These criteria were:

- Whether there is a copy of the instrument and the original article concerning its development available for review.
- The number of citations found. In the case of new instruments some care was taken to assess this criterion as it was considered that recently developed instruments may not have a high citation rate. However, for instruments developed more than 5 years previously a low citation rate might indicate limited adoption by the field.
- The amount and range of the published psychometric evidence.
- Whether the instrument is used in clinical practice (evidence from the literature and data from NEP and other surveys).
- The availability of normative and clinical reference data.
- Administration time (generally 30 minutes or less) where a shorter administration time would be preferred. It was noted that as a number of instruments assessing different aspects (e.g. symptoms, cognition, HRQOL) will need to be utilized, lengthy instruments that may be more appropriate for detailed follow-up assessment may not be appropriate for use in routine assessment and across the range of practice settings.
- Whether the instrument is applicable for people with varying levels of dementia severity.
- Proprietary considerations (e.g. prohibitive cost).
- Applicability for use in routine care. Instruments would be preferred if they did not require specialist skills for administration or if extensive training in their use was not required (e.g. as for many neuropsychological/medical assessments).

Using the criteria above the shortlist of contender instruments was reduced to 5-6 measures for each category of measures and a decision summary sheet was developed to justify the selection or non-selection of contender instruments. Further searches were then undertaken for the selected

instruments using other databases (e.g. CINAHL, Cochrane Library etc) and the comprehensive reviews of these instruments commenced.

All instrument reviews make use of the AHOC instrument review sheet (refer Appendix 3) and provide information concerning the instrument's availability, applicability, requirements for administration, psychometric properties (reliability, validity, responsiveness, sensitivity, specificity) and the availability of normative and clinical reference data.

With all instruments consideration was also given to the following aspects:

- Type and stages of dementia
- Purpose of the instrument (assessment, screening, outcomes monitoring and the evaluation of interventions)
- Self-reporting and proxy reporting
- Respondent and staff burden
- Appropriateness for CALD and Aboriginal and Torres Strait Islander groups
- Appropriateness for a range of settings (e.g. community and residential care)

Once the comprehensive review for each instrument was completed an Instrument Scoring and Weighting Sheet was also completed for each instrument as indicated in Table 1 below.

**Table 1 Table of Criteria and Weights for Instrument Ranking****Criteria and weights used to assess instruments (DOMS)\*****Instrument Name ..... Total Score = .....**

Evaluation Criteria	Scoring system	Score	Weight	Weighted Score
Availability of comparison data	1 = minimal or no comparison data available		3	
	2 = some international comparison data available			
	3 = Australian and international dementia comparison data available including normative data and clinical reference data			
Length/feasibility of instrument for inclusion in battery	1 = long instrument, 30+ items		2	
	2 = medium length instrument, 15-30 items			
	3 = short instrument, less than 15 items			
Complexity of administration (for clinician use); and cognitive burden (for self report or proxy instruments)	1 = demanding to understand or administer		2	
	2 = some difficulties to understand or administer			
	3 = easy to understand and administer			
Cultural Appropriateness (ease of use with an interpreter, client literacy, CALD criteria including Indigenous Australians)	1 = not appropriate for use by CALD or illiterate clients, or with an interpreter		1	
	2 = limited appropriateness for use by CALD or illiterate clients and interpreters			
	3 = appropriate for use by CALD or illiterate clients and interpreters			
Ease of obtaining score by the evaluator	1 = scoring complex and requires computer		2	
	2 = can be scored without computer but time consuming			
	3 = scoring easy and does not require computer			
Sensitivity to dementia	1 = not known to be sensitive to dementia status		3	
	2 = sensitive to dementia status			
	3 = good sensitivity to dementia status			
Reliability evidence available	1 = little published evidence identified		3	
	2 = evidence suggests moderate reliability			
	3 = evidence suggests good reliability			
Validity evidence available	1 = little published validity evidence identified		3	
	2 = evidence suggests moderate validity			
	3 = evidence suggests good validity			
Cost of the instrument	1 = costs charged for using instrument		2	
	2 = costs for commercial use/training costs			
	3 = instrument available free of charge			
Cost of instrument administration	1 = professional		2	
	2 = paraprofessional/ staff member			
	3 = self complete			

The instrument is given a score against each criterion and this is multiplied by the weight for this criterion. The resulting weighted score for each criterion is then added to form a total score for each instrument (refer Table 1). For each category of instruments a comparative table of scores for the

instruments is then produced and it is on this basis the recommendations for each category of instruments are formed.

## **1.4 The Recommended Measures**

Sections 4 -11 of this report provide summaries and recommendations for the instrument categories reviewed to date. These include:

- Dementia Staging and Descriptive Measures (refer Section 4 and Appendix 5)
- Health Related Quality of Life Measures (refer Section 5 and Appendix 6)
- Cognitive Assessment Measures (refer Section 6 and Appendix 7)
- Multi-Attribute Utility Measures (refer Section 7 and Appendix 8)
- Measures of Social Participation and Isolation (refer Section 8 and Appendix 9)
- Measures of the Associated Symptoms of Dementia (refer Section 9 and Appendix 10)
- Measures of Function (refer Section 10 and Appendix 11)
- Measures of Patient and Carer Satisfaction (refer Section 11 and Appendix 12)

The recommended instruments for each category of measures are outlined below.

### **1.4.1 Dementia Staging and Descriptive Measures**

An outline of the selection processes relevant to this class of instruments can be found in Section 4. Five instruments were selected for comprehensive review in this class. These were:

1. Blessed Dementia Rating Scale (BDS)
2. Clinical Dementia Rating Scale (CDR)
3. Dementia Severity Rating Scale (DSRS)
4. Global Deterioration Scale (GDS)
5. Sandoz Clinical Assessment for Geriatric (SCAG)

**Table 2 Summary of Comparative Ratings for Dementia Staging and Descriptive Instruments**

Criteria	Weight	Instrument				
		GDS	CDRS	DSRS	Blessed	Sandoz
Availability of comparison data	3	2.5	2.5	1.5	2	2
Length/feasibility of instrument for inclusion in battery	2	3	1	3	1	2
Complexity of administration/ cognitive burden	2	3	2	3	2	3
Cultural Appropriateness	1	2	3	2	2	2
Ease of obtaining score	2	3	3	2	3	3
Sensitivity to dementia	3	2	3	2	3	2
Reliability evidence	3	3	3	3	2	2
Validity evidence	3	3	3	3	3	2
Cost of the instrument	2	3	3	3	3	3
Cost of instrument administration	2	2	1	2	2	2
<b>Weighted Total</b>		<b>61.5</b>	<b>57.5</b>	<b>56.5</b>	<b>52</b>	<b>50</b>

Table 2 provides a comparison of the scores of these instruments against the review criteria. It can be seen that the highest rated instrument was the GDS followed by the CDR. Both these instruments provide a rating of the severity of dementia although the GDS is somewhat easier to use than the CDR scale and is coupled with a much shorter administration time. The GDS can also be administered by care staff as well as clinicians. The GDS is also related to the Functional Assessment Staging (FAST) instrument.

- It is recommended that the GDS would be more appropriate for use as an initial assessment instrument and CDR might be more appropriate where a more comprehensive or second assessment is required. However, both instruments have good psychometric properties and are appropriate for use in both clinical and research settings for both assessment and outcomes evaluation.
- The DSRS also performed quite well but this is a rating scale for use by the caregiver rather than a clinical rating scale per se. It is, however, often used by care staff. This scale would be recommended for use in community settings and where information needs to be obtained from the caregiver. It is also easy for care staff to administer in residential care settings.

Burns et al (2004) indicates these measures are widely used as staging measures in descriptive and intervention studies. It is noted that specialist clinicians are less likely to use these global staging instruments than other clinical or research personal. Such instruments may not be particularly useful for fine differentiation at an early stage of dementia. However, global functional scales like the GDS

and the CDR have their place in broadly describing people with dementia; particularly for research purposes and in residential care and community care settings.

### 1.4.2 Health Related Quality of Life and Health Status Measures

HRQOL and health status instruments may be generic or disease-specific. A generic measure can be used for comparisons across diseases and health conditions. Widely used examples include multi-dimensional profiles such as the SF-36, Nottingham Health Profile, and the Sickness Impact Profile, and indices for economic evaluation such as EQ-5D and AQoL (which are reviewed in Section 7). In contrast, disease or condition specific health related quality of life measures focus on those aspects of health (e.g. symptoms) and health-related quality of life that are relevant to a particular health condition such as cancer or heart disease. Dementia-specific examples include the Quality of Life in Alzheimer's disease scale or the DEM-QOL.

With regard to the assessment of health related quality of life of those experiencing dementia there are significant limitations concerning the use of generic health related quality of life scales with people with dementia. As the symptoms of dementia differ significantly from those of other illnesses, and as generic health related quality of life measures do not cover some key domains for dementia (e.g. cognition, behavioural disturbance), many researchers prefer just to use a disease specific measure to assess health related quality of life in dementia (Rabins and Black, 2007). Some items in these instruments may be inappropriate to elderly people – for example questions concerning vigorous activities or how health has affected work (McDowell, 2006). The question frames in some of the items included in these scales are complex and assume a level of cognitive function that would make them unsuitable for use with those experiencing moderate to severe cognitive impairment. Most generic HRQOL measures are also self report measures and as Rabins and Black (2007) indicate many individuals with dementia, particularly those with moderate to severe illness, lack the capacity to self rate.

A discussion concerning the capacity to self rate and the use of proxies can be found in Sections 12.2 and 12.3. Self report instruments such as the SF-36 are clearly not suitable for use with people with severe dementia (MMSE of 10 or less) and require an assisted interview administration for those with an MMSE less than 15 (Novella, et al. 2001). While such measures could possibly be used with people with mild dementia, these measures may be more appropriately used to assess the health related quality of life of carers of people with dementia.

Section 5 of this report provides a more detailed discussion of the generic health related quality of life measures and these measures were also recently reviewed by Thomas, et al. (2006). No generic health related quality of life measure is recommended for use with people with dementia. At the present time the dementia specific quality of life measures, reviewed below, would seem more appropriate measures to use with people with dementia. Dementia specific measures more adequately capture the relevant dimensions for this condition and as such are more likely to capture the way that people with dementia decline and/or improve over time and thus are likely to be more useful measures for assessing the outcomes of people with dementia.

#### 1.4.2.1 Dementia Specific Health Related Quality of Life Measures

An outline of the selection processes relevant to this class of instruments can be found in Section 5. The six leading dementia-specific HRQOL instruments identified were:

1. Alzheimer Disease-Related Quality of Life (ADRQOL)
2. Cornell Brown Scale for Quality of Life in Dementia (CBS)
3. Dementia Quality of Life Instrument (DQOL)
4. DEMQOL (this is the instrument's full name, not an abbreviation)
5. Quality of Life in Alzheimer's Disease (QOL-AD)
6. Quality of life in Late-Stage dementia (QUALID)

Table 3 (below) provides a comparison of the scores of these instruments against the selection criteria.

**Table 3 Summary of Comparative Ratings for Dementia-Specific HRQOL Instruments**

Criteria	Weight	Instrument					
		QOL-AD	DEM QOL	QUALID	DQOL	CBS	ADR QOL
Availability of comparison data	3	2	2	2	2	1	2
Length/feasibility of instrument for inclusion in battery	2	3	2	3	2	2	1
Complexity of administration /cognitive burden	2	3	3	3	3	3	3
Cultural Appropriateness	1	2	1	2	1	1	2
Ease of obtaining score	2	2	2	2	2	2	1
Sensitivity to dementia	3	3	3	2	2	2	2
Reliability evidence	3	3	3	3	3	3	3
Validity evidence	3	3	3	3	3	3	3
Cost of the instrument	2	3	2	2	2	3	1
Cost of instrument administration	2	2	2	2	2	1	2
<b>Weighted Total</b>		<b>61</b>	<b>56</b>	<b>56</b>	<b>53</b>	<b>50</b>	<b>48</b>

After considering the key attributes of the instruments, and all the evidence about their psychometric properties, the following recommendations are made:

- Three instruments are recommended for the assessment of HRQOL in dementia; the QOL-AD and the DEMQOL for mild to moderate dementia and the QUALID for late stage dementia only.

Based on current evidence, as presented in Section 5 and in Appendix 6, the QOL-AD is clearly the strongest instrument, and if only one dementia-specific HRQOL instrument were to be recommended, then it would be the one. The decision to recommend a further two instruments was based on two factors. Firstly, late stage dementia is very different to mild or moderate dementia, in terms of both the issues that define and affect quality of life and also the way HRQOL can be measured or observed. This factor led to the recommendation of QUALID, given the relevance and appropriateness to late stage dementia of its content and mode of measurement.

The second factor was the newness of the DEMQOL balanced against the world-class credentials of its development team – it is an instrument whose promise is yet to be realized. Although limited, the available evidence suggests that the psychometric properties of both DEMQOL and DEMQOL Proxy are at least as good as those of the QOL-AD.

It is noted that none of these instruments have published Australian reference data.

- It is recommended that such data be collected in an Australian field test of these instruments.

Further detail of the three recommended instruments' psychometric properties, with citation details, plus information on other practical issues such as availability, is provided in Section 5 and summarized in the instrument review sheets in Appendix 6.

### 1.4.3 Instruments for the Assessment of Cognitive Status

After consideration of a large number of contender instruments, (refer Section 6) the final five instruments selected for comprehensive review were:

1. Alzheimer's Disease Assessment Scale – Cognition (ADAS-COG)
2. General Practitioner Cognition Scale (GPCOG)
3. Modified Mini Mental State Exam (3MS)
4. Minimum Data Set – Cognition (MDS-COG)
5. Rowland Universal Dementia Assessment Scale (RUDAS)

These instruments were selected because they covered a range of settings including primary care and residential care settings.

After consideration of the appropriateness of these tools for use with Indigenous people, an additional tool, an assessment of the Kimberley Indigenous Cognitive Assessment tool was included in the Table below and it has also been discussed in Section 6.

Table 4 below provides the comparative scores for the cognitive assessment instruments.

**Table 4 Summary of Ratings for Cognitive Instruments**

Criteria	Weight	Instrument					
		MMSE (3MS)	ADAS-COG	GPCOG	RUDAS	MDS-COG	KICA-COG <sup>b</sup>
Availability of comparison data	3	3	2	1	1	1	1
Length/feasibility of instrument for inclusion in battery	2	2	1	3	3	3	2
Complexity of administration / cognitive burden	2	2	2	3	2	3	2
Cultural Appropriateness	1	2	3	1	3	1	3 <sup>b</sup>
Ease of obtaining score	2	3	3	3	3	3	3
Sensitivity to dementia	3	3	3	2	2	2	2
Reliability evidence	3	3	3	3 <sup>a</sup>	3 <sup>a</sup>	2 <sup>a</sup>	2
Validity evidence	3	3	3	3 <sup>a</sup>	3 <sup>a</sup>	3 <sup>a</sup>	1.5
Cost of the instrument	2	3	3	3	1	2	3
Cost of instrument administration	2	2	1	1	2	2	2
<b>Weighted Total</b>		<b>62</b>	<b>56</b>	<b>54</b>	<b>52</b>	<b>51</b>	<b>46.5</b>

## Notes:

<sup>a</sup> Scored as 2 or 3 because despite their being limited evidence, what there is indicates good validity and/or reliability.

<sup>b</sup> This is a new tool designed for the cognitive assessment of Indigenous people.

- The Modified MMSE (3MS) is recommended as a widely used instrument that assesses global cognitive status in older people. It is applicable in both community and institutional settings. It has superior psychometric properties and has been extensively used in large scale epidemiological studies internationally (mostly North American studies). There is also extensive normative and clinical data available. An increasing number of studies use a translated version of the 3MS to achieve cultural appropriateness and it has slightly better psychometric properties than the standard MMSE. The instrument equals or outscores all the other instruments in almost every category.
- The ADAS-Cog is recommended for second stage or more detailed assessments and/or for particular research evaluations rather than for applications in routine care settings. It is noted that the ADAS-Cog received the second highest score and it has good psychometric properties. However, the ADAS-Cog requires staff with specialist qualifications for its administration, its use requires additional training and it takes 30-45 minutes for completion of the assessment.
- The GPCOG is recommended because of its usefulness in the primary care setting. As it is a relatively new instrument, it has not been widely used in research studies, normative data is not yet available, and the instrument has not been translated in to any other languages. Despite this, the GPCOG has scored well on the psychometric criteria. In addition, anecdotal evidence suggests that GPs are using the instrument and finding it very useful.
- The MDS-COG is recommended, despite having the lowest ranking total. The reason for this is that it was felt it was important to include a cognitive rating scale that would be useful in the residential care setting. The strength of this instrument is that it enables evidence about the cognitive status of patients to be obtained without any extra effort on the part of staff. The information is routinely entered as the patient enters long term care. Despite the total score on these criteria being slightly lower than some of the other instruments, it may be useful to include a rating scale like this for people with severe dementia.
- The RUDAS is a new instrument that was designed to enable the easy translation of the items into other languages and to be culture fair. There are relatively few papers published as yet concerning its psychometric properties (especially construct validity) but in the interim it is recommended for use with those from Culturally and Linguistically Diverse backgrounds. The RUDAS, however, contains an item on judgement that may be inappropriate for remote Indigenous people (refer below).
- An interim recommendation is to use the Kimberley Indigenous Cognitive Assessment (KICA-Cog) tool for the cognitive assessment of rural and remote Indigenous people. The KICA-Cog is a new instrument and although there is little published evidence concerning this tool available as yet, and further research is required, this instrument has been designed for use with Indigenous people.

#### 1.4.4 Multi-attribute Utility Measures

Multi-attribute utility measures are health related quality of life measures that are designed for economic evaluations of treatments and health care interventions particularly when using cost utility analysis. As indicated in Section 7 there are major difficulties in using self reported multi-attribute utility measures with patients experiencing moderate to severe dementia. However, on the other hand it is generally preferred to use patient assessments rather than those of proxies, as evidence indicates these assessments can differ widely.

There is also limited evidence concerning the use of these instruments in assessing the effectiveness of treatments for dementia and regarding the sensitivity of each of these measures in relation to dementia status. As a result the recommendations of Section 7 do not support the recommendation of any one instrument for use in economic evaluations but suggest instead some further research needs

to be undertaken.

Seven multi-attribute utility instruments were identified in the initial searches. These were:

1. Assessment of Quality of Life (AQoL)
2. European Quality of Life Measure (EQ-5D formerly the EuroQol)
3. Health Utility Index 3 (HUI3)
4. 15D
5. Quality of Well-Being (QWB)
6. Rosser Index
7. SF6D

Table 5 provides the comparative scores for each of these instruments against the rating criteria.

**Table 5 Summary of Ratings for MAU Instruments**

Criteria	Weight	Instrument						
		EQ-5D	AQoL	HUI3	15D	QWB	SF6D	Rosser
Availability of comparison data	3	3	3	3	3	3	3	2
Length/feasibility of instrument for inclusion in battery*	2	3	2	2	2	1	1#	1
Complexity of administration /cognitive burden	2	3	2	1	2	1	1#	2
Cultural Appropriateness	1	2	3	2	2	2	2	2
Ease of obtaining score	2	3	3	3	3	3	3	2
Sensitivity to dementia	3	2	2	2	1	2	1	1
Reliability evidence	3	2	2	2	1	1	1	1
Validity evidence	3	2	2	2	1	1	1	1
Cost of the instrument	2	2	3	1	3	3	1	3
Cost of instrument administration	2	3	3	2	3	1	3	1
<b>Weighted Total</b>		<b>57</b>	<b>56</b>	<b>47</b>	<b>44</b>	<b>41</b>	<b>38</b>	<b>35</b>

\* As most MAU instruments are short the criteria are revised as follows: 1= Long instrument or needs interview administration, 2= moderate length self completed instrument, 3= short, self completed instrument.

# Although it only contains 10 items it requires the full administration of the SF-36 scale

The three instruments that score most highly on these criteria are the EQ-5D, the AQoL and the HUI-

3. However both the HUI-3 and the AQoL are lengthier instruments which may place considerable cognitive burden on people with dementia. It is noted that the HUI-3 does not score as highly on these criteria as the AQoL and the EQ-5D instruments for dementia settings and there are also considerable costs associated with the use of the HUI-3 which may also preclude its adoption.

- It is recommended that the EQ-5D, and the AQoL are to be the preferred instruments when undertaking economic evaluation of dementia interventions.

The obvious instrument of choice for use in dementia studies might be the EQ-5D because of the simplicity of the descriptive system. There are however very good technical reasons which provide caveats to its widespread use, including competing scoring algorithms, ceiling effects, inconsistent utility scores and poor score distribution.

- It is recommended that an Australian study be undertaken into these aspects of the EQ-5D with a view to validate and/or revise existing EQ-5D scoring algorithms.

Based on the scoring criteria, the next best-performing MAU-instrument was the AQoL. There are, however, two important caveats to recommending it as the instrument of choice. Although the AQoL's descriptive system is simple, the wording of items is stilted. The second caveat is in relation to the number of items needed to score the AQoL (12-items) which may explain higher rates of missing data when compared with the EQ-5D, and inconsistent scores for those with severe cognitive impairment. Theoretically, given the factorial structure of the AQoL it could be shortened through removal of 4 items (1 from each dimension) leaving it as an 8-item instrument.

- It is recommended that a study be undertaken to examine the effect of simplifying the AQoL items and removing four items to make it more appropriate for use in dementia research.

A single MAU-instrument could be recommended as the preferred instrument of choice for routine use at the clinician- and specialist-levels. This instrument should be short, easy to administer and score and population norms could be made available for easy reference. If such a policy was adopted, it would be in light of the limitations outlined in this report and there would be no guarantee that results obtained would be comparable with results obtained elsewhere using another instrument. Indeed, where QALYs were computed as the result of a treatment, it is likely these would reflect instrument choice as much as treatment effect.

- It is recommended that two MAU-instruments could be included in any particular research or evaluation study, and that researchers be encouraged to provide both sets of results. One of the recommended instruments should be that recommended for clinician use. This strategy would have the benefit of reducing the bias inherent in a one-instrument strategy, and it would produce a range of estimated benefits from interventions, thus acknowledging the limitations of relying upon any particular existing MAU-instrument. Given that, inevitably, comparisons will be made with dementia studies overseas, this strategy would have the further benefit of enabling cross-cultural comparisons. An important limitation of this strategy is that it would increase the cognitive burden for those with moderate to severe cognitive impairment. It may also lead to interviewer-facilitated or proxy completions, with all the implications of mixed-methods data collection.

#### **1.4.5 Measures of Social Isolation and Participation**

Following literature searches fifteen instruments were initially identified (refer Section 8). Following further consideration of their psychometric properties and applicability to dementia seven instruments were selected for a more detailed examination. They are:

- DUKE Functional Social Support Questionnaire (Broadhead, et al. 1988)

- Friendship Scale (Hawthorne, 2006)
- Loneliness Scale (De Jong Gierveld and Kamphuis, 1985; De Jong Gierveld and Tilburg, 2006)
- Medical Outcomes Study Social Support Survey (Sherbourne and Stewart, 1991)
- Norbeck Social Support Questionnaire (Norbeck, 1984, 1981; Norbeck, et al. 1983), the Social Support Questionnaire (Sarason, et al. 1987; Sarason, et al. 1983)
- UCLA Loneliness Scale (Russell, et al. 1980, 1978; Russell, 1996) and its short derivative (refer below)
- Three-item Loneliness Scale (Hughes, et al. 2004)

The instruments selected appeared to fall into two different categories concerning their focus of measurement. There were those that are concerned with reporting social participation, networks, support or social contact (e.g. Duke FSSQ, Sarason Social Support Questionnaire, Norbeck Social Support Questionnaire) and those which focus on social isolation or loneliness (Loneliness Scale, Friendship Scale etc). The MOS Social Support Scale includes items covering both dimensions.

There is also a divide between so-called objective measurement of the number of social contacts and the more subjective personal assessment of either satisfaction with social contacts or feelings of the depth of loneliness. The literature is suggestive that it is the latter that is more important although it may be desirable to tap both dimensions.

Table 6 provides the comparative scores for each of these instruments against the rating criteria.

**Table 6 Summary Assessing Social Isolation Instruments Against the Study Criteria**

Criteria	Weight	Instrument							
		DeJong	MOS	FS	Duke	Sarason	UCLA	3-IT	Norbeck
Availability of comparison data	3	1	1	1	1	1	1	1	1
Length/feasibility of instrument for inclusion in battery	2	3	2	3	3	1	2	3	1
Complexity of administration/ cognitive burden	2	2	2	3	2	2	2	2	1
Cultural Appropriateness	1	2	2	2	2	2	2	2	2
Ease of obtaining score	2	3	2	3	3	2	2	3	1
Sensitivity to dementia	3	1	1	1	1	1	1	1	1
Reliability evidence	3	3	3	2	2	3	3	1	2
Validity evidence	3	3	3	2	1	2	1	1	2
Cost of the instrument	2	3	3	3	3	3	3	3	3
Cost of instrument administration	2	3	3	3	3	3	3	3	2

<b>Weighted Total</b>		<b>54</b>	<b>50</b>	<b>50</b>	<b>45</b>	<b>45</b>	<b>43</b>	<b>42</b>	<b>36</b>
-----------------------	--	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------

Given the discussion in Section 8 and the scores in Table 6 above, none of the reviewed instruments can be given an unqualified recommendation for use in Australian studies with older adults who have cognitive impairment or dementia.

Subject to this finding, the standout instrument was the De Jong Gierveld Loneliness Scale. The reasons were that it was carefully conceived over a very substantial period of time, that it was developed in population samples (including older adults), and that there is a very substantial body of evidence supporting its reliability and validity. The reason the De Jong Gierveld Loneliness Scale, especially the short 6-item version, cannot be recommended outright is that the response categories may be inappropriate for use in Australian samples of people with cognitive impairment. However, a study can easily be completed to undertake a linguistic validation of the De Jong Gierveld Loneliness Scale instrument for Australian use and this is recommended.

The two other instruments that performed relatively well against the criteria were the Friendship Scale and the Medical Outcomes Study Social Support Survey. The Friendship Scale generally performed well on all criteria; it is short, easy to use, the scale was developed in samples of older adults and it appears to be reliable, valid and sensitive. The limitation is that it is a new scale that has been published in just one paper to date and some issues have been raised concerning the methods of item selection for this scale although this was based on a sound theoretical model. The Medical Outcomes Study Social Support Survey is a well-conceptualised and developed instrument. In general, it performed well against the study criteria, with the exception of those criteria related to instrument length (instrument length, cognitive burden, cultural appropriateness and scoring).

Given this situation, it is further recommended:

- That the three instruments which performed well (the De Jong Gierveld Loneliness Scale, the Friendship Scale and the Medical Outcomes Study Social Support Survey) be trialled in at least one large dementia study for the explicit purpose of identifying the instrument to be recommended for future use. This would enable many of the questions raised in this report regarding the validity of these instruments to be thoroughly investigated in an Australian context. It may also be possible to derive a better short measure by selecting the items with the best properties from these scales.
- That explicit modification to the De Jong Gierveld Loneliness Scale and the Medical Outcomes Study Social Support Survey be tested. These modifications are revision of the De Jong Gierveld Loneliness Scale response categories, and a reduction in the number of items in the Medical Outcomes Study Social Support Survey (which would need to be tested in the study outlined above).
- That the three instruments which performed well be tested in a trial for the effect of administration mode on scores, given that there are good reasons for limiting self-completion among those with moderate or severe cognitive impairment. Three methods of administration should be directly compared (self-completion without assistance, interviewer-assisted completion, and proxy-completion) both cross-sectionally and longitudinally in order to develop algorithms for weighting enabling score equivalence across administration mode. This would overcome issues related to the cognitive impairment of respondents and meet the need to collect outcome efficacy data relating to program evaluation.
- That from any study carried out under the recommendations above, a statistically-derived single item measure be identified for use in everyday clinical consultations.

### 1.4.6 Measures of the Associated Symptoms of Dementia

Associated symptoms of dementia relate to characteristics of dementia that are not historically considered as major features such as cognitive impairment and related functional consequences, yet have a significant impact on the well-being of the persons with dementia and their family and caregivers. Measuring outcomes of care, service, treatment and interventions related to the associated symptoms of dementia is an important aspect. For the purpose of the DOMS project, the assessment of associated symptoms of dementia comprises:

- 1) Measures of global behavioural and psychological symptoms of dementia (BPSD Global, henceforward);
- 2) Measures of delirium, which is one of the two most frequently mistaken features requiring differential diagnosis from dementia (the other commonly mistaken feature is depression); and
- 3) Measures of particular symptoms of BPSD including aggression, agitation, anxiety, apathy, and depression.

#### 1.4.6.1 Recommendations Concerning BPSD Global Instruments

A number of global measures of behavioural and psychological disturbance (Global BPSD) have been reviewed. As shown in Table 7, the examination of key attributes and psychometric properties of the five final instruments measured against the weighting criteria indicates the Neuropsychiatric Inventory (NPI) and the Behavioural Pathology in Alzheimer's Rating Scale (BEHAVE-AD) as the best measures for assessment of BPSD, followed by the Consortium to Establish a Registry for Alzheimer's Disease – Behaviour Rating Scale for Dementia (CERAD-BRSD), the Dementia Behaviour Disturbance Scale (DBDS) and the Neurological Rating Scale (NRS). Based on these reviews it is recommended that:

- The NPI and the BEHAVE-AD be used in both clinical and research settings for assessment of Global BPSD. These instruments both have well established psychometric properties.
- The CERAD-BRSD is recommended for research rather than routine practice given its cost and the time required for its administration. A 17 item abbreviated version may be considered better for clinical utility, but limited evidence on this version is currently available.

**Table 7 Summary of Ratings for BPSD Global Instruments**

Criteria	Weight	Instrument				
		NPI	BEHAVE-AD	CERAD-BRSD	DBDS	NRS
Availability of comparison data	3	2	2	2.5	1	2
Length/feasibility of instrument for inclusion in battery	2	3	2	1	2	2
Complexity of administration/ cognitive burden	2	3	3	2	2	1
Cultural Appropriateness	1	3	3	2	2	2
Ease of obtaining score	2	3	3	3	3	3
Sensitivity to dementia	3	3	3	3	2	2
Reliability evidence	3	3	3	3	3	3
Validity evidence	3	3	3	3	2	2
Cost of the instrument	2	3	3	1	2	2
Cost of instrument administration	2	2	2	2	2	1
<b>Weighted Total</b>		<b>64</b>	<b>62</b>	<b>54.5</b>	<b>50</b>	<b>49</b>

#### 1.4.6.2 Recommendations Concerning Measures of Delirium

A number of delirium measures were also assessed in order to aid in the differential diagnosis of dementia and delirium. The Confusion Assessment Method (CAM) is the most widely utilised screening/diagnostic tool for detecting delirium internationally among older people with or without dementia. Less well known, however, the Delirium Rating Scale (DRS-R-98) is also a widely recognised and well validated measure. Whilst the CAM is superior in its utility to the DRS-R-98, it does not capture severity of delirium symptoms hence is not appropriate for repeated measures of delirium severity. The DRS-R-98 is designed for assessment of both the presence and the severity of delirium symptoms. Limitations of the DRS-R-98, and the DRS, include that they are time taxing and require sufficient training, especially for those who do not have a psychiatric background. The DRS-R-98 is not appropriate for use in the community setting given its requirement for observation over a 24 hour period. However, it allows for comprehensive assessment of individuals who are at risk or suspected of developing delirium in institutional care settings. The ratings for these instruments can be found in Table 8 below.

For the purpose of the DOMS project, it is recommended both measures be included as they have two distinct, yet equally important functions.

- It is recommended that the Confusion Assessment Method is used to assess the presence of delirium across most service settings.
- It is recommended that the Delirium Rating Scale (DRS-R-98) is used where a more comprehensive assessment of both the presence and severity of delirium is required. It is noted this instrument is not appropriate for use in community settings.

**Table 8 Summary of Ratings for Delirium Instruments**

Criteria	Weight	Instrument	
		CAM	DRS-R-98
Availability of comparison data	3	3	2
Length/feasibility of instrument for inclusion in battery	2	3	2
Complexity of administration/ cognitive burden	2	3	2
Cultural Appropriateness	1	3	2
Ease of obtaining score	2	3	3
Sensitivity to dementia	3	2	2
Reliability evidence	3	3	3
Validity evidence	3	3	3
Cost of the instrument	2	2	2
Cost of instrument administration	2	2	2
<b>Weighted Total</b>		<b>62</b>	<b>54</b>

#### 1.4.6.3 Recommendations Concerning Measures of Particular Symptoms of BPSD

In many cases the use of Global BPSD measures such as the NPI may suffice for the assessment of the associated symptoms of dementia.

- It is recommended that the following instruments are used if a more detailed assessment of a particular symptom is required:

Aggression: Rating Scale for Aggressive Behaviour in the Elderly (RAGE)  
 Agitation: Cohen Mansfield Agitation Inventory (CMAI); Pittsburgh Agitation Scale (PAS)  
 Anxiety: Rating Anxiety in Dementia (RAID)  
 Apathy: Apathy Evaluation Scale (AES)  
 Depression: Cornell Scale for Depression in Dementia (CSDD)  
 Geriatric Depression Scale (GDS Yesavage) - less severe cases and in community settings)

A full discussion of these measures and their assessment can be found in Section 9 and Appendix 10 of this report.

### 1.4.7 Measures of Function

The Functional Independence Measure (FIM), the Barthel Index and the Lawton and Brody IADL and the Older Americans Resources and Services (OARS-IADL) instruments were chosen as generic measures of ADL and IADL respectively. These instruments have been reviewed elsewhere recently (Eagar, et al. 2001; Eager, et al. 2006; Thomas, et al. 2006), have good psychometric properties and have been used in geriatric settings.

With regard to the activities of daily living, the FIM is probably more appropriate for acute care and high level residential care settings but it is noted that accredited training is required for its use. However, it is already widely used in acute care rehabilitation settings within Australia. The Barthel Index is an easier to use measure and may be more appropriate for use in primary and community care settings with people with mild to moderate forms of dementia. Although the Katz ADL instrument has been quite widely used in dementia settings the review of this instrument by Thomas, et al. (2006) indicated it had weak psychometric properties and thus it is not recommended for use (refer Table 9 below).

**Table 9 Summary of Ratings for the Generic Measurement of Function Instruments**

Criteria	Weight	ADL Instruments			IADL Instruments	
		FIM	Barthel Index	Katz	OARS-IADL	Lawton & Brody IADL
Availability of comparison data	3	3	2	1	3	2
Length/feasibility of instrument for inclusion in battery	2	2	3	3	3	3
Complexity of administration/cognitive burden	2	3	3	2	3	2
Cultural Appropriateness	1	3	3	3	3	3
Ease of obtaining score	2	3	3	3	3	3
Sensitivity to dementia	3	1	1	2	2	2
Reliability evidence	3	3	2	1	3	2
Validity evidence	3	3	2	1	2	2
Cost of the instrument	2	2	3	3	3	3
Cost of instrument administration	2	1	1	1	1	1
<b>Weighted Total</b>		<b>55</b>	<b>50</b>	<b>42</b>	<b>59</b>	<b>51</b>

- The Functional Independence Measure (FIM) and the Barthel Index are recommended as the generic measures of ADL.
- The Lawton and Brody IADL and the Older Americans Resources and Services (OARS-IADL) are recommended as generic instruments for the assessment of instrumental activities of daily living (IADL). The OARS-ADL is preferred as it is an advance on the Lawton and Brody IADL scale with improved psychometric properties and less reliance on gender role stereotypes;

and it has been adapted for use in primary and community care settings in Australia (see Green, et al. 2006).

The recommended dementia specific instruments for the assessment of function (ADL and IADL) for people with dementia include both proxy measures and clinical rating scales. While it is acknowledged that proxy reports have their limitations (refer Section 12.3), they will generally be used where assessment by interview or self rating is no longer possible due to the degree of cognitive impairment of the person with dementia. Proxy measures are also useful in primary and community care settings in order to monitor the maintenance of functional status or its decline, in conjunction with drug therapy or in terms of care management as the disease progresses. The direct observation rating scale may be more appropriate for acute care and residential care settings. By recommending both proxy and direct observation rating scales different practice settings and clinical situations (e.g. a person with dementia may not have a carer) can be addressed (refer Table 10).

- The Alzheimer's Disease Co-operative Study – ADL (ADCS-ADL) and Disability Assessment for Dementia Scale (DAD) are the two proxy report instruments that are recommended.
- For the direct observation of functioning the Cleveland Scale for Activities of Daily Living (CSADL) is recommended.

The discussion of measures of functional status in Section 10 highlights a number of measurement problems with regard to the assessment of function in people with dementia. It is clear there is an urgent need for a program of research and development in this area. It is recommended that:

- In the absence of a research consensus for the measurement of function in dementia, and given a high degree of overlap in items, there is a clear need for a streamlining of the various functional instruments and items across each of the practice settings (Spector, 1997). The work of Lindeboom, et al. (2003) in the Amsterdam Liner Disability Score Project using IRT to calibrate ADL instruments in neurology could be used as a guide. A similar study with a large group of people with dementia could examine and calibrate functional items from the short-listed instruments (both generic and dementia specific) to create a comprehensive item bank. This dementia item bank could then be used to examine item redundancy and coverage across the range of severity levels and could be used to develop new tools or provide cross-calibration between the existing instruments. This project would also need to examine the relationship of these items with recommended cognitive and functional assessment staging instruments.

**Table 10 Summary of Ratings for the Dementia Specific Measurement of Function Instruments**

Criteria	Weight	Instruments		
		DAD	ADCS-ADL	CS-ADL
Availability of comparison data	3	3	3	2
Length/feasibility of instrument for inclusion in battery	2	1	2	2
Complexity of administration/ cognitive burden	2	3	3	3
Cultural Appropriateness	1	3	3	3
Ease of obtaining score	2	3	2	3
Sensitivity to dementia	3	3	3	3
Reliability evidence available	3	3	2	2
Validity evidence available	3	2	2	2
Cost of the instrument	2	3	3	3
Cost of instrument administration	2	2	2	2
<b>Weighted Total</b>		<b>60</b>	<b>57</b>	<b>56</b>

## 1.4.8 Measures of Patient and Carer Satisfaction

### 1.4.8.1 Patient Satisfaction

The patient satisfaction literature was recently reviewed by Hawthorne (2006). Theories of patient satisfaction suggest instruments should cover 7 areas:

- Access to health services and the treatment environment;
- Provision of health information;
- The relationship with care providers;
- Participation in making health care choices;
- The technical quality of care;
- Treatment effectiveness (helping the daily life of the patient); and
- General satisfaction.

Patient dissatisfaction occurs where there are multiple transgressions or where there is a catastrophic failure in one area.

Following an examination of the literature the patient satisfaction instruments selected for review were:

- Single item assessments;
- The Client Satisfaction Questionnaire (CSQ-18 and CSQ-8);
- The Consultation Satisfaction Questionnaire (CSQ, described here as the ConsultSQ);
- The La Monica-Oberst patient satisfaction scale (LOPSS);
- The Linder-Pelz satisfaction scales;
- The Medical Interview Satisfaction Scale (MISS);
- The Patient Satisfaction Index (PSI);
- The Patient Satisfaction Questionnaire (PSQ);
- The Patient Visit Rating Questionnaire (PVRQ);
- The Patient Satisfaction Questionnaire of Gonzalez et al. (2005);
- Inpatient Evaluation of Service Questionnaire (IESO); and
- The Short Assessment of Patient Satisfaction instrument (SAPS).

Based on the criteria for measuring patient satisfaction (Section 11.3) and the reviews of instruments in sections 11.4 and 11.5, it was possible to compare the instruments reviewed. This is shown below in Table 11.

**Table 11 Summary Assessing Patient Satisfaction Instruments Against the Study Criteria**

Criteria	Weight	Instruments												
		SAPS	Consult-SQ	PVRQ	LOPPS-18	Single item	CSQ-8	CSQ-18	PSI	MISS-21	IESQ	Linder-Pelz	PSQ-III	Gonzalez
Availability of comparison data	3	3	2	1	1	1	2	2	2	1	1	1	2	1
Length/feasibility of instrument for inclusion in battery	2	3	2	3	2	3	1	2	2	2	2	2	1	1
Complexity of administration/cognitive burden	2	3	2	3	2	3	3	2	2	2	2	2	1	1
Cultural Appropriateness	1	3	2	3	2	3	2	1	2	1	1	2	1	1
Ease of obtaining score	2	3	2	3	2	3	3	2	2	2	2	2	1	1
Sensitivity to dementia	3	1	1	1	1	1	1	1	1	1	1	1	1	1
Reliability evidence	3	2	3	1	3	1	2	2	2	2	2	1	2	2
Validity evidence	3	2	3	2	2	1	2	2	1	1	1	1	2	1
Cost of the instrument	2	3	3	3	3	3	1	1	3	3	3	3	3	3
Cost of instrument administration	2	3	3	3	3	3	3	3	2	2	2	2	1	1
<b>Weighted Total</b>		<b>57</b>	<b>53</b>	<b>48</b>	<b>47</b>	<b>45</b>	<b>45</b>	<b>42</b>	<b>42</b>	<b>38</b>	<b>38</b>	<b>36</b>	<b>36</b>	<b>30</b>

The two standout instruments were the SAPS and the ConsultSQ. None of the other patient satisfaction instruments reviewed could be considered truly satisfactory.

Hawthorne, et al. (2006) previously compared the attributes of three of these leading generic measures of patient satisfaction (CSQ, Consult SQ, PSI), and one continence specific questionnaire (GUTTS), in a clinical study which examined patient satisfaction with treatment for incontinence. He concluded that all these instruments had a relatively poor coverage of the different aspects of patient satisfaction and there was evidence of response bias and poor responsiveness in most instruments. The items from all these instruments were then pooled to analyse their psychometric properties and Mokken and IRT analyses were used to construct the short generic measure of patient satisfaction that would provide the best fit to the theoretical model of patient satisfaction outlined above (Hawthorne, et al. 2006). The Short Assessment of Patient Satisfaction (SAPS) scale, a generic measure of patient satisfaction, was derived from this study. The SAPS contains only seven items (one for each dimension of patient satisfaction) and was more sensitive than any other instrument to the pooled satisfaction indicator. It also had excellent internal consistency with a Cronbach's alpha of 0.86. Although SAPS needs to be further tested in other samples and populations (e.g. including dementia patients and possibly dementia carers) it is recommended as a generic measure for the assessment of patient satisfaction pending further research.

Six generic items measuring global satisfaction were also identified from the instruments and analysed concerning their appropriateness as a single item measure for immediate assessment of patient satisfaction (Hawthorne, et al. 2006). Two of these items had better psychometric properties and were less prone to differential item functioning. These items were a) how satisfied are you with the outcome of your treatment? and b) how satisfied are you with the amount of help received? Item a) was chosen as the single item for satisfaction with incontinence treatment (Hawthorne, et al. 2006) given its' better psychometric properties. However, this item may be less appropriate for dementia settings where often general care services are provided rather than specific treatment interventions per se. Item b) would seem more appropriate in this regard, however, it had disordered response thresholds which may relate to oddities in the response set utilized for this question by its' authors. It may be useful to also retest this single item, with modified response categories, in a further study.

Given the above considerations the following recommendations are made:

- It is recommended that a study be undertaken to test the SAPS and the two single patient satisfaction items identified above with samples of people with dementia and their carers. It is noted that all these items would also require minor rewording to make them suitable for use with an informant/carer.
- That a single item patient satisfaction measure should be adopted for use in Australian settings by clinicians wishing to assess the satisfaction of their patients 'on the spot'. Strategies should be put in place to encourage clinicians to adopt this measure as a common metric across Australia. Encouragement should be given to specialists and researchers to also include this common metric in their work. In this way a bank of shared understanding will be progressively established. It may be possible that a single item measure could be drawn from the generic instruments recommended above, or from those that Hawthorne, et al. (2006) examined as single items for the National Continence Management Strategy.
- That the SAPS and ConsultSQ are validated in dementia-populations. These were the two standout generic patient satisfaction instruments identified in this report. For the reasons outlined in the report, however, neither can be recommended outright because there is no evidence of their reliability, validity or responsiveness in dementia populations. It is recommended that a head-to-head validation study be undertaken in dementia populations.
- Until the recommended research is implemented and the results published, it is recommended that the SAPS be used.

### 1.4.8.2 Carer Satisfaction with Services

Carer satisfaction is addressed by the literature in a number of ways. There are studies that examine: a) carer experience with the caring role (including carer burden); b) carer satisfaction with services and c) carer health and well-being. This project focuses on the examination of carer satisfaction with services and specifically excludes an examination of instruments used to assess carer burden.

With regard to the area of carer satisfaction with services, there are studies which focus on family and carer concerns relating to the satisfaction with quality care availability, physical and psychosocial care and information giving (Hare, et al. 2006; Kristjanson, 1989, 1993). On the other hand, carer satisfaction has also been defined as an evaluative procedure for quality assurance, marketing and health care planning (Buttle, 1996; Parasuraman, et al. 1988).

In general, care quality assurance is discussed in the literature in negative terms, viz., poor facilities or infrastructure, physical abuse of the patient, his/her psychological abuse, physical and psychological neglect and exploitation (Schulz and Williamson, 1997), whereas care satisfaction is usually asked in more neutral terms, focussing on the extent to which the carer is satisfied with the care of the care recipient. This difference in perspective may well explain differences in reported assessment levels between quality assurance and carer satisfaction (Soliman, 1992). These two perspectives imply that although the assessment of the quality of caring provided by a health service provider and carer satisfaction are different constructs which should not be confused or conflated, quality of caring cannot be adequately assessed without some consideration of both – especially when a care recipient moves from being cared for at home to being cared for in an institution, or where studies compare home care with institutional care (Kessler, et al. 2005). It is a matter of emphasis as to which perspective is of greater interest to carers, clinicians, researchers and policy makers.

This review covered the first of these two perspectives (carer satisfaction) for three reasons. First, in dementia care the primary concern of a carer is that his/her care recipient is well taken care of by community-based health care clinicians, service personnel or teams where necessary, or within institutional care. Second, there is gross market failure in the Australian health care system generally, and particularly in the dementia care sector: most Australians are not fully informed consumers and most Australians do not have the opportunity to make meaningful choices regarding available services for the care of their loved ones. Third, assessments of quality assurance are a function of service provider characteristics and carer expectations and information; areas that most carers have little experience of when they begin caregiving, with the implication that immature or uninformed assessments regarding quality assurance can be easily made (Buttle, 1996; Chesterman, et al. 2001; Soliman, 1992).

Given the above considerations the following scales or items were selected for reviewed against the study criteria (refer Table 12). They are (in alphabetical order):

- The Carer Satisfaction Questionnaire;
  - The Carer Satisfaction with Community Services Questionnaire;
  - The Carer Satisfaction Survey;
  - The Consumer Expectations Perceptions and Satisfaction Scale (CEPAS);
  - The FAMCARE (Family Satisfaction with Advanced Cancer Care) scale; and
  - The Satisfaction with Care at the End of Life in Dementia Scale (SWC-EOLD).
- In addition, single item assessments were reviewed.

Table 12 below provides the comparative summary scores of the instruments that assess carer satisfaction with services.

**Table 12 Summary Assessing Carer Satisfaction Instruments Against the Study Criteria**

Criteria	Weight	Instrument					
		SWC	CSS	FAMC	CSCS	CEPAS	CSQ
Availability of comparison data	3	1	1	1	1	1	1
Length/feasibility of instrument for inclusion in battery	2	3	3	2	2	2	1
Complexity of administration/ cognitive burden	2	3	2	2	2	1	1
Cultural Appropriateness	1	1	1	1	1	1	1
Ease of obtaining score	2	3	3	2	2	2	1
Sensitivity to dementia	3	3	1	1	1	1	1
Reliability evidence	3	2	2	2	2	2	1
Validity evidence	3	1	1	2	1	1	1
Cost of the instrument	2	3	3	3	3	3	3
Cost of instrument administration	2	3	3	3	3	3	1
<b>Weighted Total</b>		<b>52</b>	<b>44</b>	<b>43</b>	<b>40</b>	<b>38</b>	<b>27</b>

Given the findings of the review in Section 11.2 and Table 12, none of the reviewed instruments can be given an unqualified recommendation for use in Australian studies with carers of older adults who have cognitive impairment or dementia. The following recommendations are made:

- The most promising instrument appears to be the SWC-EOLD (Volicer, et al. 2001), and it is recommended that this instrument is used in an Australian study specifically designed to test its measurement properties.
- The alternative would be to mount a specific carer satisfaction study, where all items from all reviewed instruments were pooled and tested. The explicit purpose would be identifying well performing items and/or the best performing instrument.

A brief discussion is provided in Section 11.3 concerning carer satisfaction with services and its relationship to the related domains of carer burden and carer wellbeing. Carer satisfaction with services has been addressed in this project but an examination of carer burden, carer appraisal and carer wellbeing was outside the scope of this project. Although a number of recent studies (Brodaty, et al. 2002; Ramsay, et al. 2006) have examined issues relating to carer burden, a comparison of the leading instruments used to assess carer burden is yet to be undertaken.

- It is recommended that a more detailed follow up project be undertaken to examine issues relating to the assessment of instruments used to assess carer burden, carer appraisal and carer wellbeing.

## 1.5 Measurement Issues

Some key measurement issues relevant to the use of these measures with people with dementia and their carers are outlined. The first of these is the issue of the use of proxies (formal and informal carers) for the assessment of the person with dementia. People with severe dementia may not be able to be assessed directly and may be unable to provide a self report where this may be required. This is followed by a discussion of the level of cognitive impairment at which people with dementia may lose the capacity to self rate. Importantly, many carers of those with dementia may suffer mild cognitive impairment themselves. These issues are most important to consider when assessing subjective phenomena from both care recipients and carers, such as health related quality of life, social isolation or satisfaction with services.

The applicability of these measures for particular population groups is also discussed. The issue of the applicability of the measures for those from Culturally and Linguistically Diverse (CALD) populations is considered, as is the applicability of these measures for use with Aboriginal and Torres Strait Islander people. The recommendations pertaining to these issues are outlined below.

### 1.5.1 Recommendations Concerning Cognitive Impairment and the Capacity to Self Rate

Section 12.2 provides a more detailed discussion of this issue. Where it is possible and feasible subjective phenomena should be assessed by patient self report rather than by proxy report. Sometimes this is not possible with people with severe dementia and thus the following recommendations are made:

- An interim recommendation (awaiting the results of further recommended research) is that self rating report (by non interview administration) should not be considered for patients with MMSE scores below 15.
- For patients with MMSE scores ranging from 10-15 an interview administration or an interview assisted administration of these self-report measures could be considered.
- For patients with an MMSE score less than 10 it is suggested that data be collected via proxy reporting. Where a specific proxy form has been developed this should be utilised.
- It is recommended that a study be undertaken to assess the recommended self report tools by self report administration, interview administration and assisted interview administration to identify the best approach for assessing the HRQOL and other subjective phenomena of people with dementia with more severe cognitive impairments.
- As the capacity for cognitively impaired patients to self rate will depend on the structure, length, design and complexity of each questionnaire it is suggested that a follow up study be undertaken to assess the MMSE-3MS scores that are required for the recommended self report questionnaires under different modes of administration.

### 1.5.2 Recommendations Concerning Proxy Assessment

Section 12.3 provides a discussion of the issues concerning the use of proxy assessment where direct assessment of the person with dementia is not possible. Below are a number of recommendations when using proxy measures:

- Proxy reports should be examined for three potential biases: (1) the cognitive status of the proxy (as many elderly people are cared for by an elderly spouse carer, who may themselves be impaired or unwell, but to a lesser degree); (2) the health status of the proxy; and (3) the level of carer burden and stress (Harper, 2000).

- There is usually a trade-off between those “with the greatest amount of contact and those with more training” (Harper, 2000, page 488). However, generally, where a proxy report is used information should be collected from the family member/carer or care staff member that is closest to the patient and has the greatest degree of interaction with the patient.
- Proxy reports should be based on usual behaviour rather than extreme or rare behaviours (Harper, 2000).
- Proxy reports should be based on observable phenomena like physical symptoms and functioning, rather than subjective phenomena like depression, social isolation and quality of life (Snow, 2005a).

### **1.5.3 Recommendations Concerning Assessment with Culturally and Linguistically Diverse (CALD) Populations**

Section 12.4 provides a discussion of the issues concerning the use of instruments with those from Culturally and Linguistically Diverse (CALD) backgrounds. The following recommendations are made:

- Use of the DOMS selected tools can be interpreted with less confidence if used by practitioners and interpreters who are not culturally competent. For an outline of the application of culturally competent assessment see the guidelines proposed by Alzheimer’s Australia – National Cross Cultural Dementia Network (Grypma, Mahajani and Tam, 2007).
- A further project is necessary to ensure a more comprehensive database intended for dementia outcome measures solely for use with CALD communities - where translated versions of the DOMS selected measures are further reviewed and made available if possible.
- Further studies analysing the measurement equivalence of the core recommended measures (e.g. GDS, NPI, MMSE-3MS) should be undertaken for major language groups within Australia.
- Research to further examine instruments developed in Australia such as the RUDAS and the GPCOG is supported to ensure their validity and reliability in different groups of CALD populations.

### **1.5.4 Recommendations Concerning Assessment with Aboriginal and Torres Strait Islander Populations**

A more detailed discussion of these issues can be found in Section 12.5 and Appendices 14 and 15 also provide useful supplementary material.

Many of the recommended scales may have applications among urban Indigenous people but this needs to be ascertained.

- It is recommended that some focus groups in urban settings are developed to discuss how appropriate the recommended scales are to members of these communities.

There is very limited application for these tools for people from remote Aboriginal communities. A notable exception to this is the Kimberley Indigenous Cognitive Assessment tool (LoGiudice, et al. 2006) which is a new tool that has been designed for use with Indigenous people in remote locations.

- There is an interim recommendation, pending further research, that the KICA is used to assess the cognitive status of rural and remote indigenous peoples rather than the MMSE-3MS.

Clinician ratings may have more application than the self report or the proxy administered forms, as some of the ratings can be made through observation, rather than attempting to elicit answers from the patient. Cognitive assessment will be extremely difficult in many remote settings, and especially if the patient speaks and understands limited English. Clinical assessments may be improved if other confounding factors are removed, such as unfamiliarity of the clinician and environment. A clinician, who is familiar to the individual and has a good knowledge of their life, may be in a position to make a more informed judgment. While it may be possible to use some of the simpler tools in a remote setting, especially with modifications to pictorially demonstrate concepts such as volume, questions will still remain about what the answers that individuals supply actually mean.

- It is recommended that there needs to be further detailed research on the meaning of dementia in Indigenous communities, and how to ask questions which capture the experience of living with dementia in an Indigenous community.
- It is recommended that a project be undertaken to examine the modifications that may need to be undertaken to the recommended tools to make them more appropriate to Indigenous peoples.
- It is recommended that further research be undertaken to assess the psychometric properties of the KICA-Cog and its' appropriateness for the assessment of cognitive impairment with both urban and remote Indigenous people.

## **1.6 Implementation Issues**

Although issues pertaining to implementation have been discussed throughout this report and particularly in Section 12.6 a number of key areas to address are identified. These are:

- The issue of mandating the recommended measures
- The application of the instruments in different settings and for different stages of dementia
- Training issues
- A dissemination strategy
- Identified research gaps

With regard to a discussion of the issue of mandating the recommended measures the reader is referred to Section 12.6.2 of this report. Advice received from the Department of Health and Ageing in August 2006 indicated there was no desire to mandate the recommended instruments at this stage. The project team was advised that mandating was not a consideration at this time as the Dementia Outcomes Measurement Suite was a first-stage project to assess key gaps and tools. It was noted that the Dementia CRCs and Study Centres may promote the use of particular tools agreed as a result of the DOMS-NEP project; however, this would be as best practice, rather than to mandate.

Given the use of the measurement tools is to be recommended rather than mandated and more comprehensive assessment produces an increased burden on staff, there may need to be some consideration by the Department of Health and Ageing of the provision of financial incentives for services that adopt the use of the recommended tools.

It would be difficult to mandate the use of the recommended measures without full consideration of the training requirements and the burden on staff time for all service settings to implement these measures. If routine data collection and analysis is desired, with a view to benchmarking the outcomes of similar services, then careful thought must be given to the design of such systems and the phased implementation of such an approach. This should include a consideration of information technology requirements and cost and resource implications. To adopt such an approach will require

a considerable financial investment by the Department of Health and Ageing as has occurred with mental health services.

Section 12.6.3 provides a discussion of the appropriate application of each of the recommended measures for different service settings and for different stages of dementia. Readers are also referred to Figure 6 within Section 12. This is supplemented by a discussion of a staged approach to assessment in Section 12.6.4.

A dissemination strategy, to facilitate the adoption of the recommended tools has been outlined in Section 12.6.6. This could include the development of an instrument toolkit, presentations at conferences, training workshops (managers, service providers, clinical and care staff) the development of web materials, brochures, training videos, and journal articles.

- It is recommended that a dissemination strategy project be undertaken to facilitate the dissemination and uptake of findings from this report.

Notwithstanding the above, the provision of more formal education and training will also be of paramount importance.

- It is recommended that a project be sponsored to a) ascertain coverage of assessment and the use of recommended tools in current curricula and b) to develop appropriate education modules for insertion in the training curricula of relevant professional and paraprofessional groups.

Throughout the course of this project a large number of research gaps have been identified. These are outlined in Section 12.6.7. These research gaps include such issues as the need for Australian reference data for some of the instruments, the need to streamline instruments in order to remove redundancy (especially in the area of functioning), the modification of some recommended tools for CALD and Aboriginal and Torres Strait Islander Groups, and the need for research concerning proxy assessment and the level of cognitive capacity required to self rate/report.

## **1.7 Conclusion**

While further research may need to be undertaken to clarify some assessment issues the report provides a useful review of the best measures to assess the status and symptoms of people with dementia. The project has identified a set of recommended measures/tools for routine use in the assessment, diagnosis, screening and outcomes monitoring of dementia conditions and the evaluation of treatments that are applicable for the Australian health care context. By developing this set of recommended measures it is hoped to standardise the assessment and evaluation procedures used in this field to enhance comparability of findings across research and practice settings.

## **References**

Bowling A (2001) *Measuring disease (2nd ed.)* Open University Press, Buckingham England.

Bowling A (2005) *Measuring health: a review of quality of life measurement scales (3rd ed.)* Open University Press, Berkshire, England.

Broadhead W (1988) *Duke-UNC functional social support questionnaire (DUFSS): instructions for use.* Duke University Medical Centre. Report nr 919-286-9896.

Brodaty H, Green A, Banerjee S, Mittelman M, Schulz R, Whitehouse P, Harvey R, Powell M, Prince M, Rios D, Zarit S (2002) *Towards harmonisation of Caregiver Outcome Measures.* *Brain Aging*. Vol. 2, No. 4, pp. 3-12.

Burns A, Lawlor B and Craig S (2004) *Assessment scales in old age psychiatry (2nd ed.)* Taylor and Francis, London.

- Buttle F (1996) *SERVQUAL - review, critique, research agenda*. European Journal of Marketing. Vol. 30, No.1, pp.8-32.
- Chesterman J, Bauld L and Judge K (2001) *Satisfaction with the care-managed support of older people: an empirical analysis*. Health and Social Care in the Community. Vol. 9, No.1, pp.31-42.
- de Jong Gierveld J and Kamphuis F (1985) *The development of a Rasch-type loneliness scale*. Applied Psychological Measurement. Vol. 9, No.3, pp.289-299.
- de Jong Gierveld J and Tilburg TV (2006) *A 6-item scale for overall, emotional and social loneliness: confirmatory tests on survey data*. Research on Aging. Vol. 28, No.5, pp.582-598.
- Eagar K, Owen A, Cromwell D, et al. (2001) *Towards a national measure of functional dependency for home and community care services in Australia - Stage 1 report of the HACC Dependency Data Items Project*. Centre for Health Service Development, University of Wollongong.
- Eagar K, Owen A, Marosszeky N, et al. (2006) *Towards a measure of function for home and community care services in Australia: Part 1 - Development of a Standard National Approach*. Australian Journal of Primary Health. Vol. 12, No.1, pp.73-81.
- Gonzalez N, Quintana JM, Bilbao A, et al. (2005) *Development and validation of an in-patient satisfaction questionnaire*. International Journal for Quality in Health Care. Vol. 17, No.6, pp.465-72.
- Green J, Eagar K, Owen A, et al. (2006) *Towards a measure of function for home and community care services in Australia: Part 2 - Evaluation of the Screening Tool and Assessment Instruments*. Australian Journal of Primary Health. Vol. 12, No.1, pp.82-90.
- Grypma R, Mahajani S and Tam E (2007) *Screening and diagnostic assessment of non-english speaking people with dementia: Guidelines and system recommendations for practitioners, service managers and policy makers*. Alzheimer's Australia.
- Hare R, Rogers H, Lester H, et al. (2006) *What do stroke patients and their carers want from community services?* Family Practice. Vol. 23, No.1, pp.131-6.
- Harper GJ (2000) *Assessing older adults who cannot communicate*. In Kane RL, Kane RA, (eds). Assessing older persons: Measures, meaning, and practical applications. Oxford University Press, Oxford.
- Hawthorne G (2006) *Measuring social isolation in older adults: development and initial validation of the Friendship Scale*. Social Indicators Research. Vol. 77, No.3, pp.521-548.
- Hawthorne G, Sansoni J, Hayes L, et al. (2006) *Measuring Patient Satisfaction with Incontinence Treatment*. Department of Psychiatry, The University of Melbourne, and the Centre for Health Service Development, University of Wollongong for the Commonwealth Department of Health and Ageing.
- Hughes ME, Waite LJ, Hawkey LC, et al. (2004) *A short scale for measuring loneliness in large surveys*. Research on Aging. Vol. 26, No.6, pp.655-672.
- Kane RL and Kane RA (2000) *Assessing older persons: measures, meaning and practical applications*. Oxford University Press, USA.
- Kessler D, Peters TJ, Lee L, et al. (2005) *Social class and access to specialist palliative care services*. Palliative Medicine. Vol. 19, No.2, pp.105-10.
- Kristjanson LJ (1989) *Quality of terminal care: salient indicators identified by families*. Journal of Palliative Care. Vol. 5, No.1, pp.21-30.
- Kristjanson LJ (1993) *Validity and reliability testing of the FAMCARE Scale: measuring family satisfaction with advanced cancer care*. Social Science and Medicine. Vol. 36, No.5, pp.693-701.
- Lezak MD (2004) *Neuropsychological Assessment* (4<sup>th</sup> ed). Oxford University Press, New York.

- Lindeboom R, Vermeulen M, Holman R, et al. (2003) *Activities of daily living instruments: Optimizing scales for neurologic assessments*. Neurology. Vol. 60, pp.738-742.
- Lincoln Centre for Ageing and Community Care Research (2004) *The review and identification of an existing, validated, comprehensive assessment tool*. Australian Institute for Primary Care, La Trobe University.
- LoGiudice D, Smith K, Thomas J, et al. (2005) *Kimberley Indigenous Cognitive Assessment tool (KICA): Development of a cognitive assessment tool for older indigenous Australians*. International Psychogeriatrics. Vol. 18, No.2, pp.269-280.
- McDowell I (2006) *Measuring Health: A Guide to Rating Scales and Questionnaires (3rd ed.)* Oxford University Press, Oxford.
- McKeith I, Cummings J, Lovestone S, Harvey R and Wilkinson D (1999) *Outcome measures in Alzheimer's disease*. Martin Dunitz, London.
- McKeith IG, Dickson DW, Lowe J, Emre M, O'Brien JT, Feldman H, et al. (2005) *Diagnosis and management of dementia with Lewy bodies: third report of the DLB Consortium. [erratum appears in Neurology. 2005 Dec 27; 65(12):1992] [summary for patients in Neurology. 2005 Dec 27; 65(12):E26-7; PMID: 16380603]*. Neurology. Vol. 65, No.12, pp.1863-1872.
- Norbeck JS (1984) *The Norbeck Social Support Questionnaire*. Birth Defects Original Article Series. Vol. 20, No.5, pp.45-57.
- Norbeck JS, Lindsey AM and Carrieri VL (1981) *The development of an instrument to measure social support*. Nursing Research. Vol. 30, No.5, pp.264-9.
- Norbeck JS, Lindsey AM and Carrieri VL (1983) *Further development of the Norbeck Social Support Questionnaire: normative data and validity testing*. Nursing Research. Vol. 32, No.1, pp.4-9.
- Novella JL, Jochum C, Ankri J, et al. (2001) *Measuring general health status in dementia: practical and methodological issues in using the SF-36*. Journal of Aging-Clinical and Experimental Research. Vol. 13, pp 362-9.
- Parasuraman A, Zeithaml VA and Berry L (1988) *SERVQUAL: a multiple-item scale for measuring consumer perceptions of service quality*. Journal of Retailing. Vol. 64, No.1, pp.12-37.
- Rabins PV and Black BS (2007) *Measuring quality of life in dementia: purposes, goals, challenges and progress. Consensus conference paper*, International Psychogeriatrics.
- Ramsay L, Owen A, Eagar K, Marosszeky N, Samsa P, Rix M, Fildes D, Willmott L (2006) *An Australian Carer Needs Assessment Instrument: The Recommended Approach*. Centre for Health Service Development, University of Wollongong.
- Rockwood K (2007) *The measuring, meaning and importance of activities of daily living (ADLs) as an outcome*. International Psychogeriatrics. Vol. 19, No.3, pp.467-482.
- Roman GC, Tatemichi TK, Erkinjuntti T, Cummings JL, Masdeu JC, Garcia JH, et al. (1993) *Vascular dementia: diagnostic criteria for research studies. Report of the NINDS-AIREN International Workshop. [see comment]*. Neurology. Vol. 43, No.2, pp.250-260.
- Russell D, Peplau L and Cutrona C (1980) *The revised UCLA Loneliness Scale: concurrent and discriminant validity evidence*. Journal of Personality and Social Psychology. Vol. 39, No.3, pp.472-480.
- Russell D, Peplau LA and Ferguson ML (1978) *Developing a measure of loneliness*. Journal of Personality Assessment. Vol. 42, No.3, pp.290-4.
- Russell DW (1996) *UCLA Loneliness Scale (Version 3): reliability, validity, and factor structure*. Journal of Personality Assessment. Vol. 66, No.1, pp.20-40.
- Sarason I, Sarason B, Shearin E, et al. (1987) *A brief measure of social support: practical and theoretical implications*. Journal of Social and Personal Relationships. Vol. 4, pp.497-510.

- Sarason IG, Levine H, Basham RB, et al. (1983) *Assessing social support: the Social Support Questionnaire*. Journal of Personality and Social Psychology. Vol. 44, No.1, pp.127-139.
- Schulz R and Williamson GM (1997) *The measurement of caregiver outcomes in Alzheimer disease research*. Alzheimer Disease and Associated Disorders. Vol. 11, Suppl 6, pp.117-24.
- Sherbourne C and Stewart A (1991) *The MOS social support survey*. Social Science and Medicine. Vol. 32, pp.705-714.
- Snow AL, Cook KF, Lin P-S, et al. (2005a) *Proxies and other external raters: Methodological considerations*. Health Services Research. Vol. 40, No.5, pp.1676-1693.
- Soliman AA (1992) *Assessing the quality of health care: a consumerist approach*. Health Marketing Quarterly. Vol. 10, No.1-2, pp.121-41.
- Spector WD (1997) *Measuring functioning in daily activities for persons with dementia*. Alzheimer Disease and Associated Disorders. Vol. 11, No.Suppl 6, pp.81-90.
- Thomas S, Nay R, Fonda D et al. (2006) *Continence Outcomes Measurement Suite Project: Final Report*, Australian Government Department of Health and Ageing, Canberra.
- Volicer L, Hurley AC and Blasi ZV (2001) *Scales for evaluation of end-of-life care in dementia*. Alzheimer Disease and Associated Disorders. Vol. 15, No.4, pp.194-200.

## 2 Introduction

The purpose of this project is to develop a set of recommended measures/tools for routine use in the assessment, diagnosis, screening and outcomes monitoring of dementia conditions and the evaluation of treatments that are applicable for the Australian health care context. By developing a set of recommended measures it is hoped to standardise the assessment and evaluation procedures used in this field to enhance comparability of findings across research and practice settings.

The project commenced in late April, a work plan was submitted to the Department of Health and Ageing in May, and the First Project Report was submitted in July. A briefing on project progress was provided to the Department of Health and Ageing in Adelaide on 27 November 2006.

The First Project Report covered the issues of initial project implementation, the establishment of the National Expert Panel (DOMS-NEP) and the Expert Measurement Group (DOMS-EMG) and the initial considerations of these groups. A draft chapter on the standardization of clinical terminology was presented for feedback and consideration. It also outlined literature search strategies and addressed issues of project scope.

Following feedback from the DOMS-NEP and DOMS-EMG and discussion with the Department of Health and Ageing it was decided that a more detailed review should be undertaken of measures addressing the associated symptoms of dementia (e.g. global measures of behavioural and psychological symptoms of Dementia (BPSD), depression, apathy, agitation etc.) than had initially been identified in the tender application. An outline concerning this project extension was submitted to the Department of Health and Ageing in August 2006 and an extension of the contract to include this work was ratified in October 2006. The inclusion of this project component necessitated a review of the overall project timelines with a revised completion date of 30 September 2007.

### **2.1 Revised Project Timelines and Reporting Requirements**

The revised timelines for this project are outlined below:

- First Progress Report: end of June 2006 (completed)
- Project Briefing: end of November 2006 (completed)
- Second Progress Report: end of January 2007 (completed)
- Draft Final Report: July 2007 (completed)
- Final EMG and NEP Meetings for Project Ratification: August 2007 (completed)
- Final Report: end of September 2007 (completed)

### **2.2 An Outline of the Second Report**

The project team submitted the Second Report in early February 2007. The completed reviews of the following categories of measures were contained within this report:

- Dementia Staging and Descriptive Measures (Section 4)
- Dementia Specific Quality of Life Measures (Section 5)
- Cognitive Assessment Measures (Section 6)
- Multi-Attribute Utility Measures (Section 7)

The section on the standardization of clinical terminology was revised and was included in Section 3 of the Second Report.

With regard to issues arising it should be noted that the Dementia Staging and Descriptive Measures tend to be somewhat global as they usually include a mix of both cognitive and behavioural symptoms and are often also used to assess the severity of Dementia and associated conditions. Thus it was noted there will sometimes be overlap between measures considered in this category and measures of cognition and/or associated symptoms. If a contender instrument, for example the Alzheimer Disease Assessment Scale (ADAS) seemed to have a primary focus on cognition rather than on the general assessment of Dementia it would be considered in the Cognitive Measures category and so forth. It was also found there were a number of batteries of instruments that were identified in this class, for example the Consortium to Establish a Registry for Alzheimer's disease (CERAD). These batteries usually include a well known measure for the general assessment of dementia (e.g. the CDR) and so components of the batteries are included in the reviews for the relevant category of instrument assessment.

Following a scoping exercise undertaken by Prof Chenoweth concerning the cognitive measures a decision was made by the DOMS-EMG that the project should focus on the instruments/tools that are available for use in routine care and this would exclude many of the more detailed neuropsychological instruments or instruments that require specialist training for their administration and interpretation. Feedback on this issue has also been obtained from other clinicians associated with the project and the DOMS-NEP. It was thought that a follow up project could undertake a more detailed assessment of the neuropsychological instruments to determine recommendations for this specialty.

### **2.3 The Final Report**

In this phase of the project (February – September 2007) the following categories of measures have been reviewed:

- Generic Quality of Life Measures (now included in Section 5)
- Measures of Social Function and Social Support (Section 8)
- Associated Symptom Measures (Section 9)
- Measures of Functional Status (Section 10)
- Measures of Patient and Carer Satisfaction (Section 11)

This Final Report also contains a section on Measurement and Implementation Issues (Section 12) which includes:

- An assessment of the recommended measures concerning their appropriateness for use with CALD and Aboriginal and Torres Strait Islander Groups
- Discussion and recommendations concerning some key measurement issues (e.g. proxy reporting)
- Recommendations concerning implementation issues

The Conclusions and Recommendations (Section 13) of this Final Report summarises the recommendations and also identifies gaps where further research work may be required.

The Draft Final Report was forwarded to DOMS-NEP and DOMS-EMG members for feedback and ratification in August 2007 prior to the Final Report being submitted to the Australian Government Department of Health and Ageing by the end of September 2007.

### **2.4 Meetings of the Dementia Outcomes Measurement Suite National Expert Panel and the Expert Measurement Group**

Three meetings were held during the course of the project and all reports were sent to the DOMS-NEP for feedback. The last meeting of the DOMS-NEP was held on the 17<sup>th</sup> August 2007 and the terms of reference and the current membership of the National Expert Panel can be found in Appendix 1 of this report.

Four major meetings were held during the course of the project and a working party meeting was also held in August 2006. A meeting of the DOMS-EMG, which included additional representation from the DOMS-NEP, was held on the 8<sup>th</sup> June 2007. This meeting discussed the Associated Symptoms Section for the Draft Final Report and the related instrument reviews. A final meeting was held to review the Draft Final Report on the 17<sup>th</sup> August 2007. The terms of reference and the current membership of EMG will be found in Appendix 2 of this report.

## **2.5 An Overview of the Literature Search and Instrument Review Processes**

An initial overall literature search was undertaken (MEDLINE, PsycINFO) on twenty key terms (e.g. dementia, cognition, memory, function, Quality of Life etc). The major texts in the field were examined which included psychometric texts containing instrument reviews (e.g. McDowell, 2006; Bowling, 2001, 2005) as well as those containing instrument reviews applicable for Dementia and assessment of the elderly (e.g. Burns, 2004; Kane and Kane, 2000; Lezak, 2004; McKeith, 1999). This process identified a list of instrument names and then searches were undertaken on all measures identified.

A database was then developed which provided comparative data for instruments for each domain / category (Associated Symptoms, Cognitive, Comprehensive, Dementia Staging and Description, Function, Health Related Quality of Life, Miscellaneous, Neuropsychological, Satisfaction, Social and Utility Measures). This database included 844 named instruments. A CD-Rom was developed for each domain / category of instruments (e.g. dementia staging and description, cognition, health related quality of life) containing relevant papers and abstracts for each of the review teams.

An impact sheet was then developed for consideration by the review teams and the DOMS-EMG. This considered MEDLINE, text and web impacts, presence in instrument databases (e.g. PROQOLID) and its use in clinical practice. The latter was based on NEP and field surveys and clinical feedback. This process usually identified the leading twelve or so instruments for consideration in each category.

Further selection criteria were then applied to reduce this to the leading 5-6 instruments in each domain / category. The additional criteria were:

- Whether there is a copy of the instrument and the original article concerning its development available for review.
- The number of citations found. In the case of new instruments some care was taken to assess this criterion as it was considered that recently developed instruments may not have a high citation rate. However, for instruments developed more than 5 years previously a low citation rate might indicate limited adoption by the field.
- The amount and range of the published psychometric evidence.
- Whether the instrument is used in clinical practice (evidence from the literature and data from NEP and other surveys).
- The availability of normative and clinical reference data.
- Administration time (generally 30 minutes or less) where a shorter administration time would be preferred. It was noted that as a number of instruments assessing different aspects (e.g. symptoms, cognition, HRQOL) will need to be utilized, lengthy instruments that may be more appropriate for detailed follow-up assessment may not be appropriate for use in routine assessment and across the range of practice settings.
- Whether the instrument is applicable for people with varying levels of dementia severity. Generally, preference would be given to measures applicable across the range of severity levels. However; consideration could be given to an instrument that is particularly applicable to one level of severity which is not addressed well by the other selected measures in that category. For example, a self report measure may only be applicable to people with dementia of mild severity and some measures of behavioural and psychological disturbance may only be applicable to those with moderate or severe dementia.

- Proprietary considerations (e.g. prohibitive cost).
- Applicability for use in routine care. Instruments would be preferred if they did not require specialist skills for administration or if extensive training in their use was not required (e.g. as for many neuropsychological/medical assessments).

Once the shortlist of contender instruments had been reduced to 5-6 measures for each category then a decision summary sheet was developed justifying the selection or non-selection of contender instruments. Further searches were then undertaken for the selected instruments using other databases (e.g. CINAHL, Cochrane Library etc) and the comprehensive reviews of these instruments commenced.

All instrument reviews make use of the AHOC review sheet and contain the following information:

- Author, publication information, availability
- Cost
- Training requirements
- Purpose of the instrument and who it was developed for
- Administration time
- Structure of the instrument
- Scoring
- Applications and availability of normative and clinical reference data
- Carer/Patient use of the instrument
- Psychometric criteria – reliability, validity, responsiveness
- Cultural applicability and cultural adaptations
- Gender and age appropriateness

With all instruments consideration was also given to the following aspects:

- Type and stages of dementia
- Purpose of the instrument (assessment, screening, outcomes monitoring and the evaluation of interventions)
- Self-reporting and proxy reporting
- Respondent and staff burden
- Appropriateness for CALD and Aboriginal and Torres Strait Islander groups
- Appropriateness for a range of settings (e.g. community and residential care)

Once the comprehensive review is completed an Instrument Scoring and Weighting Sheet is completed for each instrument as indicated in Table 13 below.

**Table 13 Table of Criteria and Weights for Instrument Ranking**
**Criteria and weights used to assess instruments (DOMS)\***
**Instrument Name ..... Total Score = .....**

Evaluation Criteria	Scoring system	Score	Weight	Weighted Score
Availability of comparison data	1 = minimal or no comparison data available		3	
	2 = some international comparison data available			
	3 = Australian and international dementia comparison data available including normative data and clinical reference data			
Length/feasibility of instrument for inclusion in battery	1 = long instrument, 30+ items		2	
	2 = medium length instrument, 15-30 items			
	3 = short instrument, less than 15 items			
Complexity of administration (for clinician use); and cognitive burden (for self report or proxy instruments)	1 = demanding to understand or administer		2	
	2 = some difficulties to understand or administer			
	3 = easy to understand and administer			
Cultural Appropriateness (ease of use with an interpreter, client literacy, CALD criteria including Indigenous Australians)	1 = not appropriate for use by CALD or illiterate clients, or with an interpreter		1	
	2 = limited appropriateness for use by CALD or illiterate clients and interpreters			
	3 = appropriate for use by CALD or illiterate clients and interpreters			
Ease of obtaining score by the evaluator	1 = scoring complex and requires computer		2	
	2 = can be scored without computer but time consuming			
	3 = scoring easy and does not require computer			
Sensitivity to dementia	1 = not known to be sensitive to dementia status,		3	
	2 = sensitive to dementia status			
	3 = good sensitivity to dementia status			
Reliability evidence available	1 = no or little published evidence identified		3	
	2 = evidence suggests moderate reliability			
	3 = evidence suggests good reliability			
Validity evidence available	1 = no published validity evidence identified		3	
	2 = evidence suggests moderate validity			
	3 = evidence suggests good validity			
Cost of the instrument	1 = costs charged for using instrument		2	
	2 = costs for commercial use/training costs			
	3 = instrument available free of charges			
Cost of instrument administration	1 = professional		2	
	2 = paraprofessional/ staff member			
	3 = self complete			

The instrument is given a score against each criterion and this is multiplied by the weight for this criterion. The resulting weighted score for each criterion is then added to form a total score for each instrument (refer Table 13). For each category of instruments a comparative table of scores for the instruments is then produced and it is on this basis the recommendations for each category of instruments are formed.

Sections 4 -11 provide summaries and recommendations for the instrument categories reviewed. These include:

- Dementia Staging and Descriptive Measures (refer Section 4 and Appendix 5)
- Health Related Quality of Life Measures (refer Section 5 and Appendix 6)
- Cognitive Assessment Measures (refer Section 6 and Appendix 7)
- Multi-Attribute Utility Measures (refer Section 7 and Appendix 8)
- Measures of Social Isolation (refer Section 8 and Appendix 9)
- Measures of the Associated Symptoms of Dementia (refer Section 9 and Appendix 10)
- Measures of Function (refer Section 10 and Appendix 11)
- Measures of Patient and Carer Satisfaction (refer Section 11 and Appendix 12)

## References

Bowling A (2001) *Measuring disease* (2nd ed.) Open University Press, Buckingham England.

Bowling A (2005) *Measuring health: a review of quality of life measurement scales* (3rd ed.) Open University Press, Berkshire, England.

Burns A, Lawlor B and Craig S (2004) *Assessment scales in old age psychiatry* (2nd ed.) Taylor and Francis, London.

Kane RL and Kane RA (2000) *Assessing older persons: measures, meaning and practical applications*. Oxford University Press, USA.

Lezak MD (2004) *Neuropsychological Assessment* (4<sup>th</sup> ed). Oxford University Press, New York.

Lindeboom R, Vermeulen M, Holman R, et al. (2003) Activities of daily living instruments: Optimizing scales for neurologic assessments. *Neurology*. Vol. 60, pp.738-742.

McDowell I (2006) *Measuring Health: A Guide to Rating Scales and Questionnaires* (3rd ed.) Oxford University Press, Oxford.

McKeith I, Cummings J, Lovestone S, Harvey R and Wilkinson D (1999) *Outcome measures in Alzheimer's disease*. Martin Dunitz, London.

## 3 The Standardization of Clinical Terminology

### 3.1 Background Issues

While the number of different types of dementia is large, the term dementia is commonly used in two different ways. The first is as a collective term, which suggests that dementia is one clinical entity, an acquired global impairment of higher cortical functioning. The second is as a variety of conditions with cognitive features, including Alzheimer's, frontotemporal, diffuse Lewy Body, Vascular and subcortical dementias. Dementia can present in a variety of ways. Thus, a flexible approach to instruments chosen to assist with screening, diagnosis and monitoring is necessary, especially as assessments need to be made in a variety of health care contexts. At the same time, consistency in measurement is also important because measurement presupposes definition. Since diagnosis, assessment of symptom-severity and on-going monitoring are essential for health service delivery and planning, employing commonly accepted definitions of clinical terminology is important. Inaccurate or misdiagnosis, and misuse of standard terms in relation to dementia and associated symptoms is more likely to occur when there is a lack of knowledge about the cognitive characteristics and the psychosocial manifestations of dementia and other conditions that mimic dementia. For example, making an accurate diagnosis of dementia can be confounded when the person is experiencing an episode of delirium, or is simultaneously depressed, unless the clinician is well-versed in the relationship, rate of progression and presenting signs and symptoms of all three conditions.

Clinical terminology and associated classification systems are the basis for identifying and addressing service need, and at the present time variation exists in the amount and type of information collected in different national data sets. A number of these data sets are currently in use across the care continuum, and rely on single or more data items to identify people with dementia and cognitive impairment. These data sets include:

- CACP (Community Aged Care Packages), EACH (Extended Aged Care in the Home) or Dementia EACH collections
- NRCP (National Respite for Carers Program)
- ALSWH (Australian Longitudinal Study on Women's Health)
- Residential Aged Care data set and MDSv2 (Minimum Data Set version 2) for HACC (Home and Community Care services)
- ACFI (Aged Care Funding Instrument)
- SDAC (Survey of Disability, Ageing and Carers) – ABS data
- PBS (Pharmaceutical Benefits Scheme)
- NHMD (National Hospital Morbidity Database)
- ACAP (Aged Care Assessment Program)
- BEACH (Bettering the Evaluation And Care of Health) Program
- DESP (Dementia Education and Support Program) data set

Given the scope of these data sets, it is vital to promote the use of a standard classification system to ensure consistency in terminology across the health continuum and within health and social care systems. This Chapter provides a review of the literature on terminologies describing various types of dementia and severity/stages of dementia. Recommendations are made based on the literature review and consultations with clinical and research experts in the field of dementia care/services. However, it should be noted that, current usages of the terminologies in research and clinical practice may differ from the recommendations developed in this report depending on specialities, disciplines and contexts within which practitioners and researchers work.

## 3.2 Detailed Examination of Health Classification Systems

This section will examine the way terms related to dementia are used in the literature has been described and compared with the international definitions mainly from the International Classification of Diseases (ICD-10) (World Health Organisations, WHO, 1992), the Diagnostic and Statistical Manual for Mental Disorders, 4th revision (DSM-IV) (American Psychiatric Association, APA, 2000), and the International Classification of Functioning, Disability and Health (ICF) (WHO, 2002). This Section also includes a review of other diagnostic criteria developed specifically for Alzheimer's disease, vascular dementia, dementia with Lewy bodies and frontotemporal dementia.

### 3.2.1 Definitions and Diagnostic Criteria for Dementia<sup>1</sup>

Dementia is predominantly caused by a group of chronic, neurodegenerative conditions, which lead to progressive cognitive and functional impairment, and which are often accompanied by mood and behavioural disturbances as well as psychotic features. More than 80 different underlying aetiologies of dementia are identified in the main diagnostic criteria references. For the purpose of the Dementia Outcome Measurement Suite (DOMS) project this report will focus on the four most common types of dementia including vascular dementia (VaD), dementia with Lewy bodies (LBD) and frontotemporal dementia (FTD) as well as dementia of the Alzheimer's type (AD). These four types, or combinations thereof, account for over 50-75% of all dementia conditions (APA, 2000; First and Tasman, 2004; Grabowski and Damasio, 2004). A recent report prepared by the Australian Institute of Health and Welfare (AIHW, 2007) confirms these four, or the combination of AD and VaD, as the most commonly occurring dementia conditions in Australia. Other well recognised types of dementia in Australia include dementias related to Parkinson's disease, alcohol, drug ingestion, head injury, Huntington's disease, human immunodeficiency virus (HIV), Creutzfeldt-Jakob disease (CJD) and, less commonly, reversible forms of dementia caused by Vitamin B12 deficiency and hypothyroidism (AIHW, 2007).

Dementia may be classified and diagnosed using either the International Statistical Classification of Diseases and Related Health Problems, Tenth revision (ICD-10) (WHO, 1992) or the Diagnostic and Statistical Manual for Mental Disorders, 4th revision (DSM-IV)<sup>2</sup> (APA, 1994), both of which are subject to continuous reviews and revisions<sup>3</sup>. In conjunction with these, particularly in North America, clinicians and researchers utilise criteria-based definitions contained in the National Institute of Neurological and Communicative Disorders and Stroke-AD and Related Disorders Association (NINCDS-ADRDA) Work Group, which was designed to be compatible with the DSM-III and ICD criteria (McKhan, et al. 1984). See Table 14 for the comparisons of the definitions of dementia using different classification systems. The NINCDS-ADRDA criteria for Alzheimer's disease require biopsy or autopsy to satisfy the criteria of 'DEFINITE'<sup>4</sup>. For the clinical diagnosis of VaD, the State of California Alzheimer's Disease Diagnostic and Treatment Centres (ADDDTC) criteria (Chui, et al. 1992)<sup>5</sup>, and an operational version of the National Institute of Neurologic Disorders and Stroke and the Association Internationale pour la Recherche et l'Enseignement en Neurosciences (NINDS-AIREN) criteria (Roman, et al. 1993)<sup>6</sup> have also been deployed widely. Neither ICD-10 nor DSM-IV-TR provides specific diagnostic criteria for LBD or FTD, although they include dementias in Parkinson's disease and Pick's disease respectively. Two other diagnostic criteria recommended in this regard are the Manchester-Lund criteria for frontotemporal lobar degeneration (FLD) (The Lund and Manchester Groups, 1994), which has been revised and updated twice since by Neary, et al.

1 Refer to the AIHW report (2006) for more detailed review of literature on definitions and diagnosis of dementia.

2 DSM-IV-TR: In 2000 the APA revised the text of DSM-IV to amend errors identified in the DSM-IV text, to ensure up-to-date knowledge and include new research information that had been developed since the literature review for DSM-IV was conducted in 1992. No substantial change was made for diagnostic criteria in general in the DSM-IV-TR, however new diagnostic codes for dementia conditions (except VaD) were developed (294.10/294.11). (APA 2000)

3 The review and revision process of the current DSM-V and the ICD-10 is due to be completed by 2011, resulting in the DSM-V and ICD-11.

4 See Appendix 4A Criteria for PROBABLE, POSSIBLE, and DEFINITE AD

5 See Appendix 4B Criteria for PROBABLE, POSSIBLE, and DEFINITE VaD

6 See Appendix 4C Criteria for PROBABLE, POSSIBLE, and DEFINITE VaD

(1998) and McKhan, et al. (2001)<sup>7</sup>, and the Consortium on Dementia with Lewy Bodies criteria for LBD (McKeith, et al. 1996 and 2005)<sup>8</sup>. Despite the pursuit of the international standardisation of the diagnostic criteria through the ICD, the most commonly used criteria for the diagnoses of dementia, both in the clinical practice and research arena, are based on the DSM-IV or its earlier versions.

---

7 See Appendix 4E  
8 See Appendix 4D

**Table 14 Definitions and Diagnostic Features of Dementia**

ICD-10 / ICD-10-AM (Australian Modification <sup>9</sup> )	DSM-IV / DSM-IV-TR	DSM-III-R	NINCDS-ADRDA
<p>A syndrome due to disease of the brain, usually of a chronic or progressive nature, in which there is disturbance of multiple higher cortical functions, including memory, thinking, orientation, comprehension, calculation, learning capacity, language, and judgement. Consciousness is not clouded. The impairments of cognitive function are commonly accompanied, and occasionally preceded, by deterioration in emotional control, social behaviour, or motivation. This syndrome occurs in Alzheimer's disease, in cerebrovascular disease, and in other conditions primarily or secondarily affecting the brain. The primary requirement for diagnosis is evidence of a decline in both memory and thinking which is sufficient to impair personal activities of daily living. The impairment of memory typically affects the registration, storage, and retrieval of new information, but previously learned and familiar material may also be lost, particularly in the later stages. Dementia is more than dysmnnesia: there is also impairment of thinking and of reasoning capacity, and a reduction in the flow of ideas. The processing of incoming information is impaired, in that the individual finds it increasingly difficult to attend to more than one stimulus at one time, such as taking part in a conversation with several persons, and to shift the focus of attention from one topic to another. If dementia is the sole diagnosis, evidence of clear consciousness is required. However, a double diagnosis of delirium superimposed upon dementia is common (F05.1). The above symptoms and impairments should have been evident for at least 6 months for a confident clinical diagnosis of dementia to be made.</p>	<p>The disorders in the Dementia section are characterised by the development of multiple cognitive deficits (including memory impairment) that are due to the direct physiological effects of a general medical condition, to the persisting effects of a substance, or to multiple aetiologies (e.g. the combined effects of cerebrovascular disease and Alzheimer's disease). Essential to the diagnosis of dementia is the presence of multiple cognitive deficits that include memory impairment and at least one of the following abnormalities of cognition: aphasia, apraxia, agnosia, or a disturbance in executive functioning. The cognitive deficits must be sufficiently severe to cause impairment in occupational or social functioning and must represent a decline from a previously higher level of functioning. A diagnosis of a dementia should not be made if the cognitive deficits occur exclusively during the course of delirium. However, a dementia and a delirium may both be diagnosed if the dementia is present at times when the delirium is not present. Dementia may be aetiologically related to a general medical condition, to the persisting effects of substance use (including toxin exposure), or to a combination of these factors.</p>	<p>The essential feature of Dementia is impairment in short- and long-term memory, associated with impairment in abstract thinking, impaired judgment, other disturbances of higher cortical function, or personality change. The disturbance is severe enough to interfere significantly with work or usual social activities or relationships with others. The diagnosis of Dementia is not made if these symptoms occur in Delirium.</p>	<p>Dementia is the decline of memory and other cognitive functions in comparison with the person with dementia's previous level of function as determined by a history of decline in performance and by abnormalities noted from clinical examination and neuropsychological tests. A diagnosis of dementia cannot be made when consciousness is impaired by delirium, drowsiness, stupor, or coma or when other clinical abnormalities prevent adequate evaluation of mental status. Dementia is a diagnosis based on behaviour and cannot be determined by computerised tomography, electroencephalography, or other laboratory instrument, although specific causes of dementia may be identified by these means.</p>

<sup>9</sup> The ICD-10-AM Mental Health Manual (National Centre for Classification in Health, NCCH, 2002a) is designed for community-based mental health services, to assist clinicians code mental illness in a simpler and easier manner than the complete ICD-10-AM

As shown in Table 14 the ICD-10, DSM-IV and DSM-IV-TR coding systems have a great deal in common. In fact, there were various attempts to produce the terminologies and codes within the DSM-IV that are fully compatible with those of the ICD-10 through consultations and collaboration with the WHO (APA, 2000). However, the full compatibility of the two systems is fairly limited due to inconsistency of the diagnostic criteria/guidelines between them. The main differences between the two classifications (ICD-10 and DSM-IV) in terms of the diagnostic criteria for dementia and its subtypes include:

- The ICD-10\* requires a minimum duration of disturbance for six months while the DSM-IV does not.
- For dementia, in addition to memory loss the ICD-10\* requires cognitive deficits be limited to a deterioration in judgement and thinking and a deterioration in “emotional control or motivation or a change in social behaviour” while the DSM-IV requires any one of aphasia, apraxia, agnosia, or disturbance in executive functioning.
- The ICD-10\*<sup>10</sup> requires a “relatively rapid onset and progression” and a characteristic type of cognitive impairment (e.g. aphasia) for early onset AD, whereas late onset cases have a very slow and gradual onset with a predominance of memory impairment over other intellectual deficits. In the ICD-10 this disorder is referred to as Dementia in Alzheimer’s disease.
- For VaD, the ICD-10 specifies that the deficits in higher cognitive functions are unevenly distributed and that there be both clinical and laboratory evidence of focal brain damage. The ICD-10 sub-types vascular dementia into acute onset, multi-infarct, subcortical, and mixed subcortical and cortical.

(First and Tasman, 2004, p. 306)

---

10 \* ICD-10 Diagnostic Criteria for Research

**Table 15 Comparisons of the APA and WHO Classifications of Dementia**

ICD-10	DSM-IV	DSM-IV-TR
<p><u>Organic, including symptomatic, mental disorders (F00-F09)</u></p> <p>F00 Dementia in Alzheimer's disease</p> <p>F00.0 Dementia in Alzheimer's Disease with Early Onset*</p> <p>F00.1 Dementia in Alzheimer's Disease with Late Onset</p> <p>F00.2 Dementia in Alzheimer's Disease, Atypical or Mixed type</p> <p>F00.9 Dementia in Alzheimer's Disease, Unspecified</p> <p>F01 Vascular dementia</p> <p>F01.0 Vascular Dementia of Acute Onset</p> <p>F01.1 Multi-Infarct Dementia</p> <p>F01.2 Subcortical Vascular Dementia</p> <p>F01.3 Mixed Cortical and Subcortical Vascular Dementia</p> <p>F01.8 Other Vascular Dementia</p> <p>F01.9 Vascular Dementia, Unspecified</p> <p>F02 Dementia in Other Diseases Classified Elsewhere</p> <p>F02.0 Dementia in Pick's disease</p> <p>F02.1 Dementia in Creutzfeldt-Jakob Disease</p> <p>F02.2 Dementia in Huntington's Disease</p> <p>F02.3 Dementia in Parkinson's Disease</p> <p>F02.4 Dementia in Human Immunodeficiency Virus [HIV] Disease</p> <p>F02.8 Dementia in Other Specified Diseases Classified Elsewhere</p> <p>Dementia in: cerebral lipidosis; epilepsy; hepatolenticular degeneration; hypercalcemia; hypothyroidism, acquired; intoxications; multiple sclerosis; neurosyphilis; niacin deficiency [pellagra]; polyarteritis nodosa; systemic lupus erythematosus; trypanosomiasis; vitamin B<sub>12</sub> deficiency</p> <p>F03 Unspecified Dementia</p> <p>Presenile: Dementia NOS*; Psychosis NOS</p> <p>Primary Degenerative Dementia NOS</p> <p>Senile: Dementia: NOS; Depressed or Paranoid type</p> <p>· Psychosis NOS</p> <p>NOS: Not otherwise specified (or unspecified)</p> <p>*Alzheimer's Australia recommends the use of 'Younger onset dementia' or 'Younger people with dementia' instead of 'early onset' to avoid confusion with early stage dementia. However, for the purpose of this project it is important the terminology remains the same as the ICD-10.</p>	<p><u>Delirium, Dementia, and Amnesic and Other Cognitive Disorders</u></p> <p>290.xx Dementia of the Alzheimer's Type, With Early Onset*</p> <p>290.10 Uncomplicated</p> <p>290.11 With Delirium</p> <p>290.12 With Early Onset, With Delusions</p> <p>290.13 With Depressed Mood</p> <p>290.xx Dementia of the Alzheimer's Type, With Late Onset</p> <p>290.00 Uncomplicated</p> <p>290.20 Delusions</p> <p>290.21 With Depressed Mood</p> <p>290.3 With Delirium</p> <p>290.xx Vascular Dementia</p> <p>290.40 Uncomplicated</p> <p>290.41 With Delirium</p> <p>290.42 With Delusions</p> <p>290.43 With Depressed Mood</p> <p>294.9 Dementia Due to Human Immunodeficiency Virus (HIV) Disease</p> <p>294.1 Dementia Due to Head Trauma</p> <p>294.1 Dementia Due to Parkinson's Disease</p> <p>294.1 Dementia Due to Huntington's Disease</p> <p>290.10 Dementia Due to Pick's Disease</p> <p>290.10 Dementia Due to Creutzfeldt-Jacob disease</p> <p>294.1x Dementia Due to [Indicate the General Medical Condition not listed above]</p> <p>----- Substance-Induced Persisting Dementia</p> <p>----- Dementia Due to Multiple Aetiologies</p> <p>(code each of the specific aetiologies)</p> <p>294.8 Dementia NOS</p>	<p>294.xx Dementia of the Alzheimer's Type, With Early Onset*</p> <p>294.10 Without Behavioural Disturbance</p> <p>294.11 With Behavioural Disturbance</p> <p>294.xx Dementia of the Alzheimer's Type, With Late Onset</p> <p>294.10 Without Behavioural Disturbance</p> <p>294.11 With Behavioural Disturbance</p> <p>290.xx Vascular Dementia</p> <p>290.40 Uncomplicated</p> <p>290.41 With Delirium</p> <p>290.42 With Delusions</p> <p>290.43 With Depressed Mood</p> <p><i>Specify if: With Behavioural Disturbance</i></p> <p><i>Code presence or absence of a behavioural disturbance in the fifth digit for Dementia due to a General Medical Condition:</i></p> <p>0 = Without Behavioural Disturbance</p> <p>1 = With Behavioural Disturbance</p> <p>294.1x Dementia Due to Human Immunodeficiency Virus (HIV) Disease</p> <p>294.1x Dementia Due to Head Trauma</p> <p>294.1x Dementia Due to Parkinson's Disease</p> <p>294.1x Dementia Due to Huntington's Disease</p> <p>290.1x Dementia Due to Pick's Disease</p> <p>290.1x Dementia Due to Creutzfeldt-Jacob disease</p> <p>294.1x Dementia Due to [Indicate the General Medical Condition not listed above]</p> <p>----- Substance-Induced Persisting Dementia</p> <p>----- Dementia Due to Multiple Aetiologies (code each of the specific aetiologies)</p> <p>294.8 Dementia NOS</p>

Application of multiple diagnostic criteria for dementia and specific types of dementia results in conflicting estimates of the prevalence and incidence of dementia as shown in the following studies that examined the degree of agreement in detecting dementia using different classification systems. Comparisons of the diagnostic criteria of dementia between the DSM 3<sup>rd</sup> edition, revised (DSM-III-R) (APA, 1987), the DSM-IV, the ICD-10 and the Cambridge Examination for Mental Disorders of the Elderly (CAMDEX) showed that the DSM-III-R, the DSM-IV and the CAMDEX systems gave prevalence rates of dementia of 47%, 41.6% and 38.2% respectively, while the ICD-10 criteria classified 29.4% in the study population aged 90 years and older as demented (Pioggiosi, et al. 2004). Similar results were found in the large population-based Canadian Study of Health and Aging which included people 65 years of age or older. The prevalence of dementia was 17.3% when the DSM-III-R criteria were used, 13.7% with the DSM-IV, 4.9% with the CAMDEX and 3.1% with the ICD-10 (Erkinjuntti, et al. 1997). Both studies found that the ICD-10 was more restrictive in its requirements for the diagnosis of dementia and less sensitive in detecting early onset dementia than other formulations. Earlier studies by Henderson, et al. (1994) and Fichter, et al. (1995) also showed somewhat similar results in that the DSM-III-R produced almost two times higher detection rates of dementia than the ICD-10 based diagnostic criteria. In addition, the use of the different classification systems resulted in different groups of people as having dementia. Major discriminating factors between the DSM-IV and the ICD-10 were long-term memory, executive function, presence or absence of aphasia, social activities and duration of symptoms (Erkinjuntti, et al. 1997).

Similarly, the comparison of different diagnostic classification systems for VaD such as DSM-III, DSM-IV, ICD-10, ADDTC and NINDS-AIREN showed significantly varying degrees of detection rates of VaD and poor concordance between the classification systems (Gold, et al. 2002; Pohjasvaara, et al. 2000; Wetterling, et al. 1996). According to Pohjasvaara, et al. (2000), factors that most often account for disagreement between the criteria for VaD relate to requirement of focal neurological signs, unequal distribution of deficits in higher cortical functions, and evidence of relevant cerebrovascular lesions based on brain imaging findings. Gold, et al. (2002) examined the sensitivity and specificity of the four major clinical criteria for VaD (i.e. ADDTC, NINDS-AIREN, DSM-IV and ICD-10) against the neuropathological diagnosis of people with dementia whose autopsy had confirmed the basis for their dementia (N=89). The study concluded that the ADDTC criteria for possible VaD were the most sensitive for the detection of VaD while the DSM-IV criteria for VaD and the NINDS-AIREN criteria for possible VaD were more effective in excluding mixed dementia. The ICD-10 criteria for VaD, along with the ADDTC and the NINDS-AIREN criteria for probable VaD, was deemed inadequate requiring further revisions (Gold, et al. 2002). What is clear is that the available clinical diagnostic criteria for VaD are not compatible with each other, and produce different frequencies and groups of people with dementia (Chui, et al. 2000; Gold, et al. 2002; Lopez, et al. 2005; Pohjasvaara, et al. 2000; Wetterling, et al. 1996). It is worth noting that the difficulty of accurately providing a diagnosis of VaD alone is largely due to the problems of differentiating between a diagnosis of VaD and mixed (AD plus VaD) dementia and the inconsistencies between a clinical history of stroke and brain imaging findings (Lopez, et al. 2005).

With regard to the use of these classification systems in current practice a survey was conducted among registrars, psychiatrists and psychogeriatricians at the Faculty of Old Age Psychiatry meeting held in September 2006. Of 41 respondents 54% indicated the DSM system as a preferred option while 16% preferred the ICD system (30% had no preference). However, the main reason for their preference for the DSM system was because psychiatrists were usually trained to use this particular classification system and hence were more familiar with it. Personnel from other specialties (e.g. general practice) are unlikely to have received training or be familiar with the DSM system. Only one survey respondent stated they considered that DSM-IV was more valid and reliable.

The use of valid and standardised diagnostic criteria is critical not only for epidemiological studies but also for prevention, early intervention and treatment of dementia conditions, and funding

allocations for relevant health care services. Having or not having an appropriate diagnosis of dementia can make a significant difference in the individuals' quality of life and relationships with their families and friends. Given that the definite diagnosis of dementia conditions is possible only through obtaining histopathologic evidence, it appears that there is insufficient evidence to determine which classification system is most valid and reliable in the diagnosis of dementia and its subtypes. Further longitudinal, prospective studies of clinico-pathological correlation of dementia criteria and dementia cases are needed. The AIHW report (2007) recommends the ICD system<sup>11</sup> in conjunction with the ICF for the standardisation of definitions and classifications of dementia and its outcomes, as it is used in the classification of mortality and morbidity in hospitals in Australia and forms the basis of health condition codes used in the Aged Care Assessment Program, the National Health Survey and the Survey of Disability, Ageing and Carers. Similarly, other countries including the United States have adopted the ICD system (e.g. ICD-9-CM, clinical modification) for the official coding system for reporting mortality and morbidity as well as for Medicare reimbursements. However, it is questionable as to how this type of data can be translated in the local and international community in comparing epidemiological studies and clinical trials of dementia where different diagnostic criteria have been used. It is worth noting that unlike DSM-IV, the ICD-10 does not include the person's lack of ability to participate in social and occupational activities in the diagnostic criteria of dementia (Pioggiosi, 2004). Furthermore, the ICD-10 has a separate clinical guideline and research criteria for mental health<sup>12</sup>, which may result in limiting the generalisability of research findings to clinical practice (First and Tasman, 2004). See Table 16 for the consensus guidelines proposed by the American Academy of Neurology (Knopman, et al. 2001) and the Scottish Intercollegiate Guidelines Network (SIGN) (2006), of which recommendations were made based on the combination of the evidence from the literature and expert consensus.

---

11 See Appendix 4F for the definitions, classification and diagnostic guidelines.

12 The ICD-10-AM (Australian Modification), however, is based on the clinical version of the ICD.

**Table 16 Examples of Clinical/Practice Guidelines for Diagnosis of Dementia (with levels of evidence classification)**

<p>Practice Parameter: Diagnosis of Dementia (an evidence-based review) by the American Academy of Neurology (AAN) (Knopman, 2001)<sup>13</sup></p>	<p>Management of People With Dementia: A national clinical guideline by the Scottish Intercollegiate Guidelines Network (SIGN) (2006)</p>
<p>1) The DSM-III-R definition of dementia is reliable and should be used (Guideline).</p> <p>2) The National Institute of Neurologic, Communicative Disorders and Stroke-AD and Related Disorders Association (NINCDS-ADRDA) or the Diagnostic and Statistical Manual, 3rd edition, revised (DSM-IIIIR) diagnostic criteria for AD and clinical criteria for Creutzfeldt–Jakob disease (CJD) have sufficient reliability and validity and should be used (Guideline). Diagnostic criteria for vascular dementia, dementia with Lewy bodies, and frontotemporal dementia may be of use in clinical practice (Option) but have imperfect reliability and validity.</p> <ul style="list-style-type: none"> <li>• The Hachinski Ischemic Scale (HIS)<sup>14</sup> criteria may be of use in the diagnosis of cerebrovascular disease in dementia (Option).</li> <li>• The Consortium for DLB diagnostic criteria may be of use in clinical practice (Option).</li> <li>• The Consensus diagnostic criteria for FTD may be of use in clinical practice (Option).</li> </ul> <p>3) Structural neuroimaging with either a noncontrast CT or MR scan in the initial evaluation of people with dementia is appropriate. Because of insufficient data on validity, no other imaging procedure is recommended (Guideline). There are currently no genetic markers recommended for routine diagnostic purposes (Guideline). The CSF 14-3-3 protein is useful for confirming or rejecting the diagnosis of CJD (Guideline).</p> <p>4) Screening for depression, B<sub>12</sub> deficiency, and hypothyroidism should be performed (Guideline). Screening for syphilis in people with dementia is not justified unless clinical suspicion for neurosyphilis is present (Guideline).</p>	<p>1 History Taking and differential Diagnosis B DSM-IV or NINCDS-ADRDA criteria should be used for the diagnosis of Alzheimer's disease. B The Hachinski Ischaemic Scale or NINDS-AIRENS criteria may be used to assist in the diagnosis of vascular dementia. C Diagnostic criteria for dementia with Lewy bodies and fronto-temporal dementia should be considered in clinical assessment.</p> <p>2 Initial Cognitive Testing B In individuals with suspected cognitive impairment, the MMSE should be used in the diagnosis of dementia. § Initial cognitive testing can be improved by the use of Addenbrooke's Cognitive Examination. § A questionnaire, such as the IQCODE<sup>15</sup>, completed by a relative or friend may be used in the diagnosis of dementia.</p> <p>3 Screening for comorbid conditions § Physical investigations including laboratory tests should be selected on clinical grounds according to history and clinical circumstances. B As part of the assessment for suspected dementia, the presence of comorbid depression should be considered.</p> <p>4 The use of imaging C Structural imaging should ideally form part of the diagnostic workup of people with suspected dementia. C SPECT may be used in combination with CT to aid the differential diagnosis of dementia when the diagnosis is in doubt.</p> <p>5 The role of cerebrospinal fluid and electroencephalography B CSF and EEG examinations are not recommended as routine investigations for dementia. § CSF and EEG examinations may be useful where CJD is suspected.</p> <p>6 neuropsychological testing</p>

<sup>13</sup> This information is current as of March 22, 2006.

<sup>14</sup> See Appendix 4H for the HIS criteria.

<sup>15</sup> The Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) is a short questionnaire filled out by someone who knows the patient and can be an adjunct to direct cognitive testing. See Appendix 4I for the questionnaire.

	<p>B Neuropsychological testing should be used in the diagnosis of dementia, especially in persons where dementia is not clinically obvious.                  § It may be useful to repeat neuropsychological testing after six to 12 months in persons where:</p> <ul style="list-style-type: none"> <li>• the diagnosis is unclear.</li> <li>• measurement of the progression of deficits in a typical pattern supports a diagnosis of dementia and helps in differential diagnosis.</li> </ul>
<p><b>DEFINITIONS FOR PRACTICE RECOMMENDATIONS BASED ON CLASSIFICATION OF EVIDENCE (Knopman, 2001)</b>  <u>Standard:</u> Principle for patient management that reflects a high degree of clinical certainty (usually this requires Class I evidence that directly addresses the clinical question, or overwhelming Class II evidence when circumstances preclude randomised clinical trials).  <u>Guideline:</u> Recommendation for patient management that reflects moderate clinical certainty (usually this requires Class II evidence or a strong consensus of Class III evidence).  <u>Practice Option:</u> Strategy for patient management for which the clinical utility is uncertain (inconclusive or conflicting evidence or opinion).  <u>Practice Advisory:</u> Practice recommendation for emerging and/or newly approved therapies or technologies based on evidence from at least one Class I study. The evidence may demonstrate only a modest statistical effect or limited (partial) clinical response, or significant cost-benefit questions may exist. Substantial (or potential) disagreement among practitioners or between payers and practitioners may exist.</p> <p><b>CLASSIFICATION OF EVIDENCE</b>  <u>I</u> Evidence provided by a well designed prospective study in a broad spectrum of persons with the suspected condition, using a “gold standard” for case definition, in which test is applied in a blinded evaluation, and enabling the assessment of appropriate tests of diagnostic accuracy.  <u>II</u> Evidence provided by a well designed prospective study of a narrow spectrum of persons with the suspected condition, or a well designed retrospective study of a broad spectrum of persons with an established condition (by “gold standard”) compared with a broad spectrum of controls, in which test is applied in blinded evaluation, and enabling the assessment of appropriate tests of diagnostic accuracy.  <u>III</u> Evidence provided by a retrospective study in which either persons with the established condition or controls are of a narrow spectrum, and in which test is applied in a blinded evaluation.  <u>IV</u> Any design in which test is not applied in blinded evaluation OR evidence provided by expert opinion alone or in descriptive case series (without controls).</p>	<p><b>GRADES OF RECOMMENDATION (SIGN, 2006)</b>  <i>Note: The grade of recommendation relates to the strength of the evidence on which the recommendation is based. It does not reflect the clinical importance of the recommendation.</i>  <b>A</b> At least one meta-analysis, systematic review of RCTs, or RCT rated as 1++ and directly applicable to the target population; <i>or</i> A body of evidence consisting principally of studies rated as 1+, directly applicable to the target population, and demonstrating overall consistency of results.  <b>B</b> A body of evidence including studies rated as 2++, directly applicable to the target population, and demonstrating overall consistency of results; <i>or</i> Extrapolated evidence from studies rated as 1++ or 1+.  <b>C</b> A body of evidence including studies rated as 2+, directly applicable to the target population and demonstrating overall consistency of results; <i>or</i> Extrapolated evidence from studies rated as 2++.  <b>D</b> Evidence level 3 or 4; <i>or</i> Extrapolated evidence from studies rated as 2+ Good practice points.                  § Recommended best practice based on the clinical experience of the guideline development group.</p> <p><b>LEVELS OF EVIDENCE</b>  <u>1++</u> High quality meta-analyses, systematic reviews of randomised controlled trials (RCTs), or RCTs with a very low risk of bias.  <u>1+</u> Well conducted meta-analyses, systematic reviews of RCTs, or RCTs with a low risk of bias.  <u>1-</u> Meta-analyses, systematic reviews of RCTs, or RCTs with a high risk of bias.  <u>2++</u> High quality systematic reviews of case control or cohort studies. High quality case control or cohort studies with a very low risk of confounding or bias and a high probability that the relationship is causal.  <u>2+</u> Well conducted case control or cohort studies with a low risk of confounding or bias and a moderate probability that the relationship is causal.  <u>2-</u> Case control or cohort studies with a high risk of confounding or bias and a significant risk that the relationship is not causal.  <u>3</u> Non-analytic studies, e.g. case reports, case series.  <u>4</u> Expert opinion.</p>

### 3.2.2 The International Classification of Functioning, Disability and Health (ICF)

Endorsed by the World Health Assembly in May 2001 the ICF is formed as part of the WHO Family of International Classifications (WHO-FIC) of which the main purpose is to provide 'a consensual, meaningful and useful framework which governments, providers and consumers can use as a common language' (WHO, 2002). While complementing the ICD, which is also a member of the WHO-FIC, the ICF is based on the biopsychosocial model and designed to conceptualise health, functioning and disability in a holistic manner that goes beyond the issues of mortality and morbidity of the population. Hence, the ICF aims to provide a means to describe and predict health service and social care needs as well as functional outcomes of illness at the individual, institutional and social levels that are essential to health planning and management (WHO, 2002). The ICF serves as both a classification system and conceptual framework, and constitutes the person's body functions and structures, activities and participation, and environmental factors. Body functions consist of the physiological and psychological functions while body structures include anatomical parts such as organs, limbs and their components. Activities are referred to the execution of a task or action by individuals while participation is defined as involvement in life situations. Environmental factors include physical (natural and man-made), social and attitudinal aspects of human life (AIHW, 2003)<sup>16</sup>. In the ICF, human functioning is classified at the level of body or body part, the whole person and the whole person in a social context while disability involves impairments of body or body part, activity limitations and participation restrictions (WHO, 2002).

The ICF has the potential to be applied in various settings/arenas, for example, research, health outcome measures, population studies, clinical assessment, policy development and education and training, and various disciplines (AIHW, 2003). One of the potential utilities in association with the DOMS project is the adoption of the ICF as a framework for dementia outcomes measurement. Whilst the ICF does not define dementia as a separate entity, the three domains and some of their components can serve as a framework in defining the individuals' functioning and disability in relation to their experience with dementia that is not sufficiently captured in the definitions of the ICD-10 or the DSM-IV alone. See Figure 1, which demonstrates how dementia can be considered as a particular type of cognitive impairment, and described within the framework of the ICF (AIHW, 2007)<sup>17</sup>. The comparisons between the ICF and the ICD-10 made by Madden (2006, cited in AIHW, 2007) are that the ICF domains including temperament and personality (b126), energy and drive functions (b130), attention (b140), psychomotor (b147), perceptual (b156) and higher level cognitive functions (b164) are not included in the ICD definition. Furthermore, Muo, et al. (2005) suggest that the ICF suitably assists the assessment of activities of daily living in the person with AD in that it provides an avenue to consider the person's ability to communicate, establish/maintain relationships with others, and participate in recreational activities that are often overlooked in other measures of activities of daily living. The ICF does not appear to include some of the features of dementia defined in the ICD-10 such as 'comprehension', 'learning capacity' and 'social behaviour' (Madden, cited in AIHW, 2007).

Nevertheless there has been a consistent movement towards a wider application of the ICF internationally and in Australia; given its relatively short history. The complex characteristics of dementia are reflected in the utility of the classification. The validity and reliability of the ICF as well as practicability in its implementation in the Australian mainstream health care industry are already being addressed through the development of an electronic data capture tool (Functioning and Related Health Outcomes Module, FRHOM), based on the ICF, for inclusion in the electronic health records. Currently in Australia, the ICF is used in various national data collections to describe support needs for people with disability such as the Survey of Disability Ageing and

---

<sup>16</sup> See Appendix 4G for details of the domains and components.

<sup>17</sup> This figure was designed by the AIHW to illustrate and provide some examples of how the ICF could be applied to dementia, hence it is not to be taken as a WHO authorised ICF based classification of dementia.

Carers, the Commonwealth-State/Territory Disability Agreement National Minimum Data Set, the National Community Services Data Dictionary and the 2006 Census of Population and Housing (AIHW, 2007). For the purpose of the DOMS project, the ICF may be used as a conceptual framework in identifying and describing outcome measures of dementia and its subtypes although its full application is yet to be fully realised given the limited evidence regarding the applicability of the ICF for people with dementia so far.

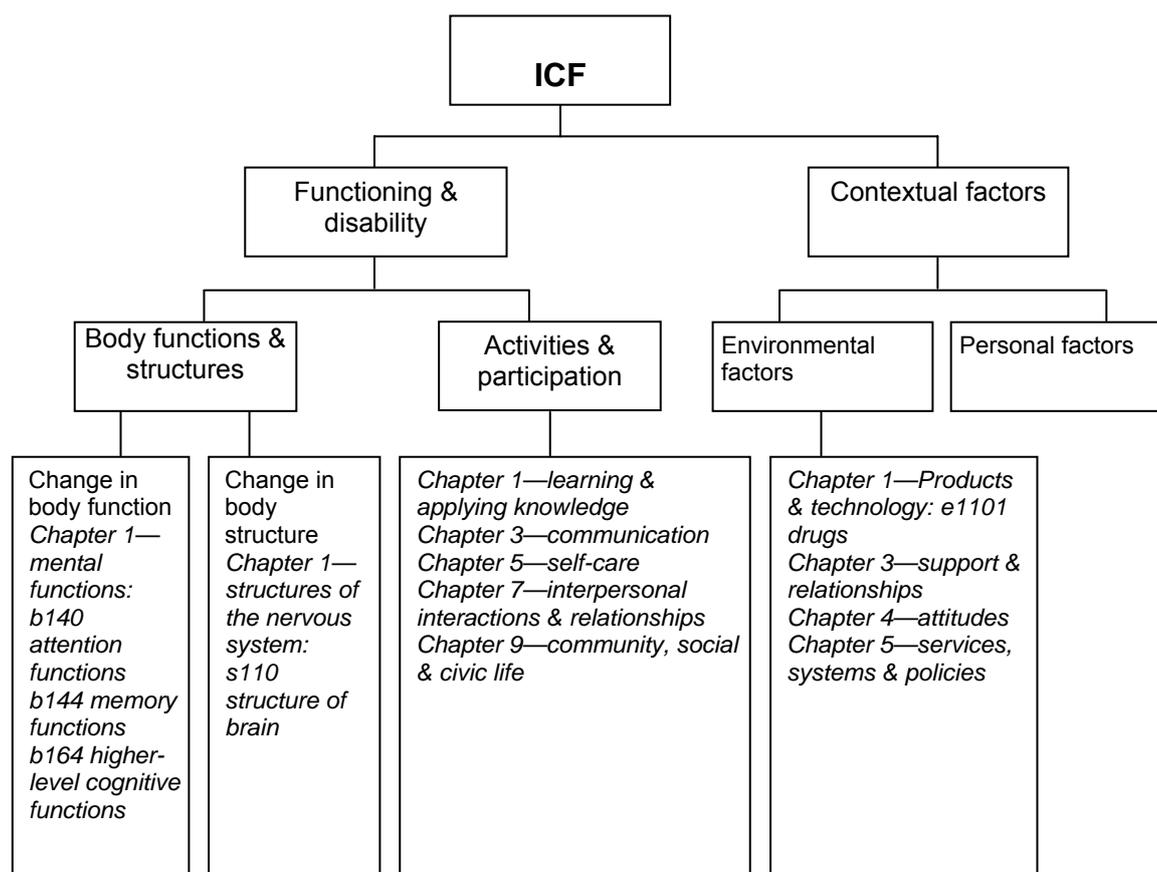
### 3.2.3 Outcome Measurement of Dementia

There are various ways of classifying health measures, depending on the type of particular frameworks applied. For example, determined by the purpose of measuring health (functional classifications), measures can be organised as diagnostic, prognostic and evaluative; or determined by the scope of the topics and the concept being examined (descriptive classifications), they can range from narrow-focused to broad spectrum measures (e.g., from measuring a particular organ system and diagnosis to broader aspects of health, overall health and quality of life) (McDowell, 2006). For the purpose of the DOMS project, 'outcome' can be defined as the effect of an intervention/care by health care professionals and health service on the person's health status (Andrews, et al. 1994), and focused in terms of evaluative and broad spectrum measures.<sup>18</sup> The outcome is examined using various measures through which comparisons are made between different timelines and/or settings to examine whether there has been a change in the person's health status that may be described as a decline, improvement or no change. Determining the type of measure depends on how one defines the outcome of dementia care or a particular intervention. Outcome measures are used as an important source for service planning and evaluation. It is also critical that the outcome measures are designed and chosen so that information obtained is meaningful to the person receiving care and his/her family. Dementia is predominantly degenerative and mostly irreversible - less than 1.5% are reversible (Boustani, et al. 2003). Consequently, evaluating health, care/service outcomes of people with dementia generally focuses on maintaining status quo and delaying further decline of symptoms of dementia as well as examining rates of recovery, mortality and institutional length of stay. Commonly used domains of outcome measures for individuals with dementia include: cognition; self-care abilities and activities of daily living; physical and functional health; quality of life; behaviour and psychological symptoms of dementia; social functioning; service user satisfaction; and service use and costs (Bamford and Bruce, 2000; Ramsay, et al. 1995). These aspects or domains of outcome measures for dementia will be framed based on the ICF as discussed above.

---

<sup>18</sup> Given the importance of diagnostic measures taking place before any form of evaluation occurs, the project will include routine diagnostic measures that are easily accessible and applicable.

**Figure 1 Dementia and its Outcomes in the Structure of the ICF (Source: AIHW, 2007)**



### 3.2.4 Differential Diagnosis of Dementia

Determining the most accurate diagnosis is important not only for the provision of adequate treatment but for the person with dementia and their family or carer so that they can anticipate and plan for their future. Defined as ‘determination of which one of two or more diseases with similar symptoms is the one from which the patient is suffering’ (The American Heritage® Stedman's Medical Dictionary, 2004), differential diagnosis involves examining the likelihood of all possible diagnoses of diseases/illnesses explaining the person’s conditions, collecting additional information including personal and family history through interviews with a family member or carer as well as with the patient, and then selecting a single most likely diagnosis. In differential diagnosis of dementia focus should be on the clinical pathway such as mode of onset and course of progression, pattern of cognitive impairment and presence of non-cognitive symptoms such as behavioural disturbance, hallucinations and delusions (SIGN, 2006). In the early stages of dementia it is important to rule out reversible/treatable conditions first (e.g. depression, acute confusional state/delirium, deficiency of vitamin B12 and niacin, hypercalcemia; hypothyroidism, and intoxications). Two of the most common illnesses that require differential diagnosis in this regard are depression and delirium - either or both may coexist with dementia - and those three are the most prevalent mental disorders among older people. Other disorders requiring consideration in the differential diagnosis include mild or moderate mental retardation, states of subnormal cognitive functioning attributable to a severely impoverished social environment and limited education, and iatrogenic mental disorders due to medication (ICD-10-AM). In addition, schizophrenia, amnesic disorder and age-related cognitive decline also need to be considered (First and Tasman, 2004).

Unlike the differential diagnosis of dementia caused by potentially reversible conditions such as Vitamin12 deficiency or hypothyroidism, differentiating between the major types of dementia can pose a significant challenge especially when diagnostic criteria are not clearly established. Kaye

(1998) summarises this into two categories: (a) dementias without distinctive neurologic signs, or evidence of medical or neurologic disease (e.g. AD, FTD), and (b) dementias with neurologic signs without obvious significant medical disorders (e.g. dementia in Parkinson's disease, VaD). Of those VaD is known to be most challenging as it is not a homogenous entity and is often accompanied by AD. Initial clinical manifestations of VaD and AD may be quite similar although VaD tends to show better function in verbal long-term memory and more impairment in frontal executive functioning when compared with AD (Looi and Sachdev, 1999). Other reasons for this problem include: (a) similar to AD, cerebrovascular disease is common amongst older people and its incidence increases with age; (b) possibly multiple lesions of infarctions may cause various syndromes; (c) older people or their family members may not understand the term 'stroke', hence using it interchangeably with 'cognitive changes' (Kaye, 1998). Studies clearly suggest that using the AD criteria with the inclusion of history of cerebrovascular disease is not adequate to detect the uniqueness of VaD (Looi and Sachdev, 1999). According to the current clinical guidelines none of the existing diagnostic criteria performs well for mixed features of dementia, however there is evidence supporting the use the Hachinski Ischaemic Score and NINDS-AIRENS criteria to discriminate AD from VaD (AAN, 2001; Kaye, 1998; SIGN, 2006). Both SIGN (2006) and AAN (2001) recommend the Lund-Manchester criteria for FTD (1994) and the consensus criteria for LBD (McKeith, et al. 1996)<sup>19</sup> as useful tools for differentiating dementia types. As shown in Table 17 for the differential diagnosis of dementia, the ICD-10 based criteria do not provide sufficient information regarding differential diagnosis of dementia.

**Table 17 Differential Diagnosis of Dementia, ICD-10-AM (2002)**

Dementia and subtypes	Guidelines for Differential Diagnosis
Dementia	Consider: a depressive disorder, which may exhibit many of the features of an early dementia, especially memory impairment, slowed thinking and lack of spontaneity; delirium; mild or moderate mental retardation; states of subnormal cognitive functioning attributable to a severely impoverished social environment and limited education; iatrogenic mental disorders due to medication. Dementia may <i>follow</i> any other organic mental disorder classified in this block, or <i>coexist</i> with some of them, notably delirium.
Dementia in Alzheimer's disease	Consider: a depressive disorder; delirium; organic amnesic syndrome; other primary dementias such as Creutzfeldt-Jacob or Huntington's disease; secondary dementias associated with a variety of physical diseases, toxic states, etc.; mild moderate or severe mental retardation. Dementia in AD may coexist with VaD, as when cerebrovascular episodes (multi-infarct phenomena) are superimposed on a clinical picture and history suggesting AD. Such episodes may result in sudden exacerbations of manifestations of dementia. According to post-mortem findings both types may co-exist in as many as 10-15% of all dementia cases.
Vascular dementia	Consider: delirium, other dementia, particularly AD; mood [affective] disorders; mild or moderate mental retardation; subdural haemorrhage. VaD may co-exist with AD, as when evidence of a vascular episode is superimposed on a clinical picture and history suggesting AD.
Dementia in Pick's disease	Frontal lobe features are more marked than temporal and parietal, unlike AD. Consider: AD; VaD; dementia secondary to other disorders such as neurosyphilis; normal pressure hydrocephalus (characterised by extreme psychomotor slowing, and gait and other sphincter disturbances); other neurological or metabolic disorders.
Dementia in Parkinson's disease	Consider: other secondary dementias; multi-infarct dementia associated with hypertensive or diabetic vascular disease; brain tumour; normal pressure hydrocephalus.

Another contemporary issue relevant to the differential diagnosis of dementia relates to age-related cognitive decline that is characterised by a persistent decline in performance in memory and/or other cognitive functions. These changes may vary in degree but are neither of a magnitude nor pattern to meet the diagnostic criteria of dementia. Commonly adopted nosology for this condition includes mild cognitive impairment, incipient dementia, mild neurocognitive disorder,

<sup>19</sup> This has been subsequently reviewed and revised (McKeith, et al. 1999; McKeith, et al. 2005)

late-life forgetfulness, possible dementia, and (benign) senescent forgetfulness, of which the first three are more severe and more likely to progress to dementia (APA, 1998). In the past decade, the term “mild cognitive impairment (MCI)” has been increasingly gaining interest and adopted as a distinctive entity among researchers and clinicians. Petersen and his colleagues (1995, 1997) have refined and developed the definition and diagnostic criteria of MCI. Defined as ‘the transitional stage between normal ageing and probable AD’ MCI is applied when a person meets all of the following criteria: complaint of defective memory, normal activities of daily living, normal general cognitive function, abnormal memory function for age, and absence of dementia (Petersen, 1995; Petersen, et al. 1997). Since then, the use of the definition and the criteria of MCI proposed by Petersen has been much criticised due to heterogeneous and unstable nature of MCI and limited explanation given to the notion of MCI initially - much focus was given to memory deficit or amnesic MCI which may lead to AD while MCI may progress to other types of dementia. Depending on how one interprets the term MCI and its subtypes<sup>20</sup>, and the type of age population under investigation the prevalence of MCI may range from 3 - 17% of older people (Portet, et al. 2006). In an attempt to draw a consensus in its application Petersen and other international expert groups such as the International Working Group on MCI and the European Alzheimer's Disease Consortium (EADC) (Petersen, et al. 2001, 2006; Winblad, et al. 2004) have further undertaken reviews and revisions of the concept of MCI. The latest consensus by the EADC describes MCI as a syndrome and provides three steps towards clear diagnosis of MCI subtypes by different experts.<sup>21</sup>

Given the lack of evidence with regards to the diagnostic criteria and tests that are still evolving, it is important first to establish that the benefits of labelling people with MCI outweigh its drawbacks, in particular for those who are diagnosed as having MCI yet never develop dementia. It is vital to note that this does not mean withholding screening for and diagnosis of dementia when suspected. Whilst evidence does not support routine screening of people who do not show any signs of cognitive impairment (USPSTF, 2003), use of efficient and practical instruments to screen those 75 years and over or those who show some form of cognitive decline for possible dementia, especially in primary health care settings (by GPs), is an important step towards early recognition of dementia (Brodaty, et al. 2006).

The review indicates that recognition of cognitive impairment is important and clinicians need to be vigilant about its further development to dementia, however there is insufficient evidence to embrace MCI as a new diagnosis and provide treatment that is beneficial to people with MCI and their family/carers. At the first meeting of the National Expert Panel (NEP), the members agreed that given MCI is not fully established as a proper diagnosis and as the DOMS project focuses on the clinical phase of diagnosis it is best not to be included in this project.

### 3.2.5 Severity of Dementia

The DSM-IV and the ICD-10<sup>22</sup> provide guidelines for describing the severity of dementia - mild, moderate and severe - these are used only after the diagnosis has been made (i.e. mild impairment as the threshold for diagnosis). As shown in Table 18, the DSM-IV uses this staging as universal to almost all other mental disorders<sup>23</sup> whereas the ICD-10 provides specific definitions for the severity of dementia in terms of: (a) a decline in memory mostly in terms of the learning of new information; (b) a decline in other cognitive abilities in terms of judgement and thinking, and the general processing of information. In this research version of the ICD, the overall severity of dementia is best expressed as the level of decline in memory or other cognitive abilities, whichever

---

20 Amnesic MCI (single domain and multiple domains), Non-Amnesic MCI (single domain and multiple domains).

21 See Appendix 4J for Different stages of the EADC MCI diagnostic procedure.

22 Of all the ICD-10 based classifications, the ICD-10 Classification of Mental and Behavioural Disorders: Diagnostic criteria for research (WHO 1993) is the only version that provides guidelines for the severity of dementia.

23 Specific criteria for defining Mild, Moderate and Severe have been made for Mental Retardation, Conduct Disorder, Manic Episode and Major Depressive Episode (APA, 2000).

is the more severe (e.g. mild decline in memory and moderate decline in cognitive abilities indicate a dementia of moderate severity) (WHO, 1993, p.30).

Two commonly used scales for the severity of dementia include the Clinical Dementia Rating (CDR) Scale (Morris, 1993)<sup>24</sup> and the Global Deterioration Scale (GDS) (Reisberg, et al. 1982)<sup>25</sup>. Unlike the DSM-IV and the ICD-10 based criteria, both scales include assessment of pre-dementia stages, and they have established good reliability when used by clinicians trained to administer these (Burns, Lawlor and Craig, 2004). They are used in both research and clinical settings to measure the global level of cognition and functioning in people who are believed to have dementia. The CDR consists of six domains of cognitive and functional performance: Memory, Orientation, Judgment and Problem Solving, Community Affairs, Home and Hobbies, and Personal Care. Each domain is scored based on information obtained through a semi-structured interview of the patient and a family or carer. The overall level of impairment is derived by standard algorithm and described as No impairment (0), Very mild (0.5), Mild (1), Moderate (2), and Severe (3) Dementia. The GDS rates the level of cognitive and daily functions for those with a presumptive diagnosis of a primary degenerative dementia in seven stages. Stages 1-3 denote the pre-dementia stages (No cognitive decline, Very mild and Mild cognitive impairment) while stages 4-7 describe the severity of dementia (Moderate, Moderately severe, Severe and Very severe cognitive decline). A review of these instruments can be found in Section 4 and Appendix 5.

**Table 18 The Severity of Dementia**

DSM-IV (APA 2000)	ICD-10 Diagnostic criteria for research (WHO 1993)
<p>Mild. Few, if any, symptoms in excess of those required to make the diagnosis are present, and symptoms result in no more than minor impairment in social or occupational functioning.</p> <p>Moderate. Symptoms or functional impairment between “mild” and “severe” are present.</p> <p>Severe. Many symptoms in excess of those required to make the diagnosis, or several symptoms that are particularly severe, are present, or the symptoms result in marked impairment in social and occupational functioning.</p>	<p>(1) The degree of memory loss</p> <p>Mild is sufficient to interfere with everyday activities, though not so severe as to be incompatible with independent living. The main function affected is the learning of new material. For example, the individual has difficulty in registering, storing, and recalling elements involved in daily living, such as where belongings have been put, social arrangements, or information recently imparted by family members.</p> <p>Moderate represents a serious handicap to independent living. Only highly learned or very familiar material is retained. New information is retained only occasionally and very briefly. Individuals are unable to recall basic information about their own local geography, what they have recently been doing, or the names of familiar people.</p> <p>Severe is characterised by the complete inability to retain new information. Only fragments of previously learned information remain. The individual fails to recognise even close relatives.</p> <p>(2) The decline in cognitive abilities</p> <p>Mild causes impaired performance in daily living, but not to a degree that makes the individual dependent on others. Complicated daily tasks or recreational activities cannot be undertaken.</p> <p>Moderate makes the individual unable to function without the assistance of another in daily living, including shopping and handling money. Within the home, only simple chores can be performed. Activities are increasingly restricted and poorly sustained.</p> <p>Severe is characterised by an absence, or virtual absence, or intelligible ideation.</p>

### 3.2.6 Behavioural and Psychological Symptoms of Dementia (BPSD)

Over the last two decades the term “Behavioural and psychological symptoms of dementia (BPSD)”<sup>26</sup> has emerged and been recognised as defining non-cognitive features of dementia, along with symptoms related to cognitive deficits (Robert, et al. 2005). According to Brodaty and his colleagues (2003) 61% - 88% of community-based populations living with dementia experience some form of BPSD - 29% of people with dementia have mild BPSD, 21% moderate BPSD, 10% severe BPSD and 1% very severe BPSD (Lyketsos, et al. 2000, cited in Brodaty, et al. 2003). In

24 See Appendix 4K

25 See Appendix 4L

26 The term BPSD has been interchangeably used with “challenging behaviours”, “difficult behaviours”, or “behaviour disturbances”. However, they contain negative connotations and do not accurately convey the meaning attributed to the symptoms of concern. “BPSD” is also a preferred term for people with dementia and their families/carers (Alzheimer’s Australia, 2004).

particular, a high prevalence of BPSD in nursing home populations, 80% - 90%, has been found in studies conducted within and outside Australia (Brodaty, et al. 2003; Finkel, 1998; Nay, et al. 2003). Often obtained by direct observation of the person with dementia and reports from families and carers, different symptoms of BPSD occur at various stages of dementia. The issue of identifying and managing BPSD is of particular interest to families and carers as well as to the person with dementia in terms of its impact on quality of life of the individual with dementia and the family/carers, stress to those who provide formal and informal care, financial costs, early nursing home admission, prognosis and the person's capacity to function in everyday activities (Brodaty, et al. 2003; IPA, 2003).

Neither the DSM-IV nor the ICD-10 provides clear definitions for BPSD while only some of the features of BPSD such as depression, hallucinations, delusion are defined in those two classifications. The expert consensus group organised by the International Psychogeriatric Association (IPA) in 1996 and 1999 provides the definition of BPSD as: 'symptoms of disturbed perception, thought content, mood or behaviour that frequently occur in patients with dementia' (Finkel and Burns, 1999, cited in IPA, 2003). Behavioural symptoms include physical aggression, screaming, restlessness, agitation, wandering, culturally inappropriate behaviours, sexual disinhibition, hoarding, cursing and shadowing. Psychological symptoms relate to anxiety, depressive mood, hallucinations, delusions and psychosis (IPA, 2003). Depending on the type of instruments used researchers and clinicians have depicted and measured BPSD under various categorisations. For example:

- Cohen-Mansfield Agitation Inventory (CMAI) (Cohen-Mansfield, et al. 1989): "agitation" consisting of aggressive, physically non-aggressive, and verbally agitated behaviours.
- Behavioural Pathology in Alzheimer's Disease Rating Scale (BEHAVE-AD) (Reisberg, et al. 1987): Delusions, Hallucinations, Activity disturbances, Aggressiveness, Diurnal rhythm disturbances, Affective disturbances, Anxieties/phobias.
- The Consortium to Establish a Registry for AD (CERAD) Behaviour Rating Scale for Dementia (Tariot, et al. 1995): Psychotic features, Aggression, Affective lability, Depressive features, Defective self-regulation, Vegetative features, Apathy, Irritability / agitation.

Given their heterogeneous nature, BPSDs need to be divided into clusters of symptoms that can provide a framework for assessment and management (Robert, et al. 2005). However, there is no consensus reached in this regard that goes beyond the IPA guidelines.

Review of the literature indicates there are no guidelines or definitions for the severity of BPSD, except for those from some of the instruments measuring BPSD, such as the Neuropsychiatric Inventory (NPI)<sup>27</sup> (Cummings, et al. 1994) and the Behavioural and Emotional Activities Manifested in Dementia (BEAM-D) (Sinha, et al. 1992), where the severity is recorded as mild, moderate and severe. The severity of BPSD is also assessed through measuring the impact of BPSD on carers (formal and informal). For example, the NPI allows examination of the level of distress (no distress, minimal, mild, moderate, moderately severe, very severe or extreme) carers experience while the BEHAVE-AD records an overall rating of the trouble that BPSD causes to carers (mild, moderate and severe). The AIHW suggests a guideline derived from Caldwell and Bird's (2004, cited in AIHW, 2007) work on challenging behaviours where the impact of BPSD is described as 'the extent of disruption to normal activities that results from the challenging behaviour' and rated as not disruptive, mildly disruptive, moderately disruptive, very disruptive and extremely disruptive (AIHW, 2007).

---

<sup>27</sup> The NPI consists of 12 areas of BPSD including: delusions, hallucinations, agitation, depression, anxiety, euphoria, apathy, disinhibition, irritability, aberrant motor behaviour, night-time behaviours, and appetite-eating disorders.

### 3.3 Recommendations

It is recommended that:

- The ICD-10-AM is used to inform the diagnostic classifications for dementia and its subtype given this system is already in place in collecting national data in Australia.
- The ICD-10-AM and ICD-10 are used for diagnostic criteria for dementia and AD. After consultations with several psychiatrists it seemed appropriate to recommend the ICD-10 instead of DSM-IV. Clinicians do not necessarily follow either of the classifications as they often rely on their clinical judgement. Given that majority of the health related information is collected based on the ICD-10 and the ICD-10-AM it is more efficient for clinicians to use one system rather than two (i.e. DSM-IV diagnostic criteria and ICD-10 for coding exercise).
- For research, the DSM-IV is preferred as it is more inclusive of mild to moderate dementia and most epidemiological studies use the DSM-IV because of ease of comparison with prior studies. However this is not mandatory, providing the study states the type of the classification used, as there is no evidence available to say the DSM-IV is superior to the ICD-10.
- In terms of differential diagnosis (DD) and diagnoses of frontotemporal dementia (FTD) and dementia with Lewy bodies (LBD), additional criteria are used: the National Institute of Neurologic Disorders and Stroke and the Association Internationale pour la Recherche et l'Enseignement en Neurosciences (NINDS-AIREN) (Roman, et al. 1993) for DD of Vascular dementia from Alzheimer's type; the Lund-Manchester criteria for FTD (1994) and the consensus criteria for LBD (McKeith, et al. 2005).
- Mild cognitive impairment (MCI) is not to be included in this project as a diagnostic entity, however screening measures for those who are suspected of cognitive impairment need to be considered.
- For assessing the severity of dementia, the CDR scale has been used for two main reasons: the AIHW recommends this and, in addition to three stages of dementia, the CDR allows room to record abnormal cognitive function without necessarily labelling it as MCI. It is well validated and widely recognised. Similarly the GDS has also been widely used to assess the severity of Dementia. A detailed review of these instruments is provided in Section 4 and in Appendix 5.
- The ICF may be used as a conceptual framework for classification of measurement scales. However, given its early developmental status as a classification system in Australia, hence its unfamiliarity among clinicians and researchers, and lack of evidence relating to validity and reliability of the classification it is deemed beyond the scope of the DOMS project to provide a definite recommendation on this subject.
- Behavioural and psychological symptoms of dementia (BPSD) are an integral part of dementia outcome measures. The guidelines provided by the International Psychogeriatric Association (IPA) are to be used for the definitions. Whilst the AIHW recommends Caldwell and Bird's guideline for the severity of BPSD, it has been suggested that a more widely recognised measure is selected for this project.

The above recommendations have been made largely based on the review of literature and have been ratified by the National Expert Panel.

### References

Alzheimer's Australia (2004) *Dementia Terminology Framework*. Position Paper 4, available at [http://www.alzheimers.org.au/upload/Terminology\\_Position\\_Paper\\_041.pdf](http://www.alzheimers.org.au/upload/Terminology_Position_Paper_041.pdf), accessed 1 June 2006.

American Psychiatric Association (1987) *Diagnostic and statistical manual of mental disorders, third edition, revised: DSM-III-R*. American Psychiatric Association, Washington DC.

American Psychiatric Association (1994) *Diagnostic and Statistical Manual of Mental Disorders DSM-IV (4th ed.)* American Psychiatric Association, Washington DC.

American Psychiatric Association (1998) *Guidelines for the Evaluation of Dementia and Age-Related Cognitive Decline*, available at <http://www.apa.org/practice/dementia.html>, accessed 10 May 2006.

American Psychiatric Association (2000) *Diagnostic and statistical manual of mental disorders, fourth edition, text revision*. American Psychiatric Association, Washington DC.

Andrews G, Peters L and Teeson M (1994) *The Measurement of Consumer Outcome in Mental Health: A Report to the National Mental Health Information Strategy Committee*. Clinical Research Unit for Anxiety Disorders, Sydney.

Australian Institute of Health and Welfare (2003) *ICF Australian User Guide Version 1.0. Disability Series (No. AIHW Cat. No. DIS 33.)* AIHW, Canberra.

Australian Institute of Health and Welfare (AIHW) (2007) *Final report – Australian Dementia Data Analysis and Development* (In press) AIHW, Canberra.

Bamford CE and Bruce E (2000) *Defining the outcomes of community care: the perspectives of older people with dementia and their carers*. *Ageing and Society*. Vol. 20, pp.543-570.

Boustani M, Peterson B, Hanson L, Harris R and Lohr K (2003) *Summary of the Evidence: Screening for dementia in primary care*, available at <http://www.ahcpr.gov/clinic/3rduspstf/dementia/demsum2.htm>, accessed 21 June 2006.

Bowling A (2001) *Measuring disease (2nd ed.)* Open University Press, Buckingham England.

Bowling A (2005) *Measuring health: a review of quality of life measurement scales (3rd ed.)* Open University Press, Berkshire, England.

Brody H, Draper BM and Low LF (2003) *Behavioural and psychological symptoms of dementia: a seven-tiered model of service delivery*. *MJA*. Vol. 178, pp.231-234.

Brody H, Low LF, Gibson L and Burns K (2006) *What is the best dementia screening instrument for general practitioners to use?* *American Journal of Geriatric Psychiatry*. Vol. 14 No.5, in press.

Burns A, Lawlor B and Craig S (2004) *Assessment scales in old age psychiatry (2nd ed.)* Taylor & Francis, London.

Chui HC, Mack W, Jackson J, Mungas D, Reed BR, Tinklenberg J, et al. (2000) *Clinical criteria for the diagnosis of vascular dementia: A multicenter study of comparability and interrater reliability*. *Archives of Neurology*. Vol. 57, No.2, pp.191-196.

Chui HC, Victoroff JI, Margolin D, Jagust W, Shankle R and Katzman R (1992) *Criteria for the diagnosis of ischemic vascular dementia proposed by the State of California Alzheimer's disease Diagnostic and Treatment Centres*. *Neurology*. Vol. 42, pp.473-480.

Cohen-Mansfield J, Marx MS and Rosenthal AS (1989) *A description of agitation in a nursing home*. *Journal of Gerontology*. Vol. 44, pp.M77–M84.

Cummings JL, Mega M, Gray K, Rosenberg-Thompson S, Carusi DA and Gornbein J (1994) *The Neuropsychiatric Inventory: comprehensive assessment of psychopathology in dementia*. *Neurology*. Vol. 44, No.12, pp.2308-2314.

Erkinjuntti T, Ostbye T, Steenhuis R and Hachinski V (1997) *The effect of different diagnostic criteria on the prevalence of dementia*. *New England Journal of Medicine*. Vol. 337, No.23, pp.1667-1674.

Fichter MM, Meller I, Schroppe H, and Steinkirchner R (1995) *Dementia and cognitive impairment in the oldest old in the community: prevalence and comorbidity*. *British Journal of Psychiatry*. Vol. 166, pp.621-629.

- Finkel SI (1998) *The signs of behavioural and psychological symptoms of dementia*. Clinician. Vol. 16, No.1, pp.33-42.
- First MB and Tasman A (eds.) (2004) *DSM-IV-TR Mental Disorders: Diagnosis, Etiology and Treatment*. John Wiley and Sons, LTD, West Sussex.
- Frank LB (2006) *Commentary on "Meta-analysis of memantine: summary and commentary on the Cochrane Collaboration's systematic review." Meaning and measurement in cognitive impairment*. Alzheimer's and Dementia. Vol. 2, pp.93-95.
- Gallagher LA (ed.) (2004) *Thesaurus of Psychological Index Terms (1974 – 2004) Tenth Edition 2004*. American Psychological Association, Washington DC.
- Gold G, Bouras C, Canuto A, Bergallo MF, Herrmann FR, Hof PR, et al. (2002) *Clinicopathological validation study of four sets of clinical criteria for vascular dementia*. [see comment]. American Journal of Psychiatry. Vol. 159, No.1, pp.82-87.
- Grabowski TJ and Damasio AR (2004) *Definition, clinical features and neuroanatomical basis of dementia*. In Esiri MM, Lee VM and Trojanowski JQ (eds.) (2004) The Neuropathology of Dementia (2nd ed.) Cambridge University Press, Cambridge.
- Grundman M, Petersen RC, Ferris SH, Thomas RG, Aisen PS, Bennett DA, et al. (2004) *Mild cognitive impairment can be distinguished from Alzheimer disease and normal aging for clinical trials*. Archives of Neurology. Vol. 61, No.1, pp.59-66.
- Henderson AS, Jorm AF, Mackinnon A, Christensen H, Scott LR, Korten AE, et al. (1994) *A survey of dementia in the Canberra population: experience with ICD-10 and DSM-III-R criteria*. Psychological Medicine. Vol. 24, No.2, pp.473-482.
- Iliffe S and Manthorpe J (2004) *The hazards of early recognition of dementia: a risk assessment*. Ageing and Mental Health. Vol. 8, No.2, pp.99-105.
- IPA (International Psychogeriatric Association) (2003) *The IPA Behavioural and Psychological Symptoms of Dementia*, available at <http://www.ipa-online.net/pdfs/bpsd.asp>, accessed 1 June 2006.
- Kane RL and Kane RA (2000) *Assessing older persons: measures, meaning and practical applications*. Oxford University Press, USA.
- Kaye JA (1998) *Diagnostic challenges in dementia*. Neurology. Vol. 51, No.1, Suppl.1, pp.S45-52; discussion pp.S65-47.
- Knopman DS, DeKosky ST, Cummings JL, Chui HC, Corey-Bloom J, Relkin N et al. (2001) *Practice parameter: Diagnosis of dementia (An evidence-based review)*. Report of the Quality Standards Subcommittee of the American Academy of Neurology. Neurology. Vol. 56, pp.1143-1153.
- Looi JC and Sachdev PS (1999) *Differentiation of vascular dementia from AD on neuropsychological tests*. Neurology. Vol. 53 No.4, pp.670-678.
- Lopez OL, Kuller LH, Becker JT, Jagust WJ, DeKosky ST, Fitzpatrick A, et al. (2005) *Classification of vascular dementia in the Cardiovascular Health Study Cognition Study*. Neurology. Vol. 64, No.9, pp.1539-1547.
- McDowell I (2006) *Measuring Health: A Guide to Rating Scales and Questionnaires (3rd ed.)* Oxford University Press, Oxford.
- McKeith I, Cummings J, Lovestone S, Harvey R and Wilkinson D (1999) *Outcome measures in Alzheimer's disease*. Martin Dunitz, London.
- McKeith IG, Dickson DW, Lowe J, Emre M, O'Brien JT, Feldman H, et al. (2005) *Diagnosis and management of dementia with Lewy bodies: third report of the DLB Consortium*. [erratum appears in Neurology. 2005 Dec 27; 65(12):1992] [summary for patients in Neurology. 2005 Dec 27; 65(12):E26-7; PMID: 16380603]. Neurology. Vol. 65, No.12, pp.1863-1872.

- McKeith IG, Galasko D, Kosaka K, Perry EK, Dickson DW, Hansen LA, et al. (1996) *Consensus guidelines for the clinical and pathologic diagnosis of dementia with Lewy bodies (DLB): report of the consortium on DLB international workshop*. Neurology. Vol. 47, No.5, pp.1113-1124.
- McKeith IG, Perry EK and Perry RH (1999) *Report of the second dementia with Lewy body international workshop: diagnosis and treatment. Consortium on Dementia with Lewy Bodies*. Neurology. Vol. 53, No.5, pp.902-905.
- McKhann G, Drachman D, Folstein M, Katzman R, Price D and Stadlan EM (1984) *Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's disease*. Neurology. Vol. 34, No.7, pp.939-944.
- McShane R, Birks J, Hermans D and Hicks K (2006) *Dementia and Cognitive Improvement Group. About The Cochrane Collaboration* (Cochrane Review Groups) 2006, Issue 2, Art No.DEMENTIA.
- Morris JC (1993) *The Clinical Dementia Rating (CDR): Current version and scoring rules*. Neurology. Vol. 43, pp.2412-2414.
- Muo R, Schindler A, Venero I, Schindler O, Ferrario E and Frisoni GB (2005) *Alzheimer's disease-associated disability: an ICF approach*. Disability and Rehabilitation. Vol. 27, No.23, pp.1405-1413.
- National Centre for Classification in Health (2002a) *ICD-10-AM Mental Health Manual. An integrated classification and diagnostic tool for community-based mental health services (1st ed.)* University of Sydney, Sydney.
- National Centre for Classification in Health (2002b) *The international statistical classification of diseases and related health problems, 10th revision, Australian modification (ICD-10-AM) (3rd ed.)* University of Sydney, Sydney.
- Nay R, Scherer S, Pitcher A, Koch S, Browning M, Flicker L, et al. (2003) *Responding to Behaviours of Concern Among People Living in Residential Aged Care*. La Trobe University, Melbourne.
- Petersen RC (1995) *Normal aging, mild cognitive impairment, and early Alzheimer's disease*. The Neurologist. Vol. 1, pp.326-344.
- Petersen RC, Smith GE, Waring SC, Ivnik RJ, Kokmen E and Tangelos EG (1997) *Aging, Memory, and Mild Cognitive Impairment*. International Psychogeriatrics. Vol. 9, suppl.1, pp.65-69.
- Petersen RC, Stevens JC, Ganguli M, Tangelos EG, Cummings JL and DeKosky ST (2001) *Practice parameter: Early detection of dementia: Mild cognitive impairment (an evidence-based review)*. Neurology. Vol. 56, No.9, pp.1133-1142.
- Pioggiosi P, Forti P, Ravaglia G, Berardi D, Ferrari G and De Ronchi D (2004) *Different classification systems yield different dementia occurrence among nonagenarians and centenarians*. Dementia and Geriatric Cognitive Disorders. Vol. 17, No.1-2, pp.35-41.
- Pohjasvaara T, Mantyla R, Ylikoski R, Kaste M and Erkinjuntti T (2000) *Comparison of different clinical criteria (DSM-III, ADDTC, ICD-10, NINDS-AIREN, DSM-IV) for the diagnosis of vascular dementia*. *National Institute of Neurological Disorders and Stroke-Association Internationale pour la Recherche et l'Enseignement en Neurosciences*. Stroke. Vol. 31, No.12, pp.2952-2957.
- Portet F, Ousset PJ, Visser PJ, Frisoni GB, Nobili J, Scheltens P, et al. (2006) *Mild cognitive impairment in medical practice: critical review of the concept and new diagnostic procedure. Report of the MCI working group of the European Consortium on Alzheimer's disease (EADC)*, available at <http://www.jnnp.bmjournals.com>, accessed 2 May 2006.
- Ramsay M, Winget C and Higginson I (1995) *Review: Measures to determine the outcome of community services for people with dementia*. Age and Ageing. Vol. 24, No.1, pp.73-83.

Reisberg B, Borenstein J, Salob SP, Ferris SH, Franssen E and Georgotas A (1987) *Behavioural symptoms in Alzheimer's disease: phenomenology and treatment*. Journal of Clinical Psychiatry. Vol. 48, Suppl, pp.9-15.

Reisberg B, Ferris SH, de Leon MJ and Crook T (1982) *The Global Deterioration Scale for assessment of primary degenerative dementia*. American Journal of Psychiatry. Vol. 139, No.9, pp.1136-1139.

Robert PH, Verhey FRJ, Byrne EJ, Hurt C, De Deyn PP, Nobili F, et al. (2005) *Grouping for behavioural and psychological symptoms in dementia: clinical and biological aspects: Consensus paper on the European Alzheimer's disease consortium*. European Psychiatry. Vol. 20, pp.490-496.

Roman GC, Tatemichi TK, Erkinjuntti T, Cummings JL, Masdeu JC, Garcia JH, et al. (1993) *Vascular dementia: diagnostic criteria for research studies. Report of the NINDS-AIREN International Workshop. [see comment]*. Neurology. Vol. 43, No.2, pp.250-260.

Scottish Intercollegiate Guidelines Network (SIGN) (2006) *Management of patients with dementia: A national clinical guideline*. SIGN, Edinburgh.

Sinha D, Zemlan FP, Nelson S, Bienenfeld D, Thienhaus O, Ramaswamy G, et al. (1992) *A new scale for assessing behavioral agitation in dementia*. Psychiatry Research. Vol. 41, No.1, pp.73-88.

Streiner DL and Norman GR (2003) *Health measurement scales: A practical guide to their development and use*. (3rd ed.) Oxford University Press, Oxford.

Tariot PN, Mack JL, Patterson MB, Edland SD, Weiner MF, Fillenbaum G, et al. (1995) *The Behavior Rating Scale for Dementia of the Consortium to Establish a Registry for Alzheimer's disease. The Behavioral Pathology Committee of the Consortium to Establish a Registry for Alzheimer's disease*. American Journal of Psychiatry. Vol. 152, No.9, pp.1349-1357.

The American Heritage Stedman's Medical Dictionary (2004) *The American Heritage® Stedman's Medical Dictionary (2nd ed.)*. Houghton Mifflin Company, USA.

The Lund and Manchester Groups (1994) *Clinical and neuropathological criteria for frontotemporal dementia*. Journal of Neurology, Neurosurgery and Psychiatry. Vol. 57, pp.416-18.

USPSTF (U.S. Preventive Service Task Force) (2003) *Recommendations and Rationale: Screening for Dementia*, available at <http://www.ahcpr.gov/clinic/3rduspstf/dementia/dementtr.htm>, accessed 21 June 2006.

Wetterling T, Kanitz RD and Borgis KJ (1996) *Comparison of different diagnostic criteria for vascular dementia (ADDC, DSM-IV, ICD-10, NINDS-AIREN)*. Stroke. Vol. 27, No.1, pp.30-36.

Winblad B, Palmer K, Kivipelto M, Jelic V, Fratiglioni L, Wahlund LO, et al. (2004) *Mild cognitive impairment – beyond controversies, towards a consensus: report of the International Working Group on Mild Cognitive Impairment*. Journal of Internal Medicine. Vol. 256, pp.240-246.

World Health Organisation (1992) *The ICD-10 Classification of Mental and Behavioural Disorders: Clinical descriptions and diagnostic guidelines*. WHO, Geneva.

World Health Organisation (1993) *The ICD-10 Classification of Mental and Behavioural Disorders: Diagnostic criteria for research*. WHO, Geneva.

## 4 Dementia Staging and Descriptive Instruments

### 4.1 *Justification for Selection of Dementia Staging and Descriptive Measures for Review*

Dementia is a syndrome, or a list of symptoms, of progressive decline in multiple areas of cognitive function, leading to significant inability to maintain occupational and social functioning. It often goes unrecognized in the older population due to its multi-faceted nature, and the commonly held belief that the symptoms are a normal part of the ageing process. This lack of recognition by families and some health care staff can impact on the person's quality of life, morbidity and mortality. Enabling caregivers and health staff to recognize and provide timely screening for dementia may result in improved outcomes for the person. Because of the increased burden and suffering that dementia can pose for the person and their family members, recommendations concerning the assessment of dementia are important.

The literature reveals that the initial presentation of dementia includes subtle, or noticeable, changes in cognition, functioning, mood, and behaviour. Studies show that screening for these changes leads to early detection with improved clinical outcomes (Patterson, et al. 1999). Screening usually includes the process of evaluating the person's mental, emotional, and social capabilities. In the clinical setting, the interpretation of clinical and psycho-social assessment conducted by health staff can be complicated by several factors: the person's age, pre-morbid intelligence, education level, cultural background, language skills, psychiatric illness, other illnesses and sensory deficits (Costa, et al. 1996) as well as the severity of the person's dementia.

Although direct assessment of the patient is generally preferred, sometimes it can be difficult to directly assess persons with severe dementia and information provided by proxies such as caregivers and family may be utilised. Family members can be an important source of information about the changes occurring in the patient, nevertheless, they will all have different relationships with the care recipient and thus their perceptions of the extent of the changes occurring can vary considerably. The person who provides the day to day care can also be a good source of information about the changes occurring that affect everyday life for the person with dementia. Their perceptions and experiences with the person with dementia are also invaluable in the screening process. This raises the issue of whether it is better to assess the person directly or use proxy ratings or some combination of both.

### 4.2 *Dementia Staging and Descriptive Measures*

For the reasons cited, the use of standardized assessment procedures and measures (also called rating scales or screening tests) that draw on the perceptions and experiences of family carers is useful in making a preliminary diagnosis of dementia, and for determining how well a patient is doing as the disease progresses. Normative assessment procedures need valid and reliable measures. The measures should be widely tested, be internally consistent, have test-retest consistency, and be both sensitive and specific to the constructs being measured. A number of validated dementia screening measures meeting these criteria are routinely used by health care staff in clinical practice (American Psychiatric Association, 1999).

Dementia staging and descriptive measures are the first level screening measure recommended for assessing the presence of dementia. To monitor the deterioration in the frequency and/or nature of the person's maladaptive behaviour in self-care, social functioning, relationships and cognition, or the recurrence of new behaviours, requires the use of consistent assessment procedures over time. As well, similar behaviours may have very different causes in different people. Therefore, dementia staging and descriptive measures can provide valuable baseline data, assist in monitoring the person's response to interventions, and can also be used to monitor changes occurring (Department of Veterans Affairs, 1997).

As such, dementia staging and descriptive measures are especially suited to harness the intimate knowledge that close family and caregivers have of the person with dementia's abilities, habits and personality to determine changes occurring. They are designed to enable informal and formal caregivers to identify specific features of cognitive decline, such as orientation, memory, attention, thinking, and perception that are implicated in the ability to engage in usual self-care activities, socialization, communication and important life skills. The benefit of their use is that they can be used readily in the community setting, as well as in a variety of health care contexts, and can be completed by non-specialist health staff.

Burns, et al. (2004) indicates these measures are widely used as staging measures in descriptive and intervention studies. It is noted that specialist clinicians are less likely to use these global staging instruments than other clinical or research personnel. Such instruments may not be particularly useful for fine differentiation at an early stage of dementia. However, global functional scales like the GDS and CDR have their place in broadly describing people with dementia; particularly for research purposes and in residential care and community care settings.

The Dementia Staging and Descriptive Measures that possessed these features were sourced from the research literature and relevant dementia measurement texts using a variety of search strategies.

### **4.3 Search Strategies**

Details of the literature search strategies used are outlined in the Introduction (refer Section 2) of this report. The initial search strategy identified 81 measures which could be classified as dementia staging and descriptive instruments. Following the search strategy (textword search), a CD-Rom was produced containing relevant papers and abstracts for each identified instrument.

Based on this work an impact sheet was developed for consideration by the review teams and the DOMS-EMG. This considered MEDLINE, text and web impacts, presence in instrument databases (PROQOLID) and its use in clinical practice. The latter was based on NEP and field surveys and clinical feedback. This process produced a list of 12 or so instruments which were regarded as leading contenders for comprehensive review.

### **4.4 Selecting the Measures for Comprehensive Review**

The process of selecting the "best" five dementia staging and descriptive measures included four distinct steps: the literature search for dementia staging and descriptive measures, assessing the impact of the identified measures, selection of those measures meeting the criteria determined by the Expert Measurement Group, undertaking a critique of the psychometric properties and functionality of the "best" five selected measures that met the selection criteria, and rank ordering these measures.

The literature Master file was first scanned to identify those cited as dementia staging and descriptive measures. This selection proved useful initially in locating global dementia measures, as distinct from cognitive and behaviour measures and measures of function. A second search was conducted for additional material using the databases MEDLINE, Psycinfo, CINAHL (Cumulative Index to Nursing and Allied Health), Proquest Health Sciences, Proquest5000, Expanded Academic Index, Web of Science, Google Academic, Google and BIOSIS. This was used to identify the unique psychometric properties of these instruments and comparisons with other measures such as the Mini Mental State Examination, and their application within clinical and community settings.

A list of key terms was identified for use when searching the literature to locate particular dementia staging and descriptive measures. The key terms included those pertinent to the type of measures required and the illness category, and included: dementia measures, dementia instruments,

dementia measurement, global dementia measures, dementia-specific screening, measuring global cognitive function, and global psycho-geriatric assessment. These key terms produced a large body of literature, generally abstracts of papers that related to dementia testing for use in clinical practice and research, including pathological testing of brain cells. Up to 196 articles were identified in the literature for each of the different global dementia screening measures used in clinical practice. Only peer-reviewed articles reporting measures of global cognitive function used in the clinical or community setting were selected for review. This process yielded 76 studies reporting the use of global dementia measures for use in clinical practice.

To gain further information about the availability, cost and application of the first cull of global dementia measures, a search was made using a number of relevant web sites, as well as personal communication with Maryann Urbashich, Associate Director, Library, Alzheimer's Association Green-Field Library.

The initial list of 19 potential dementia staging and descriptive measures cited most frequently in the literature was reviewed further, and included: Cambridge Cognitive Examination Revised (R-CAMCOG), Sandoz Clinical Assessment Geriatric (SCAG), Pfeiffer Short Portable Mental Status scale, Telephone Interview for Cognitive Status (TICS), Neuropsychiatric Inventory (NPI), CERAD-BRSD, Psychogeriatric Assessment Scales (PAS), Structured Interview for Diagnosis of Alzheimer's disease (STIDA), Dementia Care Mapping (DCM), Alzheimer's Disease Cooperative Study-Clinical Global Impression of Change (ADCS-CGIC), Canberra Interview for the Elderly (CIE), Milan Overall Dementia Assessment (MODA), Dementia Severity Rating Scale (DSRS), Mattis Dementia Rating Scale (DRS), Blessed Dementia Scale (BDS), Global Deterioration Scale (GDS), Clinical Dementia Rating Scale (CDR), and the Alzheimer's Disease Assessment Scale (ADAS).

#### **4.5 Selecting Contender Instruments for Review**

Before proceeding to the next phase of review the Expert Measurement Group developed an impact sheet from the master file which compared the instruments with regard to: accessibility, cost, number of citations in the literature, report of psychometric properties and evidence of reliability and validity, 30 minutes or less administration time, applicability as a global measure, able to be administered by a range of raters with varying levels of expertise in dementia assessment, use in clinical practice, and also to identify reports of their use in the Australian context. To locate this information on each of the dementia-specific assessment measures identified in the literature search, the literature and web-site information was reviewed in greater depth.

Many of the measures originally located were identified as dementia batteries, containing two or more scales, and allowing for assessment of many specific as well as global measures of dementia. An example of a battery is the Cambridge Mental Disorders of the Elderly Examination (CAMDEX) and Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Because of their length, these did not fit the criteria for being undertaken in 30 minutes or less. As well, many of these batteries also contained some of the measures already selected for review. On closer examination some measures were found to be primarily an assessment of cognitive symptoms, e.g. Alzheimer's Disease Assessment Scale (ADAS), rather than an overall assessment of dementia. As these instruments would be considered when examining the cognitive assessment instruments, they were not considered further in this category of instruments. Therefore, all instruments that were not reported in the literature to be dementia-specific assessment measures were removed from the first list generated.

The number found suitable for use as a single measure of global cognitive function was reduced to 12, after applying the Expert Measurement Group impact sheet criteria. The measures that met most of these criteria are listed in Table 19 below.

**Table 19 First-level Assessment of Global Dementia Measures**

Instrument	Available	No. of citations	Psychometric articles	Administration time	Applicability	Domains & Subdomains	Judgment	Clinical Practice
<b>1. Sandoz clinical Assessment Geriatric (SCAG)</b>	yes	129	reliability 2, validity 2	7min	global	18 cardinal signs/symptoms of dementia. 7 point scale of severity.	recommended	clinician use
<b>2. Structured Interview for Diagnosis of AD</b>	yes	10	reliability 3, validity 3	30 mins full, 10 mins short version	global	10 items (short version).	not recommended	community dwelling interview
<b>3. Dementia Care Mapping (DCM)</b>	yes	56	reliability 2, validity 2	30 min-5 days, normally 6 hours	activity, mood, behaviour	activities, mood, affect, behaviour.	not recommended	inpatient, day centre, residential care
<b>4. Alzheimer's Disease Cooperative Study- (ADCS-CGIC) Clinical Global Impression of Change</b>	yes	20	reliability 3, validity 3	3 sections, 20 minutes each	global/cognition, measures change	3-parts: 1. guided baseline interview admin to patient & informant.	not recommended	clinician use informed by caregiver
<b>5. Canberra Interview for the Elderly (CIE)</b>	yes	24	reliability 3, validity 3	15 mins	global	standardised- diagnostic interview-community.	not recommended	clinician & caregiver
<b>6. Milan Overall Dementia Assessment (MODA)</b>	yes	20	reliability 3, validity 3	30 mins	global	field instrument-processed by computer.	not recommended	clinician use
<b>7. Dementia Severity Rating Scale (DSRS)</b>	yes	21	reliability 3, validity 3	15 mins	global, cognitive, ADLs	multiple choice survey-deficits in 3 domains.	recommended	clinician, caregiver
<b>8. Mattis Dementia Rating Scale</b>	yes	204	reliability 3, validity 3	30-45 minutes	global, ADLs	neuropsychological test/different dementias.	not recommended (copyright and cost prohibitive)	clinician
<b>9. Dementia Rating Scale-2 (improved version of MDRS)</b>	yes	8	reliability 3, validity 3	15-30 minutes	global, ADLs	Neuropsychological test.	not recommended	clinician
<b>10. Blessed Dementia Scale (BDS)</b>	yes	91	reliability 2, validity 3	30 minutes	global & behaviour	Cognition & behaviour.	recommended	clinician
<b>11. Global Deterioration Scale (GDS)</b>	yes	274	reliability 3, validity 3	2-10 minutes	global	global ratings.	not recommended	
<b>12. Clinical Dementia Rating Scale (CDR)</b>	yes	186	reliability 2, validity 3	40 minutes	global	global ratings.	recommended	

From this first-level review, five measures were selected for in-depth critique, based on evidence in the literature of having moderate to good psychometric properties, applicable for use in the clinical and community setting, able to be used by carers, caregivers and clinicians with different levels of expertise, ease of use, time taken to administer (less than 30 minutes), low or no cost, ease of access and reported use within different cultures. Only measures with more than three quality measurement testing procedures used to establish reliability and validity were selected for the five best global measures. The “best five” meeting all of these criteria included: Sandoz Clinical Assessment – Geriatric (SCAG), Dementia Severity Rating Scale (DSRS), Blessed Dementia Scale, the Global Deterioration Scale (GDS) and the Clinical Dementia Rating Scale (CDR). The summary sheets outlining these features in the five instruments selected for further review are provided in Tables 20 and 21 below.

**Table 20 Summary Sheet - Selected Dementia Staging and Descriptive Instruments**

Instrument	Domains/Sub domains	Applicability/Stage	Patient	Proxy	Availability/Cost	Training/Manual	Admin time
SCAG Sandoz Clinical Assessment Geriatric	Cardinal signs/symptoms of dementia, including somatic (psychopathology) and self-care dimensions	Mild to moderate	Some questions answered by patient	✓clinician asks carers questions	Accessible and free	Some training required if administered by caregiver.	15-20 minutes
DSRS Dementia Severity Rating Scale	Cognitive and functional dimensions focused on how well the person with dementia functions in the home environment. Contains 11 domains	Mild to severe All stages	Caregiver Rating Scale	✓For carer use in home setting, or by clinician	Accessible and free	No training required. Can be used by caregiver alone, with detailed script for administration by clinician.	4-5 minutes
BDS Blessed Dementia Scale	BDS Incorporates Blessed Information-Memory-Concentration Test and the Dementia Scale. Measures changes in everyday performance, habits and personality over past 6 months	Mild to moderate All stages	Some questions answered by patient	✓clinician asks carers questions	Accessible and free	No training required.  Familiarity with assessment of cognitive function desirable.	20-30 minutes
GDS Global Deterioration Scale	GDS is main part of the Global Deterioration Scale Staging System. Measures stages (1-7) of cognitive and psychiatric function.	Very mild to severe All stages	Some questions answered by patient	✓clinician can ask carers questions	Accessible and free	Minimal training required. Easily used by caregiver.	5-10 minutes
CDR Clinical Dementia Rating	Assesses memory, orientation, judgment and problem solving, community affairs, home and hobbies, personal care	All stages	Some questions answered by patient	✓clinician can ask carers questions	Accessible and free	Training required. Free training provided on line, videotapes can be purchased, on-site training also offered.	40-75 minutes

**Table 21 Summary Sheet - Evaluation of Selected Dementia Staging and Descriptive Instruments**

Instrument	Citations	Psychometrics	Use in Practice (to date)	Judgements/Comments
Sandoz Clinical Assessment Geriatric (SCAG)	61 journals 2 books	Reliability is moderate  Validity is moderate	Assesses psychopathology and care ability in persons with Dementia / AD. Detects subtle levels of change and severity in 18 common symptoms associated with dementia.	Compares favourably with other measures of dementia, including the MMSE. It has moderate psychometric properties, has been extensively cited, and is easy to access and administer.  The instrument has also been used as an outcome measure in intervention studies.  Can be used by clinical staff with a range of expertise.
Dementia Severity Rating Scale (DSRS)	21 journals 2 books	Reliability is good  Validity is good	Assesses psychopathology, care ability, social functioning and community involvement in persons with Dementia/AD from the carer's perspective.  Provides carers and clinicians with a rough idea of ability, changes and severity in 11 domains.  Assists in diagnosis for those who cannot be brought in for closer examination by physician.	It has moderate to good psychometric properties.  Very limited use for research purposes as an outcome measure.  Is brief and easy to administer and is user-friendly.  Can be used by carers and clinical staff with a range of expertise.
Blessed Dementia Scale (BDS)	91 Journals 2 books	Reliability is good  Validity is moderate	Assesses psychopathology, care ability, habits and personality in persons with Dementia/AD.  Measures levels of ability and change for those who can not be assessed with detailed neuropsychological measures	Compares favourably with other measures of dementia, including the MMSE.  It has moderate to good psychometric properties, however some authors have found it does not accurately detect levels of cognitive change when evaluating individual responses.

			due to the severity of their dementia.	<p>Is easy to access and administer.</p> <p>Detects changes in persons with very low cognitive deficits.</p> <p>The instrument has also been used as an outcome measure in intervention studies.</p>
Global Deterioration Scale (GDS)	274 journals 1 book	Reliability is good  Validity is moderate	<p>Assesses psychopathology, care ability and levels of behavioural disturbance in persons with Dementia/AD.</p> <p>Provides clinicians with a measure of severity through 7 stages of dementia.</p> <p>Prognostic measure of psychiatric function disorder.</p>	<p>It has very good psychometric properties.</p> <p>Used for research purposes as an outcome measure.</p> <p>High score correlations for individual items and total score against other measures, including the MMSE.</p> <p>It is easy to access and administer.</p> <p>Most suitable for use by clinicians but can be used by care staff with a range of expertise.</p>
CDR  Clinical Dementia Rating Scale	362 journals  3 books	Very good	<p>Assesses:</p> <p>Clinically stages the severity of cognitive impairment in dementia.</p> <p>Effectiveness of drug treatments in clinical trials, and other interventions.</p> <p>Research studies of normal elderly and those with dementia.</p>	<p>Psychometric properties are very good.</p> <p>Widely used as an outcome measure in clinical trials.</p> <p>Has become one of the standard global ratings in dementia.</p> <p>Numerous translations available.</p> <p>Easy to access but administration requires training.</p>

## **4.6 The Process of Reviewing the Best Five Measures**

To obtain the best available literature on these five selected measures and to access the original measures and instruction manuals, the library was requested to locate full text of research papers not able to be accessed through databases. Some of these papers were only available on payment of a fee. Contact was also made by email with some of the researchers (overseas) to request if they would be willing to provide full papers not accessible through the library databases. As an example, email contact was made with Alzheimer's disease Disorders Association in the USA. The reference list for this Section lists the literature reviewed in detail for the five selected measures.

The selected measures and instructions for use were first read to become familiar with them. Then the most informative research articles were accessed from the available literature on each measure. These were read thoroughly to undertake an in-depth critique of their reported use and suitability as a global dementia measure. This information was summarized for each selected measure on the AHOC Instrument Review Sheet (refer Appendix 3), which had been revised by the Expert Measurement Group. Appendix 5 contains the comprehensive reviews undertaken for these instruments.

The psychometric properties of each measure were compared with the Mini-Mental State Examination (MMSE), which was reported in all literature reviewed to have variable, yet high, sensitivity. The sensitivity range of the selected measures reported in the literature against the MMSE was between .65 and .96. Reported specificity for the MMSE ranged from .74 to .99. Consequently, in the selected literature all of the "best" five measures compared favourably with the MMSE. However, there were variations in the available literature on the tests used to establish the reliability and validity of these measures, including sample size, population sampled and the designation of persons undertaking the tests/studies. Estimates of sensitivity and specificity were derived from those studies with larger sample sizes; however, not all were conducted with samples with a confirmed diagnosis of dementia, rather of possible/probable dementia.

## **4.7 Strengths and Weaknesses of the Selected Instruments**

The critique process of the selected instruments revealed a number of strengths and weaknesses, however, each was found suitable for use by clinicians and care staff in a variety of clinical settings. The Dementia Severity Rating Scale was reported as user-friendly for use by family carers in the community setting. The literature provides some evidence of their use with persons of different cultures, using the services of translators. It is worth noting that there were inconsistencies in the literature reporting the use of these instruments in clinical settings and in research, and when undertaking psychometric property testing. Reliability and validity ratings, for example, varied between good and moderate. The particular characteristics of each instrument are described, as follows.

### **4.7.1 Global Deterioration Scale (GDS)**

The Global Deterioration Scale (GDS) was developed by Reisberg, Ferris, de Leon and colleagues in 1982 to provide caregivers and clinical staff with an idea of the stage of the person's dementia by observing that individual's behavioural characteristics. The GDS is the main part of a clinical rating system called the Global Deterioration Scale Staging System. Three independent measures are included in the Staging System: the GDS, the Brief Cognitive Rating Scale (BCRS) and the Functional Assessment Staging System (FAST).

The GDS provides caregivers with an overview of the stages of cognitive function for those suffering from a primary degenerative dementia. It comprises 7 different stages. Stages 1-3 are

the pre-dementia stages in which the person is able to function quite well in their daily lives, or as usual. Stages 4-7 are the stages which reveal loss of cognitive and other functions that are needed for successful living. Beginning at stage 5, an individual can no longer survive without assistance. Bakker, et al. (2004) identified the GDS as a prognostic measure for psychiatric function disorders, including paranoia and somatic co-morbidity.

Several authors attest to the clinical usefulness of the GDS in assessing the presence of dementia. It has high score correlations for individual items and total scores when compared with other measures, like the Mini-Mental State Examination, when used to identify the presence of dementia and the level of impairment in activities of daily living and cognitive functioning. It is easy to understand and score by caregivers and clinical staff at different levels of expertise. Nevertheless, while the face validity of the GDS is high, guidelines for assigning patients ratings has not been explained in great depth and so empirical validity is not so well-documented (Kane and Kane, 2000). Eisdorfer, et al. (1992) also pointed out that in the development of the GDS, there was no explicit discussion of how Reisberg, et al. (1982) related the stages described in the scale to the progression of dementia, instead basing the scale on descriptions of observations of persons with dementia. Despite these criticisms however, the GDS is regarded by clinicians, caregivers and researchers as a very useful dementia screening instrument for use in a variety of settings.

#### **4.7.1.1 Functional Assessment Staging (FAST) (Reisberg, 1988)**

The Functional Assessment Staging Scale (FAST) is a dementia rating scale of functional changes in 7 major changes with a total of 16 successive stages and substages (Burns, et al. 1999). The stages range from 1 (no difficulties, either subjectively or objectively) to 7d (unable to hold head up). The tester is able to identify the presence as well as the number of months the person has been at the current stage of functioning in about 2 minutes. The FAST was developed in light of a considerable body of research and clinical observation that functional detriments in dementia proceed in a hierarchical ordinal pattern, reflected in the FAST scale (Sclan and Reisberg, 1992). Reliability has been demonstrated with intra-class correlations of above 0.85. Concurrent validity has been assessed against the Global Deterioration Scale (GDS) and a number of neuropsychological tests (Burns, et al. 1999). As the FAST identifies a total of 11 subscales of the later stages of the GDS, it is useful in providing detailed staging in late stage dementia. Consequently, the FAST can be used as part of the GDS.

Another favourable feature of the FAST is its utility in community and clinical settings. The well-described stages make it suitable for use with trained and untrained staff and family carers as an initial screen of cognitive and physical functioning, and consequently as an aid to care planning and function monitoring. Since care planning is based on assessment of ability to perform activities of daily living, staff will take into consideration the person's level of alertness, attention, memory, thinking ability, perception, psychomotor behaviour and higher cognitive functions. These abilities are easily identified by carers and care staff with the FAST in a very short time span- one to two minutes. Its ease of use and alignment with the GDS make it an attractive measure of function in all care settings.

#### **4.7.2 Clinical Dementia Rating Scale (CDR)**

The Clinical Dementia Rating (CDR) (Hughes, Berg, Danziger, Coben, and Martin, 1982; Morris, 1993) has become one of the standard global ratings of dementia used in studies investigating this disease. It is used in both research and clinical settings to characterize the level of cognitive and functional performance in patients at risk for, or suspected of having, Alzheimer's disease or other dementia disorders, and to clinically stage the severity of cognitive-functional impairment. The CDR is available in a number of languages (refer to the instrument review sheet in Appendix 5). A version suitable for use in chronic care facility settings is also available (Marin, et al. 2001).

The CDR was developed by a team of physicians experienced in the field and comprises a total of 75 items, 48 for the informant or collateral source (CS) and 27 for the person with dementia (PD). These items assess the level of impairment in six domains: Memory (CS = 15 items, PD = 10 items); Orientation (CS = 8 items, PD = 8 items); Judgment and Problem Solving (CS = 6 items, PD = 9 items); Community affairs (CS = 10 items); Home and hobbies (CS = 5 items) and Personal care (CS = 4 items). A validation study confirmed that all domains are adequately covered (Hughes, Berg, Danziger, Coben, and Martin, 1982).

The instrument is an interviewer administered semi-structured interview of both the patient and a reliable informant, or collateral source, who is usually a close family member. It is administered by a clinician, either a physician or other health professional and takes about 40-75 minutes depending on the level of impairment. Training is required as the administration and scoring is quite complex.

CDR ratings are 0 for no impairment (normal), 0.5 for very mild/questionable dementia and 1, 2 and 3 for mild, moderate, and severe dementia. The CDR table in Appendix 4K provides descriptive anchors that guide the clinician in making these ratings based on the interview data and clinical judgment. Items are scored as a decline from a previous level due to cognitive impairment, not impairment due to other factors or causes. The global score is achieved by first assessing each domain separately using the same levels (in the CDR Table).

The available evidence indicates the CDR has very good psychometric properties. The majority of the studies have reported considerable information to ensure the findings can be appropriately interpreted. Inter-rater reliability was assessed in several studies, mostly between physicians (or an expert physician) (Burke, et al. 1988; Haroutunian, et al. 1998; Hughes, Berg, Danziger, Coben, and Martin, 1982; Marin, et al. 2001; Morris, 1997; O'Connor, et al. 1996; Rockwood, Strang, MacKnight, Downer, and Morris, 2000; Schafer, et al. 2004; Summers, DeBoynon, Marsh, and Majovski, 1990; Tractenberg, Schafer, and Morris, 2001). Reliability between nurse ratings and / or nurse and physician ratings (McCulla, et al. 1989), and between physician and a reliable informant such as a relative (Waite, et al. 1999) was also assessed. Overall, the studies found good to very good inter-rater reliability with global kappas ranging from 0.50 to 0.91 and domain kappas from 0.27 to 1.00. ICCs reported ranged from 0.99 to 0.88. The low kappa of 0.27 was found for ratings between naïve physicians with little or no training (Tractenberg, Schafer, and Morris, 2001). The only study found reporting test-retest (at 1 month interval) reliability cited intra class correlations (ICC) for the domains ranging from 0.86 to 0.93 and a global ICC of 0.92 (Marin, et al. 2001).

No studies were found that reported factor analysis or correlations between items and or scales within the instrument. However the correlations with scales and or items of other well known tests which form part of the CDR such as the Blessed Dementia Scale (BDS), Short Portable Mental Status Questionnaire (SPMSQ) and Face-Hand Test (FHT) indicate the internal structure of the instrument is valid (Berg, et al. 1988; Botwinick, Storandt, and Berg, 1986; Botwinick, Storandt, Berg, and Boland, 1988; Davis, Morris, and Grant, 1990; Dooneief, Marder, Tang, and Stern, 1996; Hughes, Berg, Danziger, Coben, and Martin, 1982; Marin, et al. 2001; Morrism, et al. 1989; Morris, McKeel, Fulling, Torack, and Berg, 1988).

Evidence that the instrument has construct validity also comes from numerous studies citing correlations with other measures. The CDR showed expected correlations with the following measures of cognitive functioning: Face-Hands Test (FHT), Mini-Mental State Exam (MMSE), Abbreviated Mental Test (AMT), Short Portable Mental Status Questionnaire (SPMSQ), Elderly Cognitive Assessment Battery (ECAQ), Short version of Blessed Information, Memory and Concentration Test (sBIMC) and the Dementia Scale – Cognitive (DS-C). The instrument also shows expected correlations with other instruments assessing areas of functioning expected to be affected by dementia, namely: the Physical Performance test (PPT), Aphasia Battery (AB), Alzheimer's Disease Co-operative and Schwab and England (ADL), and the Visual Analogue scale

(VAS), as well as items measuring general cognitive and physical functioning, and neuropsychological and psychopathology symptoms. Finally, the CDR has been shown to correlate with another global measure of dementia, the Blessed Dementia Scale (BDS).

In terms of discriminant validity, evidence from post-mortem studies show that CDR stage is significantly associated with hippocampal volume loss and shape, neurons in the entorhinal cortex, Alzheimer's disease lesions (both gross and microscopic) density of senile plaques and neurofibrillary tangles (Berg, McKeel, Miller, Baty, and Morris, 1993; Berg, et al. 1998; Gomez-Isla, et al. 1996; Haroutunian, et al. 1998; Morris, McKeel, Fulling, Torack, and Berg, 1988; Morris, et al. 1991; Morris, et al. 1996; Wang, et al. 2003). The instrument's ability to discriminate between disease severities was also confirmed by studies reporting its predictive ability. CDR stage has been shown to be a good predictor of dementia progression, entry into nursing home placement and survival (Berg, McKeel, Miller, Baty, and Morris, 1993; Berg, et al. 1998; Berg, et al. 1988; Dooneief, Marder, Tang, and Stern, 1996; Fillenbaum, Peterson, and Morris, 1996; Gomez-Isla, et al. 1996; Haroutunian, et al. 1998; Heyman, et al. 1987; Hughes, Berg, Danziger, Coben, and Martin, 1982; Morris, McKeel, Fulling, Torack, and Berg, 1988; Morris, et al. 1991; Morris, et al. 1996; Morris, et al. 2001; Wang, et al. 2003).

The sensitivity of the instrument has been supported by evidence from clinical trials. Results from numerous trials investigating the effectiveness of drug treatment for persons with Alzheimer's disease (Burns, et al. 1999; Cortes, et al. 2005; Imbimbo, Troetel, Martell, and Lucchelli, 2000; Jones et al., 2004; Riepe, et al. 2006; Rockwood, 2004; Tariot, et al. 2001; Zemlan, 1996) show that CDR scores improved significantly after 6 months of treatment. Scores on the MMSE and the ADAS-Cog also improved, confirming the sensitivity of CDR. Studies investigating the effectiveness of Donepezil for persons with vascular dementia also show improvement in CDR scores along with improved scores on the ADAS-Cog MMSE, and the Clinician's Interview Based Impression of Change (CIBIC) (Malouf and Birks, 2004; Roman, et al. 2005). No major floor or ceiling effects were reported in these studies.

Although it can be seen that it has good psychometric properties it involves a much lengthier assessment than most other instruments reviewed in this category. It may be more suitable for a more detailed follow-up assessment rather than an initial assessment. It is also widely used in research studies.

### **4.7.3 Dementia Severity Rating Scale (DSRS)**

The Dementia Severity Rating Scale was developed by Clark and Ewbank in 1996 to identify how well the person with dementia functions in the home environment. It provides a brief multiple-choice questionnaire for caregivers to assess the mildest to the most severe stages in the major functional and cognitive domains affected in dementia. The caregiver is asked to rate the person with dementia in 11 categories. The first six address memory, orientation, judgment, social interaction, home activities and personal care. These mirror the items in the Clinical Dementia Rating (CDR) scale. The other five items address language, recognition, eating, incontinence and mobility. The total score is obtained from the summed score across all items. The range of scores for some of these categories is from 0 (normal) to 7 (very impaired ability/skill), whereas for other item there are only a possible four responses. Item 11 (mobility/walking) is sometimes skipped by informants because a pre-existing mobility impairment in the subject can make it difficult to estimate how cognitive decline affects his or her mobility in the community. Users are advised to examine completion of individual items before using the composite scores.

Although designed as a caregiver rating scale, the DSRS is also used by clinicians and researchers internationally as a first level assessment measure. Its brevity and user-friendly approach make it one of the most suitable informant questionnaires to detect dementia and measure severity. However, a comment made by Harvey, et al. (2005) was that the "DSRS uses an inconsistent format and language, thus may be too complex for the average reader". There is

also limited clinical reference data and normative data published for this instrument. As the instrument is used as a first level screen for the presence of dementia in the community setting, it may not be considered adequate for use by clinicians who wish to assess cognitive function in greater detail.

#### **4.7.4 Blessed Dementia Scale (BDS)**

The Blessed Dementia Scale (BDS) was developed by Blessed, Tomlinson and Roth in 1968. It incorporates the Blessed Information-Memory-Concentration Test (BIMC) and the Dementia Scale (DS). The BDS measures changes in everyday performance habits and personality in the person with dementia. It was designed to quantitatively assess the signs of dementia to enable comparisons to be made with pathological changes occurring in the brain. The BDS is suitable for use by nursing staff with no neuropsychological training, for example those working in nursing homes and the community. The BIMC component contains 30 items which assess clinical functions of dementia related to neuro-pathological change. Scores range between 0 (complete cognitive failure) to +37 (full cognitive capacity) The BIMC is answered, if possible, by the person with dementia but personal memory information must be obtained from a collateral source, such as a caregiver. It rates orientation, long-term memory, recall, concentration and performance, by identifying competence in personal, domestic and social activities during the preceding six months. The Dementia Scale (DS) component contains 22 items which assess changes in everyday performance, habits and personality. The scores can range between 0 (fully preserved capacity) and +28 (extreme incapacity). Given the number of items it is a quite lengthy assessment taking approximately 30 minutes to administer.

Many studies using the BDS provide convincing evidence for its utility and its sub-component the BIMC in accurately assessing the incidence and severity of dementia in a range of community and health care settings. The BDS and its sub-components are widely used as a comparative measure when testing other outcome measures. However, an article by Holmes and Lovestone (2003) found in their study of 374 Alzheimer's disease patients that the BDS has little value in detecting the rate of cognitive change when evaluating individual treatment responses. Stern (1990) also cautions against relying on the BDS to detect functional change in persons with dementia since disparate functional domains are assessed. Stern argues for the use of a multi-factorial approach to the assessment of functional capacity for this reason.

#### **4.7.5 Sandoz Clinical Assessment – Geriatric (SCAG)**

The SCAG was developed by Shader, Harmatz and Salzman in 1974 to ensure the diagnostic differentiation between early dementia and depressive disorders in the older population, by assessing early cognitive and related deterioration in the older person's ability to engage in daily life activities. It is also used to assess changes in these areas following treatment. The SCAG assesses psychopathology in the areas of mood and depression, confusion, mental alertness, motivation/initiative, irritability, hostility, being bothersome, indifference, unsociability, uncooperativeness, emotional lability, fatigue, self-care, appetite, anxiety, recent memory and disorientation. The inventory of 18 target symptoms (items) of dementia is scored by severity. Each item is rated by 7-point scale covering 4 areas: global cognition, mood and behaviour, ability to cope with activities of daily living and somatic symptoms. Scores range from 1 (not present) to 7 (severe). These areas can be assessed by caregivers and provide clinicians with a close approximation of global cognition, mood and behaviour, ability to cope with activities of daily living and somatic symptoms, such as fatigue, dizziness and poor appetite.

The SCAG is a useful tool for caregivers and clinicians to detect subtle changes in cognition and functioning by targeting 18 common symptoms associated with dementia, when compared with other measures such as the MMSE. It has also been widely used in clinical and psychopharmacological research. The major benefit of use in clinical practice lies in its utility to family caregivers and health staff with limited expertise in dementia assessment. However, the

SCAG has also been referred to by Lezak, et al. (2004) as psychometrically deficient compared to the Alzheimer Disease Assessment Scale, and McDowell (2006) also noted that the SCAG has its critics (e.g. Salzman, 1983). Furthermore, while the SCAG has been used widely for research as an outcome measure, Warburton and Rusted (1989 as cited in Curran and Wattis, 1997) felt that the SCAG should not be used alone in clinical trials as it is not precise or objective enough to pick up subtle changes in participants.

#### 4.8 Summary of Instrument Scores and the Comparative Ranking of Instruments

Once the in-depth critique of each of the “best” five measures was undertaken, this information was collated on a summary table of the criteria and weights for instrument ranking. For each of the 11 criteria the measures were ranked from 1 (indicating not meeting the criteria) to 3 (meeting the criteria). The weight for each of these criteria varied from a possible 1 - 3. The final score for each criterion is calculated by multiplying the score with the allocated weight for each criterion. Scores and weighted total scores for each of the selected instruments are listed in Table 22 below.

**Table 22 Summary of Ratings for Dementia Staging and Descriptive Instruments**

Criteria	Weight	Instrument				
		GDS	CDRS	DSRS	Blessed	Sandoz
Availability of comparison data	3	2.5	2.5	1.5	2	2
Length/feasibility of instrument for inclusion in battery	2	3	1	3	1	2
Complexity of administration/ cognitive burden	2	3	2	3	2	3
Cultural Appropriateness	1	2	3	2	2	2
Ease of obtaining score	2	3	3	2	3	3
Sensitivity to dementia	3	2	3	2	3	2
Reliability evidence	3	3	3	3	2	2
Validity evidence	3	3	3	3	3	2
Cost of the instrument	2	3	3	3	3	3
Cost of instrument administration	2	2	1	2	2	2
<b>Weighted Total</b>		<b>61.5</b>	<b>57.5</b>	<b>56.5</b>	<b>52</b>	<b>50</b>

#### 4.9 Recommendations Concerning Dementia Staging and Descriptive Instruments

The top scoring measure was the Global Deterioration Scale (GDS) followed by the Clinical Dementia Rating scale (CDR) and the Dementia Severity Rating Scale (DSRS). The GDS has good psychometric properties, is simpler to use than the CDR, can be used by both para-professional staff as well as clinicians, and the assessment is shorter and takes far less time to

complete. Thus for use across a wide variety of settings the GDS is the preferred instrument. However, if a more detailed assessment is required for either clinical or research purposes then the CDR might be preferred in these contexts.

The DSRS was designed for use by carers in the community setting, or by care staff and clinicians with varying levels of skill in dementia screening and this instrument is also suitable for first level screening particularly in community and residential care settings.

## References

- American Psychiatric Association (1999) *Practice Guidelines for the treatment of patients with Alzheimer's disease and other dementias of later life*. American Journal of Psychiatry. Vol. 154, No.5, pp.1-39.
- Auer S and Reisberg B (1996) *Reliability of the Modified Ordinal Scales of Psychological Development: a cognitive assessment batter for severe dementia*. International Psychogeriatrics. Vol. 8, No.2, pp. 225-31.
- Auer SR, Montreiro IM, et al. (1996) *The Empirical Behavioral Pathology in Alzheimer's Disease (E-BEHAVE-AD) Rating Scale*. International Psychogeriatrics. Vol. 8, No.2, pp. 247-66.
- Bakker T, Duivenvoorden H, van der Lee J, Schudel W (2004) "Life expectancy following Psychogeriatric Reactivation" *Identification of Prognostic Characteristics of Survival Assessed on Admission*. Dementia and Geriatric Cognitive Disorders. Vol. 18, pp.24-31.
- Baladi J, Bailey P, et al. (2000) *Effectiveness of Rivastigmine for Alzheimer's disease: Canadian Interpretation of intermediate outcome measures and cost implications*. Clinical Therapeutics. Vol. 22, No.12, pp.1549-61.
- Berg L, McKeel DW, Jr., Miller JP, et al. (1993) *Neuropathological indexes of Alzheimer's disease in demented and nondemented persons aged 80 years and older*. Archives of Neurology. Vol. 50, No.4, pp.349-58.
- Berg L, McKeel DWJ, Miller JP, et al. (1998) *Clinicopathologic studies in cognitively healthy aging and Alzheimer's disease: relation of histologic markers to dementia severity, age, sex, and apolipoprotein E genotype*. Archives of Neurology. Vol. 55, No.3, pp.326-35.
- Berg L, Miller JP, Storandt M, et al. (1988) *Mild senile dementia of the Alzheimer type: 2. Longitudinal assessment*. Annals of Neurology. Vol. 23, No.5, pp.477-84.
- Blessed G, Tomlinson BE and Roth R (1968) *The association between quantitative measures of dementia and senile change in cerebral grey matter of elderly subjects*. British Journal of Psychiatry. Vol.114, pp.797-811.
- Botwinick J, Storandt M and Berg L (1986) *A longitudinal, behavioral study of senile dementia of the Alzheimer type*. Archives of Neurology. Vol. 43, No.11, pp.1124-7.
- Botwinick J, Storandt M, Berg L, et al. (1988) *Senile dementia of the Alzheimer type. Subject attrition and testability in research*. Archives of Neurology. Vol. 45, No.5, pp.493-6.
- Burke WJ, Miller JP, Rubin EH, et al. (1988) *Reliability of the Washington University Clinical Dementia Rating*. Archives of Neurology. Vol. 45, No.1, pp.31-2.
- Burns A Lawler B and Craig S (1999) *Assessment Scales in Old Age Psychiatry (2<sup>ND</sup> edn.)*. London: Martin Dunitz.
- Burns A, Rosso RM, Hecker J, et al. (1999) *The effects of donepezil in Alzheimer's disease - results from a multinational trial*. Dementia and Geriatric Cognitive Disorders. Vol. 10, No.3, pp.237-44.
- Clark C and Ewbank D (1996) *Performance of the Dementia Severity Rating Scale: A Caregiver Questionnaire for Rating Severity in Alzheimer Disease*. Alz Dis and Ass Dis. Vol. 10, No.1, pp.31-39.

- Clark CM, Ewbank D, Lerner A, Doody R, Henderson VW, Panisset M, Morris JC, Fillenbaum GG and Heyman A (1997) *The relationship between extra-pyramidal signs and cognitive performance in patients with Alzheimer's disease enrolled in the CERAD study.* Neurology. Vol. 49, No.1, pp.70-75.
- Cortes F, Gillette-Guyonnet S, Nourhashemi F, et al (2005) *Recent data on the natural history of Alzheimer's disease: results from the REAL.FR Study.* Journal of Nutrition, Health and Ageing. Vol. 9, No.2, pp.86-93.
- Costa PT, Williams TE, Somerfield M. et al. (1996) *Recognition and initial assessment of Alzheimer's disease and related dementias.* Clinical Practice Guidelines. No.19. Rockville, MD.: US Department of Health and Human Services, Public Health Service, Agency for Health Care Policy and Research.
- Davis P, Morris J and Grant E (1990) *Brief screening tests versus clinical staging in senile dementia of the Alzheimer type.* Journal of the American Geriatric Society. Vol. 38, No.2, pp.129-135.
- Department of Veterans Affairs, University Health System Consortium, Federal Government Agency (1997) *Dementia identification and assessment. Guidelines for primary care practitioners.* National Guidelines Clearinghouse. (on line at [www.guideline.gov](http://www.guideline.gov))
- Dooneief G, Marder K, Tang MX, et al. (1996) *The Clinical Dementia Rating scale: community-based validation of "profound" and "terminal" stages.* Neurology. Vol. 46, No.6, pp.1746-9.
- Fillenbaum GG, Peterson B and Morris JC (1996) *Estimating the validity of the clinical Dementia Rating Scale: the CERAD experience. Consortium to Establish a Registry for Alzheimer's Disease.* Ageing-Clinical and Experimental Research. Vol. 8, No.6, pp.379-85.
- Gomez-Isla T, Price JL, McKeel Jr DW, et al. (1996) *Profound Loss of Layer II Entorhinal Cortex Neurons Occurs in Very Mild Alzheimer's disease.* Journal of Neuroscience. Vol. 16, No.14, pp.4491-4500.
- Grasel E (2002) *When Home Care Ends-Changes in the Physical Health of Informal Caregivers Caring for Dementia Patients: A Longitudinal Study.* JAGS. Vol. 50, pp.843-849.
- Haroutunian V, Perl DP, Purohit DP, et al. (1998) *Regional distribution of neuritic plaques in the nondemented elderly and subjects with very mild Alzheimer disease.* Archives of Neurology. Vol. 55, No.9, pp.1185-91.
- Harvey P, Moriarty P, et al. (2005) *The Validation of a Caregiver Assessment of Dementia.* Alz Dis Ass Disord. Vol. 9, No.4, pp.186-194.
- Herrmann W, Kurt S, Gaede K, Apeceche M (1997) *A multi-center Randomized Double-Blind Study on the Efficacy and Safety of Nicergoline in Patients with Multi-Infarct Dementia.* Dement Geriatr Cogn Disord. Vol. 8, pp.9-17.
- Heyman A, Wilkinson WE, Hurwitz BJ, et al. (1987) *Early-onset Alzheimer's disease: clinical predictors of institutionalization and death.* Neurology. Vol. 37, No.6, pp.980-4.  
<http://drugabusestatistics.samhsa.gov/SROS/sros8025.htm> provides method of administering DSRS.
- Hughes CP, Berg L, Danziger WL, et al. (1982) *A new clinical scale for the staging of dementia.* British Journal of Psychiatry. Vol. 140, pp.566-72.
- Imbimbo BP, Troetel WM, Martell IP, et al. (2000) *A 6-month, double-blind, placebo-controlled trial of eptastigmine in Alzheimer's disease.* Dementia and Geriatric Cognitive Disorders. Vol. 11, No.1, pp.17-24.
- Jacobs D, Sano M, Marder K, Bell K, Bylsma F, Laflesce G, Albert M, Brandt J and Stern Y (1994) *Age at onset of Alzheimer's disease: relation to pattern of cognitive dysfunction and rate of decline.* Neurology. Vol. 44, No.7, pp.1215-20.
- Jones RW, Soininen H, Hager K, et al. (2004) *A multinational, randomised, 12-week study comparing the effects of donepezil and galantamine in patients with mild to moderate Alzheimer's disease.* International Journal of Geriatric Psychiatry. Vol. 19, No.1, pp.58-67.

- 
- Karlawish J, Casarett D, et al. (2003) *Why Would Caregivers not want to treat their relative's alzheimer's disease?* Journal of the American Geriatrics Society. Vol. 51, pp.1391-1397.
- Landes A, Sperry S, Strauss M (2005) *Prevalence of Apathy, Dysphoria and Depression in Relation to Dementia Severity in Alzheimer's disease.* J Neuropsychiatry Clin Neurosci. Vol.17, pp.342-349.
- Malouf R and Birks J (2004) *Donepezil for vascular cognitive impairment.* Cochrane Database Systematic Review. Vol. 1, No.CD004395.
- Marin DB, Flynn S, Mare M, et al. (2001) *Reliability and validity of a chronic care facility adaptation of the Clinical Dementia Rating scale.* International Journal of Geriatric Psychiatry. Vol. 16, No.8, pp.745-50.
- McCulla MM, Coats M, Van Fleet N, et al. (1989) *Reliability of clinical nurse specialists in the staging of dementia.* Archives of Neurology. Vol. 46, No.11, pp.1210-1.
- Morris JC (1993) *The Clinical Dementia Rating (CDR): current version and scoring rules.* Neurology. Vol. 43, No.11, pp.2412-4.
- Morris JC (1997) *Clinical dementia rating: a reliable and valid diagnostic and staging measure for dementia of the Alzheimer type.* International Psychogeriatrics. Vol. 9 Suppl 1, pp.173-6; discussion 177-8.
- Morris JC, Heyman A, Mohs RC, et al. (1989) *The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease.* Neurology. Vol. 39, No.9, pp.1159-65.
- Morris JC, McKeel DW, Jr., Fulling K, et al. (1988) *Validation of clinical diagnostic criteria for Alzheimer's disease.* Annals of Neurology. Vol. 24, No.1, pp.17-22.
- Morris JC, McKeel DWJ, Storandt M, et al. (1991) *Very mild Alzheimer's disease: informant-based clinical, psychometric, and pathologic distinction from normal aging.* Neurology. Vol. 41, No.4, pp.469-78.
- Morris JC, Storandt M, McKeel DW, Jr., et al. (1996) *Cerebral amyloid deposition and diffuse plaques in "normal" aging: Evidence for presymptomatic and very mild Alzheimer's disease.* Neurology. Vol. 46, No.3, pp.707-19.
- Morris JC, Storandt M, Miller JP, et al. (2001) *Mild cognitive impairment represents early-stage Alzheimer disease.* Archives of Neurology. Vol. 58, No.3, pp.397-405.
- Newberg A, Cotter A, et al. (2003) *A Metabolic Imaging Severity Rating Scale for the Assessment of Cognitive Impairment.* Clinical Nuclear Medicine. Vol. 28, No.7, pp.565-570.
- O'Connor DW, Blessed G, Cooper B, et al. (1996) *Cross-national interrater reliability of dementia diagnosis in the elderly and factors associated with disagreement.* Neurology. Vol. 47, No.5, pp.1194-9.
- Patterson CJS, Gauthier S, Bergman H, Cohen CA, Feightner JW, Felman H, et al. (1999) *Management of dementia disorders. The recognition, assessment and management of dementing disorders- conclusions. Canadian Consensus Conference on Dementia.* Canadian Medical Association Journal, Vol. 160, No.12 (supplement) (online at [www.cma.ca](http://www.cma.ca))
- Ramirez Diaz SP, Gregorio PG, Ribera Casado JM, et al. (2005) *The need for a consensus in the use of assessment tools for Alzheimer's disease: the Feasibility Study (assessment tools for dementia in Alzheimer Centres across Europe), a European Alzheimer's Disease Consortium's (EADC) survey.* International Journal of Geriatric Psychiatry. Vol. 20, No.8, pp.744-8.
- Reisberg B (1988) *Functional Assessment Staging(FAST).* Psychopharmacology Bulletin. Vol. 24, pp.653-9.
- Reisberg B, Ferris S, de Leon M, et al. (1982) *The global deterioration scale for assessment of primary degenerative dementia.* Am J Psychiatry. Vol. 139, pp.1136-64.
-

Riepe MW, Adler G, Ibach B, et al. (2006) *Adding memantine to rivastigmine therapy in patients with mild-to-moderate Alzheimer's disease: Results of a 12-week, open-label pilot study. Primary Care Companion to the Journal of Clinical Psychiatry.* Vol. 8, No.5, pp.258-263.

Rockwood K (2004) *Size of the treatment effect on cognition of cholinesterase inhibition in Alzheimer's disease. Journal of Neurology, Neurosurgery and Psychiatry.* Vol. 75, No.5, pp.677-85.

Rockwood K, Strang D, MacKnight C, et al. (2000) *Interrater reliability of the Clinical Dementia Rating in a multicenter trial. Journal of the American Geriatrics Society.* Vol. 48, No.5, pp.558-9.

Roman GC, Wilkinson DG, Doody RS, et al. (2005) *Donepezil in vascular dementia: combined analysis of two large-scale clinical trials. Dementia and Geriatric Cognitive Disorders.* Vol. 20, No.6, pp.338-44.

Schafer KA, Tractenberg RE, Sano M, et al. (2004) *Reliability of monitoring the clinical dementia rating in multicenter clinical trials. Alzheimer Disease and Associated Disorders.* Vol. 18, No.4, pp.219-22.

Scian S, Reisberg B (1992) *FAST in Alzheimer's disease: reliability, validity and ordinality. International Psychogeriatrics.* Vol. 4 (suppl. 1), pp. 55-69.

Shader R, Harmatz J, Salzman C (1974) *A New Scale for Clinical Assessment in Geriatric Populations: Sandoz Clinical Assessment – Geriatric (SCAG). JAGS.* Vol.12, No.3, pp.107-113.

Solomon P, Adams F, Groccia M, et al. (1999) *Correlational Analysis of Five Commonly Used Measures of Mental Status/Functional Abilities in Patients with Alzheimer disease. Alz Dis and Ass Disorders.* Vol.13, No.3 pp.147-150.

Stern Y (1994) *Assessing Patient Dependence in Alzheimer 's disease. Journal of Gerontology.* Vol. 49, No.5, pp.216-225.

Stern Y, Hesdorffer D, Sano M and Mayeux R (1991) *Measurement and prediction of functional capacity in Alzheimer's disease. Neurology.* Vol. 40, No.1, pp.8-14.

Summers WK, DeBoynton V, Marsh GM, et al. (1990) *Comparison of seven psychometric instruments used for evaluation of treatment effect in Alzheimer's dementia. Neuroepidemiology.* Vol. 9, No.4, pp.193-207.

Tariot PN, Cummings JL, Katz IR, et al. (2001) *A randomized, double-blind, placebo-controlled study of the efficacy and safety of donepezil in patients with Alzheimer's disease in the nursing home setting. Journal of the American Geriatrics Society.* Vol. 49, No.12, pp.1590-9.

Tedesco M, Ratti GS, Mennella S, et al. (1999) *Comparison of losartan and hydrochlorothiazide on cognitive function and quality of life in hypertensive patients. American Journal of Hypertension.* Vol. 12, No.11, pp.1130-1134.

Tom Mesuer PhD, Associate Prof, Director of Education and Rural Outreach Alzheimer Disorder Research Centre, Washington University (2007) Personal communication. [meuser@neuro.wustl.edu](mailto:meuser@neuro.wustl.edu). The Washington University is the most important resource for the GDS.

Tractenberg RE, Schafer K and Morris JC (2001) *Interobserver disagreements on clinical dementia rating assessment: interpretation and implications for training. Alzheimer Disease and Associated Disorders.* Vol. 15, No.3, pp.155-61.

Venn, RD (1983) *The Sandoz Clinical Assessment Geriatric (SCAG) Scale. A general purpose psychogeriatric rating scale. Gerontology.* Vol. 29, pp.185-198.

Villardita LC and Lomeo (1992) *Alzheimer's disease: Correlational analysis of three screening tests and three behaviour scales. Acta Neurologica Scandinavica.* Vol. 80, No.6, pp.603-8.

Waite L, Grayson D, Jorm AF, et al. (1999) *Informant-based staging of dementia using the clinical dementia rating. Alzheimer Disease and Associated Disorders.* Vol. 13, No.1, pp.34-7.

Wang L, Swank JS, Glick IE, et al. (2003) Changes in hippocampal volume and shape across time distinguish dementia of the Alzheimer type from healthy aging. *Neuroimage*. Vol. 20, No.2, pp.667-82.

Zemlan FP (1996) *Velnacrine for the treatment of Alzheimer's disease: a double-blind, placebo-controlled trial*. The Mentane Study Group. *Journal of Neural Transmission*. Vol. 103, No.8-9, pp.1105-16.

Zillmer EA, Fowler P, Gutnick HN, Becker E (1990) *Comparison of two cognitive bedside screening instruments in nursing home residents: a factor analysis study*. *Journal of Gerontology*. Vol. 45, No.2, pp.69-74.

**Additional internet sites visited to search for dementia staging and descriptive measures:**

[http://www.alz.org/professionals\\_and\\_researchers\\_conducting\\_an\\_assessment.asp](http://www.alz.org/professionals_and_researchers_conducting_an_assessment.asp)

<http://www.neurotransmitter.net/Alzheimerscales.htm>

<http://alz.org/Services/LibraryServices.asp>

[http://www.alz.org/professionals\\_and\\_researchers\\_general\\_resources.asp](http://www.alz.org/professionals_and_researchers_general_resources.asp)

<http://www.alzforum.org/dis/dia/tes/neuropsychological.asp>  
[Greenfield@alz.org](mailto:Greenfield@alz.org)

[www.adrc.wustl.edu/adrc](http://www.adrc.wustl.edu/adrc)

<http://alzheimer.wustl.edu/cdr/default.htm>

<http://www.hnrc.ucsd.edu/publications/index.php>

[www.stroke.com](http://www.stroke.com)

<http://www.Alzheimer-insights.com/insights/vol2no3/vol2no3.htm>

<http://www.medicine.uiowa.edu/igec/tools/default.asp>

<http://library.csus.edu/guides/rogenmoserd/heas/nursing2004.pdf>

<http://libraries.uta.edu/helen/Test&Meas/testmainframe.htm>

<http://www.hartfordign.org/resources/education/tryThis.html>

## 5 Health Related Quality of Life Instruments and Dementia

### 5.1 Quality of Life in Dementia

Dementia affects many aspects of the quality of the lives of people with dementia, in particular their ability to function socially and to live independently. It also affects the quality of the lives of their families and carers. This section deals with instruments that capture the impact of dementia on the quality of life (QOL) and health related quality of life (HRQOL) of the person with dementia, including both self-report and proxy report by carers.

While the term *quality of life* is often used, it is rarely defined (Gill and Feinstein, 1994). In its broadest sense, it covers aspects of life that are beyond the scope of health care, such as living standards, housing, education, employment and the environment. It has been used in this sense in the context of economics and welfare since 1920 (Wood-Dauphinee, 1999). When used in the context of health, its meaning is often restricted to aspects of life that relate to health and health care (Schipper, Clinch, and Olweny, 1996; Ware, 1987). In this case, the term *health-related quality of life* (HRQOL) is often used to differentiate the restricted sense from the broader sense. Since the 1980s it has been used as a synonym for *health status*, *functional status* and *subjective well-being* (Patrick and Bergner, 1990; Spitzer, 1987), reflecting the conceptual heritage of QOL instruments. Throughout the 1990s, the term evolved into an ill-defined umbrella covering all aspects of the impact of disease and treatment on the bodies, minds and lives of patients. Some researchers accept that “quality of life means different things to different people, and takes on different meanings according to the area of application” (Fayers and Machin, 2000, page 3).

There is no single, concise definition of QOL as it is used in the health context. Various conceptualisations of QOL have been proposed. The *expectations* model (Calman, 1984) defines QOL as the difference between an individual’s hopes and expectations and his or her present experience. The concept of *reintegration to normal living* relates to the ability to do whatever one has to or wants to do, but does not mean being free of disease or symptoms (Wood-Dauphinee and Williams, 1987). Other models emphasise *meaning in life* (Warner and Williams, 1987), *satisfaction with life* (Pavot, Diener, Colvin, and Sandvik, 1991), patient *needs* (Coyle, Goldstein, Passik, Fishman, and Portenoy, 1996), or *spiritual* aspects of existence (Brady, Peterman, Fitchett, Mo, and Cella, 1999). The economic theory of *utility* gives rise to conceptualisations of QOL involving the relative value of dimensions of health and preferences for different states of health (Froberg and Kane, 1989); these are covered in Section 7 and Appendix 8. The authors of this section do not adopt a particular definition of QOL or HRQOL. Rather, they present the operational definitions implicit in the content and coverage of the instruments reviewed.

HRQOL and health status instruments may be generic or disease-specific. A generic measure can be used for comparisons across diseases and health conditions. Widely used examples include multi-dimensional profiles such as the SF-36, Nottingham Health Profile, and the Sickness Impact Profile, and indices for economic evaluation such as EQ-5D, AQoL (which are reviewed in Section 7). In contrast, disease or condition specific measures focus on those aspects of health (e.g. symptoms) and health-related quality of life that are relevant to a particular health condition such as cancer or heart disease. Dementia-specific examples include the Quality of Life in Alzheimer’s Disease scale or the DEM-QOL.

An integral part of most definitions of HRQOL is that it is *multidimensional*: “Although terminology may differ, there are four broad components of the quality of life construct: physical and occupational function, psychological state, social interaction and somatic sensation” (Schipper, Clinch, and Olweny, 1996, page 16). The particular dimensions that are included in a disease-specific instrument should reflect the aspects of health and life that are affected by the particular disease, in this case dementia. In other illnesses, such as cancer, asthma and arthritis, there may

be a simple association between QOL and an easily measurable clinical variable such as pain, fatigue or activity limitations, but this is not the case in dementia (Banerjee, et al. 2006). The authors of this section do not specify a set of dimensions that should be covered by disease-specific HRQOL instruments, but rather describe the dimensions covered by the instruments reviewed and the process used by the instruments' authors to determine these dimensions.

Another feature common to most definitions of HRQOL is that it is a subjective phenomenon. So when measuring HRQOL, the patient's assessment is preferred to that of a proxy such as a relative or attending nurse or doctor (Addington-Hall and Kalra, 2001; Slevin, Plant, Lynch, Drinkwater, and Gregory, 1988). HRQOL is therefore usually self-assessed. This requires a complex cognitive process of introspection and evaluation, involving several components of cognition. Common symptoms of dementia, including loss of memory, attention, comprehension, communication, insight and language skills can make self-report difficult. After a certain level of cognitive decline and language impairment, self-assessment of HRQOL may become too difficult for the respondent with dementia and therefore infeasible. Further, the nature of HRQOL may change with progressing severity, as patients increasingly withdraw from usual activities of daily living, normal social interactions and meaningful communication.

Defining and assessing HRQOL in dementia poses some unique challenges. It is more difficult to determine HRQOL in persons with dementing illness than in persons who are cognitively intact, and it is even more difficult in persons with late-stage dementing illness who cannot communicate coherently and are not involved in activities widely accepted by others as affording QOL, such as socialising or working on a hobby (Weiner, et al. 2000). An ideal measure of HRQOL in dementia could evaluate HRQOL at different stages of the disease, measure the elements of capacity that are possibly retained and valued, and enable a person besides the patient to rate their presences (Rabins and Kasper, 1997). However, because many patients with late-stage disease have impaired language, perception and judgement, self-report is not likely to be feasible for late-stage patients. In this case, measures based on externally observable elements have been suggested (Lawton, 1994). Proxy ratings are another potential solution, but they have several limitations. First, they unavoidably filter the patient's subjective state through the proxy's opinion, which may be influenced by the proxy's own state and mood and their feelings about the patient. Further, proxies may not know the person sufficiently well or spend enough time with them to observe with the necessary insight, accuracy or understanding to interpret the patient's HRQOL. Finally, they must extrapolate from behaviour to value. Yet this limitation to HRQOL measurement in dementia has no obvious solution and may be unavoidable.

In this chapter the use of generic health status and health related quality of life measures with dementia patients will be briefly discussed followed by a more detailed analysis of dementia specific health related quality of life measures.

## **5.2 Generic Health Status and Health Related Quality of Life Measures**

Generic health status and health related quality of life measures are useful if one wants to make comparisons concerning quality of life or burden of disease across different health conditions or to compare data for a particular condition in comparison to normative data that may have been obtained from the general population. For example, health status measures such as the Short Form – 36 (SF-36) or the Short Form -12 (SF-12) (Ware, et al. 2001; 2002), combined with questions on self reported morbidity, have been included in a number of Australian population surveys (ABS 1997, ABS 1998) to both assess the health status of the general population and to compare the morbidity associated with various health conditions (asthma, depression, etc). Similarly, multi-attribute utility measures (such as the EQ-5D, AQoL, and HUI etc) are generic health related quality of life measures that are used to undertake economic evaluations to compare cost effectiveness or cost utility of alternative treatments. These measures are discussed at length in Section 7.

Generic measures usually include a number of items around some core life domains. For example the SF-36 (Ware, et al. 2001) includes the domains of Physical Function, Role– Physical, Bodily Pain, Vitality, Social Function, Role–Social, Mental Health and General Health Perceptions. The other leading generic health related quality of life measures include similar domains but with some minor variations in their coverage.

These generic instruments focus on domains and items that would be relevant to any health condition and thus do not cover the symptoms or domains that may be specific to a particular condition. Generic measures are not designed to capture the particular morbidity of specific diseases/conditions and they will not always capture the range of domains in which impairments occur (e.g. cognition, sensory functioning). For these reasons many outcome studies have included both generic and disease/condition specific measures; the latter are used to capture the domains or symptoms that are central to the condition while the generic measures are used to make comparisons with other conditions.

A recent project (Thomas, et al. 2006) focussing on the development of an outcomes measurement suite for continence conditions reviewed a number of the leading generic measures used to assess health status and health related quality of life. Comprehensive reviews and comparative assessments were undertaken of the SF-36, the SF-12, the WHOQOL-Bref, the WHOQOL-100, the Sickness Impact Profile and the Nottingham Health Profile (Thomas, et al. 2006). The SF-36 Version 2 (Ware, et al. 2001) was chosen as the recommended generic instrument for assessing health related quality of life due to its better psychometric properties as compared with the other leading generic instruments. However, it must be remembered that the population of those experiencing continence conditions is quite different to those experiencing dementia. While some people with incontinence may also experience cognitive impairment many do not, whereas this is a defining attribute for a dementia population.

As the symptoms of dementia differ significantly from those of other illnesses, and generic health related quality of life measures do not cover some key domains for dementia (e.g. cognition, behavioural disturbance) many researchers prefer just to use a disease specific measure to assess health related quality of life in dementia (Rabins and Black, 2007). Most generic HRQOL measures are also self report measures and as Rabins and Black (2007) indicate many individuals with dementia, particularly those with moderate to severe illness, lack the capacity to self rate.

Some items in these instruments may be inappropriate to elderly people – for example questions concerning vigorous activities or how health has affected work (McDowell, 2006) The question frames in some of the items included in these scales are complex and assume a level of cognitive function that would make them unsuitable for use with those experiencing moderate to severe cognitive impairment. Although proxies could be used, subject to the usual limitations and caveats, these instruments have not developed particular proxy versions for such an application as has occurred with many of the dementia specific health related quality of life measures. Where proxy ratings have been compared with patient ratings there is little agreement with patient ratings and proxies have been found to be a poor substitute for obtaining the patient's perspective (Novella, et al. 2006; Novella, et al. 2001).

As many of these instruments have also been designed for use with the general population they are also prone to floor effects when used with the frail elderly (Ware, 2003). Some of these instruments are too long and would place considerable respondent burden on the person with dementia – for example the Sickness Impact Profile (Bergner, 1976, 1981) has 136 items and even the shorter version (SIP68) has 68 items. Even the shorter measures contain between 12 – 40 items. Whilst a member of the public may be able to answer the SF-36 in less than 10 minutes it has been shown that elderly and frail elderly patients take far longer than this to complete it and there are higher rates for missing data (Hayes, et al. 1995; McHorney, et al. 1996; Novella, et al. 2001; Sherbourne and Meredith, 1992). Thus one suspects the respondent burden for most people with dementia would be far too great (refer to the discussion in Section 12).

Novella, et al. (2001) using the SF-36 found that the refusal rate for persons with dementia was greater, 73% required assistance from the interviewer to complete the SF-36, and there was greater missing data associated with the increasing severity of dementia. Novella, et al. (2001) concluded that it may be possible to use an interview administration (vs. self report) of the SF-36 with those with mild to moderate dementia (MMSE > 15). However, they suggest the SF-36 is not appropriate for use with those with severe dementia (MMSE < 10) as the severity of the disease affected the feasibility, acceptability and reproducibility of the instrument.

Parker, et al. (1988) used the SF-36 with elderly hospital inpatients and outpatients and community dwelling general practice patients. They found that only 62.5% of inpatient who self completed the survey gave sufficient response to calculate a score on the mental health subscale, compared with 93.7% of general practice patients. They questioned the utility of the SF-36 as a routine health status measure for use with older hospital inpatients and concluded it was not appropriate for routine use with elderly people in hospitals.

Seymour, et al. (2001) assessed the reliability and validity of the SF-36 with cognitively normal and cognitively impaired older rehabilitation patients using an interview administration. They found that the levels of internal consistency and test-retest reliability and validity reported for the SF-36 in younger subjects were not attained for either sub-group or the sample overall. The reliability values for the cognitively normal patients were significantly higher than those for cognitively impaired patients on three of the subscales and test –retest reliability coefficients were also higher for this group. These findings would cast particular doubt on the use of the SF-36 with cognitively impaired older people.

The shortest of these instruments, the SF-12, would place the least respondent burden on patients. The SF-12 Version 1 has now been substantially revised and Version 2 was released in 2002 (Ware, et al. 2002). Version 2 now produces both profile scores (as for SF-36) as well as the physical and mental health summary scale scores and has superior psychometric properties (reliability, validity, responsiveness) when compared with Version 1 (Ware, et al. 2002). The SF-12 V2 could possibly be considered for use with patients experiencing mild to moderate dementia that retain the capacity to self rate. However, the SF-12 was designed to be used as a population survey instrument rather than as an evaluative instrument although Ware, et al. (2002) cites four studies where the SF-12 has performed as well as the SF-36 in terms of responsiveness and ability to distinguish clinically important change. However, with fewer items per domain than the SF-36 there will always be some trade off between precision and respondent burden (Ware, et al. 2002) and the performance of Version 2 needs to be further assessed.

Information on its use with cognitively impaired patients is limited, it does not cover domains such as cognition and memory which are relevant when assessing elderly populations, and although the floor effects have been lessened in the current revision when tested in general population settings, this is likely to remain a problem when it is used with frail elderly and people with dementia. Pettit, et al. (2001) assessed the reliability and validity of SF-12V1 with cognitively normal and cognitively impaired elderly people. They concluded that the SF-12V1 was not acceptable and valid for people with dementia.

It should be also noted that there are some issues concerning the scoring of the SF-12. It is scored using US population weights on the assumption that it is an 'international' measure and thus these weights are applicable in other cultures. With respect to the SF-36V2 Hawthorne, et al. (2007) has questioned this assumption and query whether it is may be appropriate to use Australian population weights here and similar issues would apply to SF-12V2. It is also noted that while Australian norms are available for SF-36V1, this version is not preferred due to its lesser psychometric properties, and as yet Australian normative data has not been collected for Version 2.

The SF family of generic measures reviewed above are largely measures of health status rather than of overall quality of life or well-being. The WHOQOL Group (1998a, 1998b) has developed a number of measures (WHOQOL-100; WHOQOL-BREF) that encompass a broader construction of Quality of Life. The WHOQOL-BREF contains 26 items covering the physical, psychological, social and environmental aspects of quality of life. The environmental domain, for example, includes items on safety, security, financial resources and the home environment which are areas not tapped by most of the generic measures mentioned above. The WHOQOL-BREF has undergone an extensive process of translation and development in order to make it comparable across languages and cultures and these instruments are primarily for use in cross-cultural research and clinical trials rather than for individual assessment. These instruments have adequate psychometric properties (Thomas, et al. 2006).

The WHOQOL group has recently developed the WHOQOL-OLD module (Power, et al. 2005) to use in conjunction with either the WHOQOL-BREF or the WHOQOL-100. The WHOQOL- OLD supplementary module has an additional 24 items covering the domains of sensory abilities, autonomy, past present and future activities, social participation, death and dying and intimacy. While this work should be commended for assessing the domains of quality of life that are most relevant to elderly people it may be difficult to administer fifty items on quality of life to people with all but the mildest dementia. A shorter scale is needed. The psychometric properties reported by Power, et al. (2005) are promising but as this is a new instrument, its validity, in particular needs to be further assessed as does its application with people with dementia.

Another simpler approach, which is thought to place less cognitive and time demands on the respondent, may be to use single item measures and pictorial methods such as that used by the Dartmouth COOP Charts (Nelson, 1987) designed for primary care settings. There are nine Dartmouth COOP charts each with a single question about health in the last month. Three charts cover function (physical, fitness, daily and social activities), three cover health perceptions (quality of life, overall health, and change in health condition) two cover symptoms and feelings (pain, emotional status) and one chart covers social support (McDowell, 2006). In the charts the responses to questions are in the form of a 5 point answer scale where the descriptor of each response level is illustrated by a picture. The nine charts are considered as separate dimensions of functioning, and thus are really a collection of nine single item measures.

Mc Dowell (2006) also provides a review of a number of other Single-Item Health Indicators which include, amongst others, the Delighted-Terrible Scale, The Faces Scale and the Ladder Scale which can also provide summary ratings of health, life satisfaction and so forth.

McDowell (2006) indicates the results so far suggest that the single-item measures can provide quite good indications of present state - they can offer a brief screen that is accurate enough to give a global impression of a patient's well-being – but they are too coarse to detect minor changes in function over time. The latter aspect would limit the application of these measures in outcomes research. It is noted that the Lancashire Quality of Life Profile (Residential) (Oliver and Mohamad, 1996) includes a number of these scales and a revision of this instrument is being undertaken to make it more suitable for use with elderly patients. This may provide further data as to the appropriateness of these scales (the Lancashire Profile itself, however, is a very lengthy instrument and was designed for the assessment of the chronically mentally ill). Given these considerations it is difficult to recommend the use of these measures at this stage. However, it may be worth investigating this method for assessing such elements as the satisfaction of people with dementia in their care.

Another issue is whether the simplicity of these single item measures really makes them less demanding for those experiencing some level of cognitive impairment. Pictorial scales appear more direct and may tap into the feelings associated with QoL without requiring the intermediary of language for the response choice (McDowell, 2006). However, some of these scales still contain quite complex language in their question stems (e.g. Dartmouth COOP Charts). A number of these

scales also have either 7 or 9 levels of response choice, with minor gradations between these levels and it is thought this level of discrimination may be difficult for those experience cognitive impairment. McDowell (2006) reports nonverbal scales might work well with children and others that may have difficulty completing a questionnaire – for example, a Faces Scale has been used in measuring pain in young children and in measuring anxiety in critically ill patients. However, there is little literature available to date concerning their use with people experiencing dementia and thus these scales are not recommended at this stage.

In conclusion Riemsma, et al. (2001) reviewed 71 studies to assess the applicability and validity of a number of health status measures when used with cognitively impaired subjects. The most commonly used measures were the SIP and the SF-36. Riemsma, et al. (2001) concluded that very few measures had been validated specifically for cognitively impaired respondents. Studies where at least 50% of the respondents were cognitively impaired generally showed poorer validity results compared with studies with fewer cognitively impaired persons. Riemsma, et al. (2001) suggest this indicates that general health status measures designed for the general population are not automatically suitable for people with cognitive impairment and advises caution with their use in these applications.

With regard to the assessment of health related quality of life of those experiencing dementia there are significant limitations concerning the use of generic health related quality of life scales with people with dementia as has been outlined above. A discussion concerning the capacity to self – rate and the use of proxies can be found in Section 12. Self report instruments such as the SF-36 are clearly not suitable for use with people with severe dementia (MMSE of 10 or less) and require an assisted interview administration for those with an MMSE less than 15 (Novella, et al. 2001). Such measures could be used to assess the health related quality of life of carers of persons with dementia but are probably not particularly useful for assessing the health related quality of life of people with dementia themselves. For this reason individual reviews for these instruments are not provided in the appendices, however, with regard to the former application readers are referred to the reviews of these instruments provided by Thomas, et al. (2006) and McDowell (2006). A possible exception to this may be the SF-12V2, which because of its brevity, may make it more applicable for use with people with mild-moderate dementia who retain the capacity to self-rate. However, as indicated above many of the concerns raised above equally apply to the SF-12V2 and initial findings with respect to its use with people with dementia are not promising (Pettit, et al. 2001).

Recently Item Response Theory (IRT) has been used to cross calibrate items that measure the same dimension across a range of different health measures (Ware, et al. 2000). A common metric/ ruler can be developed for the domain and related items (from various instruments) can be placed along it. For example, it is known that the SF-36 has very few items that assess very poor physical function whereas some ADL measures contain items that do measure these lower levels of physical function. Ware (2003) noted that about 3% of US managed care beneficiaries scored at the floor of the physical function sub-scale of the SF-36. However, by adding 3 ADL items 96% of the elderly were removed off the floor of the physical function dimension. More importantly such scales can be administered dynamically, using computerized testing approaches, which mean that items in the pool are selected and administered only if they match the respondent's level of health (Ware, et al. 2002). For example, it may only take the administration of four items to get a reliable estimate of an individuals score on the physical function 'ruler' and thus the administration of all the items is not necessary. In the future the application of this approach certainly holds the promise of reducing respondent burden and assisting with precision of measurement with regard to these generic measures of health related quality of life. Although it is unknown whether a computerised adaptive testing approach is feasible to use with people with dementia, shorter static forms more suited to this group could also be derived from the cross calibration of measures.

However, at the present time the dementia specific quality of life measures, reviewed below, would seem more appropriate measures to use with people with dementia. Dementia specific measures

more adequately capture the relevant dimensions for this condition and as such are more likely to capture the way that patients decline and/or improve over time and thus are likely to be more useful measures for assessing the outcomes of patients.

### **5.3 Dementia Specific Health Related Quality of Life Measures: Initial Literature and Impact Search**

Details of the literature search strategies used are outlined in the Introduction (refer Section 2) of this report. Following the work on dementia staging and descriptive instruments, a list of dementia specific quality of life measures was developed. Based on examination of impact sheet data (MEDLINE, text and web impact, presence on PROQOLID database and practice surveys and clinical feedback) seven instruments were identified as being contenders for comprehensive review.

An impact sheet for dementia-specific HRQOL measures was developed from searches of MEDLINE, text and web impacts, presence in PROQOLID and its use in clinical practice (as advised by the NEP) field surveys and clinical feedback. This process identified six instruments, somewhat fewer than for other categories, perhaps because the development and use of dementia-specific HRQOL instruments is a relatively new activity.

The six dementia-specific HRQOL instruments identified were:

1. Alzheimer Disease-Related Quality of Life (ADRQOL);
2. Cornell Brown Scale for Quality of Life in Dementia (CBS);
3. Dementia Quality of Life Instrument (DQOL);
4. DEMQOL (this is the instrument's full name, not an abbreviation);
5. Quality of Life in Alzheimer's Disease (QOL-AD);
6. Quality of Life in Late-Stage dementia (QUALID).

Because of the relatively small number of instruments identified at this stage, and because they each satisfied most of the criteria for inclusion in the next stage of more detailed review, further information was collected for all six instruments. This represents a slightly different procedure from that followed for the other instrument categories included in the DOMS review.

A comprehensive search of the following online bibliographic databases were conducted to identify all peer-reviewed published papers that reported the development, testing or application of these six instruments: CINAHL (Cumulative Index to Nursing and Allied Health), EMBASE, MEDLINE and PSYCINFO. The instruments' developers were also contacted by email to provide information about costs and availability of users manual, language translations and so on. Published papers and other relevant information sources were obtained and reviewed for key attributes, including: the content of the instrument in terms of the numbers items and the coverage of domains of QOL; the stage(s) of dementia that the instrument is suitable for; the availability of patient and/or proxy forms; the availability and cost of the instrument; training requirements for interviewers or observers; availability of user manuals; administration time; the number of citations; evidence about the psychometric properties; use in practice to date; and availability of CALD language translations.

This information was collated and integrated into an overall judgment, and each instrument was rated against the criteria described in Section 2 of this report to give a total weighted rating for each instrument. Consideration of these attributes and rating led to provisional recommendations, following a rationale described below. AHOC instrument review sheets were then completed for the three instruments which best satisfied the DOMS criteria.

## **5.4 Results of Detailed Review and Rating**

The key attributes of the six dementia-specific HRQOL instruments are summarized in Tables 23 and 24. Further details of the instruments' psychometric properties, with citation details, plus information on other practical issues such as availability, is provided following the tables. Instrument review sheets can be found in Appendix 5. The weighted ratings are presented in Table 25 along with an overall judgment of each instrument and other salient comments.

**Table 23 Summary of the Six Short-listed Dementia-specific HRQOL Instruments (Part 1)**

Instrument	Content: items & domains	Stage	Patient	Proxy	Availability/Cost	Training/Manual	Admin time
QOL-AD	13 items: Physical condition, Mood, Interpersonal relationship with family and friends, Ability to participate in meaningful activities, and Financial situation.	Mild to moderate	✓		Written permission required.	No formal training required.	10 -15 min (patients)
		All stages		✓	Free.	Detailed script provided for standardized administration.	5 min (proxy)
DEMQOL	28 items (patient), 31 items (proxy): Daily activities, Memory, Negative emotion, Positive emotion.	Mild to moderate	✓		Free – available on website	No formal training required.	10-20 min
		All stages		✓		User's manual available on website provides detailed instructions for standardized administration.	
QUALID	13 observable behaviours indicating: Affective state, Behavioural signs of comfort, Engagement in activities, and Interactions with others.	Late stage		✓	Written permission required.  Free for academics  Fee for commercial use, charged on individual basis.	No formal training required. However, administration by at least bachelors level technician is recommended.	5 minutes
DQOL	29 items: Self esteem,	Mild to	✓		Written permission	No formal training	10 min

	Positive affect, Negative affect, feeling of belonging, Sense of aesthetics,	moderate			required.  Free for academics.  Fee for commercial use (varies according to research – generally donation to Alzheimer's Association required).	required.  Detailed script provided for standardized administration.	
ADRQL	Social interaction, Awareness of self,  Feeling and mood, Enjoyment of activities, Response to surroundings	All stages		✓	Written permission required.  Free for academics.  Fee for commercial use (fee based on type of research and how instrument would be used).	Training required.  Users manual and training video available at cost of \$US 35.00 and \$US 50.00 respectively.	10 – 15 min
CBS	Mood related signs,  Ideational disturbances  Behavioural disturbances  Physical signs  Cyclic functions	Mild to moderate		✓	No permission required.  Free.	Manual for administration available from authors by request.	10 – 20 min

**Table 24 Summary of the Six Short-listed Dementia-specific HRQOL Instruments (Part 2)**

<b>Instrument</b>	<b>Citations</b>	<b>Psycho-metrics</b>	<b>Use in Practice to date</b>	<b>Overall judgment and other comments</b>
QOL-AD	20 journal articles  +3 reviews  2 books	Very good	Assess:  QOL in persons with Dementia/AD  Change in QOL over time  Effects of interventions for patient & or carer  Differences in patient & carer perspectives on quality of life	Instrument has both patient and carer versions which can be used separately or together.  It has very good psychometric properties, has been extensively cited, and has easy access and administration.  The instrument has also been used as an outcome measure in numerous intervention studies.
DEMQOL	2 journal articles	Good to Very good	Assess:  QOL in persons with Dementia/AD	Very new instrument developed by a well known and highly respected author in this field.  It has both patient and carer versions which can be used separately or together.  Due to its newness there are very few citations and the instrument has not been extensively used.

				<p>However the psychometric properties are good, the instrument has easy access and administration.</p> <p>The authors acknowledge more work needs to be done on validation but the instrument is very promising.</p>
QUALID	<p>2 journal articles, 1 in print 1 conference presentation</p>	Very good	<p>Assess:</p> <p>QOL in persons with Dementia</p> <p>Effects of interventions for patient</p>	<p>Although only a few citations, this is the only instrument specifically for late stage dementia.</p> <p>The instrument has very good psychometric properties.</p> <p>Is brief and easy to administer and has now been adopted by the care keys project (program to improve management of quality of life in elderly persons in Europe).</p>
DQOL	<p>8 journal articles (+ 3 reviews) 2 books</p>	Good to very good	<p>Assess:</p> <p>QOL in persons with Dementia</p> <p>Effects of interventions for patient</p> <p>Change in QOL over time</p>	<p>Instrument has several citations and has good to very good psychometric properties.</p> <p>Ease of access and administration is as good as QOLAD&lt; DEMQOL and QUALID.</p> <p>However, only patient version available and there are other alternative versions that can be used for mild to moderate dementia that have both proxy and patient versions.</p>

ADRQL	8 journal articles (+ 3 reviews)  1 book	Good	Assess:  QOL in persons with Dementia  Effects of interventions for patient	Instrument has several citations and has good psychometric properties.  Ease of access (re fees, training and manuals) is not as good as those listed above.  Also, only proxy version available and there are other alternative versions that can be used for all stages of dementia that have both proxy and patient versions.
CBS	1 journal article (+3 reviews)  1 book	Good	Assess QOL in persons with Dementia	Only a few citations.  Ease of access and administration.  Good psychometric properties but others have better properties.  However, only proxy version available and there are other alternative versions that can be used for mild to moderate dementia that have both proxy and patient versions.

#### 5.4.1 Quality of Life in Alzheimer's Disease (QOL-AD)

The QOL-AD was developed in late 1990's (Logsdon, Gibbons, McCurry, and Teri, 1999). It is the most widely cited, and therefore probably the most widely used internationally of the dementia-specific HRQOL instruments. This may be due to its brevity, free access and ease of administration, and its availability in both patient and proxy forms and in eleven languages. The QOL-AD is free with the author's permission.

Although the instrument's name includes the term "Alzheimer's disease", it is relevant to dementia from any cause. It contains 13 items: physical health, energy, mood, living situation, memory, family, marriage, friends, self as a whole, ability to do chores around their room, ability to do things for fun, money, and life as a whole. Patients, caregivers and experts were involved in item selection and item reduction to ensure an adequate coverage of the relevant domains (Logsdon, Gibbons, McCurry, and Teri, 2002; Thorgrimsen, et al. 2003). The items are summed to give a global total score ranging from 13 to 52, with higher score reflecting better HRQOL. It takes about 10 to 15 minutes for patients to complete and about five minutes for caregivers to complete.

The QOL-AD is available in patient-report form for mild to moderate dementia, and in proxy-report form for all stages. Both patient and proxy forms are self administered and rated questionnaires. An interviewer can be used to oversee the administration and provide assistance and clarification as needed. In this case the questionnaire is interviewer administered, and patient or carer rated. No formal training is required for interviewers and a detailed script is available to standardize administration. Patient and caregiver reports can also be combined, weighting the person with dementia's own HRQOL score twice as heavily as the caregiver's. Although a proxy-report form is available for late stage disease, the particular domains covered by this instrument are likely to be less relevant to patients with late stage disease than to persons with mild to moderate dementia, due to profound cognitive and functional losses that occur in late stage disease. Nevertheless, the validity and reliability of the QOL-AD has been demonstrated in late stage patients (MMSE scores 3-11), but it is unlikely to generate useful information for people with MMSE scores of < 3, where completion rates are very low (Hoe, Katona, Roch, and Livingston, 2005).

There is a considerable amount of evidence confirming excellent psychometric performance of the QOL-AD across all measurement criteria. Factor analysis supported the dimensions proposed by the instrument developers (Edelman, Fulton, Kuhn, and Chang, 2005; Thorgrimsen, et al. 2003). The instrument shows good to excellent internal reliability, with Cronbach's alpha ranging from 0.78 to 0.94 for the patient version and 0.79 to 0.88 for the proxy version (Edelman, Fulton, and Kuhn, 2004; Edelman, Fulton, Kuhn, and Chang, 2005; Fuh and Wang, 2006; Hoe, Katona, Roch, and Livingston, 2005; Logsdon, Gibbons, McCurry, and Teri, 1999, 2002; Sloane, et al. 2005; Smith, et al. 2005; Thorgrimsen, et al. 2003). QOL-AD has demonstrated good to excellent test-retest reliability. Intraclass correlations (ICC) of 0.76 have been reported for the patient version and 0.92 for the proxy version at one week re-test (Logsdon, et al. 1999). Another study reported ICCs that were "all at or above 0.60" (no further detail given) for the patient version, but gave no results for the carer version (Thorgrimsen, et al. 2003).

Most of the evidence about inter-rater reliability reflects patient-proxy comparisons. Agreement between patient and proxy scores was generally low and correlations not significant (Edelman, Fulton, and Kuhn, 2004; Edelman, Fulton, Kuhn, and Chang, 2005; Fuh and Wang, 2006; Hoe, Hancock, Livingston, and Orrell, 2006; Logsdon, Gibbons, McCurry, and Teri, 1999, 2002; Shin, Carter, Masterman, Fairbanks, and Cummings, 2005; Sloane, et al. 2005; Spector and Orrell, 2006). Only one study assessed inter-rater reliability and agreement between two interviewers' ratings of patients' HRQOL. Two interviewers were present; one asked questions while the other completed the assessment scales. It is not surprising that in this context, inter-rater reliability and agreement were good to excellent (Thorgrimsen, et al. 2003).

There is considerable evidence of construct validity in terms of the extent to which scores on the QOL-AD correlate to a number of other measures in a manner that is consistent with theoretically derived hypotheses (Edelman, Fulton, and Kuhn, 2004; Edelman, Fulton, Kuhn, and Chang, 2005; Fuh and Wang, 2006; Hoe, Hancock, Livingston, and Orrell, 2006; Hoe, Katona, Roch, and Livingston, 2005; Logsdon, Gibbons, McCurry, and Teri, 2002; Selwood, Thorgrimsen, and Orrell, 2005; Shin, Carter, Masterman, Fairbanks, and Cummings, 2005; Sloane, et al. 2005; Smith, et al. 2005; Winzelberg, Williams, Preisser, Zimmerman, and Sloane, 2005; Woods, Thorgrimsen, Spector, Royan, and Orrell, 2006). In summary, the patient version shows expected correlations with Physical and Instrumental Self Maintenance Scale – Activities of Daily Living (PIS-ADL), Alzheimer’s Disease Co-Operative Study (ACDS-ADL), Revised Memory and Behaviour Checklist (RMBPC)-depression, Cornell Scale for Depression in Dementia (CSDD), Geriatric Depression Scale (GDS Yesavage), Rating Anxiety in Dementia (RAID), Mental Outcomes Study (MOS), Pleasant Events Schedule (PES-AD), Neuropsychiatric Inventory (NPI), but evidence about correlations with the MMSE is mixed. The proxy version shows expected correlations with: PIS-ADL, Barthel’s ADL, Minimum Data Set (MDS-ADL), Clifton Assessment Procedures-behaviour rating scale (CAPE-BRS), Challenging Behaviour Scale (CBS), RMBPC- memory disruption and depression, CSDD, RAID, GDS (Yesavage), Cohen Mansfield Agitation Inventory (CMAI) MOS, MMSE and NPI.

With regard to discriminant validity, both patient and proxy versions have been shown to differentiate between patients with differing levels of depression and cognitive functioning (Fuh and Wang, 2006; Logsdon, Gibbons, McCurry, and Teri, 2002; Thorgrimsen, et al. 2003).

There is limited evidence about the sensitivity of QOL-AD to change in HRQOL over time. One study, in a sample of 201 people with dementia living in residential homes with MMSE scores between 10 and 24, reported small but statistically significant correlations between changes in self-reported QOL-AD scores and changes in clinically relevant external criterion measures in cognition (MMSE, ADAS-Cog), symptoms of depression (CSDD) and communication abilities (HCS) over an eight week period (Woods, Thorgrimsen, Spector, Royan, and Orrell, 2006). These correlations were all in the expected direction, providing good evidence that self-reports of QOL-AD are responsive to clinically important change. These data arose from a randomized trial of Cognitive Stimulation Therapy (CST), in which the intervention group had significantly improved relative to the control group on the Mini-Mental State Examination, the Alzheimer’s Disease Assessment Scale - Cognition (ADAS-Cog) and QOL-AD scales (Spector, et al. 2003). However, this evidence is countered by another RCT in which CST improved cognitive function (measured with MMSE) but not HRQOL (measured with QOL-AD) (Orrell, Spector, Thorgrimsen, and Woods, 2005).

Regarding interpretability, normative data are not available, but a considerable amount of reference data is provided by numerous studies which report QOL-AD scores from a range of settings. However, none of these studies were conducted in Australia, so there is no Australian reference data available. Applications include the use of QOL-AD for testing the psychometric properties of other HRQOL instruments, describing differences in patient and carer perspectives on HRQOL in dementia, describing change in HRQOL over time, and the effects of interventions including cognitive stimulation therapy, and drug treatment on HRQOL (Aisen, et al. 2003; Chapman, Weiner, Rackley, Hynan, and Zientz, 2004; Edelman, Fulton, and Kuhn, 2004; Edelman, Fulton, Kuhn, and Chang, 2005; Fuh and Wang, 2006; Harvey, et al. 2005; Hoe, Hancock, Livingston, and Orrell, 2006; Hoe, Katona, Roch, and Livingston, 2005; Logsdon, Gibbons, McCurry, and Teri, 2002; Orrell, Spector, Thorgrimsen, and Woods, 2005; Selwood, Thorgrimsen, and Orrell, 2005; Shin, Carter, Masterman, Fairbanks, and Cummings, 2005; Sloane, et al. 2005; Smith, et al. 2005; Spector and Orrell, 2006; Teri, McCurry, Logsdon, and Gibbons, 2005; Thorgrimsen, et al. 2003; Winzelberg, Williams, Preisser, Zimmerman, and Sloane, 2005; Woods, Thorgrimsen, Spector, Royan, and Orrell, 2006). These results provide comparative data and collectively provide a basis for interpreting QOL-AD results. No major floor or ceiling effects have been detected in these studies. However, these effects have not been

tested for late-stage only samples, where they are most likely to exist and may be masked by lower completion rates (Hoe, Katona, Roch, and Livingston, 2005).

### 5.4.2 DEMQOL

This is a very new instrument that has been developed by a team of world renowned dementia experts (Smith, et al. 2005). It was designed to address limitations and gaps in existing dementia-specific measures. Due to its newness, the instrument has not yet been widely cited. The authors acknowledge that more work needs to be done on validation, but the instrument is very promising. At this time, it is available only in English.

The DEMQOL is provided free for academic use. Costs for commercial and pharmaceutical use are determined following discussions between developers and potential user. The instrument plus a users' manual are available on the website of the Institute of Psychiatry, King's College, London.

DEMQOL has two forms: DEMQOL (self-report) and DEMQOL Proxy. Both are administered by interviewer, with standardized instructions provided in the interviewer manual. Both versions can be used in people with mild to moderate dementia (defined by the DEMQOL developers as a MMSE score of  $\geq 10$ ), and the DEMQOL Proxy can also be used for severe dementia. In mild/moderate dementia, the developers recommend use of both patient and proxy forms as they consider the two perspectives to complement one another, rather than substitute for each other. The DEMQOL takes about 10 to 20 minutes to administer.

The patient-rated version of DEMQOL contains 28 items covering 4 dimensions (Daily activities, Memory, Negative emotion, Positive emotion) plus a global item about overall QOL. Items are rated on a 4 point ordered category scale. Response options are "a lot", "quite a bit", "a little" "not at all", except for the global question, which has the options "very good", "good", "fair" and "poor". Item scores are summed to provide a global score (minimum 28, maximum 112), with a higher score indicating better QOL.

The carer-rated DEMQOL Proxy contains 31 items covering 2 domains: Functioning and Emotion. It also includes an additional global item to assess patients' feelings about their overall quality of life, as perceived by the carer. Response options are the same as for DEMQOL. Item scores are summed to provide a global score (minimum 31, maximum 124), with a higher score indicating better QOL.

Given their recent development, there is limited evidence about the psychometric properties of the DEMQOL and DEMQOL Proxy, all of which has been generated by the instruments' authors in the process of developing and field-testing the instruments. Most of this evidence is provided in a comprehensive initial report (Smith, et al. 2005), augmented by two papers (Banerjee, et al. 2006; Smith, et al. 2005b).

The content validity of the instrument was assured by careful process of item selection and item reduction to ensure an adequate coverage of the relevant domains. A conceptual framework was generated from a review of the literature, qualitative interviews with people with dementia and their carers, expert opinion and team discussion (Smith, et al. 2005b). Items for each component of the conceptual framework were drafted and piloted to produce questionnaires for the person with dementia (DEMQOL) and the carer (DEMQOL-Proxy). An extensive two-stage field testing was then undertaken of both instruments. In the first stage, an initial sample of 130 people with dementia and 126 carers representing a range of severity and care arrangements provided data. Items with poor psychometric performance were eliminated separately for DEMQOL and DEMQOL-Proxy. The internal structure was determined by factor analysis, with a four-factor solution accounting for 50% of the variance in the DEMQOL and a two-factor solution accounting for 35% of variance in the DEMQOL Proxy.

When developing the final item-reduced versions of the DEMQOL and DEMQOL Proxy forms, rigorous item selection and reduction procedures applied were independently to patient and proxy responses. The two forms have only 14 items in common, highlighting important differences in how people with dementia and their carers both conceive of and report HRQOL.

In the second field test, the final versions were evaluated alongside other measures in sample of 101 people with dementia and 99 carers for acceptability, reliability and validity (Smith, et al. 2005). In this sample, the instruments had excellent internal consistency with Cronbach's alpha of 0.87 or more for both patient and proxy versions (for sample as a whole and replicated for subgroups by severity). Test-retest reliability was good for both versions. The patient version ICCs was 0.84 when the whole sample was considered and 0.76 when only mild to moderate was considered. The proxy version ICCs was 0.75, 0.67 and 0.84 for the whole, mild to moderate and severe sample respectively.

Construct validity was also tested in the second field test in terms of expected relationships between DEMQOL and other measures. Results suggested good to excellent construct validity in relation to hypothesised relationships. Patient DEMQOL scores showed low to moderate correlations with the Geriatric Depression Scale (GDS Yesavage) and Barthel's ADL. Correlations with GDS (Yesavage) were expected to be higher. DEMQOL Proxy scores showed moderate to high correlation with the GDS (Yesavage) for people with mild to moderate dementia, but only low correlations with Barthel's ADL. For persons with severe dementia, the DEMQOL Proxy scores showed high correlation with GDS (Yesavage). Univariate and multivariate analysis showed the DEMQOL was significantly associated with the Neuropsychiatric Inventory (NPI): the total score as well as the agitation, depression, anxiety, disinhibition and irritability subscales (Banerjee, et al. 2006). The patient version was able to discriminate by age indicating some support for discriminant validity (Smith, et al. 2005).

With regard to interpretability, there are no normative data available as yet, and reference clinical data are limited to those reported from the second field test. As yet, there are no published reports of use in intervention studies or any other clinical research. The authors provide considerable information to show that major floor or ceiling effects do not exist, but they have not as yet evaluated responsiveness, noting that this needs to be done in future research (Smith, et al. 2005).

### **5.4.3 Quality of Life in Late-Stage Dementia (QUALID)**

The QUALID was designed to rate HRQOL in people with late stage Alzheimer's disease and other dementing illnesses (Weiner, et al. 2000). It is based on observable behaviours, and is administered in interview format to an informant following a set of standardized instructions. Informants may be either a family member or professional caregiver who, by having regular contact, is familiar with the subject's general behaviour. Informants must, in addition to being familiar with the subject, have spent a significant portion of at least 3 days out of the last 7 days with the subject, in order to accurately rate the items on the scale. No training is required for interviewers, but the authors recommend that interviewers have at least a bachelor's level qualification. The interview usually takes about 5 minutes.

QUALID is supplied by the author and written permission is required to use it. It is free for academics and non-profit research, and fees for commercial use are considered and charged on an individual basis. It is available in English, Swedish, Finnish, German, and Lithuanian.

QUALID contains 11 items describing observable behaviours encompassing affective state, behavioural signs of comfort, engagement in activities and interactions with others. The 11 items are: smiles, appears sad, cries, has facial expressions of discomfort, appears physically uncomfortable, verbalizations suggest discomfort, is irritable or aggressive, enjoys eating, enjoys touching/being touched, enjoys interacting with others, appears calm and comfortable. Items are

rated on a 5 point ordered category scale. The window of observation is one week. The items scores are summed to provide a global score with a minimum score of 11 and a maximum score of 55. A lower score indicates better quality of life.

There are two additional questions at the end of the questionnaire about the overall quality of the interview. These ask the interviewer to rate the informant's ability to understand the items and responses and the effort the informant put forth in answering questions, and the familiarity of the informant with the subject. These items are not included in the score, but reflect the validity and usefulness of the ratings for that subject.

QUALID was developed in a series of consensus meetings of staff with extensive experience with late stage dementia (Weiner, et al. 2000). It was not possible to involve patients due to their advanced dementia. Validation was carried out in a relatively small sample of 42 residents of a dementia care unit (MMSE mean score 11.5, SD 6.2) (Weiner, et al. 2000). Principal components analysis yielded a one factor solution. Results of this study also suggested good internal consistency (Cronbach's alpha = 0.77), good test-retest reliability (ICC = 0.81 when administered twice over a 2 – 3 day period, SEM = 0.08), and excellent inter-rater reliability (ICC = 0.83, SEM = 0.07). This study also provided some evidence of construct validity. The authors' expected correlations with Geriatric Depression Scale (GDS Yesavage) and the Neuropsychiatric Inventory (NPI) since behavioural and emotional disturbances such as agitation and depression should reflect HRQOL. These were realized, confirming convergent validity. In contrast, poor correlations with MMSE and The Physical Self Maintenance Scale (PSMS) were expected, since functional competence at this severity of dementia was unlikely to be related to HRQOL. These expectations were also realized, confirming divergent validity. In another study of 40 nursing home residents in Finland, anticipated correlations between QUALID and the Cornell Scale for Depression in Dementia (CSDD) and the Philadelphia Geriatric Morale Scale (PGMS) were observed (Luoma, Vaarama, and Hertto, 2005), providing further evidence of convergent validity. Another study showed that the QUALID could differentiate between people with mild and moderate depression (Valvanne, Luoma, Ylonen, and Vaarama, 2005), providing some evidence of discriminant validity. A study of 31 residents of long-term care facilities showed moderate correlations between change in QUALID scores and change in neuropsychiatric symptoms (psychopathology and/or behavioural disturbance) and adverse events, providing some evidence of responsiveness (Martin-Cook, Hynan, Rice-Koch, Svetlik, and Weiner, 2005).

With regard to interpretability, there are no normative data available as yet, and reference clinical data are limited to those reported in validation studies (Luoma, Vaarama, and Hertto, 2005; Valvanne, Luoma, Ylonen, and Vaarama, 2005; Weiner, et al. 2000). QUALID has not been used in any interventions that have been published to date. However, it has been adopted by the Care Keys Project, a program to improve the QOL and quality of care of elderly people in Europe. Evaluations of this project have not yet been published, but data arising from this project will provide useful reference data in the future. There is currently no information about floor or ceiling effects.

#### **5.4.4 Dementia Quality of Life Instrument (DQOL)**

The DQOL was developed to assess the health related quality of life (HRQOL) in persons with mild to moderate dementia by direct interview with the person themselves (Brod, Stewart, Sands and Walton, 1999). It consists of 29 items covering five domains: self esteem; positive affect; negative affect; feelings of belonging; and sense of aesthetics, plus an additional item to measure overall QoL. The instrument takes about 10 minutes to complete, requires minimal training and is free for academic use.

Available evidence indicates the DQOL has good to very good psychometric properties. Authors of these studies have provided considerable information to ensure the findings can be appropriately interpreted. Several studies (Brod, Stewart, Sands and Walton, 1999; Edelman,

Fulton and Kuhn, 2004; Ready, Ott and Grace, 2004; Edelman, Fulton, Kuhn and Chang, 2005; Sloane, Zimmerman, Williams, Reed, et al. 2005; Smith, Lamping, Banerjee, Harwood, et al. 2005) report very good internal reliability with Cronbach's alpha ranging from 0.66 to 0.91 for the total score and 0.37 to 0.90 for the subscales. Test-retest reliability has also been found to be very good with Pearson's correlations for the subscales ranging from 0.64 to 0.90 (Brod, Stewart, Sands and Walton, 1999).

The validity of the DQOL has also been confirmed. Evidence for construct validity, in terms of the extent to which scores on the DQOL relate to other measures in a manner consistent with theoretically derived hypotheses concerning the domains covered, comes from several studies (Brod, Stewart, Sands and Walton, 1999; Ready, Ott and Grace, 2004; Edelman, Fulton, Kuhn and Chang, 2005; Selwood, Thorgrimsen and Orrell, 2005; Smith, Lamping, Banerjee, Harwood, et al. 2005). Expected correlations were found with Geriatric Depression Scale (GDS Yesavage), Mini-Mental State Examination (MMSE), Neuropsychiatric Inventory (NPI), Activities of daily Living (ADL), Instrumental Activities of Daily Living (IADL), Rating Anxiety in Dementia (RAID) and Cornell Scale for Depression in Dementia (CSDD). Construct validity is further supported by studies reporting correlations with other instruments measuring QOL (Thorgrimsen, Selwood, Spector, Royan, et al. 2003; Edelman, Fulton and Kuhn, 2004; Edelman, Fulton, Kuhn and Chang, 2005; Selwood, Thorgrimsen and Orrell, 2005; Sloane, Zimmerman, Williams, Reed, et al. 2005; Smith, Lamping, Banerjee, Harwood, et al. 2005). DQOL correlated significantly with Quality of Life in Alzheimer's Disease (QOLAD), Alzheimer Disease Related Quality of Life (ADRQOL), Quality of Life Dementia (QOL-D), DEMQOL, EQ-5D, and SF-36. There is also evidence for discriminant validity though this is limited to findings from one study (Brod, Stewart, Sands and Walton, 1999). Results were however positive showing the scale did differentiate between people with mild to moderate dementia, and between people with differing levels of depression.

At this stage there is no evidence that DQOL meets the criteria for responsiveness. The instrument has been used as an outcome measure in two clinical studies. One assessed the effects of treatment with Donepezil for people with mild to moderate dementia attending an outpatient clinic (Mador, Hecker and Clark, 2003). Scores did improve on average as a result of the intervention but the change was not statistically significant. In the other study, which assessed change in QoL over 12 months, again scores did not change significantly (Selwood, Thorgrimsen and Orrell, 2005).

#### **5.4.5 Alzheimer Disease Related Quality of Life (ADRQOL)**

The ADRQOL was developed to assess HRQOL in persons with Alzheimer's disease (Rabins, Kasper, Kleinman and Black, 1999; Black, Rabins and Kasper, 2000). It is however relevant to dementia from any cause and is suitable for use across all stages of the disease. The instrument contains 48 items measuring social interaction, awareness of self, feelings and mood, enjoyment of activities, and response to surrounding. It is a caregiver-rated instrument that takes about 10 to fifteen minutes to administer. Training is required and a users manual describing administration and scoring and a training video are available from the authors at a cost of \$US35.00 and \$US50.00 respectively. The ADRQOL is also available in English, Spanish and Greek. Information on, or permission to use these translations are available from the author.

Evidence from several studies indicates the psychometric properties of the ADRQOL are very good. Most studies have provided considerable information to ensure the findings can be appropriately interpreted. Evidence for excellent reliability comes from several studies (Rabins, Kasper, Kleinman and Black, 1999; Edelman, Fulton and Kuhn, 2004; Edelman, Fulton, Kuhn and Chang, 2005; Sloane, Zimmerman, Williams, Reed, et al. 2005) that report Cronbach's alpha ranging from 0.80 to 0.91 for the total score and 0.29 to 0.82 for the subscales. Inter-rater reliability is also excellent with intra-class correlations (ICC's) of 0.99 for the total score and 0.90 to 1.00 for the subscales (Sloane, Zimmerman, Williams, Reed, et al. 2005). Construct validity has been confirmed in terms of expected relationships between ADRQOL and other measures.

Studies have reported good construct validity in relation to hypothesised relationships (Gonzalez-Salvador, Lyketsos, Baker, Hovanec, et al. 2000; Edelman, Fulton and Kuhn, 2004; Edelman, Fulton, Kuhn and Chang, 2005; Samus, Rosenblatt, Steele, Baker, et al. 2005). ADRQOL scores showed significant correlations with Severe Impairment Scale (SIRS), Mini Mental State Examination (MMSE), Neuropsychiatric Inventory (NPI), Psychogeriatric Dependency rating Scale – Activities of daily Living Scale (PGDRS-ADL), Cornell Scale for Depression in Dementia (CSDD) and Philadelphia Geriatric Centre Affect rating Scale ((PGC-ARS). Construct validity has also been confirmed through studies showing correlations with other QoL measures (Edelman, Fulton and Kuhn, 2004; Edelman, Fulton, Kuhn and Chang, 2005; Sloane, Zimmerman, Williams, Reed, et al. 2005). ADRQOL has been found to have significant correlations with Quality of Life in Alzheimer's Disease (QOLAD), Quality of Life Dementia (QOL-D), Dementia Care mapping (DCM), and Resident Staff Observation Checklist (RSOC).

Only one study (Lyketsos, Gonzales-Salvador, Chin, Baker, et al. 2003) was found that investigated ADRQOL in relation to responsiveness. The authors investigated the change over time of HRQOL for persons with dementia residing in a long-term care facility. Results showed a small but significant decline in ADRQOL scores over the two year study period.

#### **5.4.6 Cornell Brown Scale for Quality of Life in Dementia (CBS)**

The CBS, adapted from the Cornell Brown Scale for Depression in Dementia, provides a global assessment of quality of life in persons diagnosed with dementia (Ready, Ott, Grace and Fernandez, 2002). It is a readily available clinician rated instrument suitable for use with persons with mild to moderate dementia. The CBS comprises 19 items covering five domains: mood related signs; ideational disturbances; behavioural disturbance; physical signs and cyclic functions. It takes about 10 to 20 minutes to administer, is free to use and a manual for administration is available from the authors. A Spanish translation of the CBS is pending.

Evidence regarding the psychometric properties of the CBS is limited; however the evidence from the only study (Ready, Ott, Grace and Fernandez, 2002) available indicates the instrument has good reliability and validity. Findings from this study indicate internal consistency very good with Cronbach's alpha of 0.81 and inter-rater reliability was excellent with an intra-class correlation (ICC) of 0.90. Construct validity in terms of the extent to which scores on the CBS relate to other measures in a manner consistent with theoretically derived hypotheses concerning the domains covered was confirmed with CBS showing expected correlations with the Visual analogue Dysphoria scale (VADS) and the Mini-Mental State Examination (MMSE). Scores on the CBS were also correlated with the Clinical dementia rating Scale (CDR) indicating support for discriminant validity.

**Table 25 Summary of Comparative Ratings for Six Short-listed Dementia-Specific HRQOL Instruments**

Criteria	Weight	Instrument					
		QOL-AD	DEM QOL	QUALID	DQOL	CBS	ADR QOL
Availability of comparison data	3	2	2	2	2	1	2
Length/feasibility of instrument for inclusion in battery	2	3	2	3	2	2	1
Complexity of administration /cognitive burden	2	3	3	3	3	3	3
Cultural Appropriateness	1	2	1	2	1	1	2
Ease of obtaining score	2	2	2	2	2	2	1
Sensitivity to dementia	3	3	3	2	2	2	2
Reliability evidence available	3	3	3	3	3	3	3
Validity evidence available	3	3	3	3	3	3	3
Cost of the instrument	2	3	2	2	2	3	1
Cost of instrument administration	2	2	2	2	2	1	2
<b>Weighted Total</b>		<b>61</b>	<b>56</b>	<b>56</b>	<b>53</b>	<b>50</b>	<b>48</b>

There was evidence that each of the six dementia-specific HRQOL instruments had good psychometric properties, so this was not an attribute that distinguished well between them. There was some evidence that the QOL-AD and DEMQOL were more sensitive to differences in dementia status, as defined by external criteria such as MMSE. There were more citations, and hence more evidence, for the QOL-AD, giving it the advantage of generalisability across a range of samples and studies. Since all are relatively easy for interviewers and respondents to understand, this attribute did not distinguish between them. All except the ADRQOL are easy to administer by interview, with detailed scripts and no formal training required for interviewers. It is noteworthy that none of the instruments has Australian reference data available.

The QOL-AD was the highest ranked instrument overall, having the best profile of attributes overall and the largest number of citations. It is short, easily administered, available in patient and proxy forms, and available in ten languages in addition to English.

The DEMQOL was the second highest ranked instrument. Relative to the QOL-AD, it was disadvantaged by being longer, having a cost for commercial use and being available only in English, while the QOL-AD is available in an additional ten languages.

The QOL-AD and DEMQOL both have the advantage of being available in both patient and proxy versions, and therefore can be used across the spectrum of stages of dementia. The only qualification here is that the domains they cover may be less relevant to people with advanced or late-stage dementia, as involvement in social activities and self care becomes less probable and cognition becomes more impaired as the disease progresses.

The D-QOL was disadvantaged by not having a proxy form available, by being relatively long, and available only in English. Both the ADR-QOL and the CBS were very disadvantaged by not having patient versions available, since it is widely agreed that a fundamental aspect of HRQOL is that it should be rated by the patient whenever possible. The ADRQOL was further disadvantaged by being much longer than the other instruments, and by having significant costs associated with training and administration.

### **5.5 Patient Versus Proxy (Carer) Report of HRQOL**

In reviewing the evidence for the DEMQOL and the QOLAD, it became clear that there are important differences in how people with dementia and their carers conceive of and report HRQOL. The first point is apparent in the content of the DEMQOL (28 items) and DEMQOL Proxy (31 items), which have only 14 items in common as a result of rigorous item selection and reduction procedures applied independently to patient and proxy responses (Smith, et al. 2005). These results suggest that people with dementia are more concerned than carers with fitting into social networks and being socially accepted, while carers reflect more on deterioration in memory and self-care and lack of insight. The second point, differences in report of HRQOL, is apparent in the consistently low correlations between patient-report and carer-proxy report for the 14 common items of the DEMQOL. The QOLAD results further corroborate this, with numerous studies reporting consistently poor agreement and low correlation between patient and proxy scores (Edelman, Fulton, and Kuhn, 2004; Edelman, Fulton, Kuhn, and Chang, 2005; Fuh and Wang, 2006; Hoe, Hancock, Livingston, and Orrell, 2006; Logsdon, Gibbons, McCurry, and Teri, 1999, 2002; Shin, Carter, Masterman, Fairbanks, and Cummings, 2005; Sloane, et al. 2005; Spector and Orrell, 2006). The availability of both patient and proxy forms was therefore considered a key attribute when judging the relative value of instruments.

### **5.6 Recommendations**

After considering the key attributes of the instruments, and all the evidence about their psychometric properties, it was provisionally decided that three instruments be recommended for the assessment of HRQOL in dementia, the QOL-AD and the DEMQOL for mild to moderate dementia and the QUALID for late stage dementia only.

Based on current evidence, as presented above and in Appendix 6, the QOL-AD is clearly the strongest instrument, and if only one dementia-specific HRQOL instrument were to be allowed, then it would be the one. The decision to recommend a further two instruments was based on two factors. Firstly, late stage dementia is very different to mild or moderate dementia, in terms of both the issues that define and affect quality of life and also the way HRQOL can be measured or observed. This factor led to the recommendation of QUALID, given the relevance and appropriateness to late stage dementia of its content and mode of measurement. The second factor was the newness of the DEMQOL balanced against the world-class credentials of its development team – it is an instrument whose promise is yet to be realized. Although limited, the available evidence suggests that the psychometric properties of both DEMQOL and DEMQOL Proxy are at least as good as those of the QOL-AD.

Since none of the instruments have published Australian reference data, it is recommended that such data be collected in an Australian field test of these instruments. The greatest value would be achieved from such an exercise by administering the QOL-AD and DEMQOL concurrently, and in both patient and proxy form, in persons with mild and moderate dementia, and administering proxy versions of QOL-AD and DEMQOL along with QUALID in advanced dementia. This would provide valuable reference and comparative data for all instruments and forms. It would also provide valuable insights into systematic differences between patient and proxy reports in Australian sample. Such a reference dataset would provide normative comparator data against

which to interpret the scores from studies applications which are likely to use just one of these instruments and/or forms.

## References

Australian Bureau of Statistics (1997) *1995 National Health Survey: SF-36 Population Norms*. Australia. Cat No 4399.0. ABS, Canberra.

Australian Bureau of Statistics (1998) *Mental Health and Wellbeing: Profile of Adults Australia 1997*. Australia, Cat No 4326.0 ABS, Canberra.

Addington-Hall J and Kalra L (2001) *Who should measure quality of life?* Bmj. Vol. 322, No.7299, pp.1417-20.

Aisen PS, Schafer KA, Grundman M, et al. (2003) *Effects of rofecoxib or naproxen vs. placebo on Alzheimer disease progression: a randomized controlled trial.[see comment]*. JAMA. Vol. 289, No.21, pp.2819-26.

Banerjee S, Smith SC, Lamping DL, et al. (2006) *Quality of life in dementia: more than just cognition. An analysis of associations with quality of life in dementia*. Journal of Neurology, Neurosurgery and Psychiatry. Vol. 77, No.2, pp.146-8.

Bergner M, Bobbitt RA, Kressel S (1976) *The Sickness Impact Profile: conceptual formulation and methodology for the development of a health status measure*. Int J Health Serv. Vol. 6, pp. 393-415.

Bergner M, Bobbitt RA, Carter WB, et al. (1981) *The Sickness Impact Profile: development and final revision of a health status measure*. Med Care. Vol. 19, pp.787-805.

Black BS, Rabins PV and Kasper JD (2000) *Alzheimer disease Related Quality of Life User's manual*, Baltimore, Maryland: DEMeasure.

Brady MJ, Peterman AH, Fitchett G, et al. (1999) *A case for including spirituality in quality of life measurement in oncology*. Psycho-Oncology. Vol. 8, No.5, pp.417-28.

Brod M, Stewart AL, Sands L and Walton P (1999) *Conceptualization and measurement of quality of life in dementia: the dementia quality of life instrument (DQoL)*. Gerontologist. Vol. 39, No.1, pp.25-35.

Calman KC (1984) *Quality of life in cancer patients--an hypothesis*. Journal of Medical Ethics. Vol. 10, No.3, pp.124-7.

Chapman SB, Weiner MF, Rackley A, et al. (2004) *Effects of cognitive-communication stimulation for Alzheimer's disease patients treated with donepezil*. Journal of Speech Language and Hearing Research. Vol. 47, No.5, pp.1149-63.

Coyle N, Goldstein ML, Passik S, et al. (1996) *Development and validation of a patient needs assessment tool (PNAT) for oncology clinicians*. Cancer Nursing. Vol. 19, No.2, pp.81-92.

Edelman P, Fulton BR and Kuhn D (2004) *Comparison of dementia-specific quality of life measures in adult day centers*. Home Health Care Services Quarterly. Vol. 23, No.1, pp.25-42.

Edelman P, Fulton BR, Kuhn D, et al. (2005) *A comparison of three methods of measuring dementia-specific quality of life: perspectives of residents, staff, and observers*. Gerontologist. Vol. 45 Spec No.1, No.1, pp.27-36.

Fayers PM and Machin D (2000) *Quality of Life: Assessment, Analysis and Interpretation*. John Wiley & Sons Ltd.

Froberg DG and Kane RL (1989) *Methodology for measuring health-state preferences--I: Measurement strategies*. Journal of Clinical Epidemiology. Vol. 42, No.4, pp.345-54.

- 
- Fuh JL and Wang SJ (2006) *Assessing quality of life in Taiwanese patients with Alzheimer's disease*. International Journal of Geriatric Psychiatry. Vol. 21, No.2, pp.103-7.
- Gill TM and Feinstein AR (1994) *A critical appraisal of the quality of quality-of-life measurements*. Journal of the American Medical Association. Vol. 272, No. 8, pp. 619-26.
- Gonzalez-Salvador T, Lyketsos CG, Baker A, Hovanec L, et al. (2000) *Quality of life in dementia patients in long-term care*. International Journal of Geriatric Psychiatry. Vol. 15, No.2, pp.181-9.
- Harvey PD, Moriarty PJ, Kleinman L, et al. (2005) *The validation of a caregiver assessment of dementia: the Dementia Severity Scale*. Alzheimer Disease and Associated Disorders. Vol. 19, No.4, pp.186-94.
- Hayes V, Morris J, Wolfe C and Morgan M (1995) *The SF-36 health survey questionnaire: Is it suitable for use with older adults?* Age and Ageing. Vol. 24, pp 120-125.
- Hoe J, Hancock G, Livingston G, et al. (2006) *Quality of life of people with dementia in residential care homes*. British Journal of Psychiatry. Vol. 188, pp.460-4.
- Hoe J, Katona C, Roch B, et al. (2005) *Use of the QOL-AD for measuring quality of life in people with severe dementia--the LASER-AD study*. Age and Ageing. Vol. 34, No.2, pp.130-5.
- Lawton MP (1994) *Quality of life in Alzheimer disease*. Alzheimer Disease and Associated Disorders. Vol. 8 Suppl. 3, pp.138-50.
- Logsdon RG, Gibbons LE, McCurry SM, et al. (1999) *Quality of life in Alzheimer's disease: Patient and caregiver reports*. Journal of Mental Health and Ageing. Vol. 5, No.1, pp.21-32.
- Logsdon RG, Gibbons LE, McCurry SM, et al. (2002) *Assessing quality of life in older adults with cognitive impairment*. Psychosomatic Medicine. Vol. 64, No.3, pp.510-9.
- Luoma M-L, Vaarama M and Hertto P (2005) *Quality of life (QoL) among demented nursing home residents - Do different instruments provide different results?* 18th Congress of the International Association of Gerontology, Rio de Janeiro, Brazil.
- Lyketsos CG, Gonzales-Salvador T, Chin JJ, Baker A et al. (2003) *A follow-up study of change in quality of life among persons with dementia residing in a long-term care facility*. International Journal of Geriatric Psychiatry. Vol. 18, No.4, pp.275-81.
- Mador J, Hecker J and Clark M (2003) *Evaluation of donepezil in Alzheimer's disease*. Australasian Journal on Ageing. Vol. 22, No.3.
- McDowell I (2006) *Measuring Health: A Guide to Rating Scales and Questionnaires (3rd ed.)* Oxford University Press, Oxford.
- McHorney CA (1996) *Measuring and monitoring general health status in elderly persons: Practical and methodological issues in using the SF-36 health survey*. The Gerontologist. Vol. 36, No.5, Health Module, pp. 571.
- Martin-Cook K, Hynan LS, Rice-Koch K, et al. (2005) *Responsiveness of the quality of life in late-stage dementia scale to psychotropic drug treatment in late-stage dementia*. Dementia and Geriatric Cognitive Disorders. Vol. 19, No.2-3, pp.82-5.
- Nelson EC, Wasson J, Kirk J, et al. (1987) *Assessment of function in routine clinical practice: description of the COOP Chart method and preliminary findings*. J Chronic Dis. Vol. 40, Suppl.1, pp.55S-63S.
- Novella JL, Jochum C, Ankri J, et al. (2001) *Measuring general health status in dementia: practical and methodological issues in using the SF-36*. Journal of Ageing-Clinical and Experimental Research. Vol. 13, pp 362-9.
- Novella JL, Boyer F, Jochum C, et al. (2006). *Agreement between patients' and proxies' reports of quality of life in Alzheimers disease*. Quality of Life Research. Vol. 10, No 5, pp 443-52.
-

- Orrell M, Spector A, Thorgrimsen L, et al. (2005) *A pilot study examining the effectiveness of maintenance Cognitive Stimulation Therapy (MCST) for people with dementia.* International Journal of Geriatric Psychiatry. Vol. 20, No.5, pp.446-51.
- Parker SG, Peet SM, Jagger C, et al (1998) *Measuring health status in older patients. The SF-36 in practice.* Age and Ageing. Vol. 27, No.1, pp.13-18.
- Patrick DL and Bergner M (1990) *Measurement of health status in the 1990s.* Annual Review of Public Health. Vol. 11, pp.165-183.
- Pavot W, Diener E, Colvin CR, et al. (1991) *Further validation of the Satisfaction with Life Scale: evidence for the cross-method convergence of well-being measures.* Journal of Personality Assessment. Vol. 57, No.1, pp.149-61.
- Pettit T, Livingston G, Manela M, et al. (2001) *Validation and normative data of health status measures in older people: the Islington study.* International Journal of Geriatric Psychiatry. Vol. 16, No.11, pp.1061-70.
- Power M, Quinn K, Schmidt S and the WHOQOL-OLD Group (2005) *Development of the WHOQOL-OLD module.* Quality of Life Research. Vol. 14, pp.2197-2214.
- Rabins PV and Black BS (2007) *Measuring quality of life in dementia: purposes, goals, challenges and progress.* Consensus conference paper, International Psychogeriatrics.
- Rabins PV and Kasper JD (1997) *Measuring quality of life in dementia: conceptual and practical issues.* Alzheimer Disease and Associated Disorders. Vol. 11 Suppl.6, pp.100-4.
- Riemsma RP, Forbes CA, Glanville JM, et al. (2001) *General health status measures for people with cognitive impairment: learning disability and acquired brain injury.* Health Technology Assessment. Vol. 5, No.6, pp.1-100.
- Rabins PV, Kasper JD, Kleinman JD and Black BS (1999) *Concepts and methods in the development of the ADRQL: An instrument for assessing health-related quality of life in persons with Alzheimer's disease.* Journal of Mental Health and Ageing. Vol. 5, No.1, pp.33-48.
- Ready RE, Ott BR and Grace J (2004) *Patient versus informant perspectives of Quality of Life in Mild Cognitive Impairment and Alzheimer's disease.* International Journal of Geriatric Psychiatry. Vol. 19, No.3, pp.256-65.
- Ready RE, Ott BR, Grace J and Fernandez I (2002) *The Cornell-Brown Scale for Quality of Life in dementia.* Alzheimer Disease and Associated Disorders. Vol. 16, No.2, pp.109-15.
- Samus QM, Rosenblatt A, Steele C, Baker A, et al. (2005) *The association of neuropsychiatric symptoms and environment with quality of life in assisted living residents with dementia.* Gerontologist. Vol. 45 Spec No 1, No.1, pp.19-26.
- Schipper H, Clinch JJ and Olweny CLM (1996) *Quality of life studies: definitions and conceptual issues.* In *Spilker B (ed.) Quality of Life and Pharmacoeconomics in Clinical Trials.* (2nd ed.) Lippincott-Raven, New York. pp.11-23.
- Selwood A, Thorgrimsen L and Orrell M (2005) *Quality of life in dementia--a one-year follow-up study.* International Journal of Geriatric Psychiatry. Vol. 20, No.3, pp.232-7.
- Seymour DG, Ball AE, Russell EM, et al. (2001) *Problems in using health survey questionnaires in older patients with physical disabilities. The reliability and validity of the SF-36 and the effect of cognitive impairment.* Journal of Evaluation in Clinical Practice. Vol. 7, pp. 411-8.
- Sherbourne CD and Meredith L (1992) *Quality of self-report data: A comparison of older and younger chronically ill patients.* Journal of Gerontology. Vol. 47, pp.S204-S211.

- 
- Shin IS, Carter M, Masterman D, et al. (2005) *Neuropsychiatric symptoms and quality of life in Alzheimer disease*. American Journal of Geriatric Psychiatry. Vol. 13, No.6, pp.469-74.
- Slevin ML, Plant H, Lynch D, et al. (1988) *Who should measure quality of life, the doctor or the patient?* British Journal of Cancer. Vol. 57, No.1, pp.109-12.
- Sloane PD, Zimmerman S, Williams CS, et al. (2005) *Evaluating the quality of life of long-term care residents with dementia*. Gerontologist. Vol. 45, Spec No 1, No.1, pp.37-49.
- Smith SC, Lamping DL, Banerjee S, et al. (2005) *Measurement of health-related quality of life for people with dementia: development of a new instrument (DEMQL) and an evaluation of current methodology*. Health Technology Assessment (Winchester, England). Vol. 9, No.10, pp.1-93.
- Smith SC, Murray J, Banerjee S, et al. (2005b) *What constitutes health-related quality of life in dementia? Development of a conceptual framework for people with dementia and their carers*. International Journal of Geriatric Psychiatry. Vol. 20, No.9, pp.889-95.
- Spector A and Orrell M (2006) *Quality of Life (QoL) in Dementia: A Comparison of the Perceptions of People With Dementia and Care Staff in Residential Homes*. Alzheimer Disease and Associated Disorders. Vol. 20, No.3, pp.160-5.
- Spector A, Thorgrimsen L, Woods B, et al. (2003) *Efficacy of an evidence-based cognitive stimulation therapy programme for people with dementia: randomised controlled trial.[see comment]*. British Journal of Psychiatry. Vol. 183, pp.248-54.
- Spitzer WO (1987) *State of science 1986: Quality of life and functional status as target variables for research*. Journal of Chronic Diseases. Vol. 40, No.6, pp.465-471.
- Teri L, McCurry SM, Logsdon R, et al. (2005) *Training community consultants to help family members improve dementia care: a randomized controlled trial*. Gerontologist. Vol. 45, No.6, pp.802-11.
- Thomas S, Nay R, Fonda D, et al. (2006) *Continence Outcomes Measurement Suite Project: Final Report*, Australian Government Department of Health and Ageing, Canberra.
- Thorgrimsen L, Selwood A, Spector A, et al. (2003) *Whose quality of life is it anyway? The validity and reliability of the Quality of Life-Alzheimer's Disease (QoL-AD) scale.[see comment]*. Alzheimer Disease and Associated Disorders. Vol. 17, No.4, pp.201-8.
- Valvanne J, Luoma M-L, Ylonen L, et al. (2005) *Quality of life among clients in home care - Are there differences between demented and non-demented clients*. 18th Congress of the International Association of Gerontology. Rio de Janeiro, Brazil.
- Ware JE (1987) *Standards for validating health measures: definition and content*. Journal of Chronic Diseases. Vol. 40, No.6, pp.473-480.
- Ware JE, Kosinski M and Dewey JE (2001) *How to Score Version Two of the SF-36 Health Survey* (2nd ed.) QualityMetric Incorporated, Lincoln RI.
- Ware JE, Kosinski M, Turner-Bowker DM and Gandek B (2002) *How to Score Version 2 of the SF-12 Health Survey (With a Supplement Documenting Version 1)*. QualityMetric Incorporated, Lincoln RI
- Ware J (2003) *Standardizing health metrics: the SF-36 Health Survey and beyond*. In Sansoni J and Tilley L (eds.) Conference Proceedings, Health Outcomes 2003: The Quest for Practice Improvement. Canberra, 20-21 August 2003.
- Warner SC and Williams JI (1987) *The Meaning in Life Scale: determining the reliability and validity of a measure*. Journal of Chronic Diseases. Vol. 40, No.6, pp.503-12.
- Weiner MF, Martin-Cook K, Svetlik DA, et al. (2000) *The quality of life in late-stage dementia (QUALID) scale*. Journal of the American Medical Directors Association. Vol. 1, No.3, pp.114-116.
-

WHOQOL Group (1998a) *The World Health Organization quality of life assessment (WHOQOL): Development and general psychometric properties.* Social Science and Medicine. Vol. 12, pp.1569-1585.

WHOQOL Group (1998b) *Group development of the World Health Organization WHOQOL-BREF quality of life assessment.* Psychological Medicine. Vol. 28, pp.551-558.

Winzelberg GS, Williams CS, Preisser JS, et al. (2005) *Factors associated with nursing assistant quality-of-life ratings for residents with dementia in long-term care facilities.* Gerontologist. Vol. 45, Spec No.1, No.1, pp.106-14.

Wood-Dauphinee S (1999) *Assessing quality of life in clinical research: from where have we come and where are we going?* Journal of Clinical Epidemiology. Vol. 52, No.4, pp.355-63.

Wood-Dauphinee S and Williams JI (1987) *Reintegration to Normal Living as a proxy to quality of life.* Journal of Chronic Diseases. Vol. 40, No.6, pp.491-502.

Woods B, Thorgrimsen L, Spector A, et al. (2006) *Improved quality of life and cognitive stimulation therapy in dementia.* Ageing and Mental Health. Vol. 10, No.3, pp.219-226.

## 6 Instruments for the Assessment of Cognitive Status

### 6.1 Cognition in Dementia

Dementia has, as its central feature, impairment in cognitive function. The cognitive deficit most often manifests itself as memory problems and difficulty in the ability to retain new information. However, memory is only one of the cognitive skills affected in dementia. Other affected areas are attention, language, speech, recognition, confusion, reasoning, judgement, problem solving, and disorientation in time place and person. This impairment in cognitive function in turn affects the person's abilities to engage successfully in activities of daily living.

Problems of cognition form a spectrum, beginning from mild decline in recall and memory, or other areas of cognitive functioning such as reasoning, concentration, finding the appropriate word, all of which may be part of normal aging. At the other end of the spectrum lies dementia. Assessment and understanding of cognitive impairment in dementia is therefore crucial to any treatment of the disorder. Behavioural observation can play a limited role in the assessment of mental ability, but cognition can only be accurately assessed through the use of objective psychometric tests.

### 6.2 Measuring Cognitive Status in Dementia

Roth (1981) defined dementia as "the global deterioration of an individual's intellectual, emotional and cognitive abilities in a state of unimpaired consciousness". Three elements in this definition hold implications for the measurement of dementia.

- First, it implies a decline from a previously higher level of functioning. Measurement should therefore measure alterations in state, not just current state.
- Global deterioration implies several types of functional losses. Although memory loss is the central feature, it is not unique to dementia. As stated above, dementia also implies limitations in other functions. These include aphasia (disorders of language generally due to lesions in the left hemisphere of the brain), apraxia (disorders in performing purposeful movements, of which constructional apraxia reflects problems with visual and motor integration), and agnosia (disorders of recognition). Dementia is therefore not a single condition, but a complex of symptoms, and screening tests therefore need to have a broad content.
- "Unimpaired consciousness" relates to the knowledge that symptoms of dementia may be mimicked by reversible conditions such as depression, intoxication, delirium, or an acute confusional state. These conditions must therefore be excluded before a diagnosis of dementia can be made.

### 6.3 Measurement Instruments

Cognitive function tests can be divided into three main categories: intelligence tests, laboratory tests and clinical neuropsychological tests. This section discusses clinical neuropsychological tests, which includes mental status screening tests (including short, simple tests) as well as detailed tests of specific cognitive functions.

Clinical neuropsychological tests provide an in-depth assessment of functions such as orientation, executive function or praxis. Most mental status tests broadly assess orientation to time and place, tests of concentration and attention, and memory tests for short and long-term recall. The focus of this report was mainly on short mental status tests.

### 6.3.1 In-depth Clinical Neuropsychological Tests

A wide range of cognitive abilities are typically assessed in a comprehensive neuropsychological evaluation. These include:

- Learning and memory
- Attention and concentration
- Speech and language abilities
- Executive function (abstraction, problem solving and reasoning)
- General intellectual competence
- Visuo-spatial and visuo-constructional skills
- Sensory-perceptual abilities
- Psychomotor speed

A good in-depth measurement tool for dementia should test for the maximum number of cognitive abilities and include at least attention, expressive and receptive language, memory, constructional ability and abstract reasoning. These sorts of tests would often be used as second stage assessments.

### 6.3.2 Mental Status Tests (Including Simple Screening Tests)

These tests draw elements, used to assess specific aspects of cognitive functioning, from clinical neuropsychological tests. Many were developed by physicians because of the difficulties they experienced administering full neuropsychological test batteries to elderly patients. The focus on simplicity and practicality has resulted in the following criticisms of these tests:

- They are too narrow in scope and therefore may be insensitive to early stages of cognitive decline and unable to distinguish normal decline due to aging from pathological decline;
- They may not distinguish between the more severe levels of dementia; and
- Designing a structured test that is not affected by differing education level and cultural background is difficult.

Because of these limitations several alternatives have been developed including:

- Self report which can be reliable but not valid because people with cognitive impairment often cannot evaluate their own performance;
- Observation by clinical/nursing staff which are useful in inpatient settings but not in the community setting; and
- Observation by an informant such as a close relative.

### 6.3.3 Combination of Tests

Because dementia is a syndrome with several characteristic features all measurement instruments include separate components to assess these features. However, few instruments can discriminate across all levels and types of dementia. Therefore, some authors/test developers suggest a combination of tests in one instrument (Katzman, 1986; Shore, Overman, and Wyatt, 1983; Welsh, Butters, Hughes and Mohs, 1991), for instance, combining a test that is effective in distinguishing mild cognitive impairment from normal cognitive function with one that is suited to differentiating among more advanced stages of dementia.

## 6.4 Reviewed Instruments

Details of the literature search strategies used are outlined in the introduction (refer Section 2). The initial search strategy identified 73 measures which could be classified as measures of cognitive functioning. Following the search strategy (textword search), a CD-Rom was produced containing relevant papers and abstracts for each identified instrument.

Based on this work an impact sheet was developed for consideration by the review teams and the DOMS-EMG. This considered MEDLINE, text and web impacts, presence in instrument databases (PROQOLID) and its use in clinical practice. The latter was based on NEP and field surveys and clinical feedback. This process produced a list of 12 or so instruments which were regarded as leading contenders for comprehensive review.

Using the impact measure approach, the following cognitive instruments were regarded as contenders: Alzheimer's Disease Assessment Scale – Cognitive (ADAS-Cog) (Rosen, Mohs and Davis, 1984); Modified Mini Mental Status Exam (3MS) (Teng and Chui, 1987); Telephone Interview of Cognitive Status (TICS) (Brandt, et al. 1988); Short Portable Mental Status Questionnaire (SPMSQ) (Pfeiffer, 1975); Clock Drawing Test (CDT) (Sunderland, et al. 1989); Cognitive Capacity Screening Examination (CCSE) (Jacobs, et al. 1977); General Practitioner Assessment of Cognition (GPCOG) (Brodaty, et al. 2000); Rowland Universal Dementia Assessment Scale (RUDAS) (Storey, et al. 2004); Memory Impairment Screen (MIS) (Buschke, et al. 1999); Mini-Cog (Borson, et al. 2000); Abbreviated Mental Test Score (AMTS) (Hodkinson, 1972); MDS-Cog (Morris, et al. 1994); Brief Cognitive Rating Scale (BCRS) (Reisberg and Ferris, 1988); and the Psychogeriatric Assessment Scales – Cognition (PAS-Cog) (Jorm, et al. 1995) (*The PAS-Cog is used in the new Aged Care Funding Instrument (ACFI).*)

For the telephone administration of cognitive tests, the Telephone Interview of Cognitive Status (TICS) (Brandt, et al. 1988) could be considered for people with mild dementia. It is reliable and correlates highly with the MMSE (Burns, et al. 2004). Though, like the MMSE, it appears to be a proprietary instrument.

A major issue for discussion was the boundary between a cognitive instrument and a neuropsychological instrument. It was decided that a test requiring verbal fluency items or memory items requiring cued recall or recognition recall components, would not be able to be widely implemented because of the degree of skill and training required to administer and score these measures. These are best left to trained professionals with experience of the patient population.

This approach, therefore, excluded a number of detailed cognitive instruments on the border between neuro-psychology and cognitive testing as they required a high degree of skill in administration and training. These included: Addenbrooke's Cognitive Examination (ACE) (Mathuranath, et al. 2000), the Cambridge Cognitive Examination Revised (CAMCOG-R) (part of the Cambridge Mental Disorders of the Elderly Examination Revised [CAMDEX - R]) (Roth, et al. 1986), the Mattis Dementia Rating Scale (MDRS) (also known as the Dementia Ratings Scale-2; Extended Scale for Dementia) (Mattis, 1976), the Seven-Minute Screen (7MS) (also known as the Seven-Minute Neuro-cognitive Screening Battery) (Solomon, et al. 1998), the Severe Impairment Battery (SIB) (Saxton, et al. 1990) for severe or late stage dementia and the Test for Severe Impairment (TSI) (Albert and Cohen, 1992).

This list of contender instruments was then reviewed by the team using the selection criteria outlined in the Section 2.

The profusion of instruments assessing cognitive functioning made selection of instruments for this chapter difficult. The initial literature search and review procedure led to a short list of five instruments measuring cognitive functioning in dementia.

The final five instruments selected for review were the:

1. Alzheimer's Disease Assessment Scale – Cognition (ADAS-COG),
2. General Practitioner Cognition Scale (GPCOG),
3. Modified Mini Mental State Exam (3MS),
4. Minimum Data Set – Cognition (MDS-COG), and the
5. Rowland Universal Dementia Assessment Scale (RUDAS).

*Note: It should be noted that the cognitive component of the Blessed Dementia Scale, the Blessed Information - Memory Concentration Test (BIMCT) (also known as the Blessed Orientation Memory Concentration Test or the Short Blessed Test) is also being reviewed in the Dementia Staging and Descriptive Instruments Section of this report.*

These instruments were selected because they covered a range of settings including primary care and nursing homes. The ADAS-COG was selected because it is a highly cited instrument and its component parts cover a range of cognitive tasks. A further two instruments that did not make it to the final five were the Mini-Cog (Borson, Scanlan, Brush, Vitaliano, and Dokmak, 2000) and the Memory Impairment Screen (MIS) (Buschke, et al. 1999). While both these instruments have been shown to be suitable for use in the primary care setting, particularly for screening, they were not included in our final five for review. Reasons for this are outlined below in the section discussing the General Practitioner Assessment of Cognition.

It should also be noted that the review of terminology in Section 3 indicates that recognition of mild cognitive impairment (MCI) is important and clinicians need to be vigilant about its further development to dementia, however there is insufficient evidence as yet to embrace Mild Cognitive Impairment (MCI) as a new diagnosis. At the first meeting of the National Expert Panel (NEP), the members agreed that given MCI is not fully established as a proper diagnosis and as the DOMS project focuses on the clinical phase of diagnosis it is best not to be included in this project. It is also noted that a related project by Cherbuin, et al. (2006) has a specific focus on reviewing dementia screening instruments to facilitate early detection of dementia and Mild Cognitive Impairment.

Cherbuin, et al. (2006) have recently completed a review of dementia screening instruments (self-assessed and informant report) suitable for placement on a web site. They initially short listed 25 instruments (largely assessing cognitive impairment in dementia) against a range of psychometric criteria (reliability, validity, sensitivity, specificity, misclassification rate etc) and in relation to the benchmark of the psychometric properties of the MMSE. A number of the tools recommended in this report are included amongst their initial selection of twenty-five measures.

Cherbuin, et al. (2006) then examined these measures for their applicability for completion on a web site and also excluded instruments that had not been validated in a community or population sample. As some instruments require, for example, activities such as drawing a clock, measures including these were automatically excluded (e.g. GPCOG). They found no self-assessment measure was currently suitable for completion on the web although contender screening instruments for dementia were the Memory Impairment Screen and the Six Item Screen (Callahan, et al. 2002). However, they recommended that the applicability of these instruments for online administration and the relevant cut scores for dementia would need to be further assessed. This study, however, did recommend the IQ Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) (Jorm, 1989) for placement on the National Dementia Website.

Although there are some similarities between the projects in terms of their methodological approach, the focus of the Cherbuin, et al. (2006) study is limited to a) instruments suitable as screening measures and b) instruments that can be administered on a web site. Whereas this project has a focus which includes all stages of assessment and has an orientation to assessment

in routine clinical practice rather than on-line assessment. Given these considerations it is not surprising that there are some differences concerning the recommended instruments from these reports.

A further literature search was conducted for the five short-listed instruments to identify all peer-reviewed published papers that reported their development, testing or application - the following online bibliographic databases were used - CINAHL, EMBASE, MEDLINE and PSYCINFO. The instruments' developers were also contacted by email to provide information about costs and availability of users manual, language translations and so on. The papers were then obtained and reviewed, and the following attributes were considered: content of the instrument, in terms of the numbers items and their coverage; the stage of dementia that the instrument is suitable for; the availability and cost of the instrument; training requirements for interviewers or observers; availability of user manuals; administration time; the number of citations; evidence about the psychometric properties and use in practice to date. Each instrument was also rated against the criteria described in Table 1 at the beginning of this document (Criteria and weights used to assess instruments). Recommendations were then made following a rationale described in detail below.

A promising new instrument, the Kimberley Indigenous Cognitive Assessment (KICA-Cog) tool for the assessment of cognitive impairment of Indigenous people was also identified. Although there is limited evidence available as yet concerning its' psychometric properties; as the other identified tools are unlikely to be suitable for remote indigenous peoples, this tool has also been briefly reviewed in this section.

The attributes of each of the five short-listed instruments and the KICA-Cog are described below. Further details for the five short-listed instruments are provided in the instrument review sheets (see Appendix 7).

#### **6.4.1 Modified Mini Mental State Exam (3MS) (Teng and Chui, 1987)**

Assessing global cognitive function has been a keystone in screening for dementia and cognitive impairment, and evaluating clinical and non-clinical interventions. The most well known and widely utilised tool in both research and clinical practice is the Mini-Mental State Exam (MMSE) (Folstein, Folstein and McHugh, 1975).

Over the years, variations of the MMSE have been developed to remedy limitations of the MMSE, which include insufficient guidelines for its application, dichotomised responses disallowing credit for near misses, narrow score range (0-30), floor and ceiling effects, false positive responses due to low education, and limited sensitivity and specificity particularly for a mild form of cognitive impairment or dementia (McDowell, 2006). Two more commonly adopted variations include the Standardised Mini-Mental State Examination (SMMSE) (Molloy, Alemayehu, and Roberts, 1991), designed to improve consistency in administering the MMSE including explanatory questions, time restrictions in answering to the questions and scoring methods, and the Modified Mini-Mental State Examination (3MS) (Teng and Chui, 1987), designed to improve reliability and validity of the tool, minimise the floor/ceiling effects, and enhance discrimination of various levels of cognitive abilities among people with cognitive impairment and dementia. A more recent development and less known instrument is the Severe Mini-Mental State Examination (SMMSE) (Harrell, Marson, Chatterjee, and Parrish, 2000), modelled after the MMSE (0-30 points) and designed specifically for people with a moderate to severe form of cognitive impairment or dementia. Whilst having improved some aspects of the MMSE, neither of the SMMSE versions (Harrell, Marson, Chatterjee, and Parrish, 2000; Molloy, Alemayehu, and Roberts, 1991) appears to have overcome the shortcomings of the MMSE, and there is insufficient evidence to support their psychometric properties.

Due to these similar or same abbreviations of the variations it has been a challenging exercise in reviewing the literature for the DOMS project. Anecdotal evidence also indicates clinicians often do not differentiate between the variations, which can lead to confusion and misinterpretation and utilization of the tool.

After more than three decades the MMSE is still recommended as the standard tool for screening cognitive impairment and dementia (Boustani, Peterson, Hanson, Harris, and Lohr, 2003; McDowell, 2006) largely because it is sufficiently brief (5-10 minutes required) and easy to score, well known and familiar amongst clinicians and researchers internationally, and has psychometric properties that are within the acceptable range.

Despite it being slightly more time demanding to administer and score (minimum 10 minutes) and to master its application methods, the review indicates the 3MS is superior to the MMSE in all psychometric evaluations. Nevertheless, it is believed that the superiority of the 3MS in psychometric evaluations outweighs minor problems in mastering the application of the tool and time required. It is also notable that the MMSE is copyrighted by Psychological Assessment Resources, Inc. and costs about US\$1 per test form, whereas the 3MS is available free of charge.

Derived from the MMSE, the 3MS is an interviewer rated tool to assess cognitive function in terms of orientation, registration, recall, simple language, and construction, which includes four additional items of long-term memory, verbal fluency, abstract thinking and the recall of the three words an additional time (McDowell, 2006; Teng and Chui, 1987). The range of scores for the 3MS is broader (0-100) compared to the MMSE, which contributes to improved validity of the tool (McDowell, 2006). Cultural applicability needs further adaptation however an increasing number of translated versions are becoming available.

There are also various norms and adjustments available based on age, some ethnic groups and education levels (Bravo and Hebert, 1997a; Brown, Schinka, Mortimer, and Graves, 2003; Jones, et al. 2002; Tombaugh, McDowell, Kristjansson, and Hubley, 1996; Tschanz, et al. 2002). However, studies show inconsistent results in terms of the impact of using adjusted norms based on demographic factors when detecting cognitive impairment and dementia. Whilst adjustments for age, education and sensory impairment resulted in improved sensitivity and specificity to screen for dementia (Hayden, et al. 2003; Khachaturian, Gallo, and Breitner, 2000), findings from a large population-based study showed the use of age and education adjusted normative data resulted in reduced validity of the instrument as well as reducing sensitivity to dementia (O'Connell, Tuokko, Graves, and Kadlec, 2004). Caution needs to be taken in using adjusted norms and further research is needed to substantiate the existing claims (McDowell, 2006).

With respect to this classification of the severity of cognitive impairment Wlodarczyk, et al. (2003) suggest the most commonly accepted score ranges for classifying the severity of cognitive impairment are: scale cut-points of <10 to indicate severe cognitive impairment, 10–14 moderate cognitive impairment, 15–19 mild to moderate cognitive impairment, and 20–24 mild impairment. A similar classification of severity and guidance for interpretation are outlined by the National Institute for Health and Clinical Excellence (NICE) (2007) in the United Kingdom. These are the classifications of severity that are used in this report.

#### **6.4.2 Alzheimer's Disease Assessment Scale – Cognition (ADAS-Cog)**

##### **ADAS**

The Alzheimer's Disease Assessment Scale (ADAS) is a clinical rating scale developed specifically to assess the major cognitive, affective and behavioural deficits in persons with Alzheimer's disease (AD) (Rosen, Mohs, and Davis, 1984). It is however, relevant to dementia from any cause and can also be used in other settings where the evaluation of cognitive functioning is required.

This 21 item scale comprises two sections: an 11 item cognitive subscale that employs short psychological tests of memory, language and praxis function; and a 10 item non-cognitive subscale that rates mood, vegetative function, agitation, hallucinations, delusions, and concentration and distractibility. The total score on the ADAS ranges from 0 to 120 (cognitive section: 0 to 70, non-cognitive section: 0 to 50). A higher score indicates poorer performance.

The ADAS has been found to have excellent inter-rater reliability (Rosen, Mohs, and Davis, 1986), and good to excellent test-retest reliability (one month interval between session) (Rosen, Mohs, and Davis, 1986) for persons with AD (Weyer, Erzigkeit, Kanowski, Ihl, and Hadler, 1997). Construct and discriminative validity have also been demonstrated in several studies (Burch and Andrews, 1987; Ihl, Frolich, Dierks, Martin, and Maurer, 1992; Rosen, Mohs, and Davis, 1984; Zec, et al. 1992). The instrument has been widely used in longitudinal studies and in clinical trials including persons with AD. However, in clinical trials, the cognitive subscale only, is typically used as a primary outcome measure and the 10 item non-cognitive component is not as widely used or as highly regarded as the ADAS-Cog.

### **ADAS-Cog**

The ADAS-Cog, the cognitive section of the ADAS, is used either as part of the full ADAS, or alone, specifically to assess cognitive impairment. The instrument is widely used as an outcome measure in clinical trials (Aisen, et al. 2003; Farlow, et al. 1992; Knapp, et al. 1994; Mador, Hecker, and Clark, 2003; Riepe, et al. 2006; Rockwood, 2004; Rogers, Farlow, Doody, Mohs, and Friedhoff, 1998; Rogers and Friedhoff, 1996) and other intervention studies (Olazaran, et al. 2004; Spector, et al. 2003), as well as longitudinal studies evaluating change in cognitive impairment over time (Cortes, et al. 2005; Doraiswamy, Bieber, Kaiser, Connors, et al. 1997; Doraiswamy, Bieber, Kaiser, Krishnan, et al. 1997; Doraiswamy, Kaiser, Bieber, and Garman, 2001; Feldman, Van Baelen, Kavanagh, and Torfs, 2005; Rosen, Mohs, and Davis, 1986; Schmeidler, Mohs, and Aryan, 1998; Serra, et al. 2004; Suh, Ju, Yeon, and Shah, 2004; Weiner, Vobach, Svetlik, and Risser, 1993). As a component of the full ADAS, it has been translated into most European languages as well as Chinese, Turkish, Indian and Brazilian. Particular attention has been given to Spanish versions (Hannedottir and Snaedal, 2002; Kolibas, Korinkova, Novotny, Vajdickova, and Hunakova, 2000; Liu, et al. 2002; Manzano, Llorca, Ledesma, and Lopez-Ibor, 1994; Pascual, et al. 1997; Pena-Casanova, Aguilar, et al. 1997; Tsolaki, Fountoulakis, Nakopoulou, Kazis, and Mohs, 1997; Youn, et al. 2002).

The instrument consists of 11 items: memory (orientation to time place and person, word recall, word recognition, and recall of test instructions on word recognition); language (naming objects and fingers, spoken language, language comprehension, word finding difficulty, and following commands); and praxis (ideational and constructional). It was developed by a team of experts in the field. Items were selected from a variety of existing instruments or constructed specifically for the scale, based on clinical observations and experimental investigations. Validation studies were conducted to ensure that all domains were adequately covered (Rosen, Mohs, and Davis, 1986; Rosen, Mohs, and Davis, 1984).

ADAS-Cog is usually administered by a neuropsychologist or psychologist and takes about 30 to 45 minutes to complete, depending on the level of cognitive impairment. It can be administered by other personnel, as administration and scoring is not difficult, but training is required. As such it is more appropriate for use in specialist or research settings rather than in routine care. The total score, which indicates the level of impairment, is the sum of the scores on each of the items. As stated above, the total score can range from 0 to 70. Maximum scores for each of the domains are memory (35), language (25) and praxis (10). For the domains and for the total score, a higher score indicates greater impairment.

There is considerable evidence indicating the ADAS-Cog has very good to excellent psychometric properties. Most studies have reported considerable information to ensure the findings can be appropriately interpreted. The internal structure of the instrument has been confirmed through

factor analysis confirming the domains proposed by the authors (Kim, Nibbelink, and Overall, 1994), as well as significant item to item correlations ranging from 0.17 to 0.55 (Doraiswamy, Bieber, Kaiser, Connors, et al. 1997; Doraiswamy, Bieber, Kaiser, Krishnan, et al. 1997; Doraiswamy, Kaiser, Bieber, and Garman, 2001), and item to total correlations ranging from 0.52 to 0.90 (Liu, et al. 2002). Significant correlations (0.47 to 0.52) between domain scores have also been reported (Kim, Nibbelink, and Overall, 1994; Talwalker, Overall, Srirama, and Gracon, 1996).

The instrument has excellent internal consistency and test-retest reliability (Chu, et al. 2000; Kim, Nibbelink, and Overall, 1994; Liu, et al. 2002; Pena-Casanova, Aguilar, et al. 1997; Pena-Casanova, Meza, et al. 1997; Rosen, Mohs, and Davis, 1984; Weyer, Erzigkeit, Kanowski, Ihl, and Hadler, 1997). Studies report Cronbach's alpha ranging from 0.75 to 0.96 and intra class correlations (ICC) of 0.86 to 0.96 for the total score and 0.33 to 0.89 for the individual items. The relatively low ICC of 0.33 (for ideational praxis) was reported in one study only (Liu, et al. 2002). Test-retest was also conducted on the factor structure and an ICC of 0.78 to 0.87 was reported (Kim, Nibbelink, and Overall, 1994). The ADAS-Cog also has excellent inter-rater reliability with studies reporting ICCs for the total scores ranging from 0.91 to 0.99 and for the individual items ranging from 0.76 to 1.00 (Chu, et al. 2000; Liu, et al. 2002; Mohs and Cohen, 1988; Rosen, Mohs, and Davis, 1984).

There is evidence that the instrument has construct validity in terms of the extent to which scores on the ADAS-Cog relate to other measures in a manner that is consistent with theoretically derived hypotheses concerning the domains measured by results (Doraiswamy, Kaiser, Bieber, and Garman, 2001; Feldman, Van Baelen, Kavanagh, and Torfs, 2005; Lam, Lui, Tam, and Chiu, 2005; Suh, Ju, Yeon, and Shah, 2004). Expected correlations were found between ADAS-Cog scores and the Disability Assessment in Dementia (DAD) as well as the number of memory complaints, (Memscore).

Considerable evidence of construct validity in terms of correlations with other instruments measuring cognitive functioning is also available (Baxter, et al. 2006; Blessed, Tomlinson, and Roth, 1968; Burch and Andrews, 1987; Doraiswamy, Bieber, Kaiser, Connors, et al. 1997; Doraiswamy, Bieber, Kaiser, Krishnan, et al. 1997; Doraiswamy, Kaiser, Bieber, and Garman, 2001; Hannesdottir and Snaedal, 2002; Ihl, Frolich, Dierks, Martin, and Maurer, 1992; Ihl, Grass-Kapanke, Janner, and Weyer, 1999; Liu, et al. 2002; Pena-Casanova, Meza, et al. 1997; Serra, et al. 2004; Silvestrini, et al. 2006; Suh, Ju, Yeon, and Shah, 2004; Weyer, Erzigkeit, Kanowski, Ihl, and Hadler, 1997; Zec, et al. 1992). The ADAS-Cog has been found to have significant correlations with the following instruments: the Mini Mental State Examination (MMSE); Brief Cognitive Rating Scale (BCRS); Memory and Information Test (MIT); Cognitive Abilities Scoring Instrument (CASI); Nurses Observation Scale for Geriatric Impairment (NOSGER); Syndrom-Kurz-Test (SKT); and the Geriatric Evaluation by Relatives Rating Scale (GERRI). It also correlates with the Clinical Dementia Rating Scale (CDR), and the Computer Neuropsychological Test battery Scores, global instruments that include cognitive subscales (Chu, et al. 2000; Cutler, et al. 1993).

Discriminant validity has been confirmed through numerous studies. The instrument has been shown to discriminate between different levels of dementia severity (Doraiswamy, Bieber, Kaiser, Connors, et al. 1997; Doraiswamy, Bieber, Kaiser, Krishnan, et al. 1997; Liu, et al. 2002; Rosen, Mohs, and Davis, 1984; Schmeidler, Mohs, and Aryan, 1998; Wang, et al. 2004; Weyer, Erzigkeit, Kanowski, Ihl, and Hadler, 1997; Zec, et al. 1992) as well as between dementia and no dementia (Hannesdottir and Snaedal, 2002; Pena-Casanova, Meza, et al. 1997; Schultz, Siviero, and Bertolucci, 2001; Zec, et al. 1992). It has also been associated with changed brain structure in one study reporting a relationship between ADAS scores and a decrease in grey brain matter (Baxter, et al. 2006).

Further support for discriminative validity comes from studies reporting the instrument's predictive ability. ADAS-Cog scores have been shown to be a good predictor of mild cognitive impairment (Lam, Lui, Tam, and Chiu, 2005) as well as the level of dependency on a carer (Caro, et al. 2002).

There is considerable evidence that the ADAS-Cog meets the criteria for responsiveness. Due to the nature of dementia, it is expected that cognitive functioning will decline over time. A measurement instrument therefore needs to be sensitive to this change. Several studies have shown that the ADAS-Cog is sensitive to change over time with study results showing significant differences in scores from baseline to 6 months (Doraiswamy, Kaiser, Bieber, and Garman, 2001; Suh, Ju, Yeon, and Shah, 2004), and from baseline to 12 months (Farlow, et al. 1992; Feldman, Van Baelen, Kavanagh, and Torfs, 2005; Holford and Peace, 1992; Rosen, Mohs, and Davis, 1984; Schmeidler, Mohs, and Aryan, 1998; Serra, et al. 2004; Suh, Ju, Yeon, and Shah, 2004; Weiner, Vobach, Svetlik, and Risser, 1993).

The instrument has also been shown to be sensitive to the effects of drug treatment. Clinical trials investigating the effectiveness of drug treatment for persons with Alzheimer's disease confirmed the sensitivity of ADAS-COG (Burns, et al. 1999; Cortes, et al. 2005; Imbimbo, Troetel, Martell, and Lucchelli, 2000; Jones, et al. 2004; Riepe, et al. 2006; Rockwood, 2004; Tariot, et al. 2001; Zemlan, 1996). Scores improved significantly after 6 months treatment. Scores on the MMSE and the CDR also improved, confirming the sensitivity of ADAS-Cog. Studies investigating the effectiveness of Donepezil for persons with vascular dementia also show improvement in ADAS-Cog scores along with improved scores on the MMSE, CDR and the Clinician's Interview-Based Impression of Change (CIBIC) (Malouf and Birks, 2004; Roman, et al. 2005).

There is only limited evidence relating to sensitivity of the ADAS-Cog to interventions such as cognitive stimulation therapy. We found only one study (Spector, et al. 2003) reporting evidence for this. In this study ADAS-Cog scores for those in the intervention group improved significantly compared to those in the control group. Scores on the MMSE and QOL-AD scales also improved significantly.

### **6.4.3 General Practitioner Cognition Scale (GPCOG)**

In the primary care setting, brief screening tools for cognitive impairment are a valuable tool that can be used by general practitioners (GPs) in diagnostic investigations. The General Practitioner Cognition Scale (Brodaty, et al. 2002) is one such instrument. Two other instruments, the Mini-Cog (Borson, Scanlan, Brush, Vitaliano, and Dokmak, 2000) and the Memory Impairment Screen (MIS) (Buschke, et al. 1999) have also been shown to be suitable for use in the primary care setting. The GPCOG was chosen for review instead of these instruments for a number of reasons. While the attributes of the MIS and the Mini-Cog are comparable (psychometric attributes for the MIS are slightly better) (Brodaty, Low, Gibson, and Burns, 2006; Lorentz, Scanlan, and Borson, 2002), they do have some drawbacks. The Mini-Cog does not include cued recall and there is no available evidence about its reliability. The MIS has a narrow focus, and assesses only memory. In addition, use of the MIS in the field would be limited due to the materials needed and the degree of training required (staff would require training to give cued recall questions). Therefore the GPCOG was the instrument chosen for review in this field setting.

The GPCOG is a relatively new instrument (Brodaty, et al. 2002) that has been developed to assist GPs in detecting dementia. Items for the scale were derived from other instruments measuring cognitive, physical and psychological functioning and the geriatric population. The GPCOG is readily accessible and quick and easy to administer and score. It consists of 9 items covered in two sections: Cognitive testing (patient examination), which consists of four items (word recall, time orientation, clock drawing, reporting a recent event); and historical (informant interview), which consists of 6 items (patient's memory of recent conversations, misplacing objects, word finding difficulties, ability to manage money, ability to manage medication, and need for travel assistance).

The instrument is administered to the patient (9 items) and the informant (6 items) to obtain a more definite rating. A score of 9 (out of 9) on the patient section indicates no cognitive impairment,

while a score of 4 or lower suggests cognitive impairment. If scores are in the range of 5 – 8 then cognitive impairment is regarded as being doubtful / uncertain. The informant section should then be completed to obtain more information. A score of 3 or lower here suggests cognitive impairment.

This is a very new instrument and hence there is limited evidence regarding its psychometric properties. However, evidence that is available suggests the instrument has very good reliability and validity. Only one study reported reliability for the instrument (Brodaty, et al. 2002). Internal consistency and test-retest reliability was found to be very good with Cronbach's alpha of 0.84 for both the patient and informant section, and ICC's of 0.87 for both sections. Inter-rater reliability was good for the patient (ICC of 0.75), and satisfactory for the informant section (ICC of 0.56).

Support for construct validity comes from two studies (Brodaty, Kemp, and Low, 2004; Brodaty, et al. 2002). Results indicate that scores on the GPCOG significantly correlated with two other instruments that measure cognitive functioning: the Mini-Mental State Exam (MMSE) and the Global Deterioration Scale (GDS). The diagnostic accuracy of the instruments attests to its discriminant validity. In one study (Brodaty, et al. 2002) the area under the curve (ROC) was 0.86 and sensitivity and specificity were 82% and 70% respectively for the patient section. Positive and negative predictive values were 0.53 and 0.90. In the informant section ROC was 0.84 and sensitivity and specificity were 89% and 66% respectively. Positive and negative predictive values were 0.52 and 0.94. These findings were similar to those found for the MMSE. In another study (Thomas P, et al. 2006) sensitivity and specificity were 82% and 70% respectively, and positive and negative predictive values were 0.53 and 0.90. Patient and informant sections were not evaluated separately in this study. The instrument's ability to differentiate between disease stages provides further support for discriminant validity. Results from one study show that scores on the GPCOG differentiate between patients with and without dementia (Brodaty, Kemp, and Low, 2004). However, this evidence is limited to discriminating between no dementia and dementia. The instrument's ability to distinguish between all stages of dementia has yet to be confirmed. To date, there is no evidence available regarding the responsiveness of the GPCOG.

As stated above, in the primary care setting, brief screening tests for cognitive impairment are a valuable tool that can be used by general practitioners (GPs) for initial diagnostic assessment.

The Mini-Cog (Borson, Scanlan, Brush, Vitaliano, and Dokmak, 2000) and the Memory Impairment Screen (MIS) (Buschke, et al. 1999) have also been shown to be suitable for use in the primary care setting but they had a less comprehensive coverage of cognitive domains.

#### **6.4.4 Rowland Universal Dementia Assessment Scale (RUDAS)**

The Rowland Universal Dementia Assessment Scale (RUDAS) (Storey, Rowland, Basic, Conforti, and Dickson, 2004) is a short multicultural cognitive screening tool for the assessment of dementia. It was developed and validated in an area where 40% of the population are born in non-English speaking countries and more than 80 languages are spoken (Rowland, Basic, Storey, and Conforti, 2006). It was developed by a team of experts in the field of dementia care in consultation with representatives from 22 cultural and linguistic groups.

The RUDAS is an interviewer administered, six item questionnaire, covering the following cognitive domains: memory, visuo-spatial orientation, praxis, visuo-construction drawing (cube drawing), judgement and language. The instrument is scored out of 30 with scores below 23 suggesting dementia. Item scores are summed to give a total score. Total possible individual items scores are memory (8), visuo-spatial orientation (5), praxis (2), visuo-construction drawing (3), judgment (4), language (8). The interview takes about 10 minutes to complete. Training is required but access to training materials are easily accessible and inexpensive (\$15.00). For interviews involving persons from non-English speaking backgrounds (NESB) an interpreter is used. The

RUDAS can be simply translated into other languages, without the need to change the structure or the format of an item (Storey, Rowland, Basic, Conforti, and Dickson, 2004).

This is a relatively new instrument and evidence relating to its psychometric properties is at this stage limited. Existing data, from the original validation study does however, indicate the RUDAS has excellent test-retest, and inter-rater reliability with ICC's of 0.98 and 0.99 respectively and item-total correlations ranging from 0.35 to 0.50 (Storey, Rowland, Basic, Conforti, and Dickson, 2004). The construct validity of the instrument is supported in this study with RUDAS scores significantly correlated with the MMSE. The instrument also has good diagnostic accuracy. Evidence from three studies report area under the receiver operated curves' (ROC) figures ranging from 0.86 to 0.94, sensitivity and specificity figures ranging from 72% to 89% and 76% to 100% respectively. The sensitivity and specificity figures were better than those for the MMSE (67% and 95%) and the GPCOG (86% and 67%). The items are relevant to most cultures and the instrument can be directly translated into other languages without the need to change the structure or format of any item. The RUDAS was found not to be affected by gender or educational background (Storey, Rowland, Basic, Conforti, et al., 2004).

#### **6.4.5 Minimum Data Set – Cognition (MDS-COG)**

The Minimum Data Set – Cognition (MDS-COG) Scale is a component of the full Minimum Data Set developed as a data collection method to be used in all nursing homes in the United States (Hartmaier, Sloane, Guess, and Koch, 1994). Although cognitive instruments are available to estimate the severity of dementia they are often not feasible for use in nursing home populations due to the need for skilled personnel for administration, and excessive administration costs. The MDS-Cognitive performance scale (MDS-CPS) (Morris, et al. 1994) was therefore developed to enable MDS data to be obtained to provide a valid measure of cognitive impairment. Following the development of the MDS-CPS, the MDS-COG was developed to provide a continuous measure as opposed to the hierarchical MDS-CPS. InterRAI, an international collaborative of scientists and clinicians that provide procedures for enhancing clinical care utilising standardised clinical protocols and data collection in aged care services, has now incorporated the MDS-COG into their interRAI-LTCF (Long Term Care Facility), for persons with complex care requirements in residential care settings.

The MDS-COG is an interviewer administered questionnaire using data that is routinely collected by staff on a patient's entry to the long term care facility. Administration time is 10 to 20 minutes. It combines 8 items from the Minimum Data Set (MDS), in use at all care facilities, into a simple 10 point additive scale. Items cover the following domains: Cognitive patterns (short term memory, long term memory, location of own room, knows he/she is in a nursing home, no orientation items recalled and decision making; Communication patterns (making self understood); and Physical Functioning (dressing self performance). The instrument is scored as an additive scale that ranges from 0 for no cognitive impairment to 10 for very severe cognitive impairment.

Evidence relating to the psychometric properties of the MDS-COG is limited. Only one paper was found regarding the reliability of the instrument (Gruber-Baldini, Zimmerman, Mortimore, and Magaziner, 2000). The authors reported a Cronbach's alpha of 0.85 and item - total correlations ranging from 0.32 to 0.81, which provides some support for the internal structure of the instrument. Evidence for validity is also limited but does indicate that the MDS-COG has construct validity with scores showing expected correlations with the Psychogeriatric Dependency Rating Scale (PGDRS), orientation and behaviour scales ( $r = 0.66$  and  $0.31$ ) and the Katz Activities of Daily Living (ADL) (Hartmaier, Sloane, Guess, and Koch, 1994). Correlations with other measures of cognitive functioning (GDS, MMSE, and the MDS-CPS) were also reported (Cohen-Mansfield, Taylor, McConnell, and Horton, 1999; Gruber-Baldini, Zimmerman, Mortimore, and Magaziner, 2000; Hartmaier, Sloane, Guess, and Koch, 1994). There is also some support for discriminative validity. The instrument appears to have very good diagnostic accuracy with ROC value of 0.94

and sensitivity and specificity of 89% and 98% (Hartmaier, Sloane, Guess, and Koch, 1994). There is no evidence relating to responsiveness. (It should be noted that the structure of the Cognitive and Communication sections of the interRAI-TLCF, which includes items predominantly from the MDS-COG with some of the original MDS-CPS, is continually evolving.)

#### **6.4.6 Kimberley Indigenous Cognitive Assessment (KICA-Cog)**

As there was no validated tool to assess cognition of Indigenous Australians this instrument was designed by Lo Giudice, et al. (2005) to address this deficiency. The KICA-Cog was adapted from previous 'culture fair' instruments although the authors report that none of these instruments were completely suitable for this group as they included the use of some concepts which did not translate well into Indigenous languages. The tool comprises cognitive, informant and functional sections but the focus of this review is on the cognitive section. It is a new tool (LoGiudice, et al. 2005) and there is little published evidence concerning its psychometric properties available as yet. It has largely been tested with rural and remote Indigenous peoples and needs to be further assessed in both urban and rural and remote settings.

KICA Cog is a cognitive rating scale which has 16 questions and the total score can range from 0-39 with lower scores indicating increased cognitive impairment. It assesses orientation, free and cued recall, verbal fluency, copying sequence pattern and ideation praxis (refer Appendix 15) and thus predominantly assess memory and language skills. It has limited coverage of executive function and the authors note the need for identifying more sophisticated ways to assess executive function in this community.

Lo Giudice, et al. (2005) report that 77 subjects were assessed systematically with the KICA-Cog. Inter-rater reliability for the 16 items was very good to excellent with the kappa for most items greater than or equal to  $k=0.6$ . Internal consistency as assessed by Cronbach's alpha was 0.87. It appeared to successfully discriminate between Indigenous people with and without dementia. The psychometric properties from this instrument are promising but they need to be replicated in further studies and the tool was accepted well by the Indigenous people participating in the study.

#### **6.4.7 Other Approaches to Cognitive Assessment**

Other cognitive instruments not reviewed in depth in this report are those measures of cognition obtained by informant or proxy reports. Two noteworthy informant measures, the Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) and the AD8 are briefly discussed below.

The Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) (Jorm, 2004) is an Australian developed and widely used, informant based measure to screen for dementia. The short (and recommended version) of the questionnaire includes 16 items examining everyday cognitive abilities (e.g. remembering own telephone number and learning new things), with a few functional items (e.g. handling money for shopping). The IQCODE has excellent reliability properties (Cronbach's alpha = 0.93 – 0.97; test-retest reliability = 0.96 [timeframe = 3 days] and 0.75 [timeframe = 12 months]), and correlates well with the MMSE in the range of -0.37 to -0.78. It has well developed validity data, including comparison studies with clinical diagnosis, neuropathology, neuroimaging and other cognitive and informant tests. In terms of informant / proxy measures, McDowell acknowledges that the IQCODE is the leader in this field (McDowell, 2006 p. 454). In terms of further development, there is some debate regarding its basic uni-dimensional factor structure and how informants seem to make global judgments about cognitive decline, and how they do not seem to distinguish between different cognitive processes addressed by IQCODE items (McDowell, 2006).

While the AD8 (Galvin, et al. 2006) is a new and brief informant based screening instrument. It takes about 3 minutes to complete, asking the informant to rate changes in memory, problem

solving, orientation and daily activities (8 items), using a Yes / NO response format. It was designed to distinguish between demented and non demented individuals in a clinic sample, using the CDR as its measure of criterion validity (correlation is 0.74). The AD8 has good discriminating properties and adequate to good reliability (Cronbach's alpha = 0.86; test-retest kappa = 0.67 over a 2-3 week timeframe). It also correlates -0.41 with the MMSE. In sum, the AD8 is a promising screening tool for dementia, but it requires further psychometric development work, especially in primary care and community care settings. Because of its apparent simplicity and ease of use it may be worthwhile to undertake a study of the AD8's screening ability in the Australian community. Using a cut-off score of 2 or 3 the AD8 has good sensitivity (approx. 90%) to predict dementia in a clinic sample, though its reported specificity could be improved (approx. 46%).

Measurement issues regarding the use of informant / proxy measures are discussed in Section 12.3. A new Australian computerised measure, the CogState is also briefly discussed in Section 12.5.

## 6.5 Recommendations

Consideration of the attributes described above lead to a weighted total score for each instrument provided in Table 26. These scores, as well as considerations relating to applicability in different field settings, have led us to make the following recommendations. The recommended instruments are the MMSE (3MS), GPCOG and the MDS-COG.

**Table 26 Summary of Ratings for Cognitive Instruments**

Criteria	Weight	Instrument					
		MMSE (3MS)	ADAS-COG	GPCOG	RUDAS	MDS-COG	KICA-COG <sup>b</sup>
Availability of comparison data	3	3	2	1	1	1	1
Length/feasibility of instrument for inclusion in battery	2	2	1	3	3	3	2
Complexity of administration/ cognitive burden	2	2	2	3	2	3	2
Cultural Appropriateness	1	2	3	1	3	1	3 <sup>b</sup>
Ease of obtaining score	2	3	3	3	3	3	3
Sensitivity to dementia	3	3	3	2	2	2	2
Reliability evidence	3	3	3	3 <sup>a</sup>	3 <sup>a</sup>	2 <sup>a</sup>	2
Validity evidence	3	3	3	3 <sup>a</sup>	3 <sup>a</sup>	3 <sup>a</sup>	1
Cost of the instrument	2	3	3	3	1	2	3
Cost of instrument administration	2	2	1	1	2	2	2
<b>Weighted Total</b>		<b>62</b>	<b>56</b>	<b>54</b>	<b>52</b>	<b>51</b>	<b>45</b>

Notes:

<sup>a</sup> Scored as '2' or '3' because despite there being limited evidence, what there is indicates good reliability and or validity.

<sup>b</sup> This is a new tool designed for the cognitive assessment of Indigenous people.

The MMSE (3MS) is a widely used instrument that assesses global cognitive status in older people. It is applicable in both community and institutional settings. It has superior psychometric properties and has been extensively used in large scale epidemiological studies internationally (mostly North American studies). There is also extensive normative and clinical data available. An increasing number of studies use a translated version of the 3MS to achieve cultural appropriateness. The instrument equals or outscores all the other instruments in almost every category.

It is noted that the ADAS-Cog received the second highest score and it is clear it has excellent psychometric properties. However, some component parts of the ADAS-Cog require observational training as well as skills in psychological test administration (especially the word recall and word recognition tasks). The ADAS-Cog also requires additional test materials and it takes 30-45 minutes for completion of the assessment. Although it is widely used in clinical trials, the ADAS-Cog may be more appropriate for second stage assessments and for particular research evaluations rather than for applications in routine care settings.

It should be noted there are a range of other neuro-psychological tests and cognitive assessment batteries that are used for more in depth assessment of cognitive function. Some of these instruments have been discussed in section 6.4. A decision was made by the DOMS-EMG that the project should focus on the instruments that are suitable for use in routine care and this would exclude many of the more detailed neuropsychological instruments or cognitive instruments that require specialist training for their administration and interpretation. However, it is recommended that a further study could examine neuropsychological and cognitive assessment batteries for people with dementia, in association with the relevant professional groups. Within such a project the ADAS-Cog should be compared with newer instruments (using other memory recall and recognition items) which also provide a detailed assessment of cognitive function.

The GPCOG is recommended because of its usefulness in the primary care setting. As it is a relatively new instrument, it has not been widely used in research studies, normative data is not available, and the instrument has not been translated in to any other languages. Despite this, the GPCOG has scored highly on the ranking criteria. In addition, anecdotal evidence suggests the GPs are using the instrument and finding it very useful.

The MDS-COG is recommended, despite having the lowest ranking total. The reason for this is that it was felt it was important to include an instrument that would be useful in the residential care setting. The strength of this instrument is that it enables evidence about the cognitive status patients to be obtained without any extra effort on the part of staff. The information is routinely entered as the patient enters long term care. Despite the total score concerning its psychometric properties being slightly lower than the other instruments, the individual attributes are more than adequate.

The RUDAS is a new instrument that was designed to enable the easy translation of the items into other languages and to be culture fair. There are relatively few papers published as yet concerning its psychometric properties (especially construct validity) but in the interim it is recommended for use with those from Culturally and Linguistically Diverse backgrounds. The RUDAS, however, contains an item on judgement that may be inappropriate for remote Indigenous people (refer below).

Another instrument in this class is the Kimberley Indigenous Cognitive Assessment (KICA) tool which has been designed for use with Indigenous people. An interim recommendation, pending further research, is to use the Kimberley Indigenous Cognitive Assessment (KICA) tool for the

cognitive assessment of rural and remote Indigenous people. The KICA is a new instrument and although there is little published evidence concerning this tool available as yet, and further research is required, this instrument has been designed for use with Indigenous people.

## References

Aisen PS, Schafer KA, Grundman M, et al. (2003) *Effects of rofecoxib or naproxen vs. placebo on Alzheimer disease progression: a randomized controlled trial.* [see comment]. JAMA. Vol. 289, No.21, pp.2819-26.

Albert M and Cohen C (1992) *The Test for Severe Impairment: An instrument for the assessment of patients with severe cognitive dysfunction.* Journal of the American Geriatrics Society. Vol. 40, pp.449-453.

Baxter LC, Sparks DL, Johnson SC, et al. (2006) *Relationship of cognitive measures and gray and white matter in Alzheimer's disease.* Journal of Alzheimer's Disease. Vol. 9, No.3, pp.253-60.

Blessed G, Tomlinson BE and Roth M (1968) *The association between quantitative measures of dementia and of senile change in cerebral grey matter of elderly subjects.* British Journal of Psychiatry. Vol. 114, pp.797-811.

Borson S, Scanlan JM, Brush M, et al. (2000) *The Min-Cog: A cognitive 'vital signs' measure for dementia screening in multi-lingual elderly.* International Journal of Geriatric Psychiatry. Vol. 15, pp.1021-1027.

Boustani M, Peterson B, Hanson L, Harris R and Lohr K (2003) *Summary of the Evidence: Screening for dementia in primary care*, available at <http://www.ahcpr.gov/clinic/3rduspstf/dementia/demsum2.htm>, accessed 21 June, 2006.

Brandt J, Spencer M, Folstein M (1988) *The telephone interview for cognitive status.* Neuropsychiatry Neuropsychol Behav Neurol. Vol. 1, pp.111-117.

Bravo G and Hebert R (1997) *Age and education-specific reference values for the Mini-Mental and modified Mini-Mental State Examinations derived from a non-demented elderly population.* International Journal of Geriatric Psychiatry. Vol. 12, No.10, pp.1008-1018.

Brodaty H, Kemp NM and Low LF (2004) *Characteristics of the GPCOG, a screening tool for cognitive impairment.* International Journal of Geriatric Psychiatry. Vol. 19, No.9, pp.870-4.

Brodaty H, Low LF, Gibson L, et al. (2006) *What is the best dementia screening instrument for general practitioners to use?* American Journal of Geriatric Psychiatry. Vol. 14, No.5, pp.391-400.

Brodaty H, Pond D, Kemp NM, et al. (2000) *The GPCOG: A new screening test for dementia designed for general practice.* Journal of the American Geriatrics Society. Vol. 50, pp.530-534.

Brown LM, Schinka JA, Mortimer JA and Graves AB (2003) *3MS normative data for elderly African Americans.* Journal of Clinical and Experimental Neuropsychology. Vol. 25, No.2, pp.234-241.

Burch EA, Jr. and Andrews SR (1987) *Comparison of two cognitive rating scales in medically ill patients.* International Journal of Psychiatry in Medicine. Vol. 17, No.2, pp.193-200.

Burns A, Rosso RM, Hecker J, et al. (1999) *The effects of donepezil in Alzheimer's disease - results from a multinational trial.* Dementia and Geriatric Cognitive Disorders. Vol. 10, No.3, pp.237-44.

Buschke H, Kuslansky G, Katz M, et al. (1999) *Screening for dementia with the Memory Impairment Screen.* American Academy of Neurology. Vol. 52, No.2, pp.231-238.

Callahan CM, Unverzagt FW, Hui SL, Perkins AJ and Hendrie HC (2002) *Six-item screener to identify cognitive impairment among potential subjects for cognitive research.* Med Care, Vol. 40, No.9, pp.771-781.

Caro J, Ward A, Ishak K, et al. (2002) *To what degree does cognitive impairment in Alzheimer's disease predict dependence of patients on caregivers?* BMC Neurology. Vol. 2, pp.6.

- Cherbuin N, Anstey KJ and Lipnicki D (2006) *Dementia Screening Instruments: Recommendations for the National Dementia Website*. Report prepared for the Consortium for the National Dementia Website including PTG Global, Morphem, Alzheimer's Australia, Vision Australia and the Department of Health and Ageing.
- Chu LW, Chiu KC, Hui SL, et al. (2000) *The reliability and validity of the Alzheimer's Disease Assessment Scale Cognitive Subscale (ADAS-Cog) among the elderly Chinese in Hong Kong*. Annals of the Academy of Medicine, Singapore. Vol. 29, No.4, pp.474-85.
- Cohen-Mansfield J, Taylor L, McConnell D, et al. (1999) *Estimating the cognitive ability of nursing home residents from the minimum data set*. Outcomes Management for Nursing Practice. Vol. 3, No.1, pp.43-6.
- Cortes F, Gillette-Guyonnet S, Nourhashemi F, et al. (2005) *Recent data on the natural history of Alzheimer's disease: results from the REAL.FR Study*. Journal of Nutrition, Health and Ageing. Vol. 9, No.2, pp.86-93.
- Cutler NR, Shrotriya RC, Sramek JJ, et al. (1993) *The use of the Computerized Neuropsychological Test Battery (CNTB) in an efficacy and safety trial of BMY 21,502 in Alzheimer's disease*. Annals of New York Academy of Sciences. Vol. 695, pp.332-336.
- Doraiswamy PM, Bieber F, Kaiser L, et al. (1997a) *Memory, language, and praxis in Alzheimer's disease: norms for outpatient clinical trial populations*. Psychopharmacology Bulletin. Vol. 33, No.1, pp.123-8.
- Doraiswamy PM, Bieber F, Kaiser L, et al. (1997b) *The Alzheimer's Disease Assessment Scale: patterns and predictors of baseline cognitive performance in multicenter Alzheimer's disease trials*. Neurology. Vol. 48, No.6, pp.1511-7.
- Doraiswamy PM, Kaiser L, Bieber F, et al. (2001) *The Alzheimer's Disease Assessment Scale: evaluation of psychometric properties and patterns of cognitive decline in multicenter clinical trials of mild to moderate Alzheimer's disease.[erratum appears in Alzheimer Dis Assoc Disord 2002 Jan-Mar;16(1):48]*. Alzheimer Disease and Associated Disorders. Vol. 15, No.4, pp.174-183.
- Farlow M, Gracon S, Hershey L, et al. (1992) *A controlled trial of tacrine in Alzheimer's disease. The Tacrine Study Group. Clinical Trial. Multicenter Study. Randomized Controlled Trial*. JAMA. Vol. 268, No.18, pp.2523-2529.
- Feldman HH, Van Baelen B, Kavanagh SM, et al. (2005) *Cognition, function, and care giving time patterns in patients with mild-to-moderate Alzheimer disease: a 12-month analysis*. Alzheimer Disease and Associated Disorders. Vol. 19, No.1, pp.29-36.
- Folstein MF, Folstein SE and McHugh PR (1975) *Mini-Mental State: A practical method for grading the cognitive state of patients for clinicians*. Journal of Psychiatric Research. Vol. 12, pp.189-198.
- Galvin JE, Roe CM, Xiong C, Morris JC (2006) *Validity and reliability of the AD8 informant interview in dementia*. Neurology. Vol. 67, pp. 1942-1948.
- Gruber-Baldini AL, Zimmerman SI, Mortimore E, et al. (2000) *The validity of the minimum data set in measuring the cognitive impairment of persons admitted to nursing homes*. Journal of the American Geriatrics Society. Vol. 48, No.12, pp.1601-6.
- Hannesdottir K and Snaedal J (2002) *A study of the Alzheimer's Disease Assessment Scale-Cognitive (ADAS-Cog) in an Icelandic elderly population*. Nordic Journal of Psychiatry. Vol. 56, No.3, pp.201-6.
- Harrell LE, Marson D, Chatterjee A and Parrish JA (2000) *The Severe Mini-Mental State Examination: a new neuropsychologic instrument for the bedside assessment of severely impaired patients with Alzheimer disease*. Alzheimer Disease and Associated Disorders. Vol. 14, No.3, pp.168-175.
- Hartmaier SL, Sloane PD, Guess HA, et al. (1994) *The MDS Cognition Scale: a valid instrument for identifying and staging nursing home residents with dementia using the minimum data set.[see comment]*. Journal of the American Geriatrics Society. Vol. 42, No.11, pp.1173-9.

- Hayden KM, Khachaturian AS, Tschanz JT, Corcoran C, Nortond M, Breitner JCS, et al. (2003) *Characteristics of a two-stage screen for incident dementia*. Journal of Clinical Epidemiology. Vol. 56, No.11, pp.1038-1045.
- Hodkinson M (1972) *Evaluation of a mental test score for assessment of mental impairment in the elderly*. Age and Ageing. Vol. 1, pp.233-238.
- Holford NH and Peace KE (1992) *Results and validation of a population pharmacodynamic model for cognitive effects in Alzheimer patients treated with tacrine*. Proceedings of the National Academy of Sciences of the United States of America. Vol. 89, No.23, pp.11471-5.
- Ihl R, Frolich L, Dierks T, et al. (1992) *Differential validity of psychometric tests in dementia of the Alzheimer type* Psychiatry Research. Vol. 44, No.2, pp.93-106.
- Ihl R, Grass-Kapanke B, Janner M, et al. (1999) *Neuropsychometric tests in cross sectional and longitudinal studies - a regression analysis of ADAS - cog, SKT and MMSE*. Pharmacopsychiatry. Vol. 32, No.6, pp.248-54.
- Imbimbo BP, Troetel WM, Martell IP, et al. (2000) *A 6-month, double-blind, placebo-controlled trial of eptastigmine in Alzheimer's disease*. Dementia and Geriatric Cognitive Disorders. Vol. 11, No.1, pp.17-24.
- Jacobs JW, Bernhard MR, Delgado A, et al. (1977) *Screening for organic mental syndrome in the medically ill*. Annals of Internal Medicine. Vol. 86, pp.40-46.
- Jones RW, Soininen H, Hager K, et al. (2004) *A multinational, randomised, 12-week study comparing the effects of donepezil and galantamine in patients with mild to moderate Alzheimer's disease*. International Journal of Geriatric Psychiatry. Vol. 19, No.1, pp.58-67.
- Jones TG, Schinka JA, Vanderploeg RD, Small BJ, Graves AB and Mortimer JA (2002) *3MS normative data for the elderly*. Archives of Clinical Neuropsychology. Vol. 17, pp.171-177.
- Jorm AF and Jacomb PA (1989) *An Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE): Socio-demographic correlates, reliability, validity and some norms*. Psychological Medicine. Vol. 19, pp.1015-1022.
- Jorm AF (2004) *The Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE): A review*. International Psychogeriatrics. Vol. 16, No.3, pp.275-293.
- Katzman R (1986) *Differential diagnosis of dementing illness*. Neurol Clinic. Vol. 4, pp.329-340.
- Khachaturian AS, Gallo JJ and Breitner JC (2000) *Performance characteristics of a two-stage dementia screen in a population sample*. Journal of Clinical Epidemiology. Vol. 53, No.5, pp.531-540.
- Kim YS, Nibbelink DW and Overall JE (1994) *Factor structure and reliability of the Alzheimer's disease Assessment Scale in a multicenter trial with linopirdine*. Journal of Geriatric Psychiatry and Neurology. Vol. 7, No.2, pp.74-83.
- Knapp MJ, Knopman DS, Solomon PR, et al. (1994) *A 30-week randomized controlled trial of high-dose tacrine in patients with Alzheimer's disease. The Tacrine Study Group*. JAMA. Vol. 272, pp.985-991.
- Kolibas E, Korinkova V, Novotny V, et al. (2000) *ADAS-cog (Alzheimer's Disease Assessment Scale-cognitive subscale)--validation of the Slovak version*. Bratislavské Lekárske Listy. Vol. 101, No.11, pp.598-602.
- Lam LC, Lui VW, Tam CW, et al. (2005) *Subjective memory complaints in Chinese subjects with mild cognitive impairment and early Alzheimer's disease*. International Journal of Geriatric Psychiatry. Vol. 20, No.9, pp.876-82.
- Liu HC, Teng EL, Chuang YY, et al. (2002) *The Alzheimer's disease Assessment Scale: findings from a low-education population*. Dementia and Geriatric Cognitive Disorders. Vol. 13, No.1, pp.21-6.

- 
- LoGiudice D, Smith K, Thomas J, et al. (2005) *Kimberley Indigenous Cognitive Assessment tool (KICA): Development of a cognitive assessment tool for older indigenous Australians*. International Psychogeriatrics. Vol. 18, No.2, pp.269-280.
- Lorentz WJ, Scanlan JM and Borson S (2002) *Brief screening tests for dementia*. Canadian Journal of Psychiatry - Revue Canadienne de Psychiatrie. Vol. 47, No.8, pp.723-33.
- Mador J, Hecker J and Clark M (2003) *Evaluation of donepezil in Alzheimer's disease*. Australasian Journal on Ageing. Vol. 22, No.3.
- Malouf R and Birks J (2004) *Donepezil for vascular cognitive impairment*. Cochrane Database Systematic Review. Vol. 1, No.CD004395.
- Manzano JM, Llorca G, Ledesma A, et al. (1994) *Spanish adaptation of the Alzheimer's disease assessment scale (ADAS)*. Actas Luso Esp Neurol Psiquiatr Cienc Afines. Vol. 22, No.6, pp.64-70.
- Mathuranath PS, Nestor PJ, Berrios GE, et al. (2000) *A brief cognitive test battery to differentiate Alzheimer's disease and frontotemporal dementia*. Neurology. Vol. 55, No.11, pp.1613-1620.
- Mattis S (1976) *Mental status examination for organic mental syndrome in the elderly patient*. In Bellak L, Karasu T, (eds). Geriatric psychiatry: A handbook for psychiatrists and primary care physicians. Grune & Stratton, New York, pp.77-101.
- McDowell I (2006) *Measuring Health: A Guide to Rating Scales and Questionnaires (3rd ed.)*. Oxford University Press, Oxford.
- Mohs RC and Cohen L (1988) *Alzheimer's Disease Assessment Scale (ADAS)*. Psychopharmacology Bulletin. Vol. 24, No.4, pp.627-8.
- Molloy DW, Alemayehu E and Roberts R (1991) *Reliability of a Standardized Mini-Mental State Examination compared with the traditional Mini-Mental State Examination*. American Journal of Psychiatry. Vol. 148, No.1, pp.102-105.
- Morris JN, Fries BE, Mehr DR, et al. (1994) *MDS Cognitive Performance Scale*. Journal of Gerontology. Vol. 49, No.4, pp.M174-82.
- National Institute for Health and Clinical Excellence (NICE) (2007) *TA111 Alzheimer's disease - donepezil, galantamine, rivastigmine (review) and memantine: guidance (amended)*. National Institute for Health and Clinical Excellence, National Health Service, London. Available at <http://guidance.nice.org.uk/TA111/guidance/pdf/English/download.dspix>, accessed 27 September 2007.
- O'Connell ME, Tuokko H, Graves RE and Kadlec H (2004) *Correcting the 3MS for bias does not improve accuracy when screening for cognitive impairment or dementia*. Journal of Clinical and Experimental Neuropsychology. Vol. 26, No.7, pp.970-980.
- Olazarán J, Muniz R, Reisberg B, et al. (2004) *Benefits of cognitive-motor intervention in MCI and mild to moderate Alzheimer disease*. Neurology. Vol. 63, No.12, pp.2348-53.
- Pascual LF, Saz P, Larumbe R, et al. (1997) *Standardization of the Alzheimer's disease Assessment Scale in a Spanish population*. Neurologia. Vol. 12, No.6, pp.238-44.
- Pena-Casanova J, Aguilar M, Santacruz P, et al. (1997a) *Adaptation and normalization of the Alzheimer's disease Assessment Scale for Spain (NORMACODEM) (II)*. [UI: 9147454 Neurologia. Vol. 12, No.2, pp.69-77.
- Pena-Casanova J, Meza M, Bertran-Serra I, et al. (1997b) *Shortened version of the Barcelona test (III): criterion validity with the ADAS-Cog* Neurologia Croatica. Vol. 12, No.3, pp.117-9.
- Pfeiffer E (1975) *A Short Portable Mental Status Questionnaire for the assessment of organic brain deficit in elderly patients*. Journal of the American Geriatrics Society. Vol. 23, pp.433-441.
-

- Reisberg B and Ferris SH (1988) *Brief Cognitive Rating Scale (BCRS)*. Psychopharmacology Bulletin. Vol. 24, pp.629-636.
- Riepe MW, Adler G, Ibach B, et al. (2006) *Adding memantine to rivastigmine therapy in patients with mild-to-moderate Alzheimer's disease: Results of a 12-week, open-label pilot study*. Primary Care Companion to the Journal of Clinical Psychiatry. Vol. 8, No.5, pp.258-263.
- Rockwood K (2004) *Size of the treatment effect on cognition of cholinesterase inhibition in Alzheimer's disease. [erratum appears in J Neurol Neurosurg Psychiatry. 2004 Jul; 75(7):1086]*. Journal of Neurology, Neurosurgery and Psychiatry. Vol. 75, No.5, pp.677-85.
- Rogers SL and Friedhoff LT (1996) *The efficacy and safety of donepezil in patients with Alzheimer's disease: results of a US Multicentre, Randomized, Double-Blind, Placebo-Controlled Trial*. The Donepezil Study Group. Dementia. Vol. 7, No.6, pp.293-303.
- Rogers SL, Farlow MR, Doody RS, et al. (1998) *A 24-week, double-blind, placebo-controlled trial of donepezil in patients with Alzheimer's disease*. Donepezil Study Group. [see comment]. Neurology. Vol. 50, No.1, pp.136-45.
- Roman GC, Wilkinson DG, Doody RS, et al. (2005) *Donepezil in vascular dementia: combined analysis of two large-scale clinical trials*. Dementia and Geriatric Cognitive Disorders. Vol. 20, No.6, pp.338-44.
- Rosen WG, Mohs R and Davis KL (1986) *Longitudinal changes: Cognitive, behavioral and affective patterns in Alzheimer's disease*. In Poon L, editor. Handbook for clinical memory assessment of older adults. Washington, DC: American Psychological Association. pp.294-301.
- Rosen WG, Mohs RC and Davis KL (1984) *A new rating scale for Alzheimer's disease* American Journal of Psychiatry. Vol. 141, No.11, pp.1356-64.
- Roth M (1981) *The diagnosis of dementia in late and middle life*. In Mortimer JA, Schuman LM, (eds.) The epidemiology of dementia. Oxford University Press, New York. pp.24-61.
- Roth M, Tym E, Mountjoy CQ, et al (1986) *CAMDEX: A standardised instrument for the diagnosis of mental disorder in the elderly with special reference to the early detection of dementia*. British Journal of Psychiatry. Vol. 149, pp.698-709.
- Rowland JT, Basic D, Storey JE, et al. (2006) *The Rowland Universal Dementia Assessment Scale (RUDAS) and the Folstein MMSE in a multicultural cohort of elderly persons*. International Psychogeriatrics. Vol. 18, No.1, pp.111-20.
- Saxton J, McGonigle-Gibson K, Swihart A, et al (1990) *Assessment of severely impaired patients: Description and validation of a new neuropsychological test battery*. Psychological Assessment. Vol. 2, No.298-303.
- Schmeidler J, Mohs RC and Aryan M (1998) *Relationship of disease severity to decline on specific cognitive and functional measures in Alzheimer's disease*. Alzheimer Disease and Associated Disorders. Vol. 12, No.3, pp.146-51.
- Schultz RR, Siviero MO and Bertolucci PH (2001) *The cognitive subscale of the "Alzheimer's Disease Assessment Scale" in a Brazilian sample*. Brazilian Journal of Medical and Biological Research. Vol. 34, No.10, pp.1295-302.
- Serra L, Perri R, Carlesimo GA, et al. (2004) *ADAS-CoG and mental deterioration battery: Different instruments for different aspects of cognitive decline in Alzheimer's disease*. Neuroscience Research Communications. Vol. 35, No.3, pp.184-192.
- Shore D, Overman CA and Wyatt RJ (1983) *Improving accuracy in the diagnosis of dementia*. Journal of Clinical Psychiatry. Vol. 44, pp.207-212.
- Silvestrini M, Pasqualetti P, Baruffaldi R, et al. (2006) *Cerebrovascular reactivity and cognitive decline in patients with Alzheimer disease*. Stroke. Vol. 37, No.4, pp.1010-5.

Solomon PR, Hirschhoff A, Kelly B, et al. (1998) *A 7 minute neurocognitive screening battery highly sensitive to Alzheimer's disease.* Archives of Neurology. Vol. 55, pp.349-355.

Spector A, Thorgrimsen L, Woods B, et al. (2003) *Efficacy of an evidence-based cognitive stimulation therapy programme for people with dementia: randomised controlled trial.[see comment].* British Journal of Psychiatry. Vol. 183, pp.248-54.

Storey JE, Rowland JT, Basic D, et al. (2004) *The Rowland Universal Dementia Assessment Scale (RUDAS): a multicultural cognitive assessment scale.[erratum appears in Int Psychogeriatr. 2004 Jun;16(2):218].* International Psychogeriatrics. Vol. 16, No.1, pp.13-31.

Suh GH, Ju YS, Yeon BK, et al. (2004) *A longitudinal study of Alzheimer's disease: rates of cognitive and functional decline.* International Journal of Geriatric Psychiatry. Vol. 19, No.9, pp.817-24.

Sunderland T, Hill JL, Mellow AM, et al. (1989) *Clock drawing in Alzheimer's disease.* Journal of the American Geriatrics Society. Vol. 37, pp.725-729.

Talwalker S, Overall JE, Srirama MK, et al. (1996) *Cardinal features of cognitive dysfunction in Alzheimer's disease: A factor analytic study of the Alzheimer's Disease Assessment Scale.* Journal of Geriatric Psychiatry and Neurology. Vol. 9, pp.39-46.

Tariot PN, Cummings JL, Katz IR, et al. (2001) *A randomized, double-blind, placebo-controlled study of the efficacy and safety of donepezil in patients with Alzheimer's disease in the nursing home setting.* Journal of the American Geriatrics Society. Vol. 49, No.12, pp.1590-9.

Teng EL and Chui HC (1987) *The Modified Mini-Mental State (3MS) Examination.* Journal of Clinical Psychiatry. Vol. 48, pp.314-318.

Thomas P, Lalloue F, Preux PM, et al. (2006) *Dementia patients' caregiver' quality of life: the PIXEL study.* International Journal of Geriatric Psychiatry. Vol. 21, No.1, pp.50-6.

Tombaugh TN, McDowell I, Kristjansson B and Hubley AM (1996) *Mini-Mental State Examination (MMSE) and the Modified MMSE (3MS): A psychometric comparison and normative data.* Psychological Assessment. Vol. 8, No.1, pp.48-59.

Tschanz JT, Welsh-Bohmer KA, Plassman BL, Norton MC, Wyse BW, Breitner JCS, et al. (2002) *An adaptation of the modified mini-mental state examination: analysis of demographic influences and normative data: the cache county study.* Neuropsychiatry, Neuropsychology, and Behavioral Neurology. Vol. 15, No.1, pp.28-38.

Tsolaki M, Fountoulakis K, Nakopoulou E, et al. (1997) *Alzheimer's Disease Assessment Scale: the validation of the scale in Greece in elderly demented patients and normal subjects.* Dementia and Geriatric Cognitive Disorders. Vol. 8, No.5, pp.273-80.

Wang H, Yu X, Li S, et al. (2004) *The cognitive subscale of Alzheimer's Disease Assessment Scale, Chinese version in staging of Alzheimer disease.* Alzheimer Disease and Associated Disorders. Vol. 18, No.4, pp.231-5.

Weiner MF, Vobach S, Svetlik D, et al. (1993) *Cortisol secretion and Alzheimer's disease progression: a preliminary report.* Biological Psychiatry. Vol. 34, No.3, pp.158-61.

Welsh KA, Butters N, Hughes CP, et al. (1991) *Detection of abnormal memory decline in mild cases of Alzheimer's disease using CERAD neuropsychological measures.* Archives of Neurology. Vol. 48, pp.278-282.

Weyer G, Erzigkeit H, Kanowski S, et al. (1997) *Alzheimer's Disease Assessment Scale: reliability and validity in a multicenter clinical trial.* International Psychogeriatrics. Vol. 9, No.2, pp.123-38.

Youn JC, Lee DY, Kim KW, et al. (2002) *Development of the Korean version of Alzheimer's Disease Assessment Scale (ADAS-K).* International Journal of Geriatric Psychiatry. Vol. 17, No.9, pp.797-803.

Zec RF, Landreth ES, Vicari SK, et al. (1992) *Alzheimer disease assessment scale: useful for both early detection and staging of dementia of the Alzheimer type*. Alzheimer Disease and Associated Disorders. Vol. 6, No.2, pp.89-102.

Zemlan FP (1996) *Velnacrine for the treatment of Alzheimer's disease: a double-blind, placebo-controlled trial*. The Mentane Study Group. Journal of Neural Transmission. Vol. 103, No.8-9, pp.1105-16.

## 7 Economic Evaluation in Dementia Care and the Incorporation of the Patient Perspective

### 7.1 Economic Evaluation in Dementia

World-wide health care costs are increasing as a proportion of gross domestic product (GDP) driven by the demand for health care, the use of more expensive technologies and the changing demographic profile of society (Productivity Commission, 2005a; Productivity Commission, 2005b). Table 27 illustrates this for the OECD countries: between 1980 and 2000 there was an increase across the OECD of 2.8% in the cost of health care as a proportion of GDP. Very few countries experienced a decline in the cost of health care. Within the overall health care sector, mental health conditions are one of the largest contributors to the burden of disease; it was the largest contributor to non-fatal burden of disease in Australia in 1996 (Mathers, et al. 1999). Understandably a priority for governments is to cap health care systems.

**Table 27 Health Care Costs, OECD Countries 1980-2000, Percentage of Gross Domestic Product**

Health care costs, OECD countries, 1980-2000, percentage of Gross Domestic Product					
	1980	1985	1990	1995	2000
Australia	6.8	7.2	7.5	8.0	8.8
Austria	7.5	6.5	7.0	9.7	9.4
Belgium	6.3	7.0	7.2	8.2	8.6
Canada	7.1	8.2	9.0	9.2	8.9
Czech Republic			4.7	7.0	6.7
Denmark	8.9	8.5	8.3	8.1	8.3
Finland	6.3	7.1	7.8	7.4	6.7
France	7.0	7.9	8.4	9.4	9.2
Germany	8.7	9.0	8.5	10.3	10.4
Greece	6.6		7.4	9.6	9.9
Hungary				7.4	7.1
Iceland	6.2	7.2	7.9	8.4	9.2
Ireland	8.3	7.5	6.1 b	6.7	6.3
Italy			7.7	7.1	7.9
Japan	6.5	6.7	5.9	6.8	7.6
Korea		4.1	4.4	4.2	4.8
Luxembourg	5.2	5.2	5.4	5.6	5.8
Mexico			4.8	5.6	5.6
Netherlands	7.2	7.1	7.7	8.1	7.9
New Zealand	5.9	5.1	6.9	7.2	7.7
Norway	7.0	6.6	7.7	7.9	8.5
Poland			4.9	5.6	5.7
Portugal	5.6	6.0	6.2	8.2	9.4
Slovak Republic					5.5
Spain	5.3	5.4	6.5	7.4	7.2
Sweden	9.0	8.6	8.3	8.1	8.4
Switzerland	7.4	7.8	8.3	9.7	10.4
Turkey	3.3	2.2	3.6	3.4	6.6
United Kingdom	5.6	5.9	6.0	7.0	7.3
United States	8.8	10.1	11.9	13.3	13.3
Source: OECD (2006) (OECD, 2006)					

These issues are particularly important in dementia research. Between the ages of 75 and 95+ the dementia prevalence increases from 3.5-5.0% to 38.1-57.1% of the population, depending upon the definition used, and between 60 years and 80+ years the cost of health care increases fivefold (Riedel-Heller, et al. 2001, Productivity Commission, 2005a). The annual estimated costs of dementia care in a British study based on 1994 estimates rose from £76 (£75) million for males (females) aged 65-69 years to £373 (£2440) million for those aged 85+ years (McNamee, et al. 2001). In the US, a 1998 estimate showed that Alzheimer's disease was the third most costly illness to the US economy, costing ~US \$100 billion per annum (Meek, et al. 1998).

The obvious implication is that as the population ages and the prevalence of dementia increases there will be increased demand for dementia health care within health care sector resource constraints. Indeed, as one commentator has noted, "In the span of one generation, the perception of Alzheimer's disease has evolved from an odd and unusual presenile cause of dementia to an impending public health crisis" (Geldmacher, 2002, p63).

Consequently, the dementia health care sector will be asked to justify its costs and the benefits of care relative to other health areas, above the rule of rescue (Jonsen, 1986). The rule of rescue is where an individual has need for an immediate intervention and an intervention is put into place regardless of its cost-effectiveness. For example, the family of an elderly relative with rapid onset dementia, incontinence and inappropriate public behaviors may seek accommodation in a high-security dementia ward regardless of the costs; as noted by Jonsen (1986) this response to need benefits a few at cost to many. There is, then, a fundamental conflict between the rule of rescue, economic evaluation and the efficient use of limited public health resources. A possible solution to this conflict would be for a health care system to provide basic care for all health conditions (thus meeting the rule of rescue) and for economic evaluation to be used for providing information for resource allocation within the health care sector once the need for basic care has been met. Providing the evidence to support dementia care above the rule of rescue is the role of economic evaluation (Jonsen, 1986; McKie and Richardson, 2003).

Economic evaluation can be done either with or without the patient perspective being incorporated. As in other health areas, however, there is a strong argument that, where possible, the patient perspective should be incorporated since it is patients who live with the benefits of treatment for dementia; such benefits should be demonstrated in ways that justify the costs of treatment. The economic evaluation model which captures the patient perspective is cost-utility analysis (CUA).

CUA is a particular kind of cost-effectiveness analysis (CEA), where CEA refers to evaluations reporting the cost per natural health outcome unit gained for the intervention of interest, and then comparing this with the cost-per-outcome gained from a different intervention. For example, Stewart, et al. (1998) performed a CEA of Donepezil where they modelled the costs and potential benefits from different levels of Donepezil on the progression of Alzheimer's disease over a 5-year period. The benefits were expressed as delayed cognitive losses (the states were defined as minimal, mild, moderate, severe cognitive loss and dead). The three treatment options were placebo, 5mg and 10mg of Donepezil. When outcomes for those with mild cognitive impairment at baseline were expressed as the number of years spent in each cognitive state, 10mg of Donepezil was more cost-effective than 5mg or placebo (£25,121 versus £26,702 and £28,197, respectively).

Unlike cost-effectiveness studies, cost-utility studies express the outcome in quality-adjusted life years (QALYs) gained where QALYs are calculated from either vignettes or multi-attribute utility (MAU) instruments.

### **7.1.1 A Review of Cost-Utility Analysis (CUA) Dementia Studies**

Neumann, et al. (1999a) conducted a methodological demonstration study using a Markov model to determine the progression of transitions between Alzheimer's disease stages (mild, moderate and severe dementia). To estimate the effect of Donepezil on the transitions the results from a 24-

week clinical trial involving 5mg and 10mg doses were used. Costs were estimated for each disease stage from a previous study of Alzheimer's disease costs, and utilities were assigned from another study where the caregivers of patients in the three disease stages had completed the Health Utility Index 3 (HUI3) as proxies. The results showed that the modelled cost per QALY ratios, over an assumed 18-month period, were US\$9,300/QALY for people with mild Alzheimer's disease living in the community, and US \$76,000/QALY for moderate Alzheimer's disease patients. For moderate Alzheimer's disease patients Donepezil was not cost-effective.

Using similar methods, Ikeda, et al. (2002) used a Markov model to determine the progression of transition between Alzheimer's disease stages (mild, moderate and severe) for Donepezil compared with conventional therapy. The effect of Donepezil was taken from the Japanese Phase III trial, extrapolated from the 24-weeks of the trial to 2 years for the study. Costs were based on long-term health care insurance costs. QALYs were calculated from the HUI3, where the values were taken from a survey of Alzheimer's disease patients in Japan. The cost of conventional therapy over the modelled 2 years for patients with mild Alzheimer's disease was ¥5,098,000/QALY compared with ¥4,414,000/QALY for Donepezil; for those with moderate Alzheimer's disease the results were ¥21,217,000/QALY and ¥14,806,000/QALY. It was concluded that Donepezil was dominant over conventional therapy.

More recently, but again using a similar approach, Jonsson, et al. (2005) examined the cost-effectiveness of Memantine for dementia in Sweden. A Markov simulation model was used to model transitional probabilities over a 6-month period, based on 3 MMSE classifications (mild, moderate, severe). Efficacy data was based on a previous US trial of Memantine, the EQ-5D utility data was taken from a Danish study, and were costs from Swedish health care costs. The results were modelled over a 5-year period. The calculated cost per QALY was 551,063Kroner for Memantine and 671,582Kroner for placebo; thus Memantine dominated the placebo.

In an earlier study of accommodation for those with dementia, Wimo, et al. (1995) reported a CUA of group living for people with dementia in Sweden, where group living was an intermediate stage between home and institutional care. The study was conducted in two stages. Stage 1 was an open, nonrandomised control design where there were three cohorts: home-based living (n=39), group living (n=46) and institutional care (n=23). Costs were based on the resources used. Stage 2 provided the CUA. Cognitive impairment was based on the MMSE and Global Deterioration Scale (GDS) scores. Seven levels were identified. The fore-runner to the Quality of Well-Being (QWB), the Index of Wellbeing (Kaplan, et al. 1976) was used for utility. A Markov model was constructed to describe disease progression over 8 years of expected life. Over the 8 years, the cost/QALY gained for each of the three groups was Group Living US \$52860; home: US \$71914, institution: US \$94,413 in 1987 US dollars. Group living dominated home living which dominated institutional care.

Finally, McMahon, et al. (2000) modelled the cost of functional neuroimaging of Alzheimer's disease patients. The model was based on a modelled decision-tree regarding patient workup at specialist Alzheimer's disease clinics. The excess cost was US \$479,500/QALY gained when compared with usual workup of these patients and it was concluded that neuroimaging was not cost-effective.

\*\*\*

In short, the CUA literature to date has consisted almost exclusively of modelling studies based on the time spent in a cognitive state and the transition to the next cognitive state; i.e. studies have generally used standard Markov models. Utilities have been modelled from other studies rather than being collected from study participants, and costs have been assigned, in the main, from the health service perspective rather than the societal perspective.

The challenge, then, is for future research to move away from these kinds of modelling exercises and to embrace normal CUA as practiced in other branches of medicine. As shown below, there is no necessary reason this cannot be done, provided there is the will to do it. This review provides an overview of the leading MAU-instruments, it assesses them against issues relevant to dementia, and it discusses issues around the validity of self-report in dementia studies.

## **7.2 The Axioms of Utility Measurement**

The basic axiom of cost-utility analysis is simple: life years are weighted by the value of a given health state in such a way that the values - referred to as 'utilities' - act as an exchange rate between the quantity and quality of life. In this context, 'utilities' are assumed to be preferences for a given health state. Regarding the measurement of utilities, Torrance (1986) provides the classic text.

To understand utilities, consider the following. Most people would prefer to be healthy over a given time rather than suffer Alzheimer's disease. Utility measurement refers to valuing these preferences on a life-death scale with endpoints of 1.00 and 0.00, where 0.00 is death equivalent and 1.00 is perfect (very good) HRQoL. For example, the measured utility for mild Alzheimer's disease may be 0.60. If treatment maintains this over, say, a 1-year period during which without treatment utility would decline to 0.40, then the value of the treatment is  $0.60 - 0.40 = 0.20$ . If this utility gain is maintained over time, say for 5 years, then the gain is  $0.20 \times 5 = 1.00$  quality adjusted life year (QALY). Because utilities fall on the life-death scale, they are (in theory) common across all health states and therefore can be used to compare the effect of interventions in different health fields, or different interventions within the same field. For example, the QALYs gained from Treatment A for Alzheimer's could be compared with those gained from Treatment B for depression. Where treatment costs (including costs to the patient) are known, the treatment providing the lowest cost-per-QALY gained is preferred as this ensures society gains the greatest benefit from the health care dollar.

To allow for comparison, utility measures must be generic and must allow for respondents to report they have excellent HRQoL (full health equivalent state: 1.00); additionally they must allow those who have appalling HRQoL to report this (death equivalent state: 0.00). If an instrument does not permit this full range of responses, it cannot accurately measure the HRQoL of people who fall outside its range. For example, if an instrument only allows measurement between 0.50 and 1.00, then it is incapable of reporting the effect of treatment for people who are in a desperate health state (say, close to death). Under these circumstances, any claim to generalisability for the instrument is foregone.

The instrument must be applicable to HRQoL states deemed worse than death (i.e. the respondent indicates he/she would rather die now than continue living in his/her current HRQoL state). These negative health states are needed to allow for people who commit suicide or euthanasia; they have clearly made the decision that death is preferable to living in their current health state and any possible future health states. When determining negative utility boundaries, the developers of the EQ-5D and HUI3 adopted Torrance's symmetry argument. This states that since a person can 'lose' HRQoL value from 1.00 (full health) to 0.00 (death equivalent), they must be able to 'gain' an equivalent amount from  $-1.00$  to 0.00 (Torrance, 1986). However, since negative utility values do not possess the same interval properties as positive utility scores (Hawthorne, et al. 2000c, Richardson and Hawthorne, 2000), there are difficulties. For example, improving the HRQoL of a person from  $-0.35$  to  $-0.25$  (i.e. bringing them closer to a HRQoL death-equivalent state) does not have the same meaning as improving their HRQoL state from 0.25 to 0.35; yet both these would have a utility gain of +0.10. This is implausible. It seems likely, then, that negative values should have lower boundaries close to 0.00 (death equivalent) (Richardson and Hawthorne, 2000).

Implicit in axioms and mathematical modelling of utilities is that utility measurement must be at the interval level, where interval level refers to measurement scales that have equal-intervals between the measurement points. There are two forms of interval measurement that MAU-instruments must have if they are to do their job correctly. One is known as the “weak” interval property and the other the “strong” interval property (Richardson, 1994). The weak interval property is where a gain of 0.10 means the same thing across the range of instrument scores. For a person who has severe Alzheimer’s, their utility score might be 0.25; as a result of treatment this is maintained at 0.25 whereas without treatment this might decline to 0.15; i.e. the value of the treatment is  $0.25 - 0.15 = 0.10$ . Similarly, the value of the treatment is also 0.10 for a person with mild cognitive impairment with an initial utility of 0.70, and who maintains this after treatment whereas without treatment this declines to 0.60; thus  $0.70 - 0.60 = 0.10$ . The strong interval property is where there is a direct relationship between gains in utility and gains in life-length. Since QALY calculation represents the time spent in a given state multiplied by the quality of that state, this implies that a 0.20 utility gain multiplied by 5 years in the health state of interest equals 1.00 QALY (from  $0.2 \times 5$ ). But a gain of 1 QALY could also be the product of a 0.40 utility gain over 2.5 years (or any other combination).

### 7.2.1 Measuring Utilities Using MAU-Instruments

There are two steps to measuring utilities using MAU-instruments. First, the health state of interest is described. Second, the value or utility of the health state is assigned.

When a person completes a MAU-instrument, his/her numerical responses provide a description of his/her health. For example, consider two people completing an imaginary instrument with four dimensions each of which has four levels. This instrument’s ‘descriptive system’ would be: physical, mental, social and cognitive health dimensions, and the response levels are: 1 = normal, 2 = some impairment, 3 = major impairment, 4 = gross impairment. Person A, who is in the best of health, selects the best response to each item (i.e. ‘1’: normal,). Her health state would be described as ‘1,1,1,1’. Person B who suffered major cognitive impairment (level 3: major impairment on the cognitive dimension), some impairment in mental health (level 1), some social impairment (level 2), and normal cognitive function (level 1). Her health state would be ‘1,2,2,3’.

Valuing these health states is called ‘scaling’, and is usually carried out using general population samples (Sackett and Torrance, 1978). Five procedures have been used: time trade-off (TTO), standard gamble (SG), visual analogue rating scale (VAS), magnitude estimation (ME) and person trade-off (PTO). Brief descriptions are given.

- *Time trade-off (TTO)*. A person with a severe health state can have a treatment which will restore her to full health; but a side effect is she will live a shorter life. She is asked to choose how many years of her life she would be willing to ‘give up’ in order to be in full health. If, in her untreated condition, her life expectancy was 10 years and after the treatment this was 5 years she may reject the treatment. If after the treatment it was 9 years, she may accept it; if her life expectancy was 6 years, she may not. Her choices would continue back-and-forth like this until she indicated that she was indifferent to whether she had the treatment or not. If the point of indifference was that 8 years of full health was the equivalent of 10 years in the severe health state, then the quality of life value for her current health state is  $8/10$  or 0.80.
- *Standard gamble (SG)*. A person with a health condition is presented with a treatment option that has two possible outcomes: either full health for the remainder of his life, or death. He is free to choose either the treatment or to remain with the condition for life. If the probability of full health is 1.00 (i.e. he will be cured and there is no chance of death), then obviously he will choose to have the treatment. If the probability of full health is 0.90 and death 0.10, he may still choose the treatment. However there would be a point, for example at 0.80 for full health and 0.20 for death, where he is not clear as to whether he would want the treatment or would

choose to remain in his current health state. This point of indifference is the 'value' of his health state.

- *Visual analogue scale (VAS)*. The respondent is asked to consider a health state and then to rate this on a scale, where the endpoints are typically 0.00 (death equivalent) and 1.00 (full health equivalent). Unlike the TTO or SG, with the VAS there is no uncertainty: the respondent is not asked to 'trade' anything. Consequently many consider that VAS scores do not represent utilities because they provide a simple ranking of health states. Where VAS scores are used, a transformation is generally required, based on TTO or SG (Brazier and Deverill, 1999, Bennett and Torrance, 1996, Robinson, et al. 1997).
- *Magnitude estimation (ME)*. The respondent is asked to consider the distance of the health state of interest (e.g. severe dementia) from 1.00 (full health). Once several of these rating exercises have been carried out, the respondent is then asked to rank these in order (Gudex, et al. 1993). Because there is no uncertainty, it is uncertain if ME represents utility.
- *Person trade off (PTO)*. The respondent is asked to estimate the number of people that would have to be treated to make an intervention worthwhile. For example, a respondent might be asked to choose between extending the life of 10,000 people who were in full health by 1 year against a treatment which extended the life of N people with dementia, also for 1 year. The number of people with dementia would be varied until the respondent indicated they were indifferent between the two choices (Gudex, et al. 1993).

When these techniques are used to obtain the utility weights used in an MAU-instrument, in theory each health state described by the descriptive system can be scaled (as was done with the original Rosser Index [Rosser, 1993]), but this is impractical because MAU-instruments typically generate thousands of different health states. Instead, a limited number of health states are scaled and the values for other health states are then inferred using econometric or decision analytic techniques, typically either an additive or multiplicative model (Hawthorne and Richardson, 2001). During scoring, the health state descriptors (1,2,3, etc.) are replaced with the appropriate values. For example, if the value of suffering mild pain based on TTO is '0.70' and the response levels on an item measuring pain were '1' (no pain), '2' (mild pain), and '3' (severe pain), then a person who selected '2' would have this level replaced with the value '0.70' during scoring of the instrument.

Once item-level values have been assigned, these are combined into an index on a life-death scale. Three procedures have been used.

- *Additive models*. The substituted importance values are summed and the resulting score represents the utility index. The limitation is that for full health equivalent HRQoL states each instrument item or dimension must contribute a fixed amount. Under this model, a respondent can obtain a very poor utility score only if they report poor scores on all items or dimensions. Consider an instrument measuring two dimensions: physical and mental health. In an additive model, each may contribute 0.50 towards the utility score. In this model, appalling mental health (leading to suicide) could never, by itself, lead to a utility value lower than 0.50 because  $0.50$  (a person in good physical health) +  $0.00$  (mental health) =  $0.50$ . Thus additive models cannot explain people who commit suicide if their physical health is good or euthanasia if their mental health is good.
- *Econometric models*. The items are treated as explanatory variables to derive a regression equation predicting utilities. This method, however, suffers the same limitation as the additive model.
- *Multiplicative models*. These involve multiplying items or dimension scores together. This overcomes the limitation of the additive model as it allows any dimension to carry a person to a death equivalent value. Consider the case above. Here the person's value for mental health would be 0.00, and  $0.50$  (physical health) x  $0.00$  (mental health) =  $0.00$ .

Given these assumptions, preference independence is required to avoid double-counting, which is where the same underlying health condition contributes more than once to the MAU-instrument utility index. For example, if a person is cognitively impaired this should be counted in their utility

score once, although the effect of this health state may be measured in several different aspects of their life; i.e. on several different scales. Where these effects are measured using unidimensional scales that are orthogonal to each other there is no difficulty. Where the scales, however, are correlated the effect of cognitive impairment will be counted several times over thereby biasing the utility measurement. It is for this reason that MAU-instruments are required to possess structural independence (i.e. where the scales are unidimensional and orthogonal) (von Winterfeldt and Edwards, 1986). For example, if cognitive impairment is counted on social, physical and psychological dimensions as well as its effects being directly measured, then there is loss of preference independence as the scores on the social dimension may be a function of physical scores.

### 7.3 Utility Instrument Review

MAU-instruments (in alphabetical order) are the Assessment of QoL (AQoL) (Hawthorne, et al. 1999, Hawthorne, et al. 2000b, Hawthorne, et al. 2000d), EQ-5D (formerly the EuroQol) (EuroQol Group, 1990; Kind, 1996), Health Utility Index 3 (HUI3) (Feeny, et al. 1996a; Feeny, et al. 1996b, Torrance, et al. 1995), 15D (Sintonen and Pekurinen, 1993; Sintonen, 1995, 1994), Quality of Well-Being (QWB) (Kaplan, et al. 1993; Kaplan, et al. 1996b), Rosser Index (Rosser, 1993), and the SF6D (Brazier, et al. 2002; Brazier, et al. 1998).

#### EQ-5D (formerly the EuroQol)

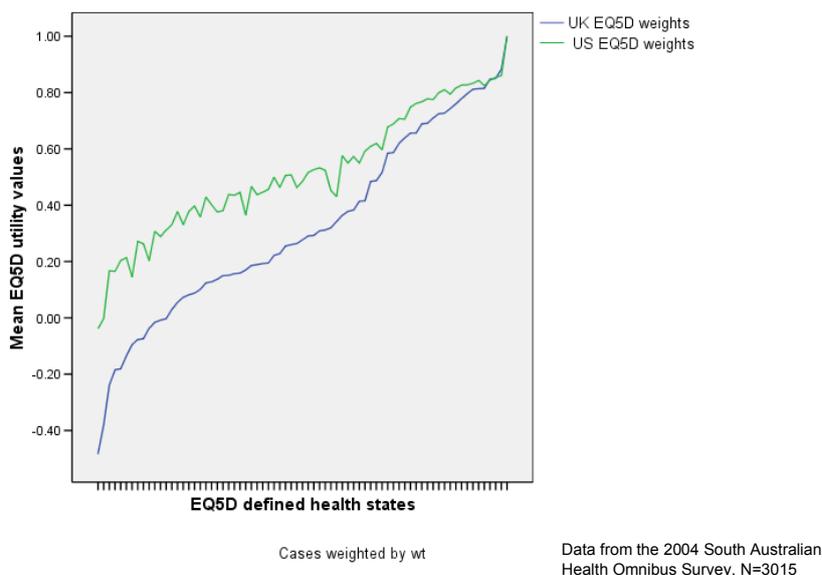
The EQ-5D (formerly the EuroQoL), developed by a team from 7 European countries (Rabin and de Charro, 2001; EuroQol Group, 1990), was based on the QWB (Kaplan and Anderson, 1988), the Sickness Impact Profile (Bergner, et al. 1981), the Nottingham Health Profile (Hunt, et al. 1981), the Rosser Index (Rosser, 1993), and group members' opinions. Designed for use in cross-cultural comparisons it has 5 items, each with 3 response levels, measuring Mobility, Self-care, Usual Activities, Pain/Discomfort and Anxiety/Depression. It takes 1-2 minutes to self-complete (Nord, 1997). The original utility weights were from a British population random sample ( $n = 3395$  respondents, response rate 56%) based on the TTO for 42 marker health states using a 10 year timeframe (Dolan, 1997). The intermediate health state values (i.e. those for which direct TTO weights were not obtained) were regression modelled (MVH Group, 1995; Dolan, 1997; Dolan, et al. 1996). The index is computed using an econometric regression model. The upper boundary is 1.00, and the lower boundary is  $-0.59$ : it permits health state values worse than death.

Recently, US weights have been published. The mean difference in health state values between the British and US weights was 10% of a life-death scale. Consequently it was recommended that when used in the US, US-derived weights should be used (Shaw, et al. 2005, Havranek and Steiner, 2005, Johnson, et al. 2005, Fryback, 2005).

Figure 2 shows the differences in utility scores that are obtained from the two different weights sets. The data, from the 2004 South Australian Health Omnibus Survey, suggest that for good health states there is little difference between the two algorithms, but that as health worsens the two algorithms provide very different estimates of utility. For example, for health state 21112 (a good health state) the UK utility is 0.78 and the US utility 0.83 (a difference of 0.05), whereas for health state 12223 (a moderate health state) the utilities are 0.15 and 0.44 respectively (a difference of 0.29), and for 22233 (a poor health state) the utilities are  $-0.18$  and  $+0.20$  respectively (a difference of 0.38).

Importantly neither set of weights (British or US) have been validated for use in Australia. Although to date Australian researchers have used the British weights, there is no particular reason this should continue to be the case. It is, however, obvious from Figure 2 that selection of weights for the EQ-5D may have a direct impact on the results from a study.

**Figure 2 Differences in Obtained Utility on the EQ-5D, by UK and US Scoring Algorithms**



Although the EQ-5D is in the public domain for public health research, the EQ-5D management group ask that researchers register their use of it. There are no costs for its use, unless it is used by commercial organisations. The EQ-5D has been translated in many languages. Further information on the EQ-5D can be obtained from: <http://www.eur.nl/bmg/imta/eq-net/EQ5d.htm>.

### Assessment of QoL (AQoL)

The Australian AQoL used the WHO's definition of health, and items describe 'handicap' as distinct from impairment and disability (Hawthorne and Richardson, 1995). The descriptive system has 15 items and 12 are used in computing the index (Hawthorne, et al. 2001b). Each item has 4 levels. There are five dimensions: Illness (not used in utility computation), Independent Living, Social Relationships, Physical Senses and Psychological Well-being (Hawthorne, et al. 1999). Designed for self-completion, Nord (1997) reported the AQoL took 5-10 minutes. A stratified sample ( $n = 350$  respondents; response rate 72%) representative of the Australian adult population completed TTOs based on a 10 year timeframe to provide the utility weights (Hawthorne, et al. 2000d). A multiplicative model is used to compute the utility index (Hawthorne, et al. 2000b). The upper boundary is 1.00, and the lower boundary is  $-0.04$ : it permits health state values worse than death. Permission to use the AQoL must be obtained, but there is no cost for its use. Further information can be obtained at: [http://www.acpmh.unimelb.edu.au/whoqol\\_aqol.html](http://www.acpmh.unimelb.edu.au/whoqol_aqol.html).

### Health Utilities Index, Mark 3 (HUI3)

The Canadian Health Utilities Index (HUI3), for general population use, is based on the HUI2 which was designed for survivors of childhood cancer. To render it generic and overcome reported difficulties, it was revised into the HUI3 (Feeny, et al. 1996a). The HUI1 has been superseded. The HUI3 measures 'within the skin' functional capacity (Feeny, et al. 1996b), a perspective adopted to enhance its use in clinical studies (Furlong, et al. 2001). Social aspects of HRQoL are not measured. Items have 4–6 response levels. Twelve of the 15 items form 8 attributes (Vision, Hearing, Speech, Ambulation, Dexterity, Emotion, Cognition and Pain). Designed for self-completion, Nord (1997) reported it took 2 minutes to complete, although 5–10 minutes is more likely given it has 15 items. The utility weights were elicited using the VAS, and scores then transformed based on four 'corner' health states valued with the SG where a 60 year timeframe was used. These results were based on stratified sampling ( $n = 256$ ; response rate 22%) of the Hamilton, Ontario, population (Furlong, et al. 1998). A multiplicative function combines the attributes into the utility score (Furlong, et al. 1998; Torrance, et al. 1995). The upper boundary is

1.00, and the lower boundary is  $-0.36$ , permitting health states worse than death. Users must be registered and the instrument is available at a cost of CAN \$4,000 per trial (at the time of writing). Copies of the HUI3 and application forms can be found at: <http://www.healthutilities.com/hui3.htm/>.

### 15D

The Finnish 15D was defined by Finnish health concerns, the WHO definition of health and medical and patient feedback (Sintonen, 2001, Sintonen and Pekurinen, 1993). It is concerned with impairment and disability of 'within the skin' functions. There are 15 items, each with 5 levels, measuring Mobility, Vision, Hearing, Breathing, Sleeping, Eating, Speech, Elimination, Usual Activities, Mental Function, Discomfort and Symptoms, Depression, Distress, Vitality and Sexual Function (Sintonen and Pekurinen, 1993). Nord (1997) reported it took 5–10 minutes for self-completion. The weights came from five random samples of the Finnish population ( $n = 1290$  respondents; response rate 51%) using VAS questions; responses were combined using a simple additive model (Sintonen, 1994, 1995). The upper boundary is 1.00, and the lower boundary is  $+0.11$ : death-equivalent and worse than death health states are not allowed. Permission must be obtained to use the instrument; however there are no costs for its use. The 15D has been translated into a number of European languages. Although there is no website devoted to the 15D, details can be obtained from <http://195.101.204.50:443/public/15D.html>.

### Quality of Well-Being (QWB or IWB)

The American QWB was designed to bridge the gap between clinical measurement, functional status and health planning policy (McDowell and Newell, 1987) and was an adaptation of US health surveys (Cadet, 1994). The early version of the QWB was the Index of Wellbeing (Kaplan, et al. 1976). The QWB has three dimensions (Mobility, Physical Activity, and Social Activity) with 3–5 levels each. There are an additional 27 illness symptoms. Combined, these provide an index of 'Well-life expectancy' of which there are 43 functioning levels (Kaplan, et al. 1976; McDowell and Newell, 1987; Kaplan, et al. 1993). This would seem to support Anderson, et al's (1989) description of it as measuring dysfunction. Mental and social health is not measured. The QWB was designed for interview administration (15–35 minutes), although a shorter self-completed version has been developed, the QWB-SA, which takes about 14 minutes (Andresen, et al. 1998). It comprises five sections covering symptoms, self-care, mobility, physical functioning and usual activities. There are 71 items altogether. Interviewer training is required for the full QWB (Bombardier and Raboud, 1991) and recommended for the QWB-SA (Kaplan, et al. 1998).

The preference weights for the QWB were elicited using VAS scores which were obtained from a sample of the San Diego population. A linear transformation was then used to place these on a 0.00–1.00 scale (Kaplan, et al. 1976; Kaplan, et al. 1996a). An additive model is used to compute the index. Extensive efforts to validate that VAS provides interval properties led to the release of a revised version (Coons and Kaplan, 1993; Kaplan, et al. 1976; Kaplan, et al. 1993). The weights for the QWB-SA were taken from the QWB, for those conditions where actual QWB weights were not available the mean QWB weight was applied to the QWB-SA health condition (Andresen, et al. 1998). The upper boundary is 1.00, and the lower boundary is 0.00 (death equivalent) and health states worse than death are not permitted. Permission must be obtained to use the QWB and there are no costs for its use. Further information on the QWB can be obtained at: <http://medicine.ucsd.edu/fpm/hoap/instruments.html>.

### Rosser Index

The British Rosser Index was designed for use in hospital settings. The original version had two dimensions measuring disability and distress, and measured 29 health states. Values were elicited using magnitude estimation from a convenience sample of 70 respondents (Rosser, 1993). A revised version was released in the early 1990s based on SG procedures and included an additional dimension of discomfort (Rosser, 1993). Administration requires a trained interviewer. The upper boundary is 1.00 and the lower boundary  $-1.49$ ; i.e. health states worse than death are permitted. The Rosser Index has given rise to two variants: the Health Measurement Questionnaire (HMQ) (Kind and Gudex, 1994) and the Utility-based Quality of Life-Heart

Questionnaire (UBQ-H) (Martin, et al. 1996). Permission must be obtained for using the instrument; however there are no costs for its use. No website was identified for the Rosser Index.

### **SF6D**

Two different algorithms have been published by Brazier, et al. for deriving preference-based values from the SF-36 (Brazier, et al. 1998, 2002). The more recent algorithm supersedes the earlier version, so only the more recent algorithm is described here. The advantage of the SF6D is that wherever SF-36 raw scores are available, the SF6D preference measure can be used.

The SF6D (Brazier, et al. 2002) uses 10 items from the SF-36: three from the physical functioning scale, one from physical role limitation, one from emotional role limitation, one from social functioning, two bodily pain items, two mental health items and one vitality item. These form 6 dimensions: Physical Functioning (PF: 6 levels), Role Limitation (RL: 4 levels), Social Functioning (SF: 5 levels), Pain (PA: 6 levels), Mental Health (MH: 5 levels) and Vitality (VI: 5 levels). Utility weights were modelled using SG values for 249 states, where each respondent valued 6 health states. Values were obtained from a random sample of the British population ( $n = 611$ ; response rate = 45%). An additive econometric model is used to compute the utility index. The endpoints for the SF6D are 1.00, and 0.30 for the worst possible health state. No website for the SF6D was identified.

Additionally, Brazier, et al. have derived a SF6D score from the SF-12 by using 7-items. Six health states were valued using the SG and a regression model used to impute intermediate values and the SF6D score. Because of difficulties with the model, Brazier, et al. note that this algorithm is unlikely to replace other utility measures (Brazier and Roberts, 2004). This model is not discussed further in this report.

## **7.4 Comparison of Instruments**

Hawthorne and Richardson (2001) outlined the axioms of utility measurement which MAU-instruments should conform to in order to possess basic validity. These axioms can be used as a checklist in instrument selection. They are:

- The use of a preference measurement to weight instrument items.
- Instruments must measure the dimensions of HRQoL deemed to be important. These are usually defined as physical, mental, social and somatic sensations (e.g. pain).
- There must be coverage of the full spectrum of HRQoL values, from full health states to values representing states worse than death.
- The combination rule for the utility index must prevent double-counting.
- There must be evidence of both the weak and strong interval measurement.
- Instruments must be sensitive to the health states of interest. This requirement is covered in the next section. For general sensitivity comparisons between the instruments the reader is referred to validation papers by Barton, et al. (2004, 2005), Conner-Spady and Suarez-Almazor (2003), Hawthorne and Richardson (2001), Hawthorne, et al. (2001b); Hawthorne, et al. (2000a), Kopec and Willison (2003), Marra, et al. (2005), or Pickard, et al. (2005).

For use in dementia studies, two additional issues are:

- There must be evidence of valid and reliable measurement; and
- The response perspective (self-completion, proxy completion) must allow for valid and reliable data collection.

### **Use of a preference measurement technique to weight instrument items**

Instruments using the SG or TTO may be regarded as possessing preference weights since both involve decisions under uncertainty. In the SG, the life outcome is uncertain (the probability of full health versus death). In the TTO, life-length is uncertain (how many life-years a person is willing to sacrifice).

There are doubts over whether ME delivers preferences because the procedure requires the respondent to estimate the divergence of a given health state from the 'full' health state (which is assigned a value of 1.00). Once several given health states have been so assigned, the respondent is then asked to rank these in order (Gudex, et al. 1993).

As reported above, there is doubt whether the VAS delivers preference measurement. Consequently it has been argued that the VAS has no place in economic theory (Brazier, et al. 1999) and that untransformed VAS scores should not be used (Robinson, et al. 2001, Torrance, et al. 2001). It is recommended that VAS data should always be transformed based on TTO or SG (Brazier, et al. 1999; Bennett and Torrance, 1996; Robinson, et al. 1997); the transformation function that has been used was developed by Torrance, et al. (1982). The preference measurement of instruments weighted with VAS scores therefore rests upon the validity of this transformation. For the EQ-5D, Dolan, et al. (1995) reported that the explanatory power of the transformations used was  $R^2 = 0.46$ , which was considered to be very good. However Sintonen (1995) reported that when applied to the 15D VAS data it assigned 12–25% of the adult population to values worse than death, a result he stated was 'implausible'. Bleichrodt and Johannesson (1997) noted that individual transformations were unstable; Robinson, et al. (2001) reported difficulties with the transformations; as did Torrance, et al. (2001).

Instruments weighted with a preference measure are the EQ-5D and SF6D Version 2 (both used the SG) and the AQoL (the TTO). The Rosser Index relies upon ME. The HUI3 relies upon transformed VAS scores; the extent to which this can claim preference weighting is dependent upon the validity of the transformations based on key SG-weights for selected health states. Nord (1993) has questioned the validity of the linear transformations for the QWB, arguing that one of the primary reasons its use in Oregon was so heavily criticised was that it lacked cardinal values. Given that the 15D uses untransformed VAS ratings there are doubts that it meets this requirement, although Martin (1996) averred that this gave the opportunity to quickly establish new weights for different populations - a procedure which Sintonen argued should be followed for each population from which study participants were drawn (Sintonen and Pekurinen, 1993).

### **Instruments must measure the dimensions of HRQoL deemed to be important**

Important areas of HRQoL are usually defined as physical, mental, social and somatic sensations (e.g. pain). Unless instruments measure all these they cannot claim to be 'generic'. It should be remembered that the measurement of utilities was explicitly developed to enable cross-condition, health state and health care comparisons. By definition MAU-instruments are supposed to be generic.

Generally there are no published formal tests of content validity (Hawthorne and Richardson, 2001). Where this is mentioned, instrument developers have reported 'face' validation, i.e. that instrument content as judged by the instrument developers 'looks about right'. For example, it has been argued the very restricted Rosser Index descriptive system makes it insensitive and provides a narrow band of responses (Hollingworth, et al. 1995; Nord, et al. 1993; Mulkay, et al. 1987; Elvik, 1995). In a study of the EQ-5D descriptive system it was reported that it only covers 39% of the concepts regarded by the public as salient to health (Buckingham, 1995). Feeny, et al. (1996a) reported that the HUI3 was valid because all levels of scores had been assigned at least once in population surveys. These various assertions do not engender confidence that the universe of HRQoL is actually measured by any of the instruments, a point which has been noted in the literature.

In three recent review articles Hawthorne, et al. (Hawthorne and Richardson, 2001; Hawthorne, et al. 2000b) mapped the content of MAU-instruments against the dimensions of 14 HRQoL instruments published between 1971 and 1993. Table 28 summarizes their work. This shows that even in the better instruments coverage of the universe is limited. Some instruments offer very narrow measurement (for example, the Rosser Index and EQ-5D), others have in-depth or duplicated measurement in particular areas (for example, the QWB, 15D and HUI3), and some offer very broad but sketchy coverage (for example, the AQoL and SF6D). Duplicated measurement may bias the obtained utility values. Two examples illustrate the problems. Despite its broad coverage, the QWB primarily measures pain and physical disability (Kaplan, et al. 1993) yet does not include either social or mental health (Anderson, et al. 1989), and analysis of the HUI3 showed it was a measure of physical impairment which did not adequately measure physical, social or mental dimensions (Richardson and Zumbo, 2000).

**Table 28 Content of Descriptive Systems of MAU-Instruments**

Content of descriptive systems of MAU-instruments (a)							
HRQoL dimensions (b)	EQ-5D	AQoL	HUI3	15D	QWB (c)	Rosser Kind (d)	SF6D
<i>Relative to the body</i>							
Anxiety/depression/distress	*	*		**	**	*	**
Bodily care	*	*					*
Cognitive ability			*	*	*		
General health							
Memory			*				
Mobility	*	*	*	*	*		**
Pain	*	*	*	*	*****		*
Physical ability/vitality/disability			*	*	*****	*	*
Rest and fatigue		*		*	**		**
Sensory functions		**	****	*****			
<i>Social expression</i>							
Activities of daily living	*	*		*			*
Communication		*	**	*	*		
Emotional fulfillment			*				
Environment					*		
Family role		*					
Intimacy/Isolation		*					
Medical aids use					**		
Medical treatment							
Sexual relationships				*	*		
Social function		*			*		*
Work function							*

**Note:**

a = Table shows only those items used in calculation of utility scores. Each asterisk represents an item. Based on item content examination.

b = Dimensions of HRQoL defined by a review of 14 HRQoL instruments, 1971–1993.

c = Excludes intoxication.

d = Areas subsumed within the two items: mobility, employment, housework.

Source: Adapted from Hawthorne, et al. (Hawthorne and Richardson, 2001).

**There must be coverage of the full spectrum of HRQoL values**

This refers to instruments providing utility values from full health states to values representing states worse than death. There are two issues. First, instruments must have combination rules

permitting very poor HRQoL, irrespective of how this is caused. Second, the range of utility scores must cover the full spectrum.

Regarding combination rules, multiplicative models are to be preferred for the reasons outlined above. Instruments with multiplicative models are the HUI3 and AQoL. The EQ-5D and SF6D rely upon regression models which are essentially additive models. The 15D is an additive instrument.

The Rosser Index, EQ-5D and HUI3 allow large negative values. Given the difficulty with the symmetry argument, these values are problematic. Hawthorne and Richardson (2001) calculated that the effect of restricting the lower boundary for the HUI3 and EQ-5D to 0.00, in population studies, would raise mean utility values by 9% and 14% respectively. This suggests the net effect of the symmetry argument is to overstate the value of interventions where people are in very poor health states. This problem does not apply to the QWB and AQoL which have lower boundaries at or near to 0.00.

The lower endpoints for the 15D (+0.11) and SF6D (+0.30) raise other questions because of the restricted utility range. Hawthorne and Richardson (2001) reported these boundaries resulted in very different QALY estimates: a 1 QALY gain from the AQoL, EQ-5D or HUI3, where a person was returned from the lowest quartile to full health for 1 year, implied a 0.50 and 0.37 QALY gain on the 15D and SF6D respectively. These contradictory results suggest that at least one of these of instrument groups is wrong.

For allowing the full range of scores, the QWB or AQoL instruments would be preferred, as would the 15D.

### **The utility combination rule must prevent double-counting**

During construction of the Rosser Index, care was taken to ensure orthogonality between the dimensions (Rosser, 1993). Brazier, et al. (1999) reported that for the QWB there is multicollinearity between the scales and symptoms. In papers describing the EQ-5D there is no mention of this issue (EuroQol Group, 1990; Kind, 1996). Based on clinicians' opinions, structural independence was claimed for the HUI3 (Furlong, et al. 2001); the factor analysis of the HUI3 published by Richardson and Zumbo (2000), which revealed a lack of independence between the attributes, challenges this claim. Sintonen claimed independence for the 15D, although no evidence was provided (Sintonen, 1995).

For the SF6D, Brazier, et al. (2002) reported that since an econometric model was used preference independence, structural independence and double-counting were unimportant. The form of the SF6D for the prediction of SG scores is

$$y_{ij} = g(\beta'x_{ij} + \theta'r_{ij} + \delta'z_j) + \varepsilon_{ij}$$

which is an additive model. Brazier's argument seems extraordinary given that orthogonality to prevent double-counting caused by multicollinearity has been axiomatic of both psychometric and decision-making theory for over 50 years (von Winterfeldt and Edwards, 1986; Cattell, 1952).

For the AQoL, during construction exploratory factor analysis was used to ensure orthogonality (Hawthorne, et al. 1999); the structure has since been confirmed by structural equation modelling (Hawthorne, et al. 2001a).

### **There must be evidence of both weak and strong interval measurement**

For meeting these criteria, all MAU-instruments rely on the presumed interval properties of the TTO, SG, or VAS. No instrument construction or validation paper has reported any formal testing of these properties and it has not been convincingly demonstrated that these properties are embedded with the TTO, SG, magnitude estimation or VAS (Brazier, et al. 1999; Rosser, 1993).

### The weak interval property

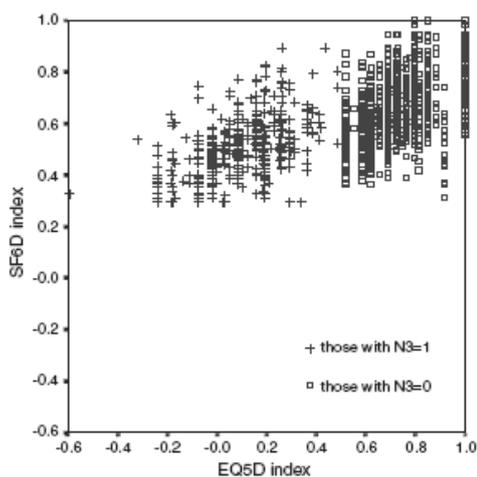
VAS responses may be functions of adaptation, context, endpoints or anchorpoints, end-aversion and rating effects. These imply VAS may produce ordinal rather than interval data (Cook, et al. 2001; Robinson, et al. 2001; Richardson, 1994; Torrance, et al. 2001). For the TTO and SG even less is known as these issues do not appear to have ever been properly investigated. Although Cook, et al. (2001) challenged the claim of interval data for all three techniques, there were methodological difficulties with the paper (Hawthorne, et al. 2003b). Subject to these caveats, Hawthorne and Richardson (2001) asserted it was likely the SG and TTO possessed interval properties given they allowed incremental probabilities (SG) or time fractions (TTO).

### The strong interval property

This means that any given incremental value in HRQoL utility was directly equivalent to the same incremental value in life-length or life-probability. This is a fundamental requirement for the correct calculation of QALYs. There is no evidence available for any of the MAU-instruments that they meet this requirement.

Finally, although the EQ-5D may theoretically meet the weak interval properties, there is evidence that it fails this requirement at the empirical level due to the 'N3 term' in the British weights. In the EQ-5D scoring process using the British weights, any person who endorses a level-3 response (the worst possible level) automatically incurs a coefficient loss of -0.269 utilities. The effect of the 'N3 term' on EQ-5D scores is shown in Figure 3 (Brazier, et al. 2004). As shown there is a large gap in scores in the region of 0.4, implying that there are areas of the utility range where scores are virtually impossible to attain. The US weights employ a 'D1 term' instead of the 'N3 term'. Since its value is -0.140 (Shaw, et al. 2005), it is likely the US weighted model may smooth out some of the lumpiness in the UK weighted scoring algorithm data distribution. Where the US scoring algorithm is used, however, it may provide different utility estimates of health states or of treatment effects (see Figure 3). Neither algorithm has been validated for Australian use.

**Figure 3 Data Distribution Issues for the EQ-5D**



Source: Brazier et al (2004)

### Valid and reliable measurement

The validity and reliability of various MAU-instruments has been assessed through either tests of concurrent validity where monotonic relationships are sought, or test-retest. Additionally, there are issues around the stability of the utility values used in the different instruments due to sample bias.

Monotonicity refers to a relationship in which the instruments of interest group or mean scores progressively increase in line with a criterion measure. For example, if a sample of people suffers symptoms of cognitive impairment from “mild” to “severe”, then an instrument measuring this underlying health condition should report manifest scores that systematically increase with the level of actual impairment. This does not imply, of course, that there will always be a 1:1 relationship between the two measures, for there will be individual variation.

Hawthorne, et al. (2000b) examined monotonicity for the EQ-5D, 15D, HUI3, AqoL and SF6D (Version 1) against health status as defined by their sample strata of community random sample, outpatients and inpatients; they also examined the same instruments by combined utility quartile (Hawthorne, et al. 2000a) and by instrument predictive power (Hawthorne and Richardson, 2001). In general their findings support monotonicity for all the instruments, although they did observe that the instruments formed two groups: those which correctly classified >50% of cases (AqoL, 15D and SF6D) and those which predicted <50% (EQ-5D and HUI3).

Data on the Rosser Index are mixed. Although Rosser Index scores have been shown to match empirical and population general health data quite well when predicting the healthy/unhealthy dichotomies (Kind and Gudex, 1994; Nord, et al. 1993), in a replication study it was shown that there are several health states where monotonicity is violated leading to difficulties with assigning logical QALY values (Gudex, et al. 1993).

For the QWB there is mixed evidence regarding monotonicity. Kaplan, et al. (1978) reported very high correlations with a number of chronic conditions, where the average was  $r = 0.96$ . Based on the revised version, similar correlations with chronic conditions have been reported (Coons and Kaplan, 1993; Kaplan, 1993). For example, Kaplan, et al. (1995) reported a monotonic relationship between QWB scores and HIV-status; similarly monotonicity has been reported for functional status of children suffering cancer (Bradlyn, et al. 1993). Against this the QWB has been criticised for producing QALY values that are non-monotonic. Thus a person wearing glasses is worse off than someone confined to a wheelchair, or curing five people with pimples would equate with saving one life (O'Connor, 1993; Nord, 1993). In a study of heart disease, non-monotonicity was reported for half the QWB scales (Visser, et al. 1994).

The Hawthorne, et al. results for the EQ-5D (see above) were particularly interesting as they indicated that the EQ-5D assigned too many cases to a utility value of 1.00, a finding consistent with earlier work by Brazier, et al. (1993). Both research groups suggested this may have been due to the insensitivity of the EQ-5D at the healthy end of the range and the consequent limited capacity to discriminate between those with full health and some health problems. At the other end of the range (very poor health states) Nord, et al. (1993), in a study comparing Norwegian and Australian populations, reported that the EQ-5D assigned excessively low values for some health states; a finding consistent with that of Hawthorne, et al. (2000b) who found that the EQ-5D assigned 4% of a population sample to health states worse than death. In a comparison with the SF-36, Brazier, et al. (1993) pointed out that the EQ-5D correlated poorly with physiological symptoms, and Andersen, et al. (1995) reported that the EQ-5D assigned non-monotonic values for people with fractures: a person with a fractured arm was assigned worse utility than someone with a fractured vertebra.

Sintonen (1995) tested the 15D for monotonicity in five population-based samples, reporting that up to 2.5% of respondents valued health states inconsistently, rising to 20% who valued ‘death’ higher than being ‘unconscious’.

For the AqoL, several papers have suggested monotonicity, including Hawthorne, et al’s work (Hawthorne and Richardson, 2001; Hawthorne, et al. 2000b), and in Alzheimer’s disease (Wlodarczyk, et al. 2004), cochlear implants (Hogan, et al. 2001), depression (Hawthorne, et al.

2003a), psychosis (Herrman, et al. 2002), suicidal ideation (Goldney, et al. 2001) and stroke (Sturm, et al. 2002).

Test-retest reliability estimates have been reported for the QWB, 15D, EQ-5D, HUI3 and AQoL. For the QWB, Kaplan, et al. (1978) reported test-retest reliability at  $r = 0.93$ – $0.98$ . In a study of chronic obstructive pulmonary disease, at 14-day separation, Stavem (1999) reported that the EQ-5D and 15D test-retest reliability using Spearman correlations were  $r = 0.73$  and  $r = 0.90$  respectively. This result for the 15D is more encouraging than that reported by Sintonen (1994), who did not give a statistical estimate but stated that the agreement was not very good. In a study of stroke patients, Dorman, et al. (1998) reported test-retest reliability estimates for the EQ-5D of ICC = 0.83; and in a Dutch population study of the EQ-5D where test-retest was carried out at 10-month interval the correlation was  $r = 0.90$  (van Agt, et al. 1994). Studies of the HUI3 (Boyle, et al. 1995; Feeny, et al. 1996b), based on a community random sample with telephone follow-up, reported test-retest reliability where  $r = 0.77$ . For the AQoL, Hawthorne (2003), using random population sampling and mail/telephone comparisons reported the test-retest ICC = 0.83. An earlier study reported test-retest reliability for the AQoL descriptive system where  $\alpha = 0.80$  (Hawthorne, et al. 1996).

Finally, and importantly, there are issues concerning the stability of the utility weights used in the various instruments. This concern stems from the fact that utility weights for most of the instruments — with the notable exception of the EQ6D where the British sample size was 3395 (for the US weights the sample size was 3773 (Shaw, et al. 2005) — were obtained from either small (e.g. 70 cases for the Rosser Index) or convenience samples (e.g. the 1290 respondents for the 15D). In most cases, this was because of the cost of data collection: face-to-face interviews where SG or TTO questions are asked are costly. Because the SG or TTO is extremely tedious, all the instrument designers eroded their sample sizes further by breaking their health states up into sub-interview routines and then administering each sub-interview to a strata within the sample. This is commonly referred to as a 'sort' procedure. The extreme case where this occurred was with the SF6D (Brazier, et al. 2002). The weights for the SF6D were obtained from a representative sample of 836 English persons of which 611 interviews were actually used. Based on a sort procedure, each respondent was asked to value 6 health states out of a possible 249 health states. Altogether 3,518 valuations were made: an average of 15 responses for each health state (the range was from 8 for health state 5,3,5,6,4,6 to 19 for health state 1,3,1,5,4,2). Similar procedures were followed for the HUI3 (Furlong, et al. 1998), AQoL (Hawthorne, et al. 1997), and 15D (Sintonen, 1995), although in each case the numbers were greater than for the SF6D. For example, for the HUI3 the numbers for each health state varied from 19 to 246; for the AQoL the range was 70 through 225). These difficulties for each instrument were compounded by the relatively low response rates (typically about 50% although the AQoL's was higher).

These wafer-thin estimates raise fundamental questions concerning the transparency of utility scores, their stability and the generalisability of the instruments. Other than for the EQ-5D, none of the instrument developers have reported validation of the obtained utility results or published an analysis of these data. Given this, it is highly likely the utility values for all instruments, other than the EQ-5D, are biased and lack transparency. Because of the restricted response rates and small sample sizes utility weights may be less than stable; a problem compounded by the fact that all instrument weights have been derived using means rather than medians. Clearly under these circumstances claims for generalisability to many health conditions, including dementia, should be interpreted cautiously.

### **The response perspective**

Implicit in the above discussion is that it is the study participant who provides the responses on an MAU-instrument. This position is consistent with that of the World Health Organization, which defined QoL as:

*...an individual's perceptions of their position in life in the context of the culture and value system in which they live and in relation to their goals, expectations, standards and concerns. It is a broad ranging concept affected in a complex way by the person's physical health, psychological state, level of independence, social relationships, and their relationship to salient features of their environment. (WHOQoL Group, 1993, p153)*

This assumes that the individual has the necessary insight into their own life to be able to provide meaningful assessments. For those with dementia, however, it may be that the level of cognitive impairment is such that insightful assessments cannot be reliably elicited. In addition to cognitive impairment, there are concerns around issues of current affective state, adaptation, lack of insight, neuroticism and emotional adjustment, and possible effects of neuroleptic therapy (Wood, et al. 1985; Diener, et al. 1999; Kring, et al. 1993; Jenkins, 1992; Awad, et al. 1995; Coucill, et al. 2001; Magaziner, 1997). It is for these reasons that proxy respondents are widely advocated and used in both mental health and dementia research (Albert, et al. 1996; Blomfeldt, et al. 2005; Kerner, et al. 1998; Magaziner, 1997; Neumann, et al. 1999b, 2000).

The issues around these arguments are briefly reviewed under the areas of missing data, reliability, proxy-completion and interviewer-facilitated data collection.

Missing data seems to be a function of instrument length and severity of cognitive impairment. In a study of Alzheimer's disease where patients were cognitively impaired (Mini Mental State Examination [MMSE]: two groups - scores 19-26 = mild dementia; scores 10-18 = moderate dementia), patients and their carers were administered the HUI3, QWB and EQ-5D. The proportion of missing data varied by instrument length and by dementia level, suggesting that the simplest instrument, the EQ-5D, was to be preferred among those with more severe impairment (Naglie, et al. 2006). In a mail-out study comparing the EQ-5D with the SF36 at 1 year after stroke, about half of the forms were completed by the patients alone and, importantly, there was an 11% difference in missing data favouring the EQ-5D suggesting that the simplicity of the EQ-5D enabled patients with poor health outcomes to participate (Dorman, et al. 1997). These results were similar to those of Holland, et al. who compared the AQoL and EQ-5D among older adults leaving intensive care, finding an overall lower response rate for the AQoL than for the EQ-5D (65% versus 81%) and a higher rate of missing data in the AQoL (Holland, et al. 2004).

Reliability of measurement, as assessed by test-retest, internal consistency or standard deviation increase or decrease, also appears to be a function of the severity of cognitive impairment. The Naglie, et al. (2006) study above reported on the test-retest reliability of the HUI3, QWB and EQ-5D at 13 days. For those with MMSE scores 19-26 the test-retests were within an acceptable range (0.70 – 0.81). For those with MMSE scores 10-18 the test-retest coefficients varied by instrument length: for the EQ-5D, HUI3 and QWB the coefficients were 0.83, 0.25 and 0.59, respectively. The conclusion was that patients with mild Alzheimer's disease could rate their own QoL, and that those with moderate impairment could do this with a facilitated interview. This finding is consistent with a study of the AQoL in those with Alzheimer's disease where the data were collected in interview, for those with MMSE scores 10+ the AQoL means and standard deviations were consistent and monotonic with MMSE scores, whereas for those with MMSE scores in the range 0-10 the standard deviation was extremely broad (Wlodarczyk, et al. 2004). The interpretation would be that for those with MMSE <10 self-completion of the AQoL was problematic. These findings are consistent with Folstein, et al's original work (1975). The mean MMSE score of dementia patients was 10 for one sample and 12 for another; at 28-day test-retest the correlation was 0.98. The implication is that, in interview, people with moderate dementia, as defined by an MMSE score of  $\geq 10$  can provide insight and complete short QoL measures, such as the EQ-5D or AQoL.

The above would suggest that proxy-completion is to be preferred where participants have moderate cognitive impairment. Generally, consistent with how MAU-instruments have been developed in the past 30 years, health professionals hold the view that:

*"...indices can be designed so that clinicians can score the patient's quality of life or health status after observing or examining a patient even without eliciting information from the patient about how he or she feels at any given point in time" (Spitzer, 1987, p469)*

Yet the research into patient self-report versus proxy report is only partly supportive of this position, and some commentators have rejected proxy-completion as being highly misleading or because there almost no evidence to support it (Jonsson, 2003; Sprangers and Aaronson, 1992; Cummins, 2002). In general, the literature suggests that patient self-report is to be preferred (Awad and Voruganti, 1999; Becchi, et al. 2004; Ankri, et al. 2003; Bullinger, et al. 2002; Tamim, et al. 2002; Naglie, et al. 2006), although there are some equivocal studies (Scocco, et al. 2005).

Where patient and proxy utility scores have been compared, the patient scores are higher than those of the proxies (Bryan, et al. 2005; Coucill, et al. 2001; Herrman, et al. 2002; Jonsson, et al. 2006; Naglie, et al. 2006; Pickard, et al. 2004; Wlodarczyk, et al. 2004; Wu, et al. 1997). The implication is that patients rate their QoL higher than do external observers, perhaps for reasons of adaptation, that external observers may not be aware of all aspects of patients' lives, or that an external observer may focus on the negative aspects of a person's life (Cummins, 2002). This judgement rests on the fact that there is, generally, highest agreement between patient and proxy assessments in those areas of more 'objective' measurement (e.g. mobility) and greater discrepancy in the subjective areas of life (e.g. social relationships) (Bryan, et al. 2005; Coucill, et al. 2001; Herrman, et al. 2002; Sainfort, et al. 1996; Voruganti, et al. 1998), although this was not supported in Naglie et al's study of the EQ-5D, HUI3 and QWB (Nagle, et al. 2006) or Wlodarczyk, et al's study of the AQoL (Wlodarczyk, et al. 2004). Ankri, et al. (2003) reported that the level of agreement was at least partly a function of cognitive impairment where the greater discrepancies were among those with the most severe impairment – a finding that was not supported by Jonsson, et al's (2006) study of the relationship between proxy and patient completion of the EQ-5D.

Regarding proxy-completion, then, it might be expected that agreement between patient and proxy assessments would be moderate, at best. In a study of proxy completion among stroke survivors of the EQ-5D and the HUI3 (Pickard, et al. 2004) the kappa agreement between patient and proxy reports ranged between 0.18 to 0.73 for the items on the EQ-5D, and for the HUI3 the percentage of cases with exact agreement varied between 33% and 81%. The differences between patient and proxy utilities varied from 0.00 (HUI3 at baseline) to 0.06 (EQ-5D at baseline and HUI3 at 1-month). The authors reported that the proxy reports were more reliable. The opposite conclusion (that patient report was to be preferred when compared with case manager report) was reached by Herrman, et al. (2002) who examined the relationship between self-report by those with long-term psychosis and the reports of their case managers on the AQoL. The criterion for making this judgement was monotonicity against a common measure of QoL status (i.e. the pooled estimate of both patient and proxy). The patient estimates were monotonic whereas the proxies' were not.

In another study of the AQoL among those with Alzheimer's disease, the correlation between patient and proxy assessments was  $r = 0.37$  (Wlodarczyk, et al. 2004). Coucill, et al. (2001) reported on the EQ-5D in persons with dementia, and their caregivers and clinicians as proxies. The kappa agreements between the three estimates were poor (items assessing usual activities, pain and anxiety) or moderate (mobility and self-care). The researchers concluded that although there was reason to discount patient reports, it was not clear who the proxies should be – a finding echoed by Ankri, et al. in their study of the EQ-5D (Ankri, et al. 2003). In another study of the EQ-5D, Bryan, et al. (2005) reported that clinician-proxies provided higher utility estimates than did carer-proxies, but that this was a function of type of item being responded to (see above). It is possible this finding reflects that the main factor affecting the QoL of those with dementia is their

ability to perform activities of daily living (Andersen, et al. 2004), or that parents' or caregivers' estimates may be biased due to the belief that their care is beneficial. It is likely that proxies, where used, should be partners or peers (Cummins, 2002).

An alternative to proxy-completion has been administration in an interviewer-facilitated setting, i.e. where the interviewer reads the MAU out to the participant (Ankri, et al. 2003). Recommendations for interview-administration are a function of impairment level (Naglie, et al. 2006, Ankri, et al. 2003). In a study of the EQ-5D in those with mild, moderate or severe dementia, Coucill, et al. (2001) concluded that the EQ-5D could be patient-completed when interviewer-administered, but that there was little evidence to support patient self-rated completion. There is, however, an important caveat to facilitated interview completion. It is often assumed that a cognitively disabled person who has difficulty reading and responding to a complex questionnaire on their own can be verbally administered a questionnaire and the results accepted as valid. This assumption, however, is challengeable for three important reasons. First, the setting (interviewer reading and respondent selecting a verbal option) may lead to acquiescent response bias, which is where a respondent provides an answer that he/she deems acceptable (usually on the grounds that he/she is trying to please the interviewer in some way) (Sigelman, et al. 1981b, 1981a; Foddy, 1993). Second, where material is poorly understood and is rephrased by the interviewer the rephrasing may represent the interviewer's beliefs about the question and what the response should be (Rapley and Antaki, 1996; Antaki and Rapley, 1996; Antaki, 1999). Third, there is some evidence that during facilitation interviewers lead the patient into particular responses (Antaki, 1999). It follows that interviewer-facilitated data collected using a non-standardized interview schedule may result in QoL data that neither represents the QoL of the respondent nor that is comparable with other data from the same study (e.g. that collected through self-report).

## **7.5 Instrument Responsiveness**

Regarding instrument responsiveness to dementia states, the literature reviewed above indicates that the AQoL, EQ-5D, HUI3 and QWB are responsive to increasing levels of dementia; no studies were found for the 15D or SF6D. The reason for this responsiveness appears to be that the losses associated with moving from one level to another level of dementia are so large that MAU instruments will pick them up. This does not mean, however, that the instrument will deliver the same scores for the same cognitive states or that the utilities are comparable.

For the AQoL, in a study of those with Alzheimer's disease, for those with mild Alzheimer's (MMSE score of 20-24) the mean utility was 0.71 compared with 0.52 for those with severe Alzheimer's (MMSE score of 10) – a difference of 0.19 utilities (Wlodarczyk, et al. 2004). A study of stroke victims showed that between those with and without dementia, the mean AQoL difference was 0.25 utilities (Sturm, et al. 2004).

Smaller differences in EQ-5D utilities by MMSE classification were reported by Anderson, et al. (2004), whose findings showed that between those with MMSE >20 and 10-19 the difference was just 0.04 utilities, and between <10 and 10-19 it was 0.11 EQ-5D utilities. These smaller differences were consistent with those reported by Jonsson, et al. (2006) who found that for the MMSE groups >25, 21-25, 15-20, 10-14 and <10 the EQ-5D differences were -0.01, 0.02, 0.10 and -0.05. In another study responses to EQ-5D items were not related to MMSE scores until these were  $\leq 10$  (which would be consistent with the Anderson, et al. study), leading to speculation that these participants may have lacked insight (Ankri, et al. 2003). Generally, these findings for the EQ-5D are inconsistent and do not engender confidence of its responsiveness to dementia level. It is possible this is because of ceiling effects; two studies reported substantial ceiling effects among patient participants that was not related to dementia level (43% at the first interview and 57% at a second interview, and 48% (Naglie, et al. 2006; Coucill, et al. 2001). In contrast with these reports, Ankri, et al. in a study where the EQ-5D was administered by an interviewer to 96% of participants with dementia, no ceiling effects were reported (15% of cases obtained ceiling scores) (Ankri, et al. 2003). It is possible that the ceiling effects on the EQ-5D are associated with

---

self-report among those with dementia. Such ceiling effects have not been reported for other instruments (AQoL, HUI3 and QWB) when used in dementia studies.

For the HUI3, Alzheimer's disease in a Japanese study revealed differences between mild, moderate and severe Alzheimer's disease of 0.17 and 0.14 utilities, respectively (Ikeda, et al. 2001). Of interest is that Neumann, et al. (2000) reported monotonic declines between dementia levels of 0.08, 0.20, 0.13, 0.14 and 0.15. For those with profound or terminal cognitive impairment the mean HUI3 scores were negative utilities, something not reported on any other instrument.

For the Index of Wellbeing, the forerunner of the QWB (Kaplan, et al. 1976), the Wimo, et al. (1995) study described above, reported that between Global Deterioration Scale classifications the decrements were 0.07, 0.11 and 0.08. In another study using the QWB, it was reported that the differences between those with Alzheimer's symptoms (e.g. asking repeatedly) and the symptom-free were ~0.10 utilities (Kerner, et al. 1998).

The results from these different studies are summarized in Table 29.

**Table 29 Responsiveness of Selected MAU-Instruments to Various Health Conditions (a)**

<b>Responsiveness of selected MAU-instruments to various health conditions (a)</b>				
<b>Study</b>	<b>Condition</b>		<b>Levels</b>	<b>Mean</b>
<b>AQoL</b>				
Wlodarczyk, et al. 2004	Alzheimer's disease	MMSE	25+	0.73
			20-24	0.71
			15-19	0.61
			10-14	0.55
			≤10	0.52
Sturm et al, 2004	Stroke victims, 2-year follow-up	Dementia status	No	0.27
			Yes	0.02
<b>EQ-5D</b>				
Andersen, et al. 2004	Dementia	MMSE	>20	0.64
			10-19	0.60
			≤9	0.49
Jonsson, et al. 2006	Alzheimer's disease	MMSE	26-30	0.84
			21-25	0.85
			15-20	0.83
			10-15	0.73
			0-9	0.78
<b>HUI3</b>				
Ikeda, et al. 2001	Alzheimer's disease		Mild	0.33
			Moderate	0.16
			Severe	0.02
Neumann, et al. 2000	Alzheimer's disease	Clinical Dementia Rating Scale classification	Questionable	0.47
			Mild	0.39
			Moderate	0.19
			Severe	0.06
			Profound	-0.08
			Terminal	-0.23
<b>QWB</b>				
Wimo, et al. 1995	Dementia	Global Deterioration Scale score	4	0.61
			5	0.54
			6	0.43
			7	0.35

**Notes:**

a = Calculated from original papers

The most obvious point from the table is that there is no consistency between the four instruments. For those with mild cognitive impairment the utility scores vary from 0.84 (EQ-5D) to 0.33 (HUI3, ignoring the stroke victims' score of 0.27 on the AQoL Sturm, et al. study which may be due to co-morbidities). Similarly, for those with severe impairment the utility scores range from -0.23 (HUI3) to 0.78 (EQ-5D).

Even between different studies for the same instrument, there are very different estimates. For example, for those with severe impairment for the AQoL the scores of 0.52 and 0.02 were reported, for the EQ-5D these values were 0.49 and 0.78, and for the HUI3 -0.23 and 0.02.

In short, although all four instruments are clearly sensitive to dementia state, there is almost no agreement between them and almost no reliable information in the literature that can be used to guide the researcher into instrument selection for any particular study.

## **7.6 Conclusions and Recommendations**

The literature reviewed above suggests that none of the existing MAU-instruments are truly suitable for use in cost-utility analyses of dementia in Australia at the present time. There are four key reasons for this judgement.

- A. There is substantial evidence in the literature that none of the existing MAU-instruments meet the axioms for valid and reliable utility measurement. The different measures' descriptive systems are based on different assumptions about the nature of QoL and its constituent parts, the technical properties of the different measures do not meet the requirements for the computation of QALYs from the scored descriptive systems and the weights used in all the instruments are, at best, circumscribed. Against this rather pessimistic view, however, is that the inclusion of MAU-instruments is the only known way of capturing the value of the patient perspective so that it can be placed alongside other information, such as clinical information, during evaluation of interventions and decision-making about resource allocation. Perhaps the results from MAU-instruments should be regarded as indicative of patient preferences rather than as absolute preferences.
- B. There are sufficient doubts in the literature in relation to self-completion by those with dementia to suggest invalidity in many situations. This conclusion rests on the research evidence suggesting that those with different levels of cognitive impairment may complete the instruments differently, and that in a study with a range of cognitively impaired participants a large number of participants may lack the insight needed to meaningfully complete an MAU-instrument. This is particularly evident among those with MMSE scores <~10 points. For those in the range 10+ MMSE points, self-completion may be possible.
- C. Although it has been widely suggested and practiced, there are good reasons to be extremely cautious of the use of proxy and interviewer-facilitated administrations. There is evidence that proxy values do not agree with those of the patient, and that different proxies produce different utilities. Additionally, the literature suggests that interviewer-facilitated completion may lead to serious bias from a variety of sources that may compromise the results.
- D. There is a very limited literature examining MAU-instruments by levels of cognitive impairment. This literature suggests that there is almost no agreement either within instruments or between instruments on the utilities for given levels of dementia.

These conclusions suggest that the results of any particular study will be a function of the study population, the research methods (including instrument administration method) and the MAU instrument chosen. In short, no MAU-instrument can be recommended as the 'gold standard' for use in dementia studies at this point of time.

\*\*\*

The recommendations below have been framed by the fact that there are three levels at which utility instruments could be used in dementia studies: (a) clinicians working in clinical practice, (b) specialists working in clinical practice, and (c) researchers or program evaluators.

At the clinical level, measurement is usually related to clinical management of individual patients and there are time and data collection issues which impact on recommended practice. Any instruments used at this level must possess sufficient nomological evidence to be used at the case level; i.e. for individual patient assessment. Additionally, at this level, data collection should be as brief as possible and there should be few data analysis demands upon clinicians.

Under (b), data collected need to be sufficient to meet the needs of specialists. Whilst these include the requirements of clinical measurement, specialists need more information and are often involved in research or evaluation.

Researchers and program evaluators' needs centre around data that are useful for answering research questions where analyses are group-based; where data collection procedures may be remote; and where findings are aimed at demonstrating the effect of new treatments or at influencing policy decisions.

MAU-instruments, by definition, were designed for use by researchers undertaking economic evaluation. However, this does not necessarily imply that they have no role to play in clinical or specialist services. At the individual level, MAU-instruments may provide HRQoL profiles based on responses to individual questions or utility scores which may be compared to group or population norms. Additionally, in a health care system committed to evidence-based practice, basic data should be collected and held at the clinician level for local analysis as well as transference to research (e.g. for inclusion in dementia services monitoring or surveillance).

### 7.6.1 Summary Comments

In general, conclusions drawn from this review should be placed in the following contexts which are germane to using MAU-instruments in dementia studies. In light of the above general comments, only the AQoL, EQ-5D and HUI3 are rated.

- *Instrument length.* Given known difficulties with dementia level and missing data (almost certainly due to higher cognitive demands on respondents), and that the chosen MAU-instrument is likely to be used in an instrument battery, it would seem that short instruments should be considered. The only truly short MAU-instrument is the EQ-5D. Both the AQoL and HUI3 are more than double the length of the EQ-5D.
- *Coverage.* There is a clear difference between the utility instruments in relation to their coverage (see Table 28). If instruments providing 'within the skin' coverage are to be preferred, the choice would be the HUI3. On the other hand, if the 'social expression' of HRQoL is desired, the AQoL would be the instrument of choice. The EQ-5D may be more suitable for use in institutions because it has items covering being confined to bed, being unable to wash or dress and being unable to perform usual activities.
- *Administration.* Recommended national instruments are likely to be used in a variety of settings, particularly in studies where data are collected through self-completion (most probably in interview situations). A key point here is the need for instruments which are insensitive to administration mode (because many interviewers may facilitate an interview, particularly with those who are suffering severe cognitive impairment). There is insufficient evidence on this topic in relation to dementia studies for any substantial recommendation to be made. A general point regarding self-completion can be made, however. In general the literature suggests self-completion is to be preferred. There is evidence that for people with severe cognitive impairment, self-completion produces spurious results. Generally, the literature seems to suggest that this is the case for people with MMSE scores  $\leq 10$  (see Table 29). There is a case for recommending that self-completion should be assisted for those with MMSE scores  $\leq 12$ .
- *Ease of use.* The instrument which respondents find simple and easy to use is the EQ-5D. The language of both the AQoL and HUI3 is more complex than that of the EQ-5D and the HUI3

has several items with 6-response levels which may pose difficulties for those with moderate or severe cognitive impairment.

- *Time to complete.* To reduce the burden on participants and the costs associated with data collection this should be as short as possible. The instrument with the shortest time to complete is the EQ-5D.
- *Translation.* Translations will almost certainly be required for some sample sub-groups given the heterogeneous Australian population and the tendency that many immigrants revert to their native language as they age. The only MAU-instrument, *per se*, that is readily available in a number of languages is the EQ-5D.
- *Scoring.* Although this does not directly impinge upon data collection, it does have some implications for data analysis where research groups may not have ready access to either a statistician or instrument technical support. Additionally, any recommended instrument must have a scoring system that is valid. The preferred instrument would be the AQoL, given it is weighted with Australian values. Neither the EQ-5D nor HUI3 have had their weights validated for Australian use. The recent literature on the EQ-5D suggests there may be competing scoring methods (see Figure 2) giving very different utility estimates for those in poor health.
- *Sensitivity.* Although all reviewed instruments for which there are published data (AQoL, EQ-5D, HUI3 and QWB) appear sensitive to dementia level, the values obtained for the EQ-5D and HUI3 are problematic (see Table 29). The AQoL and HUI3 instruments would be preferred.
- *Reliability.* All the instruments reviewed are likely to possess similar reliability characteristics. However, this has not been fully investigated for all instruments in dementia samples.
- *Validity.* All the instruments reviewed have some questions about their validity. This has not been satisfactorily established and published for any of the instruments, particularly in relation to the generalisability of the utility weights and the necessary strong interval property. Based on the reviewed literature, there is insufficient evidence to make a mature assessment of the validity of any MAU-instrument in dementia studies.
- *Utility axioms.* None of the instruments reviewed meet all the requirements for utility measurement at this time. However, the review suggests that those instruments with the better claims for meeting these axioms would be the HUI3 and AQoL, then the EQ-5D.

Table 30 provides a summary of the findings from this study. Each of the instruments reviewed was assessed against the descriptions and validity evidence presented in this report. For each of these criteria, the assessment was made on a 3-point scale where a low score indicated minimally meeting the criteria and a high score indicated mostly meeting the criteria. Additionally, each of the criteria was weighted according to its perceived importance to the Australian context. The results suggest that the instruments of choice would be the EQ-5D and AQoL. The reasons for this are:

- Based on the review and the available literature, the 15D, Rosser-Kind and SF6D are not recommended because they do not meet the basic axioms of utility theory. There is no evidence of their use in dementia studies.
- It is difficult to recommend the QWB or the more recent QWB-SA. Although this has been used in at least one study of dementia, the descriptive system suggests that the instrument is primarily concerned with dis-utility derived from having a health condition. Furthermore, it is too long to be regularly used in evaluation studies involving people with moderate or severe cognitive impairment.
- Although the HUI3 is a widely used and well respected instrument, there are doubts concerning its use in dementia studies. The descriptive system is more verbose than those of other competitor instruments. A particular concern is that the number of response options (6

levels) for some items may pose discrimination difficulties for those with cognitive impairment. The range of utility scores available is implausible – as shown by the data in Table 29 for those in terminal health states. If, on a utility scale, 0.00 represents death, then those in terminal health states should be assigned scores near 0.00. As shown in the table, this is patently not the case. The cost of the HUI3 is also a barrier to its widespread adoption, since the current cost (CAN\$4000 per trial) is likely to be beyond the financial resources available to many small institutions providing health care to those with dementia.

- The AQoL has the virtue of being weighted with Australian utility weights; at just 12-scoring items it is also the second shortest measure (along with the HUI3) after the EQ-5D. Although sensitive to cognitive impairment level, the one paper reporting the AQoL by cognitive impairment suggests invalidity of scores below MMSE ~10. Whether this is caused by the language of the descriptive system or the length of the AQoL is not known. Theoretically, given the factorial structure of the AQoL, it could be shortened through removal of 4 items (1 from each dimension) leaving it as an 8-item instrument.
- The EQ-5D is attractive because of its simple descriptive system. Against this must be placed the evidence relating to poor data distribution, that EQ-5D utility scores systematically vary by whether British or US weights are used (and that there has been no validation of either weight set for Australian use), that there is evidence of ceiling effects among those with dementia, and that reported scores are inconsistent by dementia severity, as shown in Figure 3 and Table 29.

**Table 30 Summary Assessing the Utility Instruments Against the Study Criteria**

Criteria	Weight	Instrument						
		EQ-5D	AQoL	HUI3	15D	QWB	SF6D	Rosser
Availability of comparison data	3	3	3	3	3	3	3	2
Length/feasibility of instrument for inclusion in battery	2	3	2	2	2	1	1#	1
Complexity of administration /cognitive burden	2	3	2	1	2	1	1#	2
Cultural Appropriateness	1	2	3	2	2	2	2	2
Ease of obtaining score	2	3	3	3	3	3	3	2
Sensitivity to dementia	3	2	2	2	1	2	1	1
Reliability evidence	3	2	2	2	1	1	1	1
Validity evidence	3	2	2	2	1	1	1	1
Cost of the instrument	2	2	3	1	3	3	1	3
Cost of instrument administration	2	3	3	2	3	1	3	1
<b>Weighted Total</b>		<b>57</b>	<b>56</b>	<b>47</b>	<b>44</b>	<b>41</b>	<b>38</b>	<b>35</b>

\* As most MAU instruments are short the criteria are revised as follows: 1= Long instrument or needs interview administration, 2= moderate length self completed instrument, 3= short, self completed instrument.

# Although it only contains 10 items it requires the full administration of the SF-36 scale

## 7.6.2 Research Recommendations

There are, therefore, a number of options which could be considered either individually or collectively.

1. Although there are acknowledged difficulties with the conduct of CUA studies in dementia or Alzheimer's disease, the limited evidence suggests that those with mild to moderate cognitive impairment can provide self-report utility values suitable for use in CUA studies. Where this is the case, there is no necessary reason conventional CUA cannot be performed, even though it may take longer and need some additional resources. Although they have their place in the literature, there is sufficient evidence supporting self-report to suggest that the kind of modelling exercises reported in the literature (see section 7.1) should largely give way to conventional cost utility analyses.
  - Where there is severe cognitive impairment (e.g.  $\leq 11$  MMSE points) either interview-facilitation or proxy report can be used. Where these are used, standard interview scripts should be prepared and administered and the same procedure used throughout the study for all data collection to minimise any bias arising through mixing of different data collection methods.
2. A single MAU-instrument could be recommended as the preferred instrument of choice for routine use at the clinician- and specialist-levels. This instrument should be short, easy to administer and score and population norms could be made available for easy reference. If such a policy was adopted, it would be in light of the limitations outlined in this report and there would be no guarantee that results obtained would be comparable with results obtained elsewhere using another instrument. Indeed, where QALYs were computed as the result of a treatment, it is likely these would reflect instrument choice as much as treatment effect. Where two MAU-instruments were recommended as the preferred measures, these difficulties would be compounded if some studies included one of the instruments and other studies opted for the other instrument.
3. To overcome this uncertainty, two MAU-instruments could be included in any particular research or evaluation study, and that researchers be encouraged to provide both sets of results. One of the recommended instruments should be that recommended for clinician use. This strategy would have the benefit of reducing the bias inherent in a one-instrument strategy, and it would produce a range of estimated benefits from interventions, thus acknowledging the limitations of relying upon any particular existing MAU-instrument. Given that, inevitably, comparisons will be made with dementia studies overseas, this strategy would have the further benefit of enabling cross-cultural comparisons. An important limitation of this strategy is that it would increase the cognitive burden for those with moderate to severe cognitive impairment. It may lead to interviewer-facilitated or proxy completions, with all the implications of mixed-methods data collection.
4. Several instruments could be trialled in 3 - 4 large dementia studies for the explicit purpose of identifying the instrument to be recommended for future use. Whilst this would impose an immediate burden for, say, 3 to 5 years, it would enable many of the questions raised in this report regarding the validity of MAU-instruments to be thoroughly investigated in an Australian context. This would place Australia in a position of world leadership in dementia and utility research; it would enable a fully informed decision to be made regarding instrument selection; and it is likely the Australian model would become the world standard in the immediate future given the paucity of current research in the field. Should this latter scenario eventuate, it is likely this would enhance international cooperation in the field.
5. Based on the criteria the obvious instrument of choice for use in dementia studies is the EQ-5D because of the simplicity of the descriptive system. There are however very good technical reasons which provide caveats to its widespread use, including competing

scoring algorithms, ceiling effects, inconsistent utility scores and poor score distribution. There is a prima facie case for an Australian study into these aspects of the EQ-5D with a view to validating and/or revising existing EQ-5D scoring algorithms.

6. Other than EQ-5D, the best-performing MAU-instrument was the AQoL. There are, however, two important caveats to recommending it as the instrument of choice. Although the AQoL's descriptive system is simple, the wording of items is stilted. The second caveat is in relation to the number of items needed to score the AQoL (12-items) which may explain higher rates of missing data when compared with the EQ-5D and inconsistent scores for those with severe cognitive impairment. As indicated above, theoretically, given the factorial structure of the AQoL it could be shortened through removal of 4 items (1 from each dimension) leaving it as an 8-item instrument. A study could be conducted to examine the effect of simplifying the items and removing four of them.

## References

- Albert SM, Del Castillo-Castaneda C, Sano M, Jacobs DM, Marder K, Bell K, Bylsma F, Lafleche G, Brandt J, Albert M and Stern Y (1996) *Quality of life in patients with Alzheimer's disease as reported by patient proxies*. Journal of the American Geriatrics Society. Vol. 44, No.11, pp.1342-7.
- Andersen CK, Wittrup-Jensen KU, Lolk A, Andersen K and Kragh-Sorensen P (2004) *Ability to perform activities of daily living is the main factor affecting quality of life in patients with dementia*. Health and Quality of Life Outcomes. Vol. 2, No.52.
- Andersen L, Kristiansen I, Falch J and Aursnes I (1995) *Cost-effectiveness of Alendronate for the prevention of osteoporotic fractures in Norwegian women*. Folkehelsa; Statens Institute for Folkehelsa, Oslo.
- Anderson J, Kaplan R, Berry C, Bush J and Ruben R (1989) *Interday reliability of function assessment for a health status measure: The Quality of Well-Being scale*. Medical Care. Vol. 27, No.11, pp.1076-1084.
- Andresen EM, Rothenberg BM and Kaplan RM (1998) *Performance of a self-administered mailed version of the Quality of Well-Being (QWB-SA) questionnaire among older adults*. Medical Care. Vol. 36, No.9, pp.1349-60.
- Ankri J, Beaufils B, Novella JL, Morrone I, Guillemin F, Jolly D, Ploton L and Blanchard F (2003) *Use of the EQ-5D among patients suffering from dementia*. Journal of Clinical Epidemiology. Vol. 56, No.11, pp.1055-63.
- Antaki C (1999) *Interviewing persons with a learning disability: How setting lower standards may inflate well-being scores*. Qualitative Health Research. Vol. 9, No.4, pp.437-454.
- Antaki C and Rapley M (1996) *'Quality of Life' talk: The liberal paradox of psychological testing*. Discourse and Society. Vol. 7, No.3, pp.293-316.
- Awad G, Hogan T, Voruganti L and Heslegrave R (1995) *Patients' subjective experiences on antipsychotic medications: implications for outcome and quality of life*. International Clinical Psychopharmacology. Vol. 10S, pp.123-132.
- Awad G and Voruganti P (1999) *Cost-utility analysis in schizophrenia*. Journal of Clinical Psychiatry. Vol. 60, Suppl.3, pp.22-28.
- Barton GR, Bankart J and Davis AC (2005) *A comparison of the quality of life of hearing-impaired people as estimated by three different utility measures*. International Journal of Audiology. Vol. 44, No.3, pp.157-63.
- Barton GR, Bankart J, Davis AC and Summerfield QA (2004) *Comparing utility scores before and after hearing-aid provision: results according to the EQ-5D, HUI3 and SF-6D*. Applied Health Economics and Health Policy. Vol. 3, No.2, pp.103-5.

- Becchi A, Rucci P, Placentino A, Neri G and de Girolamo G (2004) *Quality of life in patients with schizophrenia--comparison of self-report and proxy assessments*. Social Psychiatry and Psychiatric Epidemiology. Vol. 39, No.5, pp.397-401.
- Bennett K and Torrance G (1996) *Measuring health state preferences and utilities: Rating scale, time trade-off, and standard gamble techniques*. In Spilker, B (ed.) (2004) Quality of Life and Pharmacoeconomics in Clinical Trials. Lippincott-Raven Publishers, Philadelphia, pp. 253-265.
- Bergner M, Bobbitt RA, Carter WB and Gilson BS (1981) *The Sickness Impact Profile: development and final revision of a health status measure*. Med Care. Vol. 19, No.8, pp.787-805.
- Bleichrodt H and Johannesson M (1997) *An experimental test of a theoretical foundation for rating-scale valuations*. Medical Decision Making. Vol. 17, No.2, pp.208-16.
- Blomfeldt R, Tornkvist H, Ponzer S, Soderqvist A and Tidermark J (2005) *Internal fixation versus hemiarthroplasty for displaced fractures of the femoral neck in elderly patients with severe cognitive impairment*. J Bone Joint Surg Br. Vol. 87, No.4, pp.523-9.
- Bombardier C and Raboud J (1991) *A comparison of health-related quality-of-life measures for rheumatoid arthritis research. The Auranofin Cooperating Group*. Controlled Clinical Trials. Vol. 12, No.4 Suppl, pp.243S-256S.
- Boyle MH, Furlong W, Feeny D, Torrance GW and Hatcher J (1995) *Reliability of the Health Utilities Index - Mark III used in the 1991 cycle 6 Canadian General Social Survey Health Questionnaire*. Quality of Life Research. Vol. 4, pp.249-257.
- Bradlyn AS, Harris CV, Warner JE, Ritchey AK and Zaboy K (1993) *An investigation of the validity of the quality of Well-Being Scale with paediatric oncology patients*. Health Psychology. Vol. 12, No.3, pp.246-50.
- Brazier J and Deverill M (1999) *A checklist for judging preference-based measures of health related quality of life: learning from psychometrics*. Health Econ. Vol. 8, No.1, pp.41-51.
- Brazier J and Deverill M, Green C, Harper R and Booth A (1999) *A review of the use of health status measures in economic evaluation*. Health Technology Assessment. Vol. 3, No.9, pp.1-165.
- Brazier J, Jones N and Kind P (1993) *Testing the validity of the EuroQoL and comparing it with the SF-36 health survey questionnaire*. Quality of Life Research. Vol. 2, pp.169-180.
- Brazier J, Roberts J and Deverill M (2002) *The estimation of a preference-based measure of health from the SF-36*. Journal of Health Economics. Vol. 21, pp.271-292.
- Brazier J, Roberts J, Tsuchiya A and Busschbach J (2004) *A comparison of the EQ-5D and SF-6D across seven patient groups*. Health Economics. Vol. 13, No.9, pp.873-84.
- Brazier J, Usherwood T, Harper R and Thomas K (1998) *Deriving a preference-based single index from the UK SF-36 Health Survey*. Journal of Clinical Epidemiology. Vol. 51, No.11, pp.1115-1128.
- Brazier JE and Roberts J (2004) *The estimation of a preference-based measure of health from the SF-12*. Medical Care. Vol. 42, No.9, pp.851-9.
- Bryan S, Hardyman W, Bentham P, Buckley A and Laight A (2005) *Proxy completion of EQ-5D in patients with dementia*. Quality of Life Research. Vol. 14, No.1, pp.107-18.
- Buckingham K (1995) *The Research Strategy*. Methodology Workshop, EUROQoL Conference, London.
- Bullinger M, Azouvi P, Brooks N, Basso A, Christensen AL, Gobiet W, Greenwood R, Hutter B, Jennett B, Maas A, Truelle JL and von Wild KR (2002) *Quality of life in patients with traumatic brain injury-basic issues, assessment and recommendations*. Restorative Neurology and Neuroscience. Vol. 20, No.3-4, pp.111-24.
- Cadet B (1994) *History of the construction of a health indicator integrating social preference: the Quality of Well-Being Scale*. 7th Meeting, International Network on Health Expectancy (REVES). Canberra, 1994.

- Cattell R (1952) *Factor Analysis: an Introduction and Manual for the Psychologist and Social Scientist*. Harper and Row, New York.
- Conner-Spady B and Suarez-Almazor ME (2003) *Variation in the estimation of quality-adjusted life-years by different preference-based instruments*. Medical Care. Vol. 41, No.7, pp.791-801.
- Cook K, Ashton C, Byrne M, Brody B, Geraci J, Giesler R, Hanita M, Soucek J and Wray N (2001) *A psychometric analysis of the measurement level of the rating scale, time trade-off, and standard gamble*. Social Science and Medicine. Vol. 53, pp.1275-1285.
- Coons S and Kaplan R (1993) *Quality of life assessment: understanding its use as an outcome measure*. Hospital Formulary. Vol. 28, pp.486-498.
- Coucill W, Bryan S, Bentham P, Buckley A and Laight A (2001) *EQ-5D in patients with dementia: an investigation of inter-rater agreement*. Medical Care. Vol. 39, No.8, pp.760-71.
- Cummins RA (2002) *Proxy responding for subjective well-being: a review*. International Review of Research in Mental Retardation. Vol. 25, pp.183-207.
- Diener E, Suh E, Lucas R and Smith H (1999) *Subjective well-being: three decades of progress*. Psychological Bulletin. Vol. 125, pp.276-302.
- Dolan P (1997) *Modeling valuations for EuroQol health states*. Medical Care. Vol. 35, No.11, pp.1095-108.
- Dolan P, Gudex C, Kind P and Williams A (1995) *Social tariff for EUROQoL: results from a UK general population survey*. Centre for Health Economics, University of York, York.
- Dolan P, Gudex C, Kind P and Williams A (1996) *Valuing health states: a comparison of methods*. Journal of Health Economics. Vol. 15, pp.209-231.
- Dorman P, Slattery J, Farrell B, Dennis M and Sandercock P (1998) *Qualitative comparison of the reliability of health status assessments with the EuroQol and SF-36 questionnaires after stroke*. *United Kingdom Collaborators in the International Stroke Trial*. Stroke. Vol. 29, No.1, pp.63-68.
- Dorman P, Slattery J, Farrell B, Dennis MS and Sandercock PA (1997) *A randomised comparison of the EuroQol and Short Form-36 after stroke*. *United Kingdom collaborators in the International Stroke Trial*. British Medical Journal. Vol. 315, No.7106, pp.461.
- Elvik R (1995) *The validity of using health state indexes in measuring the consequences of traffic injury for public health*. Social Science and Medicine. Vol. 40, No.10, pp.1385-98.
- EuroQol Group (1990) *EuroQol: a new facility for measurement of health-related quality of life*. Health Policy. Vol. 16, pp.199-208.
- Feeny D, Furlong W and Torrance G (1996) *Health Utilities Index Mark 2 and Mark 3 (HUI2/3) 15-item questionnaire for self-administered, self-assessed usual health status*, Centre for Health Economics and Policy Analysis. McMaster University, Hamilton.
- Feeny D, Torrance G and Furlong W (1996) *Health Utilities Index*. In Spilker, B (ed.) Quality of Life and Pharmacoeconomics in Clinical Trials. Lippincott-Raven Publishers, Philadelphia.
- Foddy W (1993) *Constructing Questions for Interviews and Questionnaires: Theory and Practice in Social Research*. Cambridge University Press, Cambridge.
- Folstein M, Folstein S and McHugh P (1975) *'Mini-Mental State': a practical method for grading the cognitive state of patients for the clinician*. Journal of Psychiatric Research. Vol. 12, pp.189-198.
- Fryback DG (2005) *A US valuation of the EQ-5D*. Medical Care. Vol. 43, No.3, pp.199-200.

- Furlong W, Feeny D, Torrance G, Goldsmith C, DePauw S, Zhu Z, Denton M and Boyle M (1998) *Multiplicative Multi-attribute Utility Function for the Health Utilities Index Mark 3 (HUI3) System: A Technical Report*. Centre for Health Economics and Policy Analysis, McMaster University, Hamilton.
- Furlong WJ, Feeny DH, Torrance GW and Barr RD (2001) *The Health Utilities Index (HUI) system for assessing health-related quality of life in clinical studies*. Annals of Medicine. Vol. 33, pp.375-384.
- Geldmacher DS (2002) *Cost-effective recognition and diagnosis of dementia*. Seminars in Neurology. Vol. 22, No.1, pp.63-70.
- Goldney RD, Fisher LJ, Wilson DH and Cheek F (2001) Suicidal ideation and health-related quality of life in the community. Medical Journal of Australia. Vol. 175, No.10, pp.546-549.
- Gudex C, Kind P, van Dalen H, Durand M, Morris J and Williams A (1993) *Comparing Scaling Methods for Health State Valuations - Rosser Revisited*, Centre for Health Economics, University of York, York.
- Havranek EP and Steiner JF (2005) *Valuation of health states in the US versus the UK: two measures divided by a common language?* Medical Care. Vol. 43, No.3, pp.201-2.
- Hawthorne G (2003) *The effect of different methods of collecting data: mail, telephone and filter data collection issues in utility measurement*. Quality of Life Research. Vol. 12, pp.1081-1088.
- Hawthorne G, Cheek F, Goldney R and Fisher L (2003a) *The excess cost of depression in South Australia: a comparative study of two methods of calculating burden*. Australian and New Zealand Journal of Psychiatry. Vol. 37, No.3, pp.362-373.
- Hawthorne G, Osborne R, McNeil H and Richardson J (1996) Centre for Health Program Evaluation, Melbourne.
- Hawthorne G, Osborne RH and Elliott P (2003b) *Commentary on: a psychometric analysis of the measurement level of the rating scale, standard gamble and time-trade off, by Cook et al*. Social Science and Medicine. Vol. 56, pp.895-897.
- Hawthorne G and Richardson J (1995) *An Australian MAU/QALY Instrument: Rationale and Preliminary Results*. Proceedings Health Outcomes and Quality of Life Measurement Conference, Canberra.
- Hawthorne G and Richardson J (2001) *Measuring the value of program outcomes: a review of utility measures*. Expert Review of Pharmacoeconomics Outcomes Research. Vol. 1, No.2, pp.215-228.
- Hawthorne G, Richardson J and Day N (2000a) *A comparison of five multi-attribute utility instruments*. Australian Health Economics Society Conference: Economics and Health: 2000 Proceedings of the Twenty-First Australian Conference of Health Economists, Gold Coast, Australia, pp.151-179.
- Hawthorne G, Richardson J and Day N (2000b) *A comparison of the Assessment of Quality of Life (AQoL) with four other generic utility instruments*. XII Medical Symposium "Quality of Life Measurement in Clinical Studies", Helsinki, Finland, pp.358-3760.
- Hawthorne G, Richardson J and Day N (2001a) *A comparison of five multi-attribute utility instruments*. Australian Studies in Health Services Administration. Vol. 89, pp.151-179.
- Hawthorne G, Richardson J and Day N and McNeil H (2000c) *Life and Death: Theoretical and Practical Issues in Utility Measurement*. Centre for Health Program Evaluation, Melbourne.
- Hawthorne, Richardson J, Osborne R, Day N and McNeil H (2000d) *Construction and Utility Scaling of the Assessment of Quality of Life (AQoL) Instrument*. Centre for Health Program Evaluation. Melbourne.
- Hawthorne G, Richardson J and Day NA (2001b) *A comparison of the Assessment of Quality of Life (AQoL) with four other generic utility instruments*. Annals of Medicine. Vol. 33, No.5, pp.358-370.
- Hawthorne G, Richardson J and Osborne R (1999) *The Assessment of Quality of Life (AQoL) instrument: a psychometric measure of health related quality of life*. Quality of Life Research. Vol. 8, pp.209-224.

- Hawthorne G, Richardson J, Osborne R and McNeil H (1997) *The Assessment of Quality of Life (AQoL) instrument: construction, initial validation and utility scaling*. Centre for Health Program Evaluation. Melbourne.
- Herrman H, Hawthorne G and Thomas R (2002) *Quality of life assessment in people living with psychosis*. Social Psychiatry and Psychiatric Epidemiology. Vol. 37, No.11, pp.510-518.
- Hogan A, Hawthorne G, Kethel L, Giles E, White K, Stewart M, Plath B and Code C (2001) *Health-related quality-of-life outcomes from adult cochlear implantation: a cross-sectional study*. Cochlear Implants International. Vol. 2, No.2, pp.115-128.
- Holland R, Smith R, Harvey I, Swift L and Lenaghan E (2004) *Assessing quality of life in the elderly: a direct comparison of the EQ-5D and AQoL*. Health Economics. Vol. 13, No.8, pp.793-805.
- Hollingworth W, Mackenzie R, Todd CJ and Dixon AK (1995) *Measuring changes in quality of life following magnetic resonance imaging of the knee: SF-36, EuroQol or Rosser index?* Quality of Life Research. Vol. 4, No.4, pp.325-34.
- Hunt S, McKenna S, McEwen J, Williams J and Papp E (1981) *The Nottingham Health Profile: subjective health status and medical consultation*. Social Science and Medicine. Vol. 15A, pp.221-229.
- Ikeda S, Niwata S and Igarashi Y (2001) *Evaluating QoL of elderly people with dementia*. Japanese Journal of Pharmacoepidemiology. Vol. 5, pp.99-105.
- Ikeda S, Yamada Y and Ikegami N (2002) *Economic evaluation of donepezil treatment for Alzheimer's disease in Japan*. Dementia and Geriatric Cognitive Disorders. Vol. 13, No.1, pp.33-9.
- Jenkins C (1992) *Assessment of outcomes of health intervention*. Social Science and Medicine. Vol. 35, pp.367-375.
- Johnson JA, Luo N, Shaw JW, Kind P and Coons SJ (2005) *Valuations of EQ-5D health states: are the United States and United Kingdom different?* Medical Care. Vol. 43, No.3, pp.221-8.
- Jonsen AR (1986) *Bentham in a box: technology assessment and health care allocation*. Law, Medicine and Health Care. Vol. 14, No.3-4, pp.172-4.
- Jonsson L (2003) *Pharmacoeconomics of cholinesterase inhibitors in the treatment of Alzheimer's disease*. Pharmacoeconomics. Vol. 21, No.14, pp.1025-37.
- Jonsson L (2005) *Cost-effectiveness of memantine for moderate to severe Alzheimer's disease in Sweden*. American Journal of Geriatric Pharmacotherapy. Vol. 3, No.2, pp.77-86.
- Jonsson L, Andreassen N, Kilander L, Soininen H, Waldemar G, Nygaard H, Winblad B, Jonhagen ME, Hallikainen M and Wimo A (2006) *Patient- and proxy-reported utility in Alzheimer disease using the EuroQoL*. Alzheimer Disease and Associated Disorders. Vol. 20, No.1, pp.49-55.
- Kaplan R, Anderson J and Ganiats T (1993) *The Quality of Well-being Scale: rationale for a single quality of life index*. In Walker, S and Rosser, R (eds.) (1993). Quality of Life Assessment: Key Issues in the 1990s. Kluwer Academic Publishers, Dordrecht.
- Kaplan R, Bush J and Berry C (1976) *Health status: types of validity and the Index of Well-being*. Health Services Research. Vol. 11, No.4, pp.478-507.
- Kaplan R, Bush J and Berry C (1976) *The reliability, stability and generalizability of a health status index*. Proceedings of the Social Statistics Section. pp.704-709.
- Kaplan R, Ganiats T, Sieber W and Anderson J (1996a) *The Quality of Well-being Scale*. Medical Outcomes Trust Bulletin. Vol. 4, pp. 2-3.

- Kaplan RM (1993) *Quality of life assessment for cost/utility studies in cancer*. Cancer Treat Rev. Vol. 19, Suppl A, pp.85-96.
- Kaplan RM, Alcaraz JE, Anderson JP and Weisman M (1996b) *Quality-adjusted life years lost to arthritis: effects of gender, race, and social class*. Arthritis Care Res. Vol. 9, No.6, pp.473-82.
- Kaplan RM and Anderson JP (1988) *A general health policy model: update and applications*. Health Services Research. Vol. 23, No.2, pp.203-35.
- Kaplan RM, Anderson JP, Patterson TL, McCutchan JA, Weinrich JD, Heaton RK, Atkinson J H, Thal L, Chandler J, Grant I and HNRC Group (1995) *Validity of the Quality of Well-Being Scale for persons with human immunodeficiency virus infection*. Psychosomatic Medicine. Vol. 57, No.2, pp.138-47.
- Kaplan RM, Ganiats TG, Sieber WJ and Anderson JP (1998) *The Quality of Well-Being Scale: critical similarities and differences with SF-36*. International Journal for Quality in Health Care. Vol. 10, No.6, pp.509-20.
- Kerner DN, Patterson TL, Grant I and Kaplan RM (1998) *Validity of the Quality of Well-Being Scale for patients with Alzheimer's disease*. Journal of Ageing and Health. Vol. 10, No.1, pp.44-61.
- Kind P (1996) *The EuroQol Instrument: An index of health-related quality of life*. In Spilker, B (ed.) Quality of Life and Pharmacoeconomics in Clinical Trials. Lippincott-Raven Publishers, Philadelphia, pp.191-201.
- Kind P and Gudex C (1994) *Measuring health status in the community: a comparison of methods*. Journal of Epidemiology and Community Health. Vol. 48, pp.86-91.
- Kopec JA and Willison KD (2003) *A comparative review of four preference-weighted measures of health-related quality of life*. Journal of Clinical Epidemiology. Vol. 56, No.4, pp.317-25.
- Kring A, Kerr S, Smith D and Neale J (1993) *Flat affect in schizophrenia does not reflect diminished subjective experience of emotion*. Journal of Abnormal Psychology. Vol. 102, pp.507-517
- Magaziner J (1997) *Use of proxies to measure health and functional outcomes in effectiveness research in persons with Alzheimer disease and related disorders*. Alzheimer Disease and Associated Disorders. Vol. 11, Suppl.6, pp.168-74.
- Marra CA, Woolcott JC, Kopec JA, Shojania K, Offer R, Brazier JE, Esdaile JM and Anis AH (2005) *A comparison of generic, indirect utility measures (the HUI2, HUI3, SF-6D, and the EQ-5D) and disease-specific instruments (the RAQoL and the HAQ) in rheumatoid arthritis*. Social Science and Medicine. Vol. 60, No.7, pp.1571-82.
- Martin A, Glasziou P and Simes R (1996) *A Utility-Based Quality of Life Questionnaire for Cardiovascular Patients: Reliability and Validity of the UBQ-H(ear) Items*. NHMRC Clinical Trials Centre, University of Sydney, Sydney.
- Mathers C, Vos T and Stevenson C (1999) *The Burden of Disease and Injury in Australia*, Australian Institute of Health and Welfare, Canberra.
- McDowell I and Newell C (eds.) (1987) *Measuring health: a guide to rating scales and questionnaires*. Oxford University Press, New York.
- McKie J and Richardson J (2003) *The rule of rescue*. Social Science and Medicine. Vol. 56, No.12, pp.2407-19.
- McMahon PM, Araki SS, Neumann, PJ, Harris GJ and Gazelle GS (2000) *Cost-effectiveness of functional imaging tests in the diagnosis of Alzheimer disease*. Radiology. Vol. 217, No.1, pp.58-68.
- McNamee P, Bond J and Buck D (2001) *Costs of dementia in England and Wales in the 21st century*. British Journal of Psychiatry. Vol. 179, pp.261-6.

- 
- Meek PD, McKeithan K and Schumock GT (1998) *Economic considerations in Alzheimer's disease. Pharmacotherapy*. Vol. 18, No.2, Pt.2, pp.68-73, discussion pp.79-82.
- Mulkay M, Ashmore M and Pinch T (1987) *Measuring the quality of life: a sociological intervention concerning the application of economics to health care. Sociology*. Vol. 21, pp.541-564.
- MVHGroup (1995) *The Measurement and Valuation of Health: Final Modelling of Valuation Tariffs*, Centre for Health Economics, York.
- Naglie G, Tomlinson G, Tansey C, Irvine J, Ritvo P, Black SE, Freedman M, Silberfeld M and Krahn M (2006) *Utility-based quality of life measures in Alzheimer's disease. Quality of Life Research*. Vol. 15, No.4, pp.631-43.
- Neumann PJ, Hermann RC, Kuntz KM, Araki SS, Duff SB, Leon J, Berenbaum PA, Goldman PA, Williams LW and Weinstein MC (1999a) *Cost-effectiveness of donepezil in the treatment of mild or moderate Alzheimer's disease. Neurology*. Vol. 52, No.6, pp.1138-45.
- Neumann PJ, Kuntz KM, Leon J, Araki SS, Hermann RC, Hsu MA. and Weinstein MC (1999b) *Health utilities in Alzheimer's disease: a cross-sectional study of patients and caregivers. Medical Care*. Vol. 37, No.1, pp.27-32.
- Neumann PJ, Sandberg EA, Araki SS, Kuntz KM, Feeny D and Weinstein MC (2000). *A comparison of HUI2 and HUI3 utility scores in Alzheimer's disease. Medical Decision Making*. Vol. 20, No.4, pp.413-22.
- Nord E (1993) *Unjustified use of the Quality of Well-Being Scale in priority setting in Oregon Health Policy*. Vol. 24, No.1, pp.45-53.
- Nord E (1997) *A Review of Synthetic Health Indicators*, National Institute of Public Health for the OECD Directorate for Education, Employment, Labour and Social Affairs, Oslo.
- Nord E, Richardson J and Macarounas-Kirchmann K (1993) *Social evaluation of health care versus personal evaluation of health states. Evidence on the validity of four health-state scaling instruments using Norwegian and Australian surveys. International Journal of Technology Assessment in Health Care*. Vol. 9, No.4, pp.463-78.
- O'Connor R (1993) *Issues in the Measurement of Health-Related Quality of Life*, National Centre for Health Program Evaluation, Melbourne.
- OECD (2006) *OECD Health Data*. Organization for Economically Developed Countries, Geneva.
- Pickard AS, Johnson JA and Feeny DH (2005) *Responsiveness of generic health-related quality of life measures in stroke. Quality of Life Research*. Vol. 14, No.1, pp.207-19.
- Pickard AS, Johnson JA and Feeny DH, Shuaib A, Carriere KC and Nasser AM (2004) *Agreement between patient and proxy assessments of health-related quality of life after stroke using the EQ-5D and Health Utilities Index. Stroke*. Vol. 35, No.2, pp.607-12.
- Productivity Commission (2005a) *Economic Implications of an Ageing Australia*, Commonwealth of Australia, Canberra.
- Productivity Commission (2005b) *Impacts of Medical Technology in Australia*, Productivity Commission, Melbourne.
- Rabin R and de Charro F (2001) *EQ-5D: a measure of health status from the EuroQol Group. Annals of Medicine*. Vol. 33, pp.337-343.
- Rapley M and Antaki C (1996) *A conversation analysis of the 'acquiescence' of people with learning disabilities. Journal of Community and Applied Social Psychology*. Vol. 6, No.3, pp.207-227.
-

- Richardson C and Zumbo B (2000) *A statistical examination of the Health Utility Index-Mark III as a summary measure of health status for a general population survey*. Social Indicators Research. Vol. 51, No.2, pp.171-191.
- Richardson J (1994) *Cost utility analysis: what should be measured?* Social Science and Medicine. Vol. 39, No.1, pp.7-21.
- Richardson J and Hawthorne G (2000) *Negative Utilities and the Evaluation of Complex Health States: Issues Arising from the Scaling of a Multiattribute Utility Instrument*. Centre for Health Program Evaluation, Melbourne.
- Riedel-Heller SG, Busse A, Aurich C, Matschinger H and Angermeyer MC (2001) *Prevalence of dementia according to DSM-III-R and ICD-10: results of the Leipzig Longitudinal Study of the Aged (LEILA75+) Part 1*. British Journal of Psychiatry. Vol. 179, pp.250-4.
- Robinson A, Dolan P and Williams A (1997) *Valuing health states using VAS and TTO: what lies behind the numbers?* Social Science and Medicine. Vol. 45, No.8, pp.1289-1297.
- Robinson A, Loomes G and Jones-Lee M (2001) *Visual analog scales, standard gambles and relative risk aversion*. Medical Decision Making. Vol. 21, pp.17-27.
- Rosser R (1993) *A health index and output measure*. In Walker S and Rosser R (eds.) Quality of Life Assessment: Key Issues in the 1990s. Kluwer Academic Publishers, Dordrecht.
- Sackett D and Torrance G (1978) *The utility of different health states as perceived by the general public*. Journal of Chronic Diseases. Vol. 31, pp.697-704.
- Sainfort F, Becker M and Diamond R (1996) *Judgements of quality of life of individuals with severe mental disorders: patient self-report versus provider perspectives*. American Journal of Psychiatry. Vol. 153, pp.497-502.
- Scocco P, Fantoni G and Caon F (2005) *Role of depressive and cognitive status in self-reported evaluation of quality of life in older people: comparing proxy and physician perspectives*. Age Ageing. Vol. 35, No.2, pp.166-71.
- Shaw JW, Johnson JA and Coons SJ (2005) *US valuation of the EQ-5D health states: development and testing of the D1 valuation model*. Medical Care. Vol. 43, No.3, pp.203-20.
- Sigelman CK, Budd EC, Spanhel CL and Schoenrock CJ (1981a) *Asking questions of retarded persons: a comparison of yes-no and either-or formats*. Applied Research in Mental Retardation. Vol. 2, No.4, pp.347-57.
- Sigelman CK, Budd EC, Spanhel CL and Schoenrock CJ (1981b) *When in doubt, say yes: acquiescence in interviews with mentally retarded persons*. Mental Retardation. Vol. 19, No.2, pp.53-8.
- Sintonen H (1994) *The 15D measure of health-related quality of life: reliability, validity and sensitivity of its health state descriptive system*. National Centre for Health Program Evaluation, Melbourne.
- Sintonen H (2001) *The 15D instrument of health-related quality of life: properties and applications*. Annals of Medicine. Vol. 33, pp.328-336.
- Sintonen H and Pekurinen M (1993) *A fifteen-dimensional measure of health-related quality of life (15D) and its applications*. In Walker S and Rosser R (eds.) Quality of Life Assessment: Key Issues in the 1990s. Kluwer Academic Publishers, Dordrecht.
- Spitzer WO (1987) *State of science 1986: quality of life and functional status as target variables for research*. Journal of Chronic Diseases. Vol. 40, No.6, pp.465-71.
- Sprangers M and Aaronson N (1992) *The role of health care providers and significant others in evaluating the quality of life of patients with chronic disease: a review*. Journal of Clinical Epidemiology. Vol. 45, No.7, pp.743-760.

Stavem K (1999) *Reliability, validity and responsiveness of two multiattribute utility measures in patients with chronic obstructive pulmonary disease*. Quality of Life Research. Vol. 8, No.1-2, pp.45-54.

Stewart A, Phillips R and Dempsey G (1998) *Pharmacotherapy for people with Alzheimer's disease: a Markov-cycle evaluation of five years' therapy using donepezil*. International Journal of Geriatric Psychiatry. Vol. 13, No.7, pp.445-53.

Sturm JW, Donnan GA, Dewey HM, Macdonell RA, Gilligan AK, Srikanth V and Thrift AG (2004) *Quality of life after stroke: the North East Melbourne Stroke Incidence Study (NEMESIS)*. Stroke. Vol. 35, No.10, pp.2340-5.

Sturm JW, Osborne RH, Dewey HM, Donnan GA, Macdonell RA and Thrift AG (2002) *Brief comprehensive quality of life assessment after stroke: the Assessment of Quality of Life instrument in the north East Melbourne stroke incidence study (NEMESIS)*. Stroke. Vol. 33, No.12, pp.2888-94.

Tamim H, McCusker J and Dendukuri N (2002) *Proxy reporting of quality of life using the EQ-5D*. Medical Care. Vol. 40, No.12, pp.1186-95.

Torrance G (1986) *Measurement of health state utilities for economic appraisal: a review*. Journal of Health Economics. Vol. 5, pp.1-30.

Torrance G, Boyle M and Horwood S (1982) *Application of multi-attribute theory to measure social preferences for health states*. Operations Research. Vol. 30, pp.1043-1069.

Torrance GW, Feeny D and Furlong W (2001) *Visual analog scales: do they have a role in the measurement of preferences for health states?* Medical Decision Making. Vol. 21, No.4, pp.329-34.

Torrance GW, Furlong W, Feeny D and Boyle M (1995) *Multi-attribute preference functions*. Health Utilities Index. Pharmacoeconomics. Vol. 7, No.6, pp.503-20.

van Agt HM, Essink-Bot ML, Krabbe PF and Bonsel GJ (1994) *Test-retest reliability of health state valuations collected with the EuroQol questionnaire*. Social Science and Medicine. Vol. 39, No.11, pp.1537-1544.

Visser M, Fletcher A, Parr G, Simpson A and Bulpitt C (1994) *A comparison of three quality of life instruments in subjects with angina pectoris: the Sickness Impact Profile, the Nottingham Health Profile, and the Quality of Well Being Scale*. Journal of Clinical Epidemiology. Vol. 47, No.2, pp.157-163.

von Winterfeldt D and Edwards W (1986) *Decision Analysis and Behavioural Research*, Cambridge University Press, Cambridge.

Voruganti L, Heslegrave R, Awad A and Seeman M (1998) *Quality of life measurement in schizophrenia: reconciling the quest for subjectivity with the question of reliability*. Psychological Medicine. Vol. 28, pp.165-172.

WHOQoL Group (1993) *Study protocol for the World Health Organization project to develop a quality of life assessment instrument (WHOQOL)*. Quality of Life Research. Vol. 2, pp.153-159.

Wimo A, Mattson B, Krakau I, Eriksson T, Nelvig A and Karlsson G (1995) *Cost-utility analysis of group living in dementia care*. International Journal of Technology Assessment in Health Care. Vol. 11, No.1, pp.49-65.

Wlodarczyk JH, Brodaty H and Hawthorne G (2004) *The relationship between quality of life, Mini-Mental State Examination, and the Instrumental Activities of Daily Living in patients with Alzheimer's Disease*. Archives of Gerontology and Geriatrics. Vol. 39, No.1, pp.25-33.

Wood J, Taylor S and Lichtman R (1985) *Social comparison in adjustment to breast cancer*. Journal of Personality and Social Psychology. Vol. 49, pp.1169-1183.

---

Wu AW, Jacobson DL, Berzon RA, Revicki DA, van der Horst C, Fichtenbaum CJ, Saag MS, Lynn L, Hardy D and Feinberg J (1997) *The effect of mode of administration on medical outcomes study health ratings and EuroQol scores in AIDS*. Quality of Life Research. Vol. 6, No.1, pp.3-10.

---

## 8 Measures of Social Isolation and its Assessment in Older Adults

### 8.1 Background

There is a stereotype of the older adult with mild cognitive impairment or dementia living alone or in residential care and suffering significant perceived social isolation.<sup>28</sup> This stereotype can be traced back to the seminal work of Williamson, et al. (1964) and Townshend (1963, 1973) in the early to mid-1960s. These reports showed that older people living alone or in geriatric care, without children or other relatives, who were retired, infirm and often with mild dementia, had few social contacts. For those suffering dementia there may be additional social participation losses that cause perceived social isolation leading to considerable psychological distress.

At the global level there are two perspectives commonly described in the literature regarding social functioning; they are however the different ends of the same continuum. Table 31 presents a schematic representation of the different terms that are often used to describe these different perspectives.

---

28 The term 'perceived social isolation' is used throughout this report as a global term covering both social isolation and loneliness.

**Table 31 Definitions of the Social Functioning – Social Isolation Continuum**

<b>Definitions of the social functioning – social isolation continuum</b>			
	<b>Definition</b>	<b>Measurement</b>	
		<b>Objective</b>	<b>Subjective</b>
<i>Social functioning</i>	<i>Well-being associated with intimacy, family, friendships, social roles and institutional interactions</i>	<i>The number of social roles performed</i>	<i>Perception of well-being associated with intimacy, family, friendships, social roles and institutional interactions</i>
<i>Social participation</i>	<i>A willingness to participate and participation in social roles, activities and institutions</i>	<i>Number of social roles carried out, participated in and number of institutions a member of</i>	<i>Satisfaction with social roles</i>
<i>Social support/resources</i>	<i>The amount of social support available to an individual, including that given and received. This includes both instrumental and emotional support.</i>	<i>The frequency of social support activities</i>	<i>Satisfaction with social supports</i>
<i>Social contact/connectedness</i>	<i>The number of social contacts</i>	<i>The number of social contacts</i>	<i>Perception of satisfaction with social contacts</i>
<i>Social isolation (a)</i>	<i>The absence of social contacts, activities or participation</i>	<i>The number of social contacts</i>	<i>Perceived inadequacy of social contacts</i>
<i>Loneliness or emotional isolation</i>	<i>Feelings of being alone</i>		<i>Perceived depth of loneliness</i>
<i>Perceived social isolation</i>	<i>Living without human companionship, involving both social isolation and loneliness.</i>	<i>The number of social roles performed</i>	<i>Perceived lack of social contacts and perceived loneliness</i>

The positive end of the continuum is described in terms of social function, social participation, social support, social contacts and similar terms. Although there are differences between these terms, they all describe social networks, which have been defined as the number of social connections (i.e. those who are close, who are seen regularly and who can be relied upon for support) (Hobfoll and Walfisch, 1984; Retsinas and Garrity, 1985; Stokes and Wilson, 1984; Townsend, 1973). These social networks lead the individual to believe that he/she is cared for,

loved, esteemed or valued and that he/she belongs to a network of communication and mutual obligation (Cobb, 1976). They reflect the degree to which a person's basic needs are met through receiving instrumental aid (Procidano and Heller, 1983; Thoits, 1982). Social function is often assessed through a count of the number of social contacts or activities engaged in, and an assessment of the value of those contacts to the individual (Mendes de Leon, et al. 2003; Norbeck, et al. 1981; Sarason, et al. 1983), although there is evidence that it is the quality of these contacts that matters rather than the number (Kim, 1999b; Routasalo, et al. 2006; Victor, et al. 2000).

The negative end of the continuum is perceived social isolation. This occurs where there is a breakdown in the level of social arrangements regarded by an individual as necessary to meet his/her psychological needs (De Jong Gierveld, 1978; Marangoni and Ickes, 1989; Peplau and Perlman, 1982; Weiss, 1974); often referred to as the 'relational theory of loneliness' (van Baarsen, et al. 2001). It is described as living without companionship, social support, contact or connectedness, participation or social functioning (Tomaka, et al. 2006). It comprises two related constructs, social isolation and emotional loneliness (De Jong Gierveld and Havens, 2004; De Jong Gierveld and Tilburg, 2006; Levin, 2000; Routasalo, et al. 2006; Steptoe, et al. 2004; Tomaka, et al. 2006; Townsend, 1973; Weiss, 1973; Wenger and Burholt, 2004; Wilson, et al. 2007). As with social functioning, a variety of terms have been used to describe the constructs, such as loneliness, social loneliness and emotional isolation (Weiss, 1973).

Generally, social isolation refers to the absence of social contacts or activities. These are often assessed enumeratively through counting the number of social activities or contacts (many commentators refer to social isolation as being 'objective' for this reason). In contrast loneliness is usually defined as the emotional feelings of unmet social engagement need (often described as being 'subjective') (Townsend, 1963, 1973). The difference can be as subtle as that social loneliness can be assessed by an item such as "Do you often feel lonely" compared with "Do you experience loneliness" to indicate emotional loneliness (Holmen, et al. 2000). Both, however, are universally measured through self-report; hence they both reflect the assessments and perceptions of the respondent. To distinguish the global sense of social isolation from the enumerative sub-concept of social isolation the term 'perceived social isolation' has been adopted here.

Perceived social isolation – or the absence of social function – is associated with poorer health status and a higher consumption of health care resources (Ellaway, et al. 1999). The socially isolated have worse outcomes from acute interventions, such as cardiovascular surgery (Farmer, et al. 1996; Ruberman, et al. 1984; Williams, 1992). Those who are isolated experience compromised health-related quality of life (HRQoL), life meaning, levels of life satisfaction, wellbeing and community involvement (Cantor and Sanderson, 1999). In addition there are associations between social isolation and mental illness (particularly depression), distress, dementia, suicide and premature death (Berkman and Syme, 1979; Ellis and Hickie, 2001; Fratiglioni, et al. 2000; House, et al. 1982; Kawachi, et al. 1996; Lester and Yang, 1992; Rokach, 2000; Turner, 1981).

Three general theories have been advanced to explain these relationships. Attachment theories postulate that childhood experiences predispose adult social network behaviours (Bowlby, 1971; Fromm-Reichmann, 1959), that social networks affect responses to stressors (Cassel, 1976; Weiss, 1973) and that social support provides a 'buffer' against crises (Cobb, 1976; Peplau and Perlman, 1982). Collectively, these are consistent with the existential loneliness hypothesis; i.e. that people need to belong (Applebaum, 1978; Baumeister and Leary, 1995; Mayers and Svartberg, 2001). Because this is an internally regulated need, it can be argued that the assessment of a breakdown in social function (perceived social isolation) must reflect the perspective of the individual (this is because some individuals may choose solitude (i.e. to be alone), whereas others may lack the necessary skills to make or maintain social relationships) (Marangoni and Ickes, 1989; Sand and Strang, 2006). Where fulfilment of this need for belonging

is transgressed (the perceived discrepancy theory [De Jong Gierveld, 1978; Marangoni and Ickes, 1989]), challenging life events (e.g. relationship breakdown or partner loss, severe or life-threatening illness) may overwhelm an individual leading to the perception that he/she is both socially isolated and lonely. The resulting endogenous stressors associated with this overarching perceived social isolation may involve reciprocal causation; that is be caused by, associated with or exacerbate health symptoms, conditions or poor health care outcomes.

Given the extensive list of correlates or consequences, it would seem that the assessment of social function or perceived social isolation among those with mild cognitive impairment or dementia is important. Where it occurs, there may be justification for intervention with the specific aim of increasing social participation, and where programs aimed at alleviation are implemented it is important that these are evaluated using valid measures.

This review examines measures which may be useful for providing such assessments.

## **8.2 Method and Review Criteria**

To identify published stand-alone instruments assessing social function or perceived social isolation suitable for use with those suffering mild cognitive impairment or dementia in clinical, epidemiological and research situations in Australia a search of MEDLINE, CINAHL and PsychLIT was undertaken using the terms *friendship, loneliness, relationships, social network, friendship activity, social connectedness, social isolation, social support, social participation and community involvement*, crossed with the keywords *dementia, Alzheimer's disease and mild cognitive impairment*. The results were, in turn, crossed with *instrument, questionnaire, measure, measurement and scale*. Four hundred and eighty articles were identified in MEDLINE, 90 in CINAHL, and 2013 in PsychLIT. All titles and abstracts were searched to identify instruments, where the inclusion criteria were evidence of instrument development or reports of instrument psychometric properties. Where papers reported using a measure and its psychometric properties, the bibliography was scanned to identify the original source paper.

Fifteen instruments or scales were identified. Of these, four were scales within other instruments and so have been excluded from review because they are not stand-alone measures. These were the social relationships scale from the Assessment of Quality of Life (AQoL) utility measure (Hawthorne, et al. 1999), the social isolation scale of the Nottingham Health Profile (Hunt, et al. 1981), the social resources scale from the Older Americans Resources and Services Multi-dimensional Functional Assessment Questionnaire (OARS-MFAQ) (Fillenbaum and Smyer, 1981), and the social relationships scale from the WHOQOL-Bréf (WHOQoL Group, 1998).

Four other instruments were rejected following perusal of the articles. Procidano and Heller's (1983) Perceived Social Support from Friends and Family measure was rejected because of its length (40 items). It comprises two sub-scales, measuring social support from friends and from family, respectively. Holmen, et al. (2000) used single items to assess social loneliness ("Do you often feel lonely") and emotional loneliness ("Do you experience loneliness"). No psychometric properties were reported and it is doubtful if these two items actually form a scale.

Kristjansson, et al. (2001) developed a 6-item scale, the Indicator of Support for Community-Residing Older Canadians. It was designed to measure a lack of social support, which was defined as social isolation. The items cover the number of people lived with, the number of people available to provide help, the relationship with the main supporter, perceived closeness to this main supporter, the number of people who would help if the respondent was ill, and the time for help to arrive if the respondent was injured. The content of the items raises issues around the meaning of the scale.

Wenger (1983) developed a short scale (8-items) measuring the inadequacy of social contacts, which was later described as loneliness (Wenger and Burholt, 2004). Item content covers feeling

lonely, seeing enough of friends/relatives, meeting people, having a confidant, wishing for more friends, having real friends and spending Christmas alone. Analyses of the scale were confined to cross tabulations with socio-demographic characteristics. A second scale indicating social isolation covered living alone, having no close relatives, never visiting, having no contact with neighbours, no telephone, being alone for more than 9 hours a day, nearest neighbour more than 50 yards away and never leaving the house (Wenger and Burholt, 2004). As admitted by the researchers, these were more a collection of items than psychometric scales.

The remaining 7 scales are reviewed against the study criteria outlined below. They are (in alphabetical order): the DUKE Functional Social Support Questionnaire (Broadhead, et al. 1988), the Friendship Scale (Hawthorne, 2006), the Loneliness Scale (De Jong Gierveld and Kamphuis, 1985; De Jong Gierveld and Tilburg, 2006), the Medical Outcomes Study Social Support Survey (Sherbourne and Stewart, 1991), the Norbeck Social Support Questionnaire (Norbeck, 1984; Norbeck, et al. 1981; Norbeck, et al. 1983), the Social Support Questionnaire (Sarason, et al. 1987; Sarason, et al. 1983), the UCLA Loneliness Scale (Russell, et al. 1980, 1978; Russell, 1996) and its short derivative the Three-item Loneliness Scale (Hughes, et al. 2004).

### 8.2.1 The Review Criteria

The review criteria are those outlined in Section 2 of this report. Each criteria was weighted for its applicability to the Australian setting (refer Table 1, Executive Summary).

Although these criteria are used to rate each instrument, for ease of understanding the instrument review material has been organised to reflect basic psychometric axioms. Psychometric theory postulates that the valid and reliable measurement of a latent construct requires the construction of a manifest instrument that delivers an observed model which is isomorphic with the construct. To achieve this, the following axioms are widely accepted:

1. There should be a latent model of the construct, including an adequate description of its dimensions. For each dimension, there should be measurement items, such that the item content covers the dimension adequately. All items combined form the descriptive system of an instrument from which the manifest model is derived;
2. The resulting instrument should possess a nomological net of evidence suggesting validity (Cronbach and Meehl, 1955);
3. It should also be reliable and responsive; and
4. Instruments to be used with respondents suffering cognitive impairment, as is the case with dementia, should be short and simple to minimise response burden.

Where there is a nomological net of evidence relating to each of these criteria, it may be inferred that an instrument is valid and reliable. Since validity and reliability are functions of both the instrument itself and the respondents who complete it, these are never fixed properties but may vary from sample to sample. The important corollary is that although there may be validity and/or reliability evidence for an instrument developed in, say, the USA, that same instrument may be invalid and/or unreliable in Australia due to cultural differences. It is accepted among psychometricians that this implies basic tests of validity and reliability need to be applied each time an instrument is used with a different population.

## 8.3 Review of the Instruments

### 8.3.1 DUKE-UNC Functional Social Support Questionnaire

The US DUKE-UNC Functional Social Support Questionnaire was developed to provide a brief assessment of functional social supports of patients in a primary care setting (Broadhead, et al. 1988).

Developed among patients attending a university primary care clinic, the instrument comprises two subscales, Affective Support (3 items) and Confidant Support (5 items). The timeframe is the present, so this is a 'state' social support scale. The response categories are Guttman-type, with 5 categories of which only the endpoints are labelled (as much as I would like (5)/.../..../.../ much less than I would like (1)). Scoring is by simple summation for each of the scales. A high score indicates social support (Broadhead, 1988; Broadhead, et al. 1988).

#### Evidence of a latent construct

The origins of the DUKE-UNC Functional Social Support Questionnaire can be found in a review of the epidemiologic evidence linking social support and health, essentially framed within the buffering hypothesis (Broadhead, et al. 1983). According to Broadhead, et al. (1988), based on the review, four areas of support were determined a priori, being the quantity of support, confidant support, affective support and instrumental support. An item pool was developed with either 3 or 4 items representing each area. Fourteen items were then administered to patients (n = 401) attending a primary care clinic, where patients were randomly time-frame chosen. Following test-retest at 13 days for 22 of the patients, 3 items were eliminated. The average test-retest correlation for the remaining 11 items was  $r = 0.66$ . Factor analysis was used to examine the structure of these 11 items, and 3 more items were removed. The remaining 8 items loaded on 2 factors which were labelled Confidant Support and Affective Support.

#### Validity evidence

##### *Content validity*

Content validity evidence is poor. The construction of the DUKE-UNC Functional Social Support Questionnaire seems to have been entirely pragmatic. Although the buffering hypothesis is briefly referred to, there seems to be no connection between this, the literature review upon which the items were based, and the final set of items. No substantive connection is mentioned in the seminal paper (Broadhead, et al. 1988).

The Confidant Support scale has five items (chance to talk to someone about problems at work or housework/ chances to talk to someone about personal and family matters/ chance to talk about money matters/ invitations to go out with other people/ advice about important things in life) and the Affective Support scale three (people care about me/ love and affection/ help when sick in bed).

Review of the 8 items against the four areas of support identified from the literature (the quantity of support, confidant support, affective support and instrumental support) reveals that two areas are not measured by the DUKE-UNC Functional Social Support Questionnaire (quantity of support and instrumental support). Regarding this matter, Broadhead, et al. (1988, p718) commented that "the content of the overall dimensions may be inadequately sampled".

Furthermore, in each of the scales there are 'odd' items. In the Confidant Support scale the item 'invitations to go out and do things with other people' appears to have little in common with the other scale items. Broadhead, et al. (1988) commented that this may be due to these invitations coming from sources of confidant support, or that it was a statistical artefact arising from the fact that the greater the number of social contacts the greater the opportunity to talk about things (which is at the core of this scale). The odd item in the Affective Support scale is the item receiving

'help when sick in bed'; Broadhead, et al. explained this by hypothesizing that this kind of help may be interpreted as affective in quality.

#### *Construct validity*

Construct validity was claimed on the grounds that most of the 8 items were significantly correlated with the dimensions of general health measured by the DUKE-UNC Health Profile, of which it was asserted that "...all of which measure known correlates of social support" (Broadhead, et al. 1988, p718). The authors postulated that there were four aspects of social support, and developed the small item bank described above for testing. They then eroded this item bank by removal of 3 items based on poor test-retest reliability among 22 respondents. It is not surprising that the 11 items used in the factor analysis failed to replicate their hypothesized model. The factor model described above (that the 8 items loaded on 2 factors), however, does suggest a degree of robustness for the limited scales, Confidant and Affective.

#### *Criterion validity*

Criterion validity was assessed by observing the relationships between the DUKE-UNC Functional Social Support Questionnaire and various demographic characteristics (gender, race, relationship status, living situation, employment status, age, education level, socioeconomic status). The only variables on which scores significantly varied were race (black/white) and living situation (alone/with someone). Scores did, however systematically vary by social contacts, group participation, social function and socializing with other people (Broadhead, et al. 1988) – situations that all involve talking – although the correlations were very modest (e.g. for Confidant Support they ranged from 0.08 to 0.35, and for Affective Support from 0.08 to 0.22). As well as these correlations showing modest relationships at best, the correlations between three of these variables (social contacts, social function and socializing with other people) were all higher for the Confidant Support scale than for the Affective Support scale. This is a counterintuitive finding which provides little confidence regarding what the DUKE-UNC Functional Social Support Questionnaire scales actually measure.

#### **Reliability**

The reliability coefficient was 0.88 among HIV-infected women (Bova, 2001). The Cronbach  $\alpha$  for the Confidant Support scale was reported to be 0.79, and for the Affective Support scale 0.70 among Swiss university students. Reliability was assessed by item-rest of instrument correlations, which showed that the range was 0.52 to 0.72 (Broadhead, et al. 1988).

#### **Responsiveness**

Responsiveness was shown in a study of medical care utilization. Confidant Support scale scores systematically varied by the number of clinician visits, the length of these visits, and employment status (Broadhead, et al. 1989). The instrument has also been shown to be responsive to cancer therapy wellbeing (Nelson, et al. 2002), family cohesion (Williams, et al. 1990), mental health in HIV (Ruiz Perez, et al. 2005), physical health (Bovier, et al. 2004; Ruiz Perez, et al. 2005), postpartum depression (Watt, et al. 2002), and sexuality in climeractic women (Blumel, et al. 2004).

Studies where the DUKE-UNC Functional Social Support Questionnaire has failed to discriminate or be responsive to condition include depression in climeractic women (Blumel, et al. 2004), terminally ill cancer patients, HIV/AIDS (Bova, 2001; Nelson, et al. 2002), living location for those with newly diagnosed cancer (Howat, et al. 2006), mental health state after adjustment for internal resources and stress (Bovier, et al. 2004), social worker support for postnatal women (Morrell, et al. 2000), and voluntary support for those with psychosocial problems (Grant, et al. 2000).

#### **Assessment against the study criteria**

Regarding comparative data, no studies were identified comparing the DUKE -UNC Functional Social Support Questionnaire with another social support or isolation instrument. At 8 items the DUKE-UNC Functional Social Support Questionnaire is a short scale and it would be feasible to

include it in an instrument battery. The cognitive burden of administration or completion is likely to be light, i.e. it is an easy instrument to use, although it may be difficult to use in an interview situation or with a translator because the response scales do not have fully labelled anchorpoints (only the endpoints are labelled). Scoring is through simple summation. No studies were identified where the instrument had been used in a dementia sample.

As reviewed above, some reliability evidence has been published showing that reliability seems to be a function of the sample since some of the estimates fall below the accepted reliability standards (Nunnally and Bernstein, 1994), suggesting at best moderate reliability.

The evidence for the validity of the scale is extremely mixed; overall it would appear to suggest that the scale probably has poor validity evidence.

No costs were identified for using the DUKE-UNC Functional Social Support Questionnaire and the administration and scoring costs are likely to be low as it has been designed for self-completion.

### **8.3.2 Friendship Scale**

The Australian Friendship Scale was published in 2006 (Hawthorne, 2006), following its development in population-based samples of older adults. As such it is the most recent of the instruments reviewed: there is just the one article in the literature on it. Following a literature review of social isolation measures, it was developed to be a short, user-friendly measure of perceived social isolation; it has just 6 items.

The Guttman-type response categories are Almost always/ Most of the time/ About half the time/ Occasionally/ Not at all. The timeframe is the past four weeks, suggesting the Friendship Scale is between a state and trait scale. Half the items are negative to prevent response bias.

Scoring of the Friendship Scale is through reversing the negative items and then summing item scores. Based on response criteria (i.e. the meaning of different response categories) logical cut points were suggested classifying respondents into those who were socially isolated, isolated, with a low level of social support, socially connected, and very socially connected (Hawthorne, 2006).

#### **Evidence of a latent construct**

The descriptive system was based on transgression theories which postulate social behaviour is a function of childhood, that social support is a buffer against life's vicissitudes and that the social milieu affects responses to stress. Seven dimensions were identified from the literature: an absence of intimacy, an inability to relate to others, being unable to ask for support, having no social networks, being separate from others and unable to fulfil social roles, being isolated and feeling alone. Items were developed measuring each dimension and refined through consultation with older adults and academic colleagues.

#### **Validity evidence**

##### *Content validity*

This 6-item instrument covers self-report of both social isolation (ease of relating to others, having someone to share feelings with, finding it easy to get in touch with others) and loneliness (feeling isolated, feeling separate from other people, and being alone and friendless).

Content validity appears to be quite good; there is a good match between the descriptive system and the latent model. There is some evidence of ecological validity, since older adults were involved during item construction. Four samples of older adults were recruited, representing the sick and elderly living in nursing homes, hospital outpatients, veterans and their spouses, and a random population sample. The response rate overall was 63% (n = 829). Against this, however,

the original item pool consisted of just 7 items – one for each of the dimensions (one item was removed following principal components analysis of the item pool).

#### *Construct validity*

There is also evidence of construct validity. Based on the construction sample exploratory factor analysis was used with a random half of the sample to construct the descriptive system, and, with the second sample half, structural equation modelling used to confirm this. The fit statistics were CFI = 0.99 and the RMSEA = 0.02.

An important finding was that the three positively worded items were statistically related to each other. Hawthorne noted that there were two possible explanations for this: it could be due to these items measuring social isolation whereas the others measured loneliness in Weiss's schema (Weiss, 1973), or be a measurement artefact relating to positive and negative items.

#### *Criterion validity*

Criterion validity was assessed by comparing the Friendship Scale against the social relationship scales embedded in the AQoL and WHOQOL- Bréf instruments; the correlations were 0.61 and 0.44, respectively.

#### **Reliability**

Because of the skewness of items, the non-parametric Mokken  $\rho$  (rho) was used to assess reliability (0.81).

#### **Responsiveness**

Scores were found to systematically vary by known correlates of social isolation (accommodation, work, participation in social activities, health, relationship status and depression).

When examined against participants with mild cognitive impairment (MMSE scores <24), it was reported that there were no response bias issues.

#### **Assessment against the study criteria**

Regarding comparative data, the Friendship Scale was moderately correlated with the social relationships scales from the AQoL and the WHOQOL-Bréf (Hawthorne, 2006). No other studies were identified comparing the Friendship Scale with another social support or isolation instrument. At 6 items it is a short scale and it would be easy to include it in an instrument battery. The cognitive burden of administration or completion is likely to be light, i.e. it is an easy instrument for self-completion, interview administration or administration through a translator. No particular difficulties were reported when the scale was examined with those with mild cognitive impairment.

The only published study reported satisfactory reliability. Based on a single study, the evidence suggests it may have good validity, although clearly further evidence is needed.

No costs were identified for using the Friendship Scale, and the administration and scoring costs are likely to be low as it has been designed for self-completion.

### **8.3.3 De Jong Gierveld Loneliness Scale <sup>29</sup>**

There are two versions of the Dutch De Jong Gierveld Loneliness Scale: the original 11 item version (De Jong Gierveld and Kamphuis, 1985) and a recently released 6 item version (De Jong Gierveld and Tilburg, 2006). The original 11 item unidimensional scale was developed in response

---

<sup>29</sup> The review of the De Jong Geirveld Loneliness Scale is necessarily circumscribed because many of the papers describing it and its use are in Dutch. These papers were inaccessible to the reviewer.

to the need for a short, valid measure of loneliness,<sup>30</sup> whereas the 6 item version was developed in view of the need for a short scale suitable for use in large surveys. It possesses the same structure as the full 11 item version (De Jong Gierveld and Tilburg, 2006).

The structure of the Loneliness Scale is that the scale measures, overall, general loneliness (De Jong Gierveld and Kamphuis, 1985; De Jong Gierveld and Tilburg, 2006). There are two sub-scales measuring emotional loneliness (all negative items; 6 items) and social loneliness (all positive items, 5 items). For the 6 item version, there are 3 items on each of the sub-scales (De Jong Gierveld and Tilburg, 2006).

The timeframe for responders is "...the way you feel now". The implication of the present timeframe is that the scale was conceived as a 'state' loneliness scale.

Scoring the De Jong Gierveld Loneliness Scale is recommended through reversing positive items, dichotomizing the item responses (yes!/ yes/ more or less/ no/ no!) using the category 'no' on the ground that the interrogation point (the point of indifference; 'more or less') was not neutral – essentially it was considered a positive response. Thus yes!, yes, more or less/ no, no!, and then summing the resultant values for each item.

Where the scale is telephone administered or used with older adults the recommended response scale is yes/ more or less/ no, with dichotomization giving yes, more or less/ no (Dykstra, et al. 2005). It is reported that the correlation between the 5-point response items (range 11-55) and dichotomized items (range 0-11) was  $r = 0.87$  (De Jong Gierveld and Tilburg, 1999). An SPSS computer algorithm for scoring the scale is included in the user manual. Very basic normative data are available (De Jong Gierveld and Tilburg, 1999).

### **Evidence of a latent construct**

The Loneliness Scale had a long gestation period and the forerunner was developed in 1971. It consisted of 6 items measuring solving your own problems, true friends being hard to come by, feeling lonely, difficulty with developing lasting relationships, being on your own, and no one seeming to care. The items were developed from interviews with both lonely and non-lonely people, and from written compositions of their experiences of loneliness. The response categories were a Likert scale from strongly agree to strongly disagree (De Jong Gierveld, 1971). From these beginnings, a theoretical model of loneliness was progressively shaped during the 1970s and 1980s (De Jong Gierveld, 1984, 1987, 1989; De Jong Gierveld and Raadschelders, 1982). At the base of this model lay the concept that "The essence of the loneliness concept we used is found in the phrase: an experiencing of a lack of desired relationships as disagreeable or unacceptable" (De Jong Gierveld, 1978, p222).

Following factor analysis of data collected during interviews with Amsterdammers, three key components of loneliness were identified: deprivation type, time perspective and emotional characteristics (De Jong Gierveld and Kamphuis, 1985; De Jong Gierveld and Raadschelders, 1982). It was the deprivation type that was considered the essence of loneliness (De Jong Gierveld and Kamphuis, 1985). Three aspects of deprivation loneliness were identified: the absence of intimate attachment, feelings of emptiness and feelings of abandonment. An item bank was administered to a stratified sample of Dutch people ( $n=566$ ) and cluster and discriminant analyses used to identify 'loneliness profiles'. Three groups of lonely people were identified: the 'hopeless' lonely who were dissatisfied with their relationships, the 'periodic' lonely who believed

---

30 In addition de Jong Gierveld and Raadschelders developed a 9 item scale assessing deprivation loneliness. Comparison of the items with the Loneliness Scale shows obvious similarities, yet also some striking differences. It seems the deprivation scale was part of the development in her thinking rather than a fully developed scale in its own right. See de Jong Gierveld J and Raadschelders J (1982) *Types of loneliness*. In Peplau LA, Perlman D, (eds). *Loneliness: A Sourcebook of Current Theory, Research and Therapy*. New York: Wiley. p 105-119.

their loneliness was temporary and who were socially active, and the 'resigned' lonely who felt severely lonely and who had no hope of escaping their circumstances (De Jong Gierveld and Raadschelders, 1982).

Based on this work, a model of loneliness, also described by De Jong Gierveld, (1987) as 'subjective isolation', was theorized, consisting of four components: (a) predisposing situation factors (e.g. sociodemographics such as age and gender, living arrangements such as partner status, and personality characteristics such as concept of self); (b) social network characteristics (e.g. the number of close relationships and contacts with friends, family, neighbours or colleagues); (c) the subjective evaluation of the social network (e.g. dissatisfaction with relationships, desire for new relationships); and (d) the intensity of loneliness as measured by the loneliness scale. Loneliness was considered a subjective personal experience which was not directly related to situational factors (De Jong Gierveld, 1987).

## **Validity evidence**

### *Content validity*

The design parameters for developing a scale from this theoretical base were to adopt the loneliness perspective, to include items that covered the range of feelings (from less intense to intense), that fitted the Rasch model<sup>31</sup>, to include both negative and positive items (5 of each type), that would be easy to use in survey research and that would be suitable for both the lonely and non-lonely to complete (De Jong Gierveld and Kamphuis, 1985). Forty new items were constructed from open-ended interviews with lonely people. These items were then sorted into equal-distance spacing along the loneliness continuum by academics. At completion 28 items were administered to 1,201 disabled, unemployed and employed Dutch residents. To make the items suitable for Rasch modelling, the response categories which were 5-point Likert scales (yes!/ yes/ more or less/ no/ no!) were dichotomized using the category 'no' on the ground that the interrogation point (the point of indifference; 'more or less') was not neutral (it was considered a positive response). Rasch modeling resulted in the construction of an 11-item unidimensional loneliness scale (De Jong Gierveld and Kamphuis, 1985).

The 6 item version was developed from the full instrument. The design criteria were to preserve the two subscales (emotional and social loneliness), to have balanced sub-scales consisting of 3 items each. Item selection was based on close correlation with the appropriate subscale, coverage of the broad range of IRT item difficulties, and optimum item wording. Factor analysis was used to complement these criteria (De Jong Gierveld and Tilburg, 2006).

That the timeframe is 'now', it is likely the scale is a measure of 'state' loneliness.

The content validity of the De Jong Gierveld Loneliness Scale appears to be excellent. There is evidence that the items were based on a sound theoretical model of social isolation, based on the experiences of the isolated (De Jong Gierveld and Kamphuis, 1985). Importantly, the construction and validation samples were population-based samples from the Dutch community, stratified by loneliness level. This suggests that the content of the scale is probably reflective of the concerns of the lonely and socially isolated. That the items were empirically drawn from a larger pool of items using modern test theory (Rasch modelling) along with logical criteria to ensure fidelity and coverage of the theoretical model is strong evidence of content validity (De Jong Gierveld and Kamphuis, 1985).

The evidence for the 6 item version is probably just as strong since the construction sample was based on a stratified sample of older adults from three regions in the Netherlands (n = 3,987, response rate 62%). Care was taken during construction to maintain fidelity to the structure of the original scale through selection of items that met both logical and psychometric criteria (De Jong

---

31 Rasch models are item response theory models that require the use of dichotomous items.

Gierveld and Tilburg, 2006). The manifest instrument was then tested in a population sample using a mail survey. The response rate was 72%.

#### *Construct validity.*

Construct evidence for the Loneliness Scale is mixed. Based on the results of the original Rasch modelling, it was concluded that there were "... no theoretical grounds for bidimensionality" and that "There is sufficient evidence to treat the scale in practice as a Rasch scale" (De Jong Gierveld and Kamphuis, 1985, p295 and 297); i.e. the Loneliness Scale was unidimensional at both the theoretical level and the manifest level.

Yet at the same time it was also observed that under factor analysis the items loaded on positive and negative factors. This finding was explained by De Jong Gierveld and Kamphuis as a methodological artefact 'response set' problem such that the obtained Rasch model fit – which assumes unidimensionality – could not be claimed as proof of unidimensionality for the scale; i.e. the homogeneity of the scale was not very strong (De Jong Gierveld and Kamphuis, 1985; De Jong Gierveld and Tilburg, 1999). The evidence for this was that the Loevinger H was in the range 0.30 to 0.50 for different samples (De Jong Gierveld and Tilburg, 1999). This point is important, because in the user manual (De Jong Gierveld and Tilburg, 1999), the 1999 paper comparing the living arrangements of older adults in Italy and the Netherlands (De Jong Gierveld and van Tilburg, 1999), the 2004 paper examining gender and relationship status in older adults (Dykstra and De Jong Gierveld, 2004), and in her 2006 paper describing the development of the 6-item Loneliness Scale version, De Jong Gierveld abandoned the unidimensional claim for her scale and labelled these two factors the Emotional (negative items) and Social (positive items) sub-scales after Weiss. She stated that these could be used separately to assess emotional loneliness and social loneliness, respectively – perhaps this shift in position was not surprising given that the correlation between the two subscales was between 0.37 and 0.43 for different national samples (De Jong Gierveld and Tilburg, 1999). In this way, the model of the Loneliness Scale was made to fit with Weiss's 1973 typology and terminology (Weiss, 1973). The scale was described as measuring "Overall, emotional and social loneliness..." (De Jong Gierveld and Tilburg, 2006, p582).

Other researchers have also offered limited support for the scale's unidimensionality, and it is possible that these empirical reports influenced the 2006 paper. Moorer and Suurmeijer (1993), in a sample of older adults (n = 723) examined unidimensionality using Mokken scale analysis for polytomous items. They reported that 10 of the 11 items formed a unidimensional scale, although even these 10 items formed a weak scale (Loevinger H = 0.37).

A more searching examination was carried out by van Baarsen, et al. (1999) using Mokken and factor analyses. The results suggested the presence of two subscales concerned with emotional and social loneliness, respectively. The Loevinger H for the two subscales was 0.48 and 0.43, respectively. The factor analysis also confirmed this finding. The correlation between the two subscales was between 0.41 and 0.65. It was concluded that although two factors had been identified because each consisted of either positive or negative items only, this internal structure "may be an artefact" (van Baarsen, et al. 2001, p451). Despite this cautionary warning, they presented the results of their study under the two factors, which they labelled Emotional Loneliness and Social Loneliness. Given this uncertainty, van Baarsen, et al. (2001) carried out a second study investigating the structure of loneliness among older adults (n = 4,494, 62% response rate) using Rasch modelling. The scale was tested under both one and two dimensional models. The findings showed that the one dimension model did not fit the data whereas the two dimension model did. The correlation between the two scales was  $r = 0.55$ . Examination of the two resulting subscales suggested that the item parameters were not equal for males and females, i.e. gender differential item functioning may be present.

#### *Criterion validity*

Criterion validation evidence for the De Jong Gierveld Loneliness Scale has been published in Dutch, but this evidence was not accessible to the reviewer. In the user manual this was reported

as being an area where there is insufficient data due to the absence of research (De Jong Gierveld and Tilburg, 1999).

### **Reliability**

As reviewed above, tests of homogeneity suggest that the Loneliness Scale is a weak scale under the axioms of Mokken analysis, and consistent with this the reported range of Loevinger H values fall between 0.30 to 0.50 (De Jong Gierveld and Tilburg, 1999; Dykstra, et al. 2005). It could be expected that examination of each of the subscales would reveal that they were stronger scales. In a comparison of older adults living in Italy and the Netherlands, it was reported that for the Emotional Loneliness scale the Loevinger H was 0.40 for the Dutch and 0.34 for the Italian samples. Similarly, for the Social Loneliness scale the values were 0.48 and 0.39, respectively (De Jong Gierveld and van Tilburg, 1999). These reports may suggest the scale is culturally-specific.

Across different samples the reliability (whether Cronbach's  $\alpha$  or Mokken's  $\rho$ ) falls within the range 0.80 to 0.90 (De Jong Gierveld and Tilburg, 1999; Dykstra, et al. 2005). For example, Moorer and Suurmeijer (Moorer and Suurmeijer, 1993) reported reliability to be  $\rho = 0.88$  in a sample of older adults. Reliability for the two subscales (emotional and social loneliness) was reported to be  $\rho = 0.84$  and  $0.77$ , respectively (van Baarsen, et al. 1999). The Cronbach  $\alpha$  for the two subscales has been reported to be between 0.71-0.85 and 0.80-0.84, respectively (van Baarsen, et al. 1999).

### **Responsiveness**

Regarding responsiveness, the Loneliness Scale was sensitive to differences among older adults' increasing age (Dykstra, et al. 2005), depression (Jongenelis, et al. 2004); gender (Dykstra and De Jong Gierveld, 2004), household composition (De Jong Gierveld and Tilburg, 1999; Havens, et al. 2004), physical condition and number of chronic illnesses (Dykstra, et al. 2005; Havens, et al. 2004; van Baarsen, et al. 1999), relationship status (De Jong Gierveld and Tilburg, 1999; Dykstra, et al. 2005; Havens, et al. 2004; Sadler, et al. 2006; Stevens and Westerhof, 2006; van Baarsen, et al. 2001; van Tilburg, et al. 2004), self esteem or life satisfaction (Havens, et al. 2004; van Baarsen, et al. 2001), self reported loneliness (van Baarsen, et al. 2001), social anxiety (van Baarsen, et al. 1999), social network size (De Jong Gierveld and Tilburg 1999; Dykstra, et al. 2005) and social participation/support (Dykstra and De Jong Gierveld, 2004; van Baarsen, et al. 1999; van Baarsen, et al. 2001). National differences have also been reported among older adults between Canada, Italy and the Netherlands (De Jong Gierveld and Tilburg, 1999; van Tilburg, et al. 2004).

However, the scale was not sensitive to visual impairment in adolescents (Kef, 2002).

Responsiveness over time among older adults ( $n=139$ ), who between administrations lost their partner by death, was reported by van Baarson, et al. (1999). Significant increases in the global De Jong Loneliness Scale score were reported ( $t = -5.72$ ,  $p<0.01$ ) as were increases on the Emotional Loneliness subscale ( $t=-8.57$ ,  $p<0.01$ ), but not on the Social Loneliness subscale. Kremers, et al. (2006) reported that scores over time significantly changed for both the treatment and control cohorts in a study of the effectiveness of self-management for older women, whereas in a study of a friendship enrichment program there was a significant decline in scores (Martina and Stevens, 2006).

When different administration methods have been compared, the user manual reports results from four different studies, each comparing some combination of self-completion, interview and telephone administration. It was concluded that administration mode affected scores consistent with previous theory that in the privacy of self-administration respondents are less likely to present themselves in a favourable light (De Jong Gierveld and Tilburg, 1999).

### **Assessment against the study criteria**

Regarding comparative data, no studies were identified comparing the De Jong Gierveld Loneliness Scale with another social support or isolation instrument. At 11 items for the full scale or 6 items for the recently published scale it is a short scale which would be reasonably easy to

include in an instrument battery. The cognitive burden of administration or completion is likely to be light, although there may be some difficulties because of the response scales (Yes!/ Yes / More or less/ No/ No!); different versions of these response scales have been used for face-to-face, telephone and mail completion raising questions of equivalence. These issues have not been sufficiently explored, leading to caution where the instrument is used with a translator or over the telephone.

No studies were identified where the instrument had been used in a dementia sample. There is considerable reliability and validity evidence, although a caveat is that the instrument may be both culturally-bound and subject to administration mode effects.

No costs were identified for using the De Jong Gierveld Loneliness Scale, and the administration and scoring costs are likely to be low as it has been designed for self-completion.

### **8.3.4 Medical Outcomes Study Social Support Survey**

The US MOS Social Support Survey was developed for the Medical Outcomes Study, a 2-year longitudinal study of the process and outcomes of care for patients with chronic health conditions (Sherbourne and Stewart, 1991). The decision to develop this scale was because the researchers failed to identify a short, valid and sensitive measure of social support.

The perspective of the instrument is that of perceived availability of functional support; it consists of 20 items. The first item asks about the number of close relatives and friends the respondent has. The other 19 items ask the respondent to rate the frequency with which contact is made with others. The Guttman response categories are none of the time/ a little of the time/ some of the time/ most of the time/ all of the time. The 19 items cover helping if the respondent is confined to bed, a person who will listen, someone who can give good advice in a crisis, a person who will transport the respondent to the doctor, a person who shows the respondent love, a person to have a good time with, someone who will give information to understand a situation, a person to confide in, a person who hugs the respondent, a person to relax with, a person who will provide good advice, a person to help get the respondent's mind off things, a person to help with daily chores if the respondent is ill, a person to share private worries and fears with, a person to turn to for help with personal problems, a person who understands private problems, and a person to love the respondent.

These items are combined into four subscales: emotional/information, tangible, affectionate and social interactions. Scoring is through each item being scored on a 1–5 point scale, scores within the dimensions are summed and then transformed to a 0-100 point linear scale.

#### **Evidence of a latent construct**

The MOS Social Support Survey (Sherbourne and Stewart, 1991) was developed from functional support theories of social relationships. The perspective was that of functional support for the respondent. The reason for adopting this perspective was the researchers' belief that a person's perceptions about available support were important. They noted that received support is confounded with need and that it may not reflect the level of support that is actually available to the respondent.

A literature review guided the development of the conceptual model, which was based on the most commonly reported aspects of social support. The initial items (n=50) were designed to cover the areas within the model, and to be as simple as possible to reduce response burden. A judging panel was asked to review draft items and to allocate them the model dimensions. The final item bank (n=37) was then administered to patients visiting health clinics (n=2,987). These data were analysed using multitrait analysis. Items that did not discriminate, or that were not internally consistent with their hypothesized dimensions, were removed. Twenty items survived, in the four subscales described above.

## Validity evidence

### *Content validity*

Generally, coverage appears to be good. Considerable care was taken to define a latent concept and its dimensions, and then to operationalise this through standard psychometric procedures. The researchers had hypothesized five dimensions, but the multitrait analysis revealed overlap between emotion and information so these were collapsed. The correlation between items and their dimensions were all  $>0.70$ .

The content validity of the MOS Social Support Survey was assessed through matching the items against the model using multitrait analysis. In general, the instrument has representation of emotional support, information support, tangible support, social interaction and companionship.

### *Construct validity*

Construct validity was assessed through confirmatory factor analysis of the structure of the instrument, which suggested that the four subscales were distinguishable (no further details of this model are provided in the Sherbourne and Stewart paper [Sherbourne and Stewart, 1991]). The structure of the MOS Social Support Survey was confirmed in a convenience sample of Chinese patients, using confirmatory factor analysis (Yu, et al. 2004), but in a sample of black South African patients with diabetes mellitus, factor analysis identified just 2 factors (socio-emotional support and tangible support) – each of which was extremely reliable in its own right (Westaway, et al. 2005). In the same study it was also reported that the single item number of close friends and relatives was not related to any of the social support dimensions of the instrument. A similar finding was reported for a Taiwanese sample of cancer caregivers (Shyu, et al. 2006). Factor analysis identified two factors – labelled emotional support and tangible support respectively. The interfactor correlation was  $r = 0.71$ .

### *Criterion validity*

Criterion validity evidence was reported by correlations with a range of other health measures also developed for the Medical Outcomes Study, including loneliness ( $r = -0.67$  overall), family functioning (0.53), marital functioning (0.56), mental health (0.45) and social activity (0.30) (Sherbourne and Stewart, 1991). Significant correlations between the MOS Social Support Survey and the summary scales of the Mental Health Inventory in a sample of women with breast cancer have been reported ( $r_{\text{range}} = 0.50-0.58$ ) (Kornblith, et al. 2001), and it correlated 0.82 with the Multidimensional Perceived Social Support Survey (Yu, et al. 2004). The correlation with the Hospital Anxiety and Depression Scale was reported by Yu et al to be  $r = -0.58$  (Yu, et al. 2004).

## Reliability

The Cronbach  $\alpha$ s of the four scales in the construction sample were between 0.91 and 0.97 (Sherbourne and Stewart, 1991). In other samples it has been reported to be between 0.85 to 0.98 (Grace, et al. 2004; Heinonen, et al. 2001a; McQuellon, et al. 1998; Shyu, et al. 2006; Westaway, et al. 2005; Yu, et al. 2004).

Test-retest reliability over a 1-year period was reported by Sherbourne and Stewart (Sherbourne and Stewart, 1991) to be within the range 0.72 to 0.78. In convenience sample of Chinese patients at 2-week interval as measured by the intraclass correlation was ICC = 0.84 (Yu, et al. 2004).

## Responsiveness

The MOS Social Support Survey has been responsive to age (Sherbourne and Stewart, 1991), bone marrow transplantation (Heinonen, et al. 2001a; McQuellon, et al. 1998), type of cancer patient (Lehto-Jarnstedt, et al. 2004), gender (Heinonen, et al. 2001a; Sherbourne and Stewart, 1991), relationship status (Burgoyne and Saunders, 2000; Sherbourne and Stewart, 1991), and social wellbeing (Heinonen, et al. 2001a). Regarding its sensitivity to HIV/AIDS patients when compared with Sherbourne and Stewart's patients' norms, the MOS Social Support Survey was sensitive only on the Tangible Support scale (Burgoyne and Saunders, 2000).

The instrument was not sensitive over time in a longitudinal study of bone marrow transplantation (Heinonen, et al. 2001b), referral to cardiac rehabilitation (Grace, et al. 2004), or to respite care for caregivers (Nicoll, et al. 2002).

### **Assessment against the study criteria**

Regarding comparative data, the MOS Social Support Survey correlated 0.82 with the Multidimensional Perceived Social Support Survey (Yu, et al. 2004). No other studies were identified comparing the MOS Social Support Survey with another social support or isolation instrument.

At 20 items it is a moderately long scale to be included in an instrument battery and it should be borne in mind that for those with dementia or cognitive impairment this may be a long instrument to complete. The cognitive burden of administration or completion is likely to be moderate because of the conditional tense of several items (e.g. Someone to help with daily chores if you were sick) which might pose problems with translator-administration or use among the illiterate or those with cognitive impairment. No studies were identified where the instrument had been used in a dementia sample.

There is considerable reliability and validity evidence, although it should be noted that the internal structure of the instrument may not hold cross-culturally.

No costs were identified for using the MOS Social Support Survey and the administration and scoring costs are likely to be low as it has been designed for self-completion.

### **8.3.5 Norbeck Social Support Questionnaire**

The US Norbeck Social Support Questionnaire was developed to enable the assessment of social support, primarily for use in nursing or clinical settings (Norbeck, 1984; Norbeck, et al. 1981; Norbeck, et al. 1983).

It consists of two parts. In the first part the respondent provides a list of people he/she knows (up to 24 persons) and judges to be in his/her personal social network. In the second part, each person (person X) listed in the personal social network is rated for his/her affect, affirmation and aid provided to the respondent. There are two items for each of these components. For affect the items cover how much person X makes you feel liked and loved, and how much respected and admired. For affirmation the items cover how much you can confide in this person, and whether person X agrees with your actions and thoughts. For aid the two items cover financial or practical help person X would provide, and how much aid he/she would provide if the respondent was confined to bed. The Guttman ratings for each of these six items, on a 5 point scale, are not at all/ a little/ moderately/ quite a bit/ a great deal.

There are three additional items covering the length of time person X has been known, how often the respondent is in contact with person X, and whether the respondent has lost an important relationship in the previous year.

The entire Norbeck Social Support Questionnaire is presented in a booklet format whereby on one side (and always visible) is the list of persons in the respondent's convoy, and on the facing page the items laid out such that for each person in the convoy there is a corresponding space for the assessments. Each of the three components is on a separate page. The three additional items are presented on a separate page.

Scoring is through rating response levels for each person on a scale of 0-4, and then summing across all persons in the convoy for that particular item. This procedure is repeated for each item. The number of people in the respondent's social network is the number listed. The final item (loss

of an important relationship) is scored dichotomously, and the quality of that loss is scored on a 5-point rating scale. Once scored like this, three aggregate scores are computed. These are Total Function (the sum of affect, affirmation and aid), Total Network (the sum of the number of persons in the convoy, the duration of these relationships, and the frequency of contact), and Total Loss (the sum of the number of persons lost and the rated amount of support lost). Subscales can also be computed for each of affect, affirmation and aid by simply summing the scores on each of the two relevant items (Norbeck, 1995).

Very preliminary normative data were reported by Norbeck, et al. (1983), based on a random sample of the staff of a large university medical centre (n = 136, participation rate 31%).

### **Evidence of a latent construct**

The Norbeck Social Support Questionnaire is based on the definition of social support advanced by Kahn (1979, p85), which was

“...interpersonal transactions that include one or more of the following: the expression of positive affect of one person toward another; the affirmation or endorsement of another person’s behaviours, perceptions or expressed views, the giving of symbolic or material aid to another.”

Additionally he advanced the concept of an individual’s ‘convoy’ which was defined as the set of persons whom the individual relied up for support (or supported in turn). Norbeck, et al. (1981) used this conceptual background to define that social support consisted of three components: affect, affirmation and aid, each of which could apply to each person in an individual’s convoy. The three additional items described above account for changes of personnel in the convoy.

None of Norbeck’s three seminal papers (Norbeck, 1984; Norbeck, et al. 1981; Norbeck, et al. 1983) describe how the instrument was actually developed – it is presented to the reader as a *fait accompli*.

### **Validity evidence**

#### *Content validity*

Because there is no explanation of how the instrument was developed, and for the technical reasons outlined below, it is difficult to establish whether the latent concept is actually covered by the six items in the Norbeck Social Support Questionnaire. Norbeck, et al. presented very little evidence on this point (Norbeck, 1984; Norbeck, et al. 1981; Norbeck, et al. 1983). A concern is that reported scores significantly varied by education attainment (Chan, et al. 2004), which may reflect the cognitive demands of completing the questionnaire as much as real differences in social networks by education level.

Content validity for the Norbeck Social Support Questionnaire has not been satisfactorily reported. As noted above, the scale was presented by Norbeck, et al. as a *fait accompli*, and the seminal study of the instrument involved college nursing students (Norbeck, et al. 1981). No information is available on its development other than that it was based on Kahn’s work. There is no information on how the items were developed, whether they were tested on a construction sample or if they were the best items from a competing pool of potential items. The correlations between the items are very high ( $r = 0.72$  to  $0.97$  [Norbeck, et al. 1981]), and examination of the underlying structure suggests that the content of the items (the descriptive system) does not match the theoretical model postulated by Norbeck. Additionally, there is no evidence that the views of users were taken into account during instrument construction; thus it may lack ecological validity evidence. There is also some evidence that the scoring system is confounded. Norbeck, et al. (1983) reported that when regressed on a measure of life stress, the social support scales explained ~20% of the variance, but that when the combined Total Function or Total Network summary variables were entered the explanatory power dropped to ~2% of the variance. It was recommended that the scales should not be combined into the summary scores where there was an effect due to one of the scales or an interaction between the scales.

### *Construct validity*

Regarding the construct validity of the scale, there are two technical issues which limited Norbeck, et al's ability to examine this thoroughly. First, the items that make up the three components are all highly correlated. The range of correlations for all six items was reported to be between  $r = 0.72$  to  $0.97$  (Norbeck, et al. 1981). Norbeck reported the results of a factor analysis (varimax rotation) which identified just two factors, which she labelled Emotional Support and Tangible Support. Although the two affect and two aids items loaded on Emotional and Tangible, the two items representing affirmation cross-loaded on both factors  $>0.30$  (Norbeck, 1995).

The second problem relates to the practice of including in the score the actual number of persons nominated in the convoy. This directly confounds enumeration (the quantity of contacts) and the functional content the items are supposed to be measuring (the quality of the contact). Part of the problem here is that a large number of nominations will always lead to higher scores, regardless of the quality of contacts (Norbeck, 1995). This aspect of scoring the Norbeck Social Support Questionnaire was challenged by Seckel, et al. (1996) who reported that social support strength was more highly correlated with stress than the network size.

The structure of the scale was examined by Gigliotti (2002, 2006) using structural equation modelling. The results confirmed the presence of a 3-factor model with excellent fit properties (RMSEA = 0.03) in one sample (Gigliotti, 2002) and a 2-factor model in another sample (Gigliotti, 2006).

### *Criterion validity*

Criterion validity evidence was obtained from a student sample ( $n=42$ ) by correlation with Cohen and Lazarus' Social Support Questionnaire which has three dimensions: Tangible, Informational and Emotional support (Norbeck, et al. 1981). Affect, Affirmation and Aids all correlated significantly with Emotional support ( $r = 0.51, 0.56$  and  $0.44$ , respectively). Affirmation correlated significantly with Information, but none of the scales correlated significantly with Tangible support (the correlation for Aid was  $-0.03$ , for example). The Total Functional support summary was significantly correlated with Emotional and Information, Total Network with Emotional, and Total Loss with Tangible. Later the relationship with other measures of social support was examined showing that it was significantly related to measures of the need for inclusion and need for affection (Norbeck, et al. 1983). Nelson (1989), in a study of the non-institutionalized elderly reported that it correlated with depression.

### **Reliability**

The reliability of the Norbeck Social Support Questionnaire has been reported to be very high in the range  $\alpha = 0.89$  to  $0.98$  for the three scales and/or 6 items (Chan, et al. 2004; Connelly, 1998; Gigliotti, 2002, 2006; May, 1992; Miller, 1997). These findings are not surprising given that the correlations between the items have been reported to be all over  $r = 0.70$  (Gigliotti, 2002; Norbeck, 1984).

Norbeck, et al. (1981) reported that among college students the 1-week test-retest correlations of each of the items ranged from  $r = 0.87$  to  $0.92$ ; in a later paper they reported that over a 7-month period the test-retest correlations were in the range  $r = 0.58$  to  $0.78$  (Norbeck, 1984).

### **Responsiveness**

It has been shown to be sensitive to anxiety (Muller and Lemieux, 2000), asthma management (Sin, et al. 2005), breast self-examination frequency (Wagle, et al. 1997), care adequacy for pregnant women and health behaviours (Schaffer and Lia-Hoagberg, 1997), coping strategies in cancer surgery patients (Chan, et al. 2004), depression in pregnant women (McKee, et al. 2001), dietary changes among older adults (Murphy, et al. 2001), immigrant status (May, 1992), mental health status in pregnant women (McKee, et al. 2001), relationship status (Norbeck, et al. 1983), and self-image (Muller and Lemieux, 2000).

The scale was insensitive to the number of hospital admissions for ischemic heart disease (Stewart, et al. 1997) or pregnancy status among teenage women (Connelly, 1998).

Regarding responsiveness over time, among college students over a 7-month period the only scale to show significant score change was the Total Network variable. Between baseline and follow-up the correlations were in the range  $r = 0.58$  to  $0.78$  (Norbeck, et al. 1983).

### **Assessment against the study criteria**

Regarding comparative data, the Norbeck Social Support Questionnaire has been correlated with Cohen and Lazarus' Social Support Questionnaire; the correlations between the various scales were between 0.44 and 0.56 (Norbeck, et al. 1981). No other studies were identified comparing the Norbeck Social Support Questionnaire with another social support or isolation instrument.

Although nominally an 11-item instrument (question 9 has two parts), its length is actually a function of the number of persons in the respondent's convoy nominated: for 1 person there would be 11 items, for 2 there would be 18, and so on. Given that among adults, the average number of persons nominated was 12 (Norbeck, et al. 1983), this implies that >50 assessments would need to be made even where there was a small nominated convoy. In practical terms, then, this is a long instrument. It is also likely to be a very demanding instrument to complete, particularly for those with cognitive impairment, because of the need to keep track across booklet pages of each nominated person and to assess that person on multiple criteria. It is unlikely this could be successfully done by those with dementia, or where the instrument was translator administered. No studies were identified where the instrument had been used in a dementia sample.

The reliability evidence as assessed by internal consistency for the Norbeck Social Support Questionnaire is consistently high across different samples, suggesting that the 6 items forming the central core of the instrument are measuring very similar constructs, if not the same construct. However, over a longer period of time (7 months) the test-retest reliability was below accepted standards, probably reflecting a real change in the respondents' social networks.

Regarding validity, the evidence is generally unsatisfactory. There is almost no information on how it was constructed, there are issues relating to the scoring system used, and the factorial structure of the measure appears to be confounded.

No costs were identified for using the Norbeck Social Support Questionnaire, but the administration and scoring costs are likely to be high relative to the other instruments reviewed due to the inherent difficulties of completion and scoring.

### **8.3.6 Sarason Social Support Questionnaire**

The US Sarason Social Support Questionnaire was developed to quantify the perceived availability of and satisfaction with social support (Sarason, et al. 1983).

Based on a review of the literature, 61 items were written sampling situations where social support might be important to people (e.g. Whom could you really count on to help you out in a crisis situation, even though they would have to go out of their way to do so?). The items were administered to college students, and based on item-correlations those with low correlations were eliminated. Correlation with the number of supportive people was also used as a criterion for item retention. Twenty-seven items were retained. Each item consists of two parts. Respondents are asked (a) to provide a list of people to whom they can turn to when support is needed, and (b) to indicate their level of satisfaction with these social supports.

Scoring each item is a two-step process. First, for each item the number of people available for support (the SSQ Number or Perceived Availability score (Sarason, et al. 1987)) and the SSQ

Satisfaction score (from 1 to 6 for each item, based on response scales with 6 options from very dissatisfied to very satisfied) is computed through simple summation. Then the overall score for each of the Number and Satisfaction scales is obtained by dividing the sum by 27, the number of items. Sarason, et al. reported that the correlation between the two scores was in the range of 0.30-0.40 (Sarason, et al. 1987; Sarason, et al. 1983).

In addition to the full 27-item version, Sarason, et al. (Sarason, et al. 1987) developed 6-item (SSQ6) and 3-item versions. These descriptors are somewhat misleading, however, since each item has two parts. Thus the SSQ6, for example, actually has 12 questions to be answered.

The SSQ6 items refer to persons able to distract the respondent when the respondent feels under stress, persons the respondent can turn to when feeling under pressure or being tense, persons who accept the respondent, persons who care about the respondent regardless of what is happening to the respondent, persons who can help the respondent when the respondent is feeling poorly, and persons who can help the respondent when the respondent is upset (Sarason, et al. 1987).

### **Evidence of a latent construct**

The scale was based on a literature review of social support, and standard psychometric practices generally followed during construction. However, Sarason, et al. did not explicate the model from the literature review in any detail. Although it was reported that the items were written to cover the universe of social support situations, no description of these is actually given (Sarason, et al. 1983). Additionally, there are two areas of concern. The construction sample was composed entirely of college students and item selection was, at least partly, based on the correlation with the number of social contacts.

Items in the SSQ6 were selected on the basis of factor analysis of three samples of student data; the highest loading items across both samples and both scales were selected. No item selection procedures were reported for the development of the 3- item version (Sarason, et al. 1987).

### **Validity evidence**

#### *Content validity*

The consequence of the construction procedures outlined above is that the Sarason Social Support Questionnaire is likely to primarily reflect the concerns of first year college students enrolled in psychology courses. This is particularly a concern for the SSQ6 since this was derived from the pivotal items in two samples of college students. The content of the instrument, again particularly the 6-item version, reflects a concern with support from others when the respondent needs this support. This is a very narrow perspective on social support.

#### *Construct validity*

Construct validity evidence was reported by Sarason, et al. (1983) using factor analysis. This revealed a single factor for each of the Number (or Availability) and Satisfaction scales. The correlation between the scales was in the range  $r = 0.21$  to  $0.34$ . Sarason, et al. (1987) concluded on the basis of correlations with other measures of psychological constructs (see below) that low social support is related to an external locus of control, limited ability to persist with demanding tasks, increased levels of cognitive interference, and a relative dissatisfaction with life.

These findings are strongly suggestive that scores on the Sarason Social Support Questionnaire are a function of the mental health state of the respondent.

#### *Criterion validity*

Criterion validity evidence for the scale was presented in the seminal paper by comparing the Sarason Social Support Questionnaire with the several other instruments assessing psychological constructs. Significant correlations were reported for males for the Number scale with depression ( $-0.24$ ) and hostility ( $-0.23$ ), and for the Satisfaction scale with depression ( $-0.22$ ). For females,

both scales correlated significantly with anxiety (-0.30, -0.39 for Number and Satisfaction, respectively), depression (-0.31, -0.43), hostility (-0.26, -0.36), lack of protection (-0.32, -0.22), and for the Number scale with extraversion (0.35) and for the Satisfaction scale with neuroticism (-0.37).

In three other samples of college students, Sarason, et al. (1987) compared the SSQ6 and SSQ3 with various psychological constructs (e.g. measures of anxiety, depression); there were 52 comparisons altogether. Across all 52 comparisons there were 51 significant effects reported for the SSQ6 Satisfaction scale or SSQ6 Number scales. The highest correlations were with the Interpersonal Support Evaluation List, Mother Care, Loneliness, and Family Relationships. The correlation with the Inventory of Social Supportive Behaviours was 0.28 for the Number scale and 0.24 for the Satisfaction scale. Two different measures of depression were used, and the correlations were -0.19 and -0.47. Elsewhere, the correlation between the Sarason Social Support Questionnaire and depression was  $r = 0.20$ . (Alpass and Neville, 2003) and the correlation between the Sarason and the revised UCLA Loneliness Scale was  $r = 0.40$  (Barron, et al. 1994). Among Chinese family caregivers the correlation between the SSQ6 and the SF36 was 0.52 and -0.40 with the Family Assessment Device (Chien, et al. 2007).

### Reliability

The internal consistency of the Sarason Social Support Questionnaire was reported in a sample of college students to be Cronbach  $\alpha = 0.97$  (Sarason, et al. 1983). For the SSQ6 construction samples it was reported to be in the range  $\alpha = 0.90 - 0.93$  (Sarason, et al. 1987).

In a sample of visually-impaired persons aged 65 or more years, the Cronbach  $\alpha = 0.67$  and 0.81 for the Number and Satisfaction scales, respectively (Barron, et al. 1994); among a sample of elderly Korean female immigrants to the USA the Cronbach  $\alpha = 0.87$  and 0.93 (Kim, 1999b) and 0.70 and 0.83 for the Chinese version, respectively (Pang, et al. 2001), and in a sample of adolescents the reliability of the Number score for the 6-item version was reported at 0.73 (Bal, et al. 2003).

Among college students the test-retest correlation at 4-week interval was reported to be 0.83 (Sarason, et al. 1983); higher values were reported for the French translation at 0.89 and 0.84, respectively (Rasclé, et al. 2005).

### Responsiveness

Studies reporting responsiveness of the full 27-item or 6-item versions include suicide potential among adolescents (D'Attilio, et al. 1992), the burden of care giving among Chinese caregivers (Chien, et al. 2007); the quality of life among those with AIDS (Swindells, et al. 1999), the network satisfaction scale to loneliness among vision-impaired older adults (Barron, et al. 1994), separation in childhood from the mother by adults with psychiatric symptoms (Furukawa, et al. 1999), loneliness among older female Korean immigrants (Kim, 1999b), gender differences in bone marrow transplant patients (Heinonen, et al. 2001a), extroversion and depression among French students and unemployed males (Rasclé, et al. 2005), weight gain among Chinese patients undergoing haemodialysis (Pang, et al. 2001), disclosure of HIV status (Pettrak, et al. 2001), with Sickness Impact Profile scores following stroke (Mackenzie and Chang, 2002), and an interaction was reported between dichotomised scores (at the median) and age for perceived control in a study of housing density (Sinha and Nayyar, 2002).

The full 27-item version was not responsive to relationship status in a sample of older adults with vision impairment (Barron, et al. 1994), relationship orientations (Alemi, et al. 2003), bone marrow transplantation (Heinonen, et al. 2001b), suicidal behaviour in adolescents (Morano, et al. 1993), postnatal support (Reid, et al. 2002), and depression in older males (Alpass and Neville, 2003).

The short 6-item version (SSQ6) was not responsive to adults with psychiatric symptoms reporting the childhood loss of a parent (Furukawa, et al. 1999), to adolescents with trauma experiences

(although family support was a significant positive predictor and friend support a significant negative predictor) (Bal, et al. 2003), or to bone marrow transplant patients at 12-month follow-up (Heinonen, et al. 2001b).

### **Assessment against the study criteria**

Regarding comparative data, the Sarason Social Support Questionnaire has been correlated with the revised UCLA Loneliness Scale; the correlation was  $r = 0.40$  (Barron, et al. 1994). No other studies were identified comparing it with another social support or isolation instrument.

Although the full instrument is nominally a 27-item instrument, each item has two parts (the number of social contacts (1 through 9) and a rating of these (very satisfied to very dissatisfied)); its real length is thus 54-items. The SSQ6 similarly has 12 items rather than the nominal 6. In practical terms, then, the full Sarason is a very long instrument which is likely to be demanding for those with cognitive impairment. No studies were identified where the instrument had been used in a dementia sample.

The reliability evidence as assessed by internal consistency is consistently high across different samples. Given that reliability is a function of the number of items in a scale, the mean covariance between items and the sum of all the elements in the variance/covariance matrix (Cortina, 1993), it follows that the high reliability may be a function of the instrument length more than anything else. Within this framework it is not surprising that the reliability of the SSQ6 was considerably lower in a non-construction sample.

Regarding validity, the evidence is generally unsatisfactory. It was developed among first-year American college students and the correlations with measures of mental health suggest it may be related to mental health concerns of adolescents rather than a more mature understanding of social support. The, at best, moderate correlations between the Number and Satisfaction scales are another source of concern. There are two reasons for this. First, the Number scale is a simple enumeration of social contacts (whether they are important or not). This implies that rating these collectively may lead to systematic measurement error. Second, the literature suggests that it is the quality of social relationships that is important rather than the number, and that the number of social contacts declines across the lifespan (Carstensen, 1992; Charles, et al. 2001; DiTommaso, and Spinner, 1997; Iliffe, et al. 1991; Revenson and Johnson, 1984; Stevens and Westerhof, 2006). That there is mixed evidence on its responsiveness is also a concern.

No costs were identified for using the Sarason Social Support Questionnaire, but the administration and scoring costs are likely to be high relative to the other instruments reviewed due to its length.

### **8.3.7 UCLA Loneliness Scale**

The American UCLA (University of California Los Angeles) Loneliness Scale has been through three iterations. Originally published in 1978, it was revised in 1980 and again in 1996 (Russell, et al. 1980; Russell, et al. 1978; Russell, 1996). The first revision, Version 2, was undertaken to prevent response bias, including social desirability, which had been identified in the original scale. Additionally, there were concerns that it was confounded by depression and low self-esteem. To ameliorate these concerns the original scale plus an additional 19 new items written by Russell, et al. was administered to 162 students. Following data analysis, 6 of the original items were replaced with new items, and 10 items were reversed so they became positive. The criterion for item replacement was higher correlation with a self-labelling loneliness index. Russell, et al. (1980) described the positive and negative sub-scales as measuring satisfaction and dissatisfaction with social relationships, respectively.

The third version, Version 3, was published in response to identified problems with the Version 2 items (e.g. double-barrelled item stems, difficult words such as 'superficial') when administered to

older adults, to reinforce that the timeframe is the present (the items stems all read “How often to you feel...” whereas in earlier versions this timeframe was given only in the instructions) and to change the instrument voice from first to third person (from ‘I’ to ‘you’). The implication of the present timeframe is that the UCLA Loneliness Scale was conceived as a ‘state’ loneliness scale and (Russell, et al. 1980, p473) described the scale as being about “feelings of social dissatisfaction”. Regarding the issue of timeframe, at least two research teams have changed the timeframe to reflect lifelong (i.e. trait) loneliness through use of the instruction “looking back over your lifetime” (Gerson and Perlman, 1979; Hector-Taylor and Adams, 1996).

Scoring of Version 3 of the scale is through reversal of the nine positive items, then summing of all items. Items are scored on 4 point Guttman-type scales, never/ rarely/ sometimes/ always. Higher scores indicate greater loneliness (Russell, 1996).

The UCLA Loneliness Scale has 20 items. In addition to the standard UCLA Loneliness Scale, there are several shorter versions, including an 11-item version which was specially constructed for use with older adults (Perlman, et al. 1978), a 10-item version for mail administration to teachers (Russell, 1996), an 8-item version (Hays and DiMatteo, 1987) and a 4-item version (Russell, et al. 1980). None of these versions appears to have been widely adopted or used. Wilson, et al. (1992) and Hays and DiMatteo (1987) both reported that the correlation between the full 20-item version and the 8-item version was between 0.82-0.91; slightly lower correlations were reported for the 4-item version. Although Wilson, et al. (1992) reported that neither the 8- or 4-item versions were deemed particularly reliable (for both these versions the reliabilities were  $\leq 0.60$ ), this was not confirmed by Hays and DiMatteo (1987). Based on multitrait analysis, Hays and DiMatteo (1987) reported that the 8-item version performed as well as the full 20-item version.

### **Evidence of a latent construct**

The UCLA Loneliness Scale was developed from items drawn from Sisenwein’s 1964 PhD thesis scale, which had been designed to measure the intensity of feelings of loneliness (Russell, 1982). Items in the Sisenwein scale were written by a team of 20 psychologists describing the experience of loneliness (Russell, et al. 1978). The UCLA Loneliness Scale was designed to be a psychometrically adequate, easily administered, general loneliness scale. In sampling from Sisenwein’s 75 items, the criterion was, in Russell’s words, “...unsystematic; the only criterion was to eliminate very extreme statements” (Russell, 1982, p90). Twenty-five items were selected. Following administration to two student samples (two focus groups of volunteers (n=47), and psychology course students (n=192) items which correlated  $\geq 0.50$  were retained for the UCLA Loneliness Scale (20 items) (Russell, 1982; Russell, et al. 1978). Subsequent testing showed that all 20 items correlated with each other  $\geq 0.40$  (Russell, 1982). Russell (1982) has argued that the UCLA Loneliness Scale is consistent with Weiss’ (1973) loneliness theory (see above).

### **Validity evidence**

#### *Content validity*

Since no latent model of loneliness was postulated during development of the UCLA Loneliness Scale, it is difficult to assess its coverage other than through examination of item content. Essentially, in this reversal of conventional psychometric practice the role of effect indicators (the items) is to become causal indicators which define what is being measured; these models are known as reflective models (Pedhazur and Schmelkin, 1991; Streiner and Norman, 2006). Using this criterion, the UCLA Loneliness Scale appears to be measuring state loneliness, depression and poor self-esteem arising from an absence of companionship. This emphasis suggests that the scale may be mainly orientated towards the concerns of the young (college students) and the core components of establishing friendships (Solano, 1980). Twenty years later, Steptoe, et al. (2004) came to the same conclusion, reporting that loneliness as measured by Version 2 was primarily related to poor self-image and maladaptive methods of psychological coping.

Perhaps the point should be made that the item reversal and language simplification undertaken in Versions 2 and 3 do not, by themselves, change the meaning of a measure. This point is clearly

illustrated by Russell, et al's (1980) report that the correlation between the original scale and the revised Version 2 was 0.91, suggesting that the revisions had not changed the fundamental nature of the scale (indeed, this correlation would suggest that the two scales were equivalent measures). The items cover being unhappy (2 items), having no-one to talk to, unable to tolerate being alone, lacking companionship, having no-one who understands (2 items), waiting for people to contact me, no-one to turn to, no longer close to anyone, a lack of shared interests, feeling left out (2 items), feeling alone or isolated, being unable to communicate, having superficial relationships, being starved for company, finding it difficult to make friends and being separate from others when with others.

In Version 1 all the items were negative items, as described above. In Version 2, half the items were reversed, and in Version 3 one more item was also reversed.

The content validity of the UCLA has not been well established. As described above, there was no underlying theory of loneliness behind its development, and the selection of items was based on convenience and correlations between items. Russell, et al. (1978) described the scale as having face validity as shown by the content of the items. The high correlations with depression (ranging from  $r = 0.38$  to  $0.62$  among samples of college students) (Russell, et al. 1980; Russell, et al. 1978), anxiety ( $r = 0.35$  to  $0.36$ ) and with various measures of self-satisfaction (between  $-0.36$  to  $0.58$  among the same students) suggests that the scale may be measuring a general psychological distress construct rather than loneliness per se. Additionally, there is no evidence that the views of users were taken into account during construction; thus it may lack ecological evidence.

#### *Construct validity*

No construct validity evidence for the original UCLA Loneliness Scale was published by the instrument developers (Russell, et al. 1978). Russell, et al. (1980) argued that the validity of the scale was shown by its high correlation with a self-labelling loneliness item ( $r = 0.71$  to  $0.79$  for different samples) and, for Version 3, with the NYU Loneliness Scale ( $r = 0.65$ ) and the Differential Loneliness Scale ( $r = 0.72$ ) (Russell, 1996). In a study of older adults, Perlman, et al. (1978) reported that, on the 11-item UCLA Loneliness Scale the correlation with the self-labelling loneliness item was  $0.72$ . However, Version 1 correlated just  $r = 0.12$  between the UCLA Loneliness Scale and the number of social activities (Russell, et al. 1978). In a study of British middle-aged civil servants, Steptoe, et al. (2004) reported that UCLA Loneliness Scale scores systematically varied with social isolation and emotional support scales. Barron, et al. (1994) reported that the UCLA Loneliness Scale correlated  $0.40$  with the Sarason Social Support Questionnaire in a sample of 87 older adults with visual impairment.

Regarding the structure of the UCLA Loneliness Scale, (Russell, et al. 1980; Russell, et al. 1978) reported that it was unidimensional. Other researchers, however, have consistently showed that it is not. Zakahi and Duran (1982) reported two factors for the UCLA Loneliness Scale, each of 10 items. These were labelled Intimate Other and Social Network. Wilson, et al. (1992) in a Zimbabwean study of two samples (adolescents and adults) identified the same 2-factor structure, as did Mahon and Yarcheski (1990) and Miller and Cleary (1993). Importantly, all these researchers (including the Austin study described below) reported that negatively worded items loaded on one factor and positive items on the other (or were split between two factors). This suggests that respondents react differently to positively and negatively worded items. If so, the consistent UCLA Loneliness Scale factor structure results may have been determined by the different item response directions rather than item content. Miller and Cleary (1993) interpreted the UCLA Loneliness Scale factor structure as evidence of response bias – a point acknowledged by Russell in his 1996 paper (Russell, 1996).

Other researchers, however, have found a different internal structure. Data from college students analyzed by Hays and DiMatteo (1987) produced a 5-factor structure. Hartshorne (1993), using a confirmatory factor analysis approach, reported a single-factor model given that the data from the

UCLA Loneliness Scale were non-normally distributed. Austin (1983) using confirmatory factor analysis reported three factors, labelled Intimate Others, Social Others, and Belonging and Affiliation. A 3-factor solution was also reported by Hawkey, et al. (2005) who labeled the factors Isolation, Relational Connectedness and Collective Connectedness.

#### *Criterion validity*

Criterion validity for the UCLA Loneliness Scale was initially assessed by correlation with a self-report question on loneliness ( $r = 0.79$ ,  $n = 45$ ). There were large differences in scores between students participating in two focus groups, one of which was composed of volunteer lonely students and the other a comparison group (Russell, et al. 1978). The UCLA Loneliness Scale, in college students, correlated with the Bradley loneliness measure  $r = 0.74$  (Solano, 1980). The correlation between loneliness scores and suicidal ideation scores on the Geriatric Suicide Ideation Scale was 0.66, and it was 0.67 with the Life Satisfaction Index, and 0.63 with the Center for Epidemiological Studies Depression Scale (CES-D) (Chou, et al. 2005). Alpass and Neville (2003) reported a correlation between the UCLA and depression of  $r = 0.63$ ; importantly, regarding the interpretation of the UCLA Loneliness Scale referred to above, they also reported that the correlation between the Sarason Social Support Questionnaire and depression was just 0.20, while between the UCLA Loneliness Scale and the Sarason it was  $r = 0.30$ . Correlations with depression (the Beck Depression Inventory;  $r = 0.42$  to  $0.53$  for different age groups) were reported by Nolen-Hoeksema, et al. (2002). Swami, et al. (2007) reported that in a sample of Malaysian students it correlated with general health  $r = -0.50$ , life satisfaction  $-0.40$  and with depression  $0.38$ . In a study of overseas students in the USA, it was concluded that the scale was highly associated with depression, but that emotional loneliness was identified whereas social loneliness was not (Hsu, et al. 1987). These findings are consistent with an earlier study (Solano, 1980) which reported that the unidimensional UCLA Loneliness Scale specifically identified loneliness due to a lack of social interaction (which was not measured by the scale!). In a recent study, the UCLA Loneliness Scale was used as the measure of loneliness and the Sarason Social Support Questionnaire as the measure of social support; the Sarason satisfaction scale explained 52% of the variance in the UCLA Loneliness Scale (Kim, 1999b).

#### **Reliability**

Russell, et al. (1978) reported that among the construction sample of 239 students the reliability of the UCLA Loneliness Scale was  $\alpha = 0.96$ . The reliability of Version 2 has been reported in the range  $\alpha = 0.90$  to  $0.94$  (Cacioppo, et al. 2006; Chou, et al. 2005; Hartshorne, 1993; Hughes, et al. 2004; Nolen-Hoeksema and Ahrens, 2002; Russell, 1982; Storch, et al. 2004). For Version 3 the Cronbach  $\alpha = 0.89$  to  $0.95$  across different samples (adolescents, students, nurses, teachers and the elderly) (Chipuer, et al. 2003; Kim, 1999a; Kim, 1999b; Russell, 1996). Elsewhere reliability has been reported among retirees and nursing home residents to be Cronbach  $\alpha = 0.86$  to  $0.90$  (Adams, et al. 2004; Bergman-Evans, 2004).

Test-retest reliability among college students, over a 2-month period, was reported to be between 0.62 and 0.73 for Version 2 for different samples (Russell, 1982; Russell, et al. 1978). Based on a college student sample, Hartshorne (1993) reported a 2-week test-retest correlation of  $r = 0.85$ . Hector-Taylor and Adams (1996), in a New Zealand sample of adults aged over 60 years, reported that the correlation between repeat administrations of the UCLA Loneliness Scale at 2-week interval, with a changed timeframe ('looking back over your life' for the second administration), was 0.86. In a sample of older adults with cerebral palsy at 3-4 week test-retest the intraclass correlation for Version 3 was 0.83 (Balandin, et al. 2006). Test-retest reliability of Version 3 over a 12-month period in a sample of older adults was reported to be  $r = 0.73$  (Russell, 1996), and for a sample of adults aged  $\geq 55$  the test-retest reliability was between 0.73 and 0.84 (Cacioppo, et al. 2006).

#### **Responsiveness**

Regarding the responsiveness of the scale, scores have been found to systematically vary by age (Geller, 2004; Nolen-Hoeksema and Ahrens, 2002; Russell, 1982), depression (Adams, et al.

2004; Cacioppo, et al. 2006; Russell, et al. 1980; Russell, 1996), gender (in college student samples for Version 3) (Russell, 1996), relationship status (Russell, 1982), self-labelling loneliness (Russell, 1982), and traumatic experiences (Springer, et al. 2003).

Elsewhere, however, age (Adams, et al. 2004; Springer, et al. 2003; Steptoe, et al. 2004), employment (Steptoe, et al. 2004); gender (Adams, et al. 2004; Hector-Taylor and Adams, 1996; Steptoe, et al. 2004) and relationship status (Geller, 2004) have not been predictive of scores.

Responsiveness over time has been shown in a study of animal-assisted therapy for older adults (Banks and Banks, 2002), but there was no significant change in UCLA Loneliness Scale Version 3 scores over time for nursing home residents participating in a care trial (Bergman-Evans, 2004).

### **Assessment against the study criteria**

Regarding comparative data, the UCLA Loneliness Scale has been correlated with several other scales measuring either social support or loneliness. These include with the NYU Loneliness Scale ( $r = 0.65$ ), the Differential Loneliness Scale ( $r = 0.72$ ) (Russell, 1996), the Sarason Social Support Questionnaire ( $r = 0.40$ ) (Barron, et al. 1994) and the Bradley loneliness measure ( $r = 0.74$ ) (Solano, 1980).

At 20 items the UCLA Loneliness Scale is a medium length instrument, which may limit its usefulness in a constrained instrument battery for use with those with cognitive impairment. Additionally, the use of double-negative item stem and response sets (e.g. My interests and ideas are not shared by those around me, combined with the response scale of Never) is likely to be confusing for those with limited cognitive capacity, although it is acknowledged that this judgement must be tempered by an awareness that the instrument has been used among the elderly and/or nursing home residents (Adams, et al. 2004; Bergman-Evans, 2004; Calvert, 1989). No studies were identified where the instrument had been used in a dementia sample.

The reliability evidence as assessed by internal consistency is consistently high across different samples. The test-retest reliability estimates vary considerable among different samples, and range from the unacceptable to the acceptable. The responsiveness evidence is, again, variable suggesting that it may be sample and condition specific rather than a stable scale.

Regarding validity, the evidence is unsatisfactory. The selection of items was unsatisfactory and further development and validation was among American college students. The implication is that it may be mainly orientated towards the concerns of the young (college students), including issues around establishing friendships, depression, poor self-image and maladaptive methods of psychological coping (Solano, 1980; Steptoe, et al. 2004). Although claimed to unidimensional, researchers have consistently reported that it is at least bi-dimensional along the lines of positive and negative items, findings which have been interpreted by one research team as evidence of response bias (Miller and Cleary, 1993) – a point acknowledged by Russell (1996).

No costs were identified for using the UCLA Loneliness Scale, but the administration and scoring costs may be high relative to the other instruments reviewed due to its length.

### **8.3.8 Three-item Loneliness Scale**

The Three-item Loneliness Scale was developed from the UCLA Loneliness Scale (Hughes, et al. 2004). The purpose was to produce a short scale that took less than 3-minutes to complete over the telephone for inclusion in the US Health and Retirement Study.

Factor analysis of UCLA Loneliness Scale ( $n = 1,255$  respondents) revealed the presence of 3 factors. The three items with the highest loading on the first factor were selected. The three items were then re-worded to make them suitable for telephone administration through use of 'you' instead of 'I' and by reducing the response scales from never/ rarely/ sometimes/ often to hardly

ever/ some of the time/ often. The item responses are coded 1, 2, and 3, respectively, and scores are obtained through simple summation. The three items are: How often do you feel you lack companionship?, How often do you feel left out? and How often you feel isolated from others?

### **Evidence of a latent construct**

The content of the items suggest that it is measuring state loneliness, and poor self-esteem arising from an absence of companionship. An interesting suggestion arising from the wording is that the items almost carry a sense of social exclusion where the focus is on others excluding the respondent. This hypothesis is consistent with the correlations on social involvement and neighbourhood safety (see below). For example, there was a monotonic relationship between scores and neighbourhood safety from no relationship among those living where safety was considered excellent to a correlation of 0.61 for those living in areas they reported to have poor safety (Hughes, et al. 2004).

### **Validity evidence**

#### *Content validity*

Like the full UCLA Loneliness Scale, this is a reflective model instrument since there was no underlying theory behind its development (Pedhazur and Schmelkin, 1991; Streiner and Norman, 2006).

Generally, then, the same content validity limitations that were discussed for the full UCLA Loneliness Scale also apply to the Three-item Loneliness Scale; viz., the lack of ecological evidence, and the substantial correlation with depression (reported to be 0.48 and 0.49 for two different samples). It also correlated 0.44 and 0.40 with stress (Hughes, et al. 2004). In the USA Health and Retirement Study (adults aged 54 years and older) the correlation between the Three-item Loneliness Scale and depression was 0.43 (Cacioppo, et al. 2006).

#### *Construct validity.*

Regarding construct evidence, this can be assumed from the selection of the items which were the three highest loading items on the first factor of the full UCLA Loneliness Scale. That the three items were selected from a single factor of the UCLA Loneliness Scale suggests that the Three-item Loneliness Scale lacks content validity; at best it could be described as assessing 1/3 of the construct assessed by the UCLA.

#### *Criterion validity*

Regarding criterion validity, the correlation with the full UCLA Loneliness Scale was 0.82 (Hughes, et al. 2004).

### **Reliability**

The Cronbach  $\alpha$  was reported to be 0.72, which Hughes stated was good for a 3 item scale (Hughes, et al. 2004).

### **Responsiveness**

Responsiveness was assessed by relationship status ( $r = -0.42$ ), accommodation arrangements ( $r = -0.11$  to 0.50 for various indicators), voluntary community participation/providing help to others ( $r = -0.15$  and 0.19) and rating of neighbourhood safety ( $r = -0.10$  to 0.61).

Responsiveness over time was assessed in the Health and Retirement Study by test-retest reliability at 12 month intervals over 36 months (Cacioppo, et al. 2006). The reported reliabilities were between 0.73 and 0.84. As reported by the authors, loneliness in Year 1 predicted loneliness in the subsequent years.

### **Assessment against the study criteria**

Regarding comparative data, the Three-item Loneliness Scale has been correlated with the full UCLA Loneliness Scale ( $r = 0.82$ ).

At just 3 items the Three-item Loneliness Scale is the shortest of any of the scales reviewed, making it particularly easy to administer. No studies were identified where the instrument had been used in a dementia sample.

The reliability evidence is fine for a 3-item scale, but is subject to the caveat that the test-retest reliability estimate over 36-months ( $r = 73-0.84$ ) may suggest non-responsiveness.<sup>32</sup>

Regarding validity, the evidence is unsatisfactory. The selection of items was based entirely on the highest loading on the first factor of the UCLE Loneliness Scale, and these items were then re-written into a form suitable for telephone administration together with simplification of the response sets to just 3 options one of which was “often”. Given the skewness of the scale, that good fit statistics on a confirmatory analysis were obtained is hardly surprising and does not constitute sound validity evidence.

No costs were identified for using the Three-item Loneliness Scale, and administration and scoring costs will be very low.

#### **8.4 Discussion and Recommendations**

It is widely reported that there is a continuum from perceived full social functioning to complete social isolation, and that those who are isolated are at risk of a range of health problems.

The instruments measuring this continuum can be divided into two types: those that are concerned with reporting social participation, networks, support or social contact and those which focus on social isolation or loneliness. In general, there is also a divide between so-called objective measurement of the number of social contacts and the more subjective personal assessment of either satisfaction with social contacts or feelings of the depth of loneliness. The literature is suggestive that it is the latter that is more important.

This review has focussed on instruments that are primarily concerned with identifying those at risk of social isolation so that appropriate interventions can be devised. It is the socially isolated who are, from a health perspective, at risk. This study, then, is concerned with perceived social isolation because of the emotional impact it has upon people suffering dementia or cognitive impairment (e.g. people with dementia who have loving caregivers may, on objective measures, have a high number of social contacts, but they may feel isolated if they forget these contacts).

The purpose of this review, within this general framework, was to assess and recommend instruments for use in Australian studies of geriatric care for those with cognitive impairment, usually from dementia. Within the broad study perspective above, there are two possible approaches.

On the one hand, it would be possible to recommend instruments that measured the number of social contacts, the number of events or activities participated in or the number of people in a person’s social network. Whilst this approach is intuitively appealing, the theoretical model behind it assumes that the quantity of social contacts or activities makes for good social functioning

---

32 High reliability and high responsiveness do not necessarily go together. Think of an instrument with 2 items: do you have friends (Yes/No), and are you friendless (Yes/No). After recoding the second item, we might expect a very high estimate of reliability since there will be almost no variance associated with the responses. But this will clearly be a non-responsive instrument because people will not change their response sets due to the fact that most people will already be at the ceiling (Yes) and would only change their endorsement if there was a catastrophic change in their circumstances. It is likely this situation applied to the Three-Item Loneliness Scale for the authors report that in their first study where the 3-category response set was used the mean score was 3.89 (sd = 1.34) where the scale range was 3.00 to 9.00. The implication is that ~56% of cases were at the floor and scored 3.

among the elderly. The literature is, at best, only partly supportive of this position for two reasons. In general, it suggests that mental health improves over the lifespan due to a decrease in negative affect and that ageing involves socio-emotional selectivity (Ganz, et al. 1998; Hjermstad, et al. 1998; Michelson, et al. 2000; Schwarz and Hinz, 2001; Sullivan, et al. 1995), leading to increased importance on the quality of social relationships rather than the number (Carstensen, 1992; Charles, et al. 2001; Revenson and Johnson, 1984). Thus it is the type of relationship that affects perceived social isolation (DiTommaso and Spinner, 1997; Iliffe, et al. 1991; Stevens and Westerhof, 2006). For those living in residential care, this is a particularly important issue because the number of social contacts may be artificially large since people may be thrown together into a community not necessarily of their own making or under their control (indeed there is ample evidence showing that many older people resist moving into residential care and find it disorienting). The second difficulty with this approach is that many individuals report being alone even when with others.

The other approach is to focus more on the subjective assessment of the individual regarding the adequacy of his/her social life. The theoretical model here is that where a person's needs are not being met, he/she will suffer perceived social isolation. This is the position that has been taken in this review. The difficulty with this perspective, however, is that it assumes sufficient insight and cognitive awareness for a person with mild dementia to be able to make meaningful self-assessments and reports. The literature generally suggests that those with severe dementia (say, <10 on the MMSE) may lack this insight (Baro, et al. 2006; Mozley, et al. 1999; Włodarczyk, et al. 2004).

Following a careful search of the literature, seven scales were identified for review: the DUKE Functional Social Support Questionnaire (Broadhead, et al. 1988), the Friendship Scale (Hawthorne, 2006), the Loneliness Scale (De Jong Gierveld and Kamphuis, 1985; De Jong Gierveld and Tilburg, 2006), the Medical Outcomes Study Social Support Survey (Sherbourne and Stewart, 1991), the Norbeck Social Support Questionnaire (Norbeck, 1984; Norbeck, et al. 1981; Norbeck, et al. 1983), the Social Support Questionnaire (Sarason, et al. 1987; Sarason, et al. 1983), the UCLA Loneliness Scale (Russell, et al. 1980, 1978; Russell 1996) and its short derivative the Three-item Loneliness Scale (Hughes, et al. 2004).

The literature for each scale was extracted and reviewed against the study criteria. The results are summarized in Table 32 where rankings against the criteria are weighted by their assessed importance in Australian settings. Brief comments are provided on each of the criteria.

**Table 32 Summary Assessing Social Isolation Instruments Against the Study Criteria**

Criteria	Weight	Instrument							
		DeJong	MOS	FS	Duke	Sarason	UCLA	3-IT	Norbeck
Availability of comparison data	3	1	1	1	1	1	1	1	1
Length/feasibility of instrument for inclusion in battery	2	3	2	3	3	1	2	3	1
Complexity of administration/ cognitive burden	2	2	2	3	2	2	2	2	1
Cultural Appropriateness	1	2	2	2	2	2	2	2	2
Ease of obtaining score	2	3	2	3	3	2	2	3	1
Sensitivity to dementia	3	1	1	1	1	1	1	1	1
Reliability evidence	3	3	3	2	2	3	3	1	2
Validity evidence	3	3	3	2	1	2	1	1	2
Cost of the instrument	2	3	3	3	3	3	3	3	3
Cost of instrument administration	2	3	3	3	3	3	3	3	2
<b>Weighted Total</b>		<b>54</b>	<b>50</b>	<b>50</b>	<b>45</b>	<b>45</b>	<b>43</b>	<b>42</b>	<b>36</b>

**Availability of comparison data**

No head-to-head comparative studies were identified; i.e. there is no information concerning whether the different instruments are measuring the same underlying construct or whether any one instrument outperforms other instruments.

**Instrument length**

Parsimony is important for reasons of enabling inclusion in instrument batteries and for psychometric reasons related to clarity of what is being measured. The length of instruments reviewed ranged from just 3-items to over 50 items. The three shortest instruments were the Three-item Loneliness Scale, the Friendship Scale and the De Jong Gierveld Loneliness Scale.

**Complexity and cognitive burden**

Complex instruments should be avoided, especially in studies involving those with limited cognitive capacity. This suggests that instruments should have simple and straightforward items and scoring systems for both ease of administration and to minimise cognitive burden. Several of the short instruments in this review, had difficult items or item responses. For example, although the DUKE Functional Social Support Questionnaire is generally easy to understand, some item stems are long and complex (e.g. I get chances to talk to someone about problems at work or with my housework). Similarly, although De Jong Gierveld Loneliness Scale is easy to understand and

administer, there may be some difficulties with the response categories being understood (Yes!, Yes, More or less, No, No!). The opposite problem was encountered with the Three-item Loneliness Scale due to the limited response set leading to over-endorsement of values at the floor (~56% of cases obtained these extreme scores). The simplest instrument was the Friendship Scale.

### **Cultural appropriateness**

There is no reported research involving any of these instruments examining whether the construct of social support/social isolation is culturally bound in any way; there does not appear to have been any cross-cultural validation work done on any of the measures reviewed. In short, there was no evidence for any instrument on this criterion referring to appropriate use by CALD or illiterate clients or with an interpreter. All instruments were therefore ranked similarly.

### **Ease of scoring**

Scoring ease will assist with instrument acceptance in the field by clinicians. Accordingly, those instruments with simple scoring algorithms which, if necessary, can be applied during interview are preferred. These were the DUKE Functional Social Support Questionnaire, the Friendship Scale, the De Jong Gierveld Loneliness Scale, and the Three-item Loneliness Scale.

### **Sensitivity to dementia**

No studies were identified which reported the use of any of the instruments in samples of people with dementia or cognitive impairment. The exception was the Friendship Scale which had been used with the mildly cognitively impaired (MMSE <24). This is an important limitation of this review because it implies that the recommendations from this study are speculative.

### **Reliability evidence**

All the instruments reviewed had published evidence suggesting they met the general criteria for reliability. The Three-item Loneliness Scale was downgraded on this criterion because its reliability would have been a function of endorsement of extreme values by the majority of respondents. The instruments with the most published reliability evidence, across several different samples, were the De Jong Gierveld Loneliness Scale, the Medical Outcomes Study Social Support Survey, the Sarason Social Support Questionnaire and the UCLA Loneliness Scale. Whether popularity, as measured by the number of publications, should be taken into account is, however, doubtful evidence.

### **Validity evidence**

As shown in the detailed instrument reviews above, there was considerable variation in the available validity evidence. What is striking about the literature is the, generally, limited approach adopted by instrument designers to the importance of this aspect of their work. For example, several of the instruments were developed among US college students; whether the concerns of 16-20 year olds represent those of the general population or of older adults with cognitive impairment or dementia is highly doubtful.

Similarly, as the detailed reviews show, in several instances validity evidence was claimed where instruments were correlated against psychological constructs rather than against known measures of the construct of interest. These observations suggest that although there may be a considerable amount of published material on a particular instrument, this may not reflect substantial validity evidence. The only instrument reviewed for which there is evidence of careful conceptualisation, construction using community samples, and careful testing over time in different populations, is the De Jong Gierveld Loneliness Scale.

### **Instrument costs**

None of the instruments reviewed appears to have been commercialized; no commercial websites were identified for any of the instruments and no copyright costs were identified. All the instruments appear to be available free to users, subject to journal copyright permissions.

### **Instrument administration costs**

All the instruments were designed for self-completion. This judgement, however, is subject to the caveat that in samples with dementia or mild cognitive impairment, the longer and more difficult instruments may require interviewer-administration thereby adding to research costs. This is particularly likely with the Norbeck Social Support Questionnaire.

### **Administration mode**

An additional issue pertinent to the use of instruments in dementia studies is the variation in assessments between self-completion, interviewer-administered completion and proxy-completion. Long-standing research suggests that administration mode and interviewer reinforcement can lead to increases of 20-30% in health conditions reported (Marquis, 1970; Sigelman, et al. 1981a; Sigelman, et al. 1981b) and it has been shown that proxy-report compared with self-assessment understates health and quality of life (Herrman, et al. 2002) and may be invalid altogether (Cummins, 2002). Importantly, none of the instruments reviewed here appear to have been tested for the effects of administration mode or proxy-report. The exception is the De Jong Gierveld Loneliness Scale where some research was reported in the user manual suggesting that scores systematically varied by administration mode.

## **8.4.1 Recommendations**

Given the review findings, none of the reviewed instruments can be given an unqualified recommendation for use in Australian studies with older adults who have cognitive impairment or dementia.

Subject to this finding, the standout instrument was the De Jong Gierveld Loneliness Scale. The reasons were that it was carefully conceived over a very substantial period of time, that it was developed in population samples (including older adults), and that there is a very substantial body of evidence supporting its reliability and validity. The reason the De Jong Gierveld Loneliness Scale, especially the short 6-item version, cannot be recommended outright is that the response categories may be inappropriate for use in Australian samples of people with cognitive impairment. However, a study could easily be completed to undertake a linguistic validation of this instrument for Australian use and this is recommended

The two other instruments that performed relatively well against the criteria were the Friendship Scale and the Medical Outcomes Study Social Support Survey. The Friendship Scale generally performed well on all criteria; it is a short, easy to use and score. It is a scale that was developed in samples of older adults that appears to be reliable, valid and sensitive. The limitation is that it is a new scale that has been published in just one paper to date. The Medical Outcomes Study Social Support Survey is a well-conceptualised and developed instrument. In general, it performed well against the study criteria, with the exception of those criteria related to instrument length (instrument length, cognitive burden, cultural appropriateness and scoring).

Given this situation, it is further recommended:

1. That the three instruments which performed well (the De Jong Gierveld Loneliness Scale, the Friendship Scale and the Medical Outcomes Study Social Support Survey) be trialled in at least one large dementia study for the explicit purpose of identifying the instrument to be recommended for future use. Whilst this would impose an immediate burden for, say, 3 to 5 years, it would enable many of the questions raised in this report regarding the validity of these instruments to be thoroughly investigated in an Australian context.
2. That explicit modification to the De Jong Gierveld Loneliness Scale and the Medical Outcomes Study Social Support Survey be tested. These modifications are revision of the De Jong Gierveld Loneliness Scale response set (which would need to be tested in #1), and a reduction in the

number of items in the Medical Outcomes Study Social Support Survey (which could be undertaken with the #1 data).

3. That the three instruments which performed well be tested in a trial for the effect of administration mode on scores given that there are good reasons for limiting self-completion among those with moderate or severe cognitive impairment. Three methods of administration should be directly compared (self-completion without assistance, interviewer-assisted completion, and proxy-completion) both cross-sectionally and longitudinally in order to develop algorithms for weighting enabling score equivalence across administration mode. This would overcome issues related to the cognitive impairment of respondents and meet the need to collect outcome efficacy data relating to program evaluation.

4. That from any study carried out under recommendations #1, #2 or #3, a statistically-derived single item measure be identified for use in everyday clinical consultations.

## References

Adams KB, Sanders S and Auth EA (2004) *Loneliness and depression in independent living retirement communities: risk and resilience factors*. Ageing and Mental Health. Vol. 8, No.6, pp.475-85.

Alemi F, Stephens R, Llorens S, et al. (2003) *The Orientation of Social Support measure*. Addictive Behaviors. Vol. 28, No.7, pp.1285-98.

Alpass FM and Neville S (2003) *Loneliness, health and depression in older males*. Ageing and Mental Health. Vol. 7, No.3, pp.212-216.

Applebaum F (1978) *Loneliness: a taxonomy and psychodynamic view*. Clinical Social Work Journal. Vol. 6, No.1, pp.13-20.

Austin BA (1983) *Factorial structure of the UCLA Loneliness Scale*. Psychological Reports. Vol. 53, No.3, Pt 1, pp.883-9.

Bal S, Crombez G, Van Oost P, et al. (2003) *The role of social support in well-being and coping with self-reported stressful events in adolescents*. Child Abuse and Neglect. Vol. 27, No.12, pp.1377-95.

Balandin S, Berg N and Waller A (2006) *Assessing the loneliness of older people with cerebral palsy*. Disability and Rehabilitation. Vol. 28, No.8, pp.469-79.

Banks MR and Banks WA (2002) *The effects of animal-assisted therapy on loneliness in an elderly population in long-term care facilities*. J Gerontol A Biol Sci Med Sci. Vol. 57, No.7, pp.M428-32.

Baro E, Ferrer M, Vazquez O, et al. (2006) *Using the Nottingham Health Profile (NHP) among older adult inpatients with varying cognitive function*. Quality of Life Research. Vol. 15, No.4, pp.575-85.

Barron CR, Foxall MJ, Von Dollen K, et al. (1994) *Marital status, social support, and loneliness in visually impaired elderly people*. Journal of Advanced Nursing. Vol. 19, No.2, pp.272-80.

Baumeister RF and Leary MR (1995) *The need to belong: desire for interpersonal attachments as a fundamental human motivation*. Psychological Bulletin. Vol. 117, No.3, pp.497-529.

Bergman-Evans B (2004) *Beyond the basics. Effects of the Eden Alternative model on quality of life issues*. Journal of Gerontological Nursing. Vol. 30, No.6, pp.27-34.

Berkman L and Syme L (1979) *Social networks, host resistance and mortality: a nine-year follow-up study of Alameda County residents*. American Journal of Epidemiology. Vol. 109, No.2, pp.186-204.

Blumel JE, Castelo-Branco C, Cancelo MJ, et al. (2004) *Relationship between psychological complaints and vasomotor symptoms during climacteric*. Maturitas. Vol. 49, No.3, pp.205-10.

- Bova C (2001) *Adjustment to chronic illness among HIV-infected women*. Journal of Nursing Scholarship. Vol. 33, No.3, pp.217-23.
- Bovier PA, Chamot E and Perneger TV (2004) *Perceived stress, internal resources, and social support as determinants of mental health among young adults*. Quality of Life Research. Vol. 13, No.1, pp.161-70.
- Bowlby J (1971) *Attachment*. Pelican, London.
- Broadhead W (1988) *Duke-UNC functional social support questionnaire (DUFSS): instructions for use*. Duke University Medical Centre. Report nr 919-286-9896.
- Broadhead WE, Gehlbach SH, de Gruy FV, et al.(1988) *The Duke-UNC Functional Social Support Questionnaire. Measurement of social support in family medicine patients*. Medical Care. Vol. 26, No.7, pp.709-23.
- Broadhead WE, Gehlbach SH, deGruy FV, et al. (1989) *Functional versus structural social support and health care utilization in a family medicine outpatient practice*. Medical Care. Vol. 27, No.3, pp.221-33.
- Broadhead WE, Kaplan BH, James SA, et al (1983) *The epidemiologic evidence for a relationship between social support and health*. American Journal of Epidemiology. Vol. 117, No.5, pp.521-37.
- Burgoyne RW and Saunders DS (2000) *Perceived support in newly registered HIV/AIDS clinic outpatients*. AIDS Care. Vol. 12, No.5, pp.643-50.
- Cacioppo JT, Hughes ME, Waite LJ, et al. (2006) *Loneliness as a specific risk factor for depressive symptoms: cross-sectional and longitudinal analyses*. Psychology and Ageing. Vol. 21, No.1, pp.140-51.
- Calvert MM (1989) *Human-pet interaction and loneliness: a test of concepts from Roy's adaptation model*. Nursing Science Quarterly. Vol. 2, No.4, pp.194-202.
- Cantor N and Sanderson C (1999) *Life task participation and well-being: the importance of taking part in daily life*. In Kahneman D, Diener E, Schwarz N, (eds). Well-Being: The Foundations of Hedonic Psychology. New York: Russell Sage Foundation. pp.230-243.
- Carstensen LL (1992) *Social and emotional patterns in adulthood: support for socioemotional selectivity theory*. Psychology and Ageing. Vol. 7, No.3, pp.331-8.
- Cassel J (1976) *The contribution of the social environment to host resistance: the Fourth Wade Hampton Frost Lecture*. American Journal of Epidemiology. Vol. 104, No.2, pp.107-23.
- Chan CW, Hon HC, Chien WT, et al. (2004) *Social support and coping in Chinese patients undergoing cancer surgery*. Cancer Nursing. Vol. 27, No.3, pp.230-6.
- Charles ST, Reynolds CA and Gatz M (2001) *Age-related differences and change in positive and negative affect over 23 years*. Journal of Personality and Social Psychology. Vol. 80, No.1, pp.136-51.
- Chien WT, Chan SW and Morrissey J (2007) *The perceived burden among Chinese family caregivers of people with schizophrenia*. Journal of Clinical Nursing. Vol. 16, No.6, pp.1151-61.
- Chipuer HM, Bramston P and Pretty G (2003) *Determinants of subjective quality of life among rural adolescents: a developmental perspective*. Social Indicators Research. Vol. 61, No.1, pp.79-95.
- Chou KL, Jun LW and Chi I (2005) *Assessing Chinese older adults' suicidal ideation: Chinese version of the Geriatric Suicide Ideation Scale*. Ageing and Mental Health. Vol. 9, No.2, pp.167-71.
- Cobb S (1976) *Social support as a moderator of life stress*. Psychosomatic Medicine. Vol. 38, pp.300-314.
- Connelly CD (1998) *Hopefulness, self-esteem, and perceived social support among pregnant and nonpregnant adolescents*. Western Journal of Nursing Research. Vol. 20, No.2, pp.195-209.
- Cortina J (1993) *What is coefficient alpha? Examination of theory and applications*. Journal of Applied Psychology. Vol. 78, No.1, pp.98-104.

- 
- Cronbach J and Meehl P (1955) *Construct validity in psychological tests*. Psychological Bulletin. Vol. 52, No.4, pp.281-302.
- Cummins RA (2002) *Proxy responding for subjective well-being: a review*. International Review of Research in Mental Retardation. Vol. 25, pp.183-207.
- D'Attilio JP, Campbell BM, Lubold P, et al. (1992) *Social support and suicide potential: preliminary findings for adolescent populations*. Psychological Reports. Vol. 70, No.1, pp.76-8.
- de Jong Gierveld J (1971) *Social isolation and the image of the unmarried*. Sociologia Neerlandica. Vol. 7, pp.1-14.
- de Jong Gierveld J (1978) *The construct of loneliness: components and measurement*. Essence. Vol. 2, No.4, pp.221-237.
- de Jong Gierveld J (1984) *The experience of loneliness: a test of a theoretical model of loneliness*. Nederlands Tijdschrift voor de Psychologie en haar Grensgebieden. Vol. 39, No.7, pp.376-384.
- de Jong Gierveld J (1987) *Developing and testing a model of loneliness*. Journal of Personality and Social Psychology. Vol. 53, No.1, pp.119-28.
- de Jong Gierveld J (1989) *Personal relationships, social support, and loneliness*. Journal of Social and Personal Relationships. Vol. 6, No.2, pp.197-221.
- de Jong Gierveld J and Havens B (2004) *Cross-national comparisons of social isolation and loneliness: introduction and overview*. Canadian Journal on Ageing. Vol. 23, No.2, pp.109-13.
- de Jong Gierveld J and Kamphuis F (1985) *The development of a Rasch-type loneliness scale*. Applied Psychological Measurement. Vol. 9, No.3, pp.289-299.
- de Jong Gierveld J and Raadschelders J (1982) *Types of loneliness*. In Peplau LA, Perlman D, (eds.) Loneliness: A Sourcebook of Current Theory, Research and Therapy. New York: Wiley. pp.105-119.
- de Jong Gierveld J and Tilburg TV (1999) *Manual of the Loneliness Scale*. Vrije Universiteit Amsterdam, Faculty of Social Sciences, Department of Sociology. p.18.
- de Jong Gierveld J and Tilburg TV (2006) *A 6-item scale for overall, emotional and social loneliness: confirmatory tests on survey data*. Research on Ageing. Vol. 28, No.5, pp.582-598.
- de Jong Gierveld J and van Tilburg T (1999) *Living arrangements of older adults in the Netherlands and Italy: coresidence values and behaviour and their consequences for loneliness*. Journal of Cross Cultural Gerontology. Vol. 14, pp.1-24.
- DiTommaso E and Spinner B (1997) *Social and emotional loneliness: a re-examination of Weiss' typology of loneliness*. Personality and Individual Differences. Vol. 22, No.3, pp.417-427.
- Dykstra P, van Tilburg TG and de Jong Gierveld J (2005) *Changes in older adult loneliness: results from a seven-year longitudinal study*. Research on Ageing. Vol. 27, No.6, pp.725-747.
- Dykstra PA and de Jong Gierveld J (2004) *Gender and marital-history differences in emotional and social loneliness among Dutch older adults*. Can J Ageing. Vol. 23, No.2, pp.141-55.
- Ellaway A, Wood S and MacIntyre S (1999) *Someone to talk to? The role of loneliness as a factor in the frequency of GP consultations*. British Journal of General Practice. Vol. 49, pp.363-367.
- Ellis P and Hickie I (2001) *What causes mental illness?* In Bloch S, Singh B, (eds.) Foundations of Clinical Psychiatry. Melbourne: Melbourne University Press. pp.43-62.
- Farmer I, Meyer P, Ramsey D, et al. (1996) *Higher levels of social support predict greater survival following acute myocardial infarction: the Corpus Christi Heart Project*. Behavioural Medicine. Vol. 22, pp.59-66.
- Fillenbaum GG and Smyer MA (1981) *The development, validity, and reliability of the OARS multidimensional functional assessment questionnaire*. Journal of Gerontology. Vol. 36, No.4, pp.428-34.
-

- Fratiglioni L, Wang HX, Ericsson K, et al. (2000) *Influence of social network on occurrence of dementia: a community-based longitudinal study*. Lancet. Vol. 355, No.9212, pp.1315-9.
- Fromm-Reichmann F (1959) *Loneliness*. Psychiatry. Vol. 22, No.1, pp.1-15.
- Furukawa T, Yokouchi T, Hirai T, et al. (1999) *Parental loss in childhood and social support in adulthood among psychiatric patients. Group for Longitudinal Affective Disorders Study (GLADS)*. Journal of Psychiatric Research. Vol. 33, No.2, pp.165-9.
- Ganz PA, Rowland JH, Desmond K, et al. (1998) *Life after breast cancer: understanding women's health-related quality of life and sexual functioning*. Journal of Clinical Oncology. Vol. 16, No.2, pp.501-14.
- Geller JS (2004) *Loneliness and pregnancy in an urban Latino community: associations with maternal age and unscheduled hospital utilization*. Journal of Psychosomatic Obstetrics and Gynaecology. Vol. 25, No.3-4, pp.203-9.
- Gerson AC and Perlman D (1979) *Loneliness and expressive communication*. Journal of Abnormal Psychology. Vol. 88, No.3, pp.258-61.
- Gigliotti E (2002) *A confirmation of the factor structure of the Norbeck Social Support Questionnaire*. Nursing Research. Vol. 51, No.5, pp.276-84.
- Gigliotti E (2006) *A confirmatory factor analysis of situation-specific Norbeck Social Support Questionnaire items*. Nursing Research. Vol. 55, No.3, pp.217-22.
- Grace SL, Evindar A, Kung TN, et al. (2004) *Automatic referral to cardiac rehabilitation*. Medical Care. Vol. 42, No.7, pp.661-9.
- Grant C, Goodenough T, Harvey I, et al. (2000) *A randomised controlled trial and economic evaluation of a referrals facilitator between primary care and the voluntary sector*. British Medical Journal. Vol. 320, No.7232, pp.419-23.
- Hartshorne TS (1993) *Psychometric properties and confirmatory factor analysis of the UCLA loneliness scale*. Journal of Personality Assessment. Vol. 61, No.1, pp.182-95.
- Havens B, Hall M, Sylvestre G, et al. (2004) *Social isolation and loneliness: differences between older rural and urban Manitobans*. Canadian Journal on Ageing. Vol. 23, No.2, pp.129-40.
- Hawkley LC, Browne MW and Cacioppo JT (2005) *How can I connect with thee? Let me count the ways*. Psychological Science. Vol. 16, No.10, pp.798-804.
- Hawthorne G (2006) *Measuring social isolation in older adults: development and initial validation of the Friendship Scale*. Social Indicators Research. Vol. 77, No.3, pp.521-548.
- Hawthorne G, Richardson J and Osborne R (1999) *The Assessment of Quality of Life (AQoL) instrument: a psychometric measure of health related quality of life*. Quality of Life Research. Vol. 8, pp.209-224.
- Hays RD and DiMatteo MR (1987) *A short-form measure of loneliness*. Journal of Personality Assessment. Vol. 51, No.1, pp.69-81.
- Hector-Taylor L and Adams P (1996) *State versus trait loneliness in elderly New Zealanders*. Psychological Reports. Vol. 78, No.3, Pt.2, pp.1329-30.
- Heinonen H, Volin L, Uutela A, et al. (2001a) *Gender-associated differences in the quality of life after allogeneic BMT*. Bone Marrow Transplant. Vol. 28, No.5, pp.503-9.
- Heinonen H, Volin L, Uutela A, et al. (2001b) *Quality of life and factors related to perceived satisfaction with quality of life after allogeneic bone marrow transplantation*. Annals of Hematology. Vol. 80, No.3, pp.137-43.
- Herrman H, Hawthorne G and Thomas R (2002) *Quality of life assessment in people living with psychosis*. Social Psychiatry and Psychiatric Epidemiology. Vol. 37, No.11, pp.510-518.

- Hjermstad MJ, Fayers PM, Bjordal K, et al. (1998) *Health-related quality of life in the general Norwegian population assessed by the European Organization for Research and Treatment of Cancer Core Quality-of-Life Questionnaire: the QLQ=C30 (+ 3)*. Journal of Clinical Oncology. Vol. 16, No.3, pp.1188-96.
- Hobfoll SE and Walfisch S (1984) *Coping with a threat to life: a longitudinal study of self-concept, social support, and psychological distress*. American Journal of Community Psychology. Vol. 12, No.1, pp.87-100.
- Holmen K, Ericsson K and Winblad B (2000) *Social and emotional loneliness among non-demented and demented elderly people*. Archives of Gerontology and Geriatrics. Vol. 31, No.3, pp.177-192.
- House J, Robbins C and Metzner H (1982) *The association of social relationships and activities with mortality: prospective evidence from the Tecumseh Community Health Study*. American Journal of Epidemiology. Vol. 116, No.1, pp.123-140.
- Howat A, Veitch C and Cairns W (2006) *A descriptive study comparing health attitudes of urban and rural oncology patients*. Rural and Remote Health. Vol. 6, No.4, pp.563.
- Hsu LR, Hailey BJ and Range LM (1987) *Cultural and emotional components of loneliness and depression*. Journal of Psychology. Vol. 121, No.1, pp.61-70.
- Hughes ME, Waite LJ, Hawkey LC, et al. (2004) *A short scale for measuring loneliness in large surveys*. Research on Ageing. Vol. 26, No.6, pp.655-672.
- Hunt SM, McKenna SP, McEwen J, et al. (1981) *The Nottingham Health Profile: subjective health status and medical consultations*. Social Science and Medicine. Vol. 15A, No.3, Pt.1, pp.221-9.
- Iliffe S, Haines A, Stein A, et al. (1991) *Assessment of elderly people in general practice. 3. Confiding relationships*. British Journal of General Practice. Vol. 41, No.352, pp.459-61.
- Jongenelis K, Pot AM, Eisses AMH, et al. (2004) *Prevalence and risk indicators of depression in elderly nursing home patients: The AGED study*. Journal of Affective Disorders. Vol. 83, No.2-3, pp.135-142.
- Kahn RL (1979) *Ageing and social support*. In Riley MW (ed.) Ageing from Birth to Death: Interdisciplinary Perspectives. Boulder: Westview Press. pp.77-91.
- Kawachi I, Colditz G, Ascherio A, et al. (1996) *A prospective study of social networks in relation to total mortality and cardiovascular disease in men in the USA*. Journal of Epidemiology and Community Health. Vol. 50, No.3, pp.245-251.
- Kef S (2002) *Psychosocial adjustment and the meaning of social support for visually impaired adolescents*. Journal of Visual Impairment and Blindness. Vol. 96, No.1, pp.22-37.
- Kim O (1999a) *Mediation effect of social support between ethnic attachment and loneliness in older Korean immigrants*. Research in Nursing and Health. Vol. 22, No.2, pp.169-75.
- Kim O (1999b) *Predictors of loneliness in elderly Korean immigrant women living in the United States of America*. Journal of Advanced Nursing. Vol. 29, No.5, pp.1082-8.
- Kornblith AB, Herndon JE, 2nd, Zuckerman E, et al. (2001) *Social support as a buffer to the psychological impact of stressful life events in women with breast cancer*. Cancer. Vol. 91, No.2, pp.443-54.
- Kremers IP, Steverink N, Albersnagel FA, et al. (2006) *Improved self-management ability and well-being in older women after a short group intervention*. Ageing and Mental Health. Vol. 10, No.5, pp.476-484.
- Kristjansson B, Breithaupt K and McDowell I (2001) *Development and validation of an indicator of support for community-residing older Canadians*. Int Psychogeriatr. Vol. 13 Supp.1, pp.125-35.
- Lehto-Jarnstedt US, Ojanen M and Kellokumpu-Lehtinen P (2004) *Cancer-specific social support received by newly diagnosed cancer patients: validating the new Structural-Functional Social Support Scale (SFSS) measurement tool*. Support Care Cancer. Vol. 12, No.5, pp.326-37.
- Lester D and Yang B (1992) *Social and economic correlates of the elderly suicide rate*. Suicide and Life-Threatening Behavior. Vol. 22, No.1, pp.36-47.

- 
- Levin C (2000) *Social functioning*. In Kane RL, Kane RA (eds.) Assessing Older Persons. Oxford University Press, Oxford. pp.170-199.
- Mackenzie AE and Chang AM (2002) *Predictors of quality of life following stroke*. Disability and Rehabilitation. Vol. 24, No.5, pp.259-65.
- Mahon NE and Yarcheski A (1990) *The dimensionality of the UCLA Loneliness Scale in early adolescents*. Research in Nursing and Health. Vol. 13, No.1, pp.45-52.
- Marangoni C and Ickes W (1989) *Loneliness: a theoretical review with implications for measurement*. Journal of Social and Personal Relationships. Vol. 6, pp.93-128.
- Marquis KH (1970) *Effects of social reinforcement on health reporting in the household interview*. Sociometry. Vol. 33, pp.205-215.
- Martina CMS and Stevens NL (2006) *Breaking the cycle of loneliness? Psychological effects of a friendship enrichment program for older women*. Ageing and Mental Health. Vol. 10, No.5, pp.467-475.
- May KM (1992) *Middle-Eastern immigrant parents' social networks and help-seeking for child health care*. Journal of Advanced Nursing. Vol. 17, No.8, pp.905-12.
- Mayers AM and Svartberg M (2001) *Existential loneliness: a review of the concept, its psychosocial precipitants and psychotherapeutic implications for HIV-infected women*. British Journal of Medical Psychology. Vol. 74, No.Pt 4, pp.539-53.
- McKee MD, Cunningham M, Jankowski KR, et al. (2001) *Health-related functional status in pregnancy: relationship to depression and social support in a multi-ethnic population*. Obstetrics and Gynecology. Vol. 97, No.6, pp.988-93.
- McQuellon RP, Russell GB, Rambo TD, et al. (1998) *Quality of life and psychological distress of bone marrow transplant recipients: the 'time trajectory' to recovery over the first year*. Bone Marrow Transplant. Vol. 21, No.5, pp.477-86.
- Mendes de Leon CF, Glass TA and Berkman LF (2003) *Social engagement and disability in a community population of older adults: the New Haven EPESE*. American Journal of Epidemiology. Vol. 157, No.7, pp.633-42.
- Michelson H, Bolund C, Nilsson B, et al. (2000) *Health-related quality of life measured by the EORTC QLQ-C30--reference values from a large sample of Swedish population*. Acta Oncologica. Vol. 39, No.4, pp.477-84.
- Miller PA (1997) *Support systems of nurses recovering from chemical dependency: a pilot study*. Holistic Nursing Practice. Vol. 11, No.4, pp.56-70.
- Miller TR and Cleary TA (1993) *Direction of wording effects in balanced scales*. Education and Psychological Measurement. Vol. 53, No.1, pp.51-60.
- Moorer P and Suurmeijer TPBM (1993) *Unidimensionality and cumulativeness of the Loneliness Scale using Mokken Scale Analysis for polychotomous items*. Psychological Reports. Vol. 73, No.3, Pt.2, pp.1324-1326.
- Morano CD, Cisler RA and Lemerond J (1993) *Risk factors for adolescent suicidal behavior: loss, insufficient familial support, and hopelessness*. Adolescence. Vol. 28, No.112, pp.851-65.
- Morrell CJ, Spiby H, Stewart P, et al. (2000) *Costs and benefits of community postnatal support workers: a randomised controlled trial*. Health Technology Assessment. Vol. 4, No.6, pp.1-100.
- Mozley CG, Huxley P, Sutcliffe C, et al. (1999) *'Not knowing where I am doesn't mean I don't know what I like': cognitive impairment and quality of life responses in elderly people*. International Journal of Geriatric Psychiatry. Vol. 14, No.9, pp.776-83.
- Muller RT and Lemieux KE (2000) *Social support, attachment, and psychopathology in high risk formerly maltreated adults*. Child Abuse and Neglect. Vol. 24, No.7, pp.883-900.
-

- Murphy PA, Prewitt TE, Bote E, et al. (2001) *Internal locus of control and social support associated with some dietary changes by elderly participants in a diet intervention trial*. Journal of the American Dietetic Association. Vol. 101, No.2, pp.203-8.
- Nelson CJ, Rosenfeld B, Breitbart W, et al. (2002) *Spirituality, religion, and depression in the terminally ill*. Psychosomatics. Vol. 43, No.3, pp.213-20.
- Nelson PB (1989) *Social support, self-esteem, and depression in the institutionalized elderly*. Issues in Mental Health Nursing. Vol. 10, No.1, pp.55-68.
- Nicoll M, Ashworth M, McNally L, et al. (2002) *Satisfaction with respite care: a pilot study*. Health Soc Care Community. Vol. 10, No.6, pp.479-84.
- Nolen-Hoeksema S and Ahrens C (2002) *Age differences and similarities in the correlates of depressive symptoms*. Psychology and Ageing. Vol. 17, No.1, pp.116-24.
- Norbeck JS (1984) *The Norbeck Social Support Questionnaire*. Birth Defects Original Article Series. Vol. 20, No.5, pp.45-57.
- Norbeck JS (1995) *Scoring Instructions for the Norbeck Social Support Questionnaire*. University of California.
- Norbeck JS, Lindsey AM and Carrieri VL (1981) *The development of an instrument to measure social support*. Nursing Research. Vol. 30, No.5, pp.264-9.
- Norbeck JS, Lindsey AM and Carrieri VL (1983) *Further development of the Norbeck Social Support Questionnaire: normative data and validity testing*. Nursing Research. Vol. 32, No.1, pp.4-9.
- Nunnally JC and Bernstein IH (1994) *Psychometric Theory*. McGraw-Hill, New York. 752 p.
- Pang SK, Ip WY and Chang AM (2001) *Psychosocial correlates of fluid compliance among Chinese haemodialysis patients*. Journal of Advanced Nursing. Vol. 35, No.5, pp.691-8.
- Pedhazur E and Schmelkin L (1991) *Measurement, Design and Analysis: An Integrated Approach*. Lawrence Erlbaum, Hillsdale.
- Peplau LA and Perlman D, (eds.) (1982). *Loneliness: A Sourcebook of Current Theory, Research and Therapy*. Wiley, New York.
- Perlman D, Gerson AC and Spinner B (1978) *Loneliness among senior citizens: an empirical report*. Essence. Vol. 2, pp.239-248.
- Petrak JA, Doyle AM, Smith A, et al. (2001) *Factors associated with self-disclosure of HIV serostatus to significant others*. British Journal of Health Psychology. Vol. 6, No.Pt 1, pp.69-79.
- Procidano ME and Heller K (1983) *Measures of perceived social support from friends and from family: three validation studies*. American Journal of Community Psychology. Vol. 11, No.1, pp.1-24.
- Rasclé N, Bruchon-Schweitzer M and Sarason IG (2005) *Short form of Sarason's Social Support Questionnaire: French adaptation and validation*. Psychological Reports. Vol. 97, No.1, pp.195-202.
- Reid M, Glazener C, Murray GD, et al. (2002) *A two-centred pragmatic randomised controlled trial of two interventions of postnatal support*. BJOG. Vol. 109, No.10, pp.1164-70.
- Retsinas J and Garrity P (1985) *Nursing home friendships*. Gerontologist. Vol. 25, pp.376-381.
- Revenson TA and Johnson JL (1984) *Social and demographic correlates of loneliness in late life*. American Journal of Community Psychology. Vol. 12, No.1, pp.71-85.
- Rokach A (2000) *Loneliness and the life cycle*. Psychological Reports. Vol. 86, pp.629-642.
- Routasalo PE, Savikko N, Tilvis RS, et al. (2006) *Social contacts and their relationship to loneliness among aged people - a population-based study*. Gerontology. Vol. 52, No.3, pp.181-7.

- Ruberman W, Weinblatt E, Goldberg J, et al. (1984) *Psychosocial influences on mortality after myocardial infarction*. New England Journal of Medicine. Vol. 311, pp.552-559.
- Ruiz Perez I, Rodriguez Bano J, Lopez Ruz MA, et al. (2005) *Health-related quality of life of patients with HIV: impact of sociodemographic, clinical and psychosocial factors*. Quality of Life Research. Vol. 14, No.5, pp.1301-10.
- Russell D (1982) *The measurement of loneliness*. In Peplau L, Perlman D, (eds.) Loneliness: A Sourcebook of Current Theory, Research and Therapy. John Wiley, New York. pp.81-104.
- Russell D, Peplau L and Cutrona C (1980) *The revised UCLA Loneliness Scale: concurrent and discriminant validity evidence*. Journal of Personality and Social Psychology. Vol. 39, No.3, pp.472-480.
- Russell D, Peplau LA and Ferguson ML (1978) *Developing a measure of loneliness*. Journal of Personality Assessment. Vol. 42, No.3, pp.290-4.
- Russell DW (1996) *UCLA Loneliness Scale (Version 3): reliability, validity, and factor structure*. Journal of Personality Assessment. Vol. 66, No.1, pp.20-40.
- Sadler EA, Braam AW, Broese van Groenou MI, et al. (2006) *Cosmic transcendence, loneliness, and exchange of emotional support with adult children: a study among older parents in The Netherlands*. European Journal of Ageing. Vol. 3, No.3, pp.146.
- Sand L and Strang P (2006) *Existential loneliness in a palliative home care setting*. Journal of Palliative Medicine. Vol. 9, No.6, pp.1376-87.
- Sarason I, Sarason B, Shearin E, et al. (1987) *A brief measure of social support: practical and theoretical implications*. Journal of Social and Personal Relationships. Vol. 4, pp.497-510.
- Sarason IG, Levine H, Basham RB, et al. (1983) *Assessing social support: the Social Support Questionnaire*. Journal of Personality and Social Psychology. Vol. 44, No.1, pp.127-139.
- Schaffer MA and Lia-Hoagberg B (1997) *Effects of social support on prenatal care and health behaviors of low-income women*. Journal of Obstetric, Gynecologic, and Neonatal Nursing. Vol. 26, No.4, pp.433-40.
- Schwarz R and Hinz A (2001) *Reference data for the quality of life questionnaire EORTC QLQ-C30 in the general German population*. European Journal of Cancer. Vol. 37, No.11, pp.1345-51.
- Seckel MM and Birney MH (1996) *Social support, stress, and age in women undergoing breast biopsies*. Clinical Nurse Specialist. Vol. 10, No.3, pp.137-43.
- Sherbourne C and Stewart A (1991) *The MOS social support survey*. Social Science and Medicine. Vol. 32, pp.705-714.
- Shyu YI, Tang WR, Liang J, et al. (2006) *Psychometric testing of the social support survey on a Taiwanese sample*. Nursing Research. Vol. 55, No.6, pp.411-7.
- Sigelman CK, Budd EC, Spanhel CL, et al. (1981a) *Asking questions of retarded persons: a comparison of yes-no and either-or formats*. Applied Research in Mental Retardation. Vol. 2, No.4, pp.347-57.
- Sigelman CK, Budd EC, Spanhel CL, et al. (1981b) *When in doubt, say yes: acquiescence in interviews with mentally retarded persons*. Mental Retardation. Vol. 19, No.2, pp.53-8.
- Sin MK, Kang DH and Weaver M (2005) *Relationships of asthma knowledge, self-management, and social support in African American adolescents with asthma*. International Journal of Nursing Studies. Vol. 42, No.3, pp.307-13.
- Sinha SP and Nayyar P (2002) *Social support and self-control as variables in attitude toward life and perceived control among older people in India*. Journal of Social Psychology. Vol. 142, No.4, pp.527-40.
- Solano CH (1980) *Two measures of loneliness: a comparison*. Psychological Reports. Vol. 46, No.1, pp.23-8.

- Springer MB, Newman A, Weaver AJ, et al. (2003) *Spirituality, depression, and loneliness among Jewish seniors residing in New York City*. Journal of Pastoral Care and Counseling. Vol. 57, No.3, pp.305-18.
- Steptoe A, Owen N, Kunz-Ebrecht SR, et al. (2004) *Loneliness and neuroendocrine, cardiovascular, and inflammatory stress responses in middle-aged men and women*. Psychoneuroendocrinology. Vol. 29, No.5, pp.593-611.
- Stevens N and Westerhof GJ (2006) *Marriage, social integration, and loneliness in the second half of life: a comparison of Dutch and German men and women*. Research on Ageing. Vol. 28, No.6, pp.713.
- Stewart MJ, Hirth AM, Klassen G, et al. (1997) *Stress, coping, and social support as psychosocial factors in readmissions for ischaemic heart disease*. International Journal of Nursing Studies. Vol. 34, No.2, pp.151-63.
- Stokes JP and Wilson DG (1984) *The inventory of socially supportive behaviors: dimensionality, prediction, and gender differences*. American Journal of Community Psychology. Vol. 12, No.1, pp.53-69.
- Storch EA, Roth DA, Coles ME, et al. (2004) *The measurement and impact of childhood teasing in a sample of young adults*. Journal of Anxiety Disorders. Vol. 18, No.5, pp.681-94.
- Streiner D and Norman G (2006) *Health Measurement Scales: A Practical Guide to their Development and Use*. Oxford Medical Publications, Oxford.
- Sullivan M, Karlsson J and Ware J (1995) *The Swedish SF-36 Health Survey--I. Evaluation of data quality, scaling assumptions, reliability and construct validity across general populations in Sweden*. Social Science and Medicine. Vol. 41, No.10, pp.1349-58.
- Swami V, Chamorro-Premuzic T, Sinniah D, et al. (2007) *General health mediates the relationship between loneliness, life satisfaction and depression : A study with Malaysian medical students*. Social Psychiatry and Psychiatric Epidemiology. Vol. 42, No.2, pp.161-6.
- Swindells S, Mohr J, Justis JC, et al. (1999) *Quality of life in patients with human immunodeficiency virus infection: impact of social support, coping style and hopelessness*. International Journal of STD and AIDS. Vol. 10, No.6, pp.383-91.
- Thoits P (1982) *Conceptual, methodological and theoretical problems in studying social support as a buffer against life stress*. Journal of Health and Social Behavior. Vol. 23, pp.145-159.
- Tomaka J, Thompson S and Palacios R (2006) *The relation of social isolation, loneliness, and social support to disease outcomes among the elderly*. Journal of Ageing and Health. Vol. 18, No.3, pp.359-84.
- Townsend P (1963) *The Family Life of Old People*. Penguin Books, London.
- Townsend P (1973) *Isolation and loneliness in the aged*. In Weiss RS (ed.) Loneliness: The Experience of Emotional and Social Isolation. Cambridge, MA: Massachusetts Institute of Technology. pp.175-188.
- Turner RJ (1981) *Social support as a contingency in psychological wellbeing*. Journal of Health and Social Behavior. Vol. 22, pp.357-367.
- van Baarsen B, Smit JH, Sniijders TAB, et al. (1999) *Do personal conditions and circumstances surrounding partner loss explain loneliness in newly bereaved older adults*. Ageing and Society. Vol. 19, No.4, pp.441-469.
- van Baarsen B, Sniijders TAB, Smit JH, et al. (2001) *Lonely but not alone: emotional isolation and social isolation as two distinct dimensions of loneliness in older people*. Educational and Psychological Measurement. Vol. 61, No.1, pp.119-135.
- van Tilburg T, Havens B and de Jong Gierveld J (2004) *Loneliness among older adults in the Netherlands, Italy, and Canada: a multifaceted comparison*. Canadian Journal on Ageing. Vol. 23, No.2, pp.169-180.
- Victor C, Scambler S, Bond J, et al. (2000) *Being alone in later life: loneliness, social isolation and living alone*. Reviews in Clinical Gerontology. Vol. 10, No.4, pp.407-417.

- Wagle A, Komorita NI and Lu ZJ (1997) *Social support and breast self-examination*. Cancer Nursing. Vol. 20, No.1, pp.42-8.
- Watt S, Sword W, Krueger P, et al. (2002) *A cross-sectional study of early identification of postpartum depression: implications for primary care providers from The Ontario Mother and Infant Survey*. BMC Family Practice. Vol. 3, pp.5.
- Weiss R (1974) *The provisions of social relationships*. In Rubin Z, (ed.) Doing Unto Others: Joining, Molding, Conforming, Helping, Loving. Prentice Hall, Englewood Cliffs. pp.17-26.
- Weiss RS (1973) *Loneliness: The Experience of Emotional and Social Isolation*. MIT Press, Cambridge, MA.
- Wenger C (1983) *Loneliness: a problem of measurement*. In Jerrome D, (ed.) Ageing in Modern Society: Contemporary Approaches. Croom Helm, London. pp.145-167.
- Wenger GC and Burholt V (2004) *Changes in levels of social isolation and loneliness among older people in a rural area: a twenty-year longitudinal study*. Canadian Journal on Ageing. Vol. 23, No.2, pp.115-27.
- Westaway MS, Seager JR, Rheeder P, et al. (2005) *The effects of social support on health, well-being and management of diabetes mellitus: a black South African perspective*. Ethnicity and Health. Vol. 10, No.1, pp.73-89.
- WHOQoL Group (1998) *Development of the World Health Organization WHOQOL-BREF Quality of Life Assessment*. Psychological Medicine. Vol. 28, No.3, pp.551-558.
- Williams A (1992) *Review article: measuring functioning and well-being*. In Stewart A, Ware J, (eds) The Medical Outcomes Study Approach. Durham: Duke University Press.
- Williams PD, Williams AR and Griggs C (1990) *Children at home on mechanical assistive devices and their families: a retrospective study*. Maternal-Child Nursing Journal. Vol. 19, No.4, pp.297-311.
- Williamson J, Stokoe IH, Gray S, et al. (1964) *Old people at home. Their unreported needs*. Lancet. Vol. 13, pp.1117-20.
- Wilson D, Cutts J, Lees I, et al. (1992) *Psychometric properties of the revised UCLA Loneliness Scale and two short-form measures of loneliness in Zimbabwe*. Journal of Personality Assessment. Vol. 59, No.1, pp.72-81.
- Wilson RS, Krueger KR, Arnold SE, et al. (2007) *Loneliness and risk of Alzheimer disease*. Archives of General Psychiatry. Vol. 64, No.2, pp.234-40.
- Wlodarczyk JH, Brodaty H and Hawthorne G (2004) *The relationship between quality of life, Mini-Mental State Examination, and the Instrumental Activities of Daily Living in patients with Alzheimer's Disease*. Archives of Gerontology and Geriatrics. Vol. 39, No.1, pp.25-33.
- Yu DS, Lee DT and Woo J (2004) *Psychometric testing of the Chinese version of the medical outcomes study social support survey (MOS-SSS-C)*. Research in Nursing and Health. Vol. 27, No.2, pp.135-43.
- Zakahi WR and Duran RL (1982) *All the lonely people: the relationship among loneliness, communicative competence and communication anxiety*. Communication Quarterly. Vol. 30, No.3, pp.203-209.

## 9 Measures of the Associated Symptoms of Dementia

### 9.1 Introduction

'Associated symptoms of dementia' refer to characteristics of dementia that have not been historically considered major features, such as cognitive impairment and related functional consequences, yet have a significant impact on the well-being of the person with dementia and their family and caregivers. Measuring outcomes of care, service, treatment and interventions related to the associated symptoms of dementia is an important aspect. For the purpose of the DOMS project, the assessment of associated symptoms of dementia comprise: 1) measures of global behavioural and psychological symptoms of dementia (BPSD Global, henceforward); 2) measures of delirium, which is one of the two most frequently mistaken features requiring differential diagnosis from dementia (the other commonly mistaken feature is depression); and 3) measures of particular symptoms of BPSD including aggression, agitation, anxiety, apathy, and depression. This section provides a set of comprehensive reviews for each of the three categories and recommendations.

#### 9.1.1 Initial Search Strategies

The initial overarching literature search strategy for dementia instruments (refer Section 2.5) identified 138 instruments that assessed the associated symptoms of dementia. Following the search strategy (text-word search), a CD-Rom was produced containing an Endnote database for each of the identified instruments with abstracts as well as relevant papers, which were then distributed among the review teams. Based on this, an Impact sheet was developed for consideration by the review teams and the DOMS-EMG. This considered the MEDLINE, text and web impacts, presence in instrument databases and its use in clinical practice for each instrument. The latter was based on literature searches, NEP and field surveys as well as clinical feedback.

A teleconference among the review teams and the DOMS-EMG was held to discuss the initial categories for associated symptoms (e.g., BPSD Global, Delirium, Aggression, Agitation, etc.) and the initial list of measures for each category. This process produced a list of 29 BPSD global measures, 11 delirium measures, 9 Aggression measures, 12 Agitation measures, 19 Anxiety measures, 2 Apathy measures, and 25 Depression measures. Further discussions via emails and teleconference took place among the review teams and the DOMS-EMG, and the leading contenders were identified for each category based on more detailed examination of the literature.

Decisions as to how leading contenders were reduced to the final review list are discussed individually throughout the relevant sections below. However, all contender instruments were examined in terms of domains/subdomains, applicability/stages, self-completed/proxy, availability/cost, training/manual, administration time, number of citations in the literature, report of the psychometric properties with evidence of reliability and validity, use in the practice, and overall judgment about each instrument.

### 9.2 Behavioural and Psychological Symptoms of Dementia (BPSD)

Behavioural and Psychological Symptoms of Dementia (BPSD), also known as Neuropsychiatric symptoms (NPS), have been considered as most upsetting and disconcerting not only for the person who is experiencing the conditions but also for those living with or providing care for the person, and are considered a leading cause of institutionalisation (International Psychogeriatric Association; IPA, 2003). Whilst described under 'associated symptoms' in the DOMS project, BPSD are now established as defining elements of dementia, along with cognitive and functional impairment, as evidenced by more recent studies reporting BPSD as common in almost all people with dementia. Steinberg, et al. (2006, cited in Lyketsos, 2007) report about 98% of people with

dementia developed one or more BPSD. The latest IPA consensus statement on defining and measuring treatment benefits in dementia acknowledges the importance of assessing BPSD as meaningful outcomes of dementia interventions, along with caregiver outcomes and quality of life measures, that were traditionally neglected (Katona, et al. 2007).

However, as discussed earlier in the Section 3.2.6, no diagnostic nomenclature has provided clear descriptions or information on the severity, course or types of BPSD. The only major advance was made when the DSM-IV (APA, 1994) was released in which the ‘specifier’ phrase ‘with behavioural disturbance’ was added to the 4th revision, and yet without any guidance how one might interpret ‘behavioural disturbance’ (Caine, 1996), and assess its types, causes, prognoses and outcomes appropriately. This poses a difficult situation when validating an instrument for BPSD, as there is no gold standard to measure against. The definition of BPSD by the IPA (2003) is by far the most commonly recognised in the discipline of psychogeriatrics. The IPA provides the definition of BPSD in two groups, behavioural symptoms and psychological symptoms as follows (IPA, 2003, p.5):

**Behavioral symptoms:**

Usually identified on the basis of observation of the patient, including physical aggression, screaming, restlessness, agitation, wandering, culturally inappropriate behaviors, sexual disinhibition, hoarding, cursing and shadowing.

**Psychological symptoms:**

Usually and mainly assessed on the basis of interviews with patients and relatives; these symptoms include anxiety, depressive mood, hallucinations and delusions.

The most commonly covered areas of measuring BPSD Global include: delusions, hallucinations, aggression, agitation, depression, anxiety, apathy, disinhibition, irritability, and aberrant behaviours in terms of sleeping, eating and sexual behaviours. Whilst most symptoms such as hallucinations, delusions, depression, apathy and anxiety denote common definitions across literature there appear mixed methods of categorising BPSD and defining some of the individual symptoms. For example, Cohen-Mansfield and Billing (1986, cited in IPA, 2003, p.10) define agitation as “inappropriate verbal, vocal or motor activity that is not judged by an outside observer to result directly from the needs or confusion of the person”, and propose four types of agitation including physically and verbally non-aggressive behaviours, and physically and verbally aggressive behaviours<sup>33</sup>, which clearly assign aggression under agitation. Lyketsos (2007) argues that classifying BPSD based on observational information with a focus on explicit behaviours may not provide accurate assessment of BPSD. For example agitation defined in Cohen-Mansfield’s measures can often be associated with underlying mental conditions of delusions, hallucinations, depression, anxiety, and so on. Lyketsos, Breitner and Rabins (2001) suggest it may be more appropriate and meaningful to measure specific symptom groups such as affect, apathy and psychosis rather than measuring global scores of BPSD. More research is needed to confirm this new classifying method.

Cummings (1996) appropriately argues that choice of best tools depends largely on the type(s) of behaviours to be measured while considering various characteristics of instruments available. He states both BPSD Global measures and measures focusing on specific aspects of BPSD are useful in different ways, in that global measures provide information on the overall condition of the person with dementia while specified measures can provide more explicit information about

---

<sup>33</sup> Physically non-aggressive behaviours: general restlessness, repetitive mannerisms, pacing, trying to get to a different place, handling things inappropriately, hiding things, inappropriate dressing or undressing, repetitive sentences  
Verbally non-aggressive behaviours: negativism, does not like anything, constant requests for attention, verbal bossiness, complaining or whining, relevant interruptions, irrelevant interruptions.  
Physically aggressive behaviors: hitting, pushing, scratching, grabbing things, grabbing people, kicking and biting.  
Verbally aggressive behaviors: screaming, cursing, temper outbursts, making strange noises (Cohen-Mansfield and Billing 1986, cited in IPA 2003, p.11).

particular attributes of BPSD that clinicians or researchers aim to evaluate. The choice of the best tools also depends on the best available sources of information.

Sources of information for measuring BPSD can be divided into four categories (Cummings, 1996): informal carers such as family members or friends of the person with dementia who provide care at home in on-going bases; professional carers; direct observations by physicians; and the person with dementia by self report. Ideally self-reporting from the person with dementia is likely to produce accurate information about their own condition; however, this method is limited to the patient who is still in the early stage of dementia. For those with dementia living in the community informal carers may be the best source of information. However, findings need to be considered in light of the level of carer's understanding about the terminologies used in the tool and the carer's level of stress and mood, as these factors may impact on the carer's judgment. Tools that assess the level of caregiver distress in conjunction with severity of BPSD are hence more useful than measuring a single aspect. Professional carers such as nurses and care workers may provide more accurate information on the grounds that they are better educated in understanding the terminologies and phenomenology of BPSD and have sufficient knowledge about the person for whom they provide care. This may not be always possible if the professional carer works on a casual basis, has recently started the position, or works in one shift (day or night) always. Physician's direct observation may be more reliable but it is not feasible for the physician to observe the person with dementia continuously and information is only a snapshot of the person's condition (Cummings, 1996).

It is beyond the scope of the DOMS project to provide an in-depth exploration of diverse definitions of each individual BPSD, or an analysis of diverse methods of grouping, however it is important to acknowledge consensus on groupings of BPSD is yet to be established and there may need to be new development of better outcome measures to assess BPSD if different approaches of grouping are introduced. The reviews of BPSD Global in this chapter are based on currently best available instruments.

### 9.2.1 Decision Making Strategies

As shown in Table 33 below, the nine strong/leading contenders for BPSD global measures include:

- Behavioral Pathology in Alzheimer's Disease Rating Scale (BEHAVE-AD) (Reisberg, Borenstein, et al. 1987);
- CERAD Behavior Rating Scale for Dementia (BRSD) (Tariot, 1996; Patterson, Mack, Mackell, Thomas, et al. 1997; Mack, Patterson and Tariot, 1999);
- Columbia University Scale for Psychopathology in Alzheimer's Disease (CUSPAD) (Devanand, et al. 1992);
- Dementia Behavior Disturbance Scale (DBDS) (Baumgarten, Becker, et al. 1990);
- Manchester and Oxford Universities Scale for the Psychopathological Assessment of Dementia (MOUSEPAD) (Allen, Gordon, Hope and Burns, 1996);
- Neurobehavioral Rating Scale (NRS) (Levin, High, et al. 1987);
- Nursing Home Behavior Problem Scale (NHBPS) (Ray, Taylor, Lichtenstein and Meador, 1992);
- Neuropsychiatric Inventory (NPI) (Cummings, 1994); and
- Revised Memory and Behavior Problem Checklist (RMBPC) (Teri, et al. 1992).

The review teams paid particular attention to the number of citations, the coverage of BPSD, application in a range of settings, and administration time. The final five review list was developed after a teleconference, which included the NPI, BEHAVE-AD, CERAD-BRSD, NRS and DBDS.

CUSPAD was not selected for detailed review due to its overly detailed emphasis on delusional aspects of BPSD and its insufficient coverage of behavioural issues; the MOUSEPAD for the lengthy time required for administration; NHBPS for its focus on difficulties perceived by nursing staff; and RMBPC for its limited coverage on the various aspects of BPSD, and limited applicability in the nursing home setting (e.g. staff may not have sufficient knowledge about residents with dementia to answer some of the questions).

Despite the relatively small number of citations identified compared to the other four selected instruments, the DBDS was chosen considering its proven applicability in both community and nursing home settings, easy availability and implementation, and short time required for administration.

**Table 33 Decision Summary of the BPSD Global Leading Contenders**

	Domains/Sub domains	Applicability/ Stage	Patient	Proxy	Availability/ Cost	Training/Manual	Admin time
<b>BEHAVE-AD</b>	25 items grouped into 7 major categories. Paranoid and delusional ideation, hallucinations, activity disturbance, aggressiveness, diurnal rhythm disturbances, affective disturbances and anxieties and phobias.	All stages		✓	Free	No formal training, though psychiatric language in the scale means that it should be used by a person with some health training.	20 minutes (interview). May be faster if self-completed by residential care staff.
<b>CERAD-BRSD</b>	46 items grouped into 6 domains. Depressive features, inertia, psychotic features, vegetative features, Irritability/aggression, behavioural dysregulation.	All stages		✓	Written permission required. Cost of \$US85.00 for instrument plus instruction manual.	Provided with instrument as per availability/cost section.	20 to 30 minutes.
<b>CUSPAD</b>	28 items grouped into 5 domains. Delusions, hallucinations, illusions, behavioural disturbance and depression.	All stages		✓	Free, can be reproduced with permission from the American Medical Association, available in the original paper.	No formal training required.	15-20 minutes.
<b>DBDS</b>	28 items. Sub-domains not defined.	All stages		✓	Free	No formal training required.	15 minutes. May be faster if self-complete by carer or residential care staff.
<b>MOUSEPAD</b>	59 items consisting of domains such as delusions, hallucinations, misidentifications, reduplications, and behavioural changes in dementia (walking, eating, sleep, sexual behaviour, aggression, other types of behaviour in the last month).	All stages		✓	Free, available in the original paper, permission required from the author.	No formal training required although some training should be provided for those who do not have psychiatric training background.	15-30 minutes.

<b>NRS</b>	27 items consisting of cognitive deficits, psychiatric symptoms and behavioural disturbances.	All stages		✓	Free	No formal training required, though originally completed after a structured interview and clinical observation.	15-20 minutes.
<b>NHBPS</b>	29 item inventory, six sub-scales include: uncooperative/aggressive, irritational/restless, sleep problems, annoying behaviour, inappropriate behaviour and dangerous behaviour.	All stages		✓	Free	No formal training required.	3-5 minutes.
<b>NPI</b>	12 domains with 5-8 items per domain. 10 behavioural symptom sub-sections: delusions, hallucinations, agitation, depression, anxiety, euphoria, apathy, disinhibition, irritability, aberrant behaviours. 2 neuro-vegetative change sub-sections: night time behaviours, appetite changes. One item on caregiver distress engendered by behavioural symptoms.	All stages		✓	Free for all users	Training not required, but training pack available if required.	10-15 minutes.
<b>RMBPC</b>	24 item checklist, provide one total score and three sub-scores for memory related, depression and disruptive behaviours.	All stages		✓	Free	Training not required.	15-20 minutes.

	Citations	Psycho-metrics	Use in Practice (to date)	Judgments/Comments
<b>BEHAVE-AD</b>	More than 120 journal articles.	Good	Assessment of BPSD in the community, outpatients and residential care. Measurement of change in pharmacological and non-pharmacological trials.	This was one of the first instruments to measure BPSD, and has been improved with the addition of a frequency scale. It has been used extensively in clinical studies and trials.
<b>CERAD-BRSD</b>	26 journals 1 book	Very good	Assess/evaluate: Behavioural disturbance in persons with dementia or cognitive impairment. Effectiveness of drug treatments or other non-pharmacological interventions.	Instrument is quite long, taking about 20-30 minutes to complete. Training and financial cost involved. Instrument has been used in both the clinical research settings. Normative data is available. Instrument is part of the CERAD battery, which may be advantageous in terms of future translations and ongoing research.
<b>CUSPAD</b>	41 articles	Good	Assessment of non-cognitive symptoms of dementia in psychiatric and psychogeriatric, memory clinics.	Simple and easy to administer, designed specifically for people with AD. Focuses more on psychoses, in particular delusion. Many of the components are measured 'yes' or 'no', hence of limited value for rating the severity of dementia.
<b>DBD</b>	18 articles	Good	Assessment of behavioural disturbance in clinical settings and residential care.	There have been relatively few research studies with this instrument, although psychometric properties are promising.
<b>MOUSEPAD</b>	7 articles	Moderate to good	Assessment of psychiatric symptoms and behavioural changes in people with dementia. Use both in in-patient and home care setting. Only a few studies available that examine the tool.	Administered by experienced clinician. Requires a lengthy time to implement and yet does not cover all BPSD. It does not measure depression and is recommended to use the Cornell Scale for Depression in Dementia as an additional measure.
<b>NRS</b>	84 articles	Good	Used extensively to assess psychiatric symptoms in head injury patients. Less frequently used for persons with dementia, though has been used in a few pharmaceutical trials.	This scale is not a pure measure of behavioural disturbance and was not developed specifically for persons with dementia.
<b>NHBPS</b>	23 articles	Good	Assesses behavioural problems encountered in nursing homes and other chronic care facilities, in particular to examine predictors of	Administered by nurses and nursing assistants. Quick and easy to implement. Measures general behavioural disturbance only, does not cover all BPSD, such as passive behaviours. Rating (frequency) is based on observation of behaviours during the last three days to

			chemical and physical restraint. Also monitors the severity of the behavioural problems. Both research and clinical utility.	improve reliability, however this may not detect more infrequent behaviours that may impact significantly on the resident's well-being.
<b>NPI</b>	158 Journal articles 2 books 1 conference presentation	Very good to excellent	Assesses: Psychopathology in dementia to distinguish different causes and include symptoms rare in Alzheimer's disease but characteristic of fronto-temporal dementias. Caregiver distress associated with behavioural symptoms. Able to be used in the community and in all health care settings.	Frequently used in all health care settings and in research as a comprehensive measure for a range of behavioural dysfunctions in dementia. Has excellent psychometric properties in comparison to other measures. Has been adapted for use in the nursing home and GP surgery without compromising its validity and reliability, however the shortened version NPI-Q cannot be relied on for intervention studies. Some researchers suggest NPI should be complemented by clinical assessment of specific features of behavioural dysfunction to avoid the possibility of a halo effect when NPI is employed by regular caregivers alone. Translated for use in several languages without compromising validity and reliability.
<b>RMBPC</b>	70 articles	Good	Assesses both the frequency of behaviours observed among people with dementia and caregiver reactions.	One of the most widely used instruments for BPSD. It can be used in both institutional and community settings. However, it does not cover the full range of BPSD. It is difficult to obtain accurate scores in repeated measures during a clinical trial as a different observer may produce different interpretations of behaviours.

## 9.2.2 Neuropsychiatric Inventory (NPI)

The Neuropsychiatric Inventory (NPI) was developed and validated initially by Cummings (1994) to assess psychopathology in the person with dementia, to help distinguish between different causes of dementia, and includes items pertaining to symptoms known to be rare in Alzheimer's disease, but are characteristic of fronto-temporal dementias. As a caregiver informed rating scale, it also assesses the level of caregiver distress engendered by each of the neuropsychiatric disorders. It was initially developed by for use in the in-patient clinical setting, but is also routinely used in the community setting by a General Medical Practitioner. The nursing home version is used by formal caregivers, such as nurses, and is also used in research, including drug and treatment trials (see below).

The NPI contains 12 domains. These comprise 10 sub-sections examining behavioural areas (delusions; hallucinations; agitation; depression; anxiety; euphoria; apathy; disinhibition; irritability; aberrant behaviours, night-time behaviours) and 2 types of neuro-vegetative change (appetite and eating disorders), each with 5-8 items. These items fall into five factors. A screening question is asked first for each item, followed by sub-questions if the response in the screening question suggests the presence of abnormalities involving the neuropsychiatric domain. If neuropsychiatric abnormalities have been present over the past four weeks, the caregiver rates the frequency and severity of each abnormality.

The Neuropsychiatric Inventory with Caregiver Distress Scale adds an additional question on each domain specifically addressing the level of distress caused to caregivers by each specific symptom.

The Neuropsychiatric Inventory-Nursing Home (NPI-NH) instrument is a modified version of the NPI and designed for care staff to measure psychiatric symptoms in persons with dementia living in residential care. Changed wording on each NPI question enables care staff to act as the informant, rather than obtaining the information from the informal carer.

The NPI-Questionnaire (NPI-Q) is a shorter version of the NPI and is useful for surveying the surface of neuropsychiatric symptoms in dementia, and therefore, is considered suitable for use for caregivers and care staff, as well as General Medical Practitioners.

### Scoring system

Both the frequency and severity of behavioural symptoms are rated with scores ranging from 0 to a possible 144. For each domain there are four scores: frequency, severity, total (frequency x severity), and caregiver distress.

Frequency is rated as 1 (occasionally-less than once per week) to 4 (very frequently (daily or essentially continuously present)). Severity is rated as: 1 (mild-produce little distress in the person) to 3 (severe-very disturbing to the person and difficult to redirect).

On the NPI Caregiver Distress Scale the distress occasioned for the caregiver for each of the behaviours is scored as: 0 (no distress) to 5 (very severe, or extreme).

### Clinical applications

The NPI and NPI-Q are helpful tools for caregivers and primary healthcare professionals, whereas the NPI-NH is aimed at assisting staff with care planning and monitoring the effect of treatment in dementia. Levy, et al. (1996) and clinicians used the NPI to measure behavioural distinctions associated with frontotemporal dementia and Alzheimer's disease. Craig, et al. (1996) reports using the NPI to measure behaviour correlates of cerebral blood flow in Alzheimer's disease. Litvan, et al. (1996) and Mega, et al. (1996) employed the NPI to track the neurological disease process, White, et al. (2004) used the NPI-NH to identify the link between behavioural disturbance

in dementia and body mass index, and Kaufer, et al. (1996; 1998) and other researchers have used the NPI to evaluate the efficacy of pharmacological interventions. The NPI, therefore, is effective at measuring change in relationship to drugs, nutrition and treatments/therapies. The NPI-Q, however, is not suitable for use in medication trials because of its brevity and reliance on caregiver report. Translations of the English version of the NPI prove to be reliable and valid measures of neuropsychiatric disturbances in persons with dementia from non-English speaking countries, including Greece, Italy, Japan, Korea, Mexico, Poland, Spain and Holland.

### **Reliability and validity**

As no gold standard measure of neuropsychiatric disturbance in dementia existed at the time the NPI was developed by Cummings, et al. (1994), an expert panel of clinicians with relevant training and experience participated in a Delphi method to develop and rate the scale items. Internal consistency was established (0.75-0.89) for each item/sub-scale of the NPI. This finding was supported by interviewing 40 family caregivers/spouses of persons diagnosed with Alzheimer's disease (mean MMSE of 28.4), to identify behaviours occurring in their family member that were different to usual behaviours observed in persons without dementia. The NPI items were influenced very little by the normal ageing process, and elevated scores on the NPI were found to be evidence of the presence of psychopathology for 88% of the 50 persons with Alzheimer's disease. Two domains not covered in similar instruments of psychopathology in dementia were added as result: night time behaviours and eating behaviours. Five factors account for 63% of the variance in behavioural disturbance in dementia: agitation, mood, psychosis, sleep/motor activity and elevated behaviour.

Other researchers/clinicians have subsequently shown high levels of internal consistency (0.88) for the NPI sub-scales using Cronbach's coefficient (Mega, et al. 1994; Cummings, et al. 1997; Choi, et al. 2000; Politis, et al. 2004). Internal consistency reliability of the NPI-NH in nursing home people with dementia, using Cronbach's alpha was found to be 0.67 (Lange, et al. 2004). In the Greek translated versions of the NPI and the NPI-NH, Cronbach's alpha for the total NPI score (across 12 domains) was 0.76 and varied from 0.69 to 0.76 for individual domains (Politis, et al. 2004). Cronbach's alpha on the English version was 0.88, with a range of 0.87 to 0.88 (Cummings, et al. 1997). The estimate of the coefficient of the NPI is indicative of a high degree of reliability. Internal consistency of the Polish translated version of the NPI-NH was satisfactory (Cronbach's alpha of 0.85 for both the frequency and severity of symptoms) (Bidzan and Bidzan, 2005). These findings are similar to other tests of internal consistency in English and translated versions of the NPI, NPI-NH, the Caregiver Distress scale and the NPI-Q (Cummings, et al. 1994; Mega, et al. 1994; Cummings, 1996, 1997; Choi, et al. 2000).

Test-retest reliability of NPI, NPI-NH and the Caregiver Distress Scale is high. This was initially established by Cummings, et al. (1994) by conducting a second round of NPI interviews within three weeks of the first, with half of the interviews conducted face to face and half by telephone. Test-retest scores of all items were significantly correlated, with overall correlations of 0.79 for frequency ( $p=0.0001$ ), and .86 for severity ( $p=0.0001$ ). The reliability of telephone interviews did not differ from clinic-based interviews. Subsequently, 20 assessments were used to establish test-retest reliability of over 0.79 for frequency and severity of all neuropsychiatric symptom items at a second interview after 3 weeks. Other researchers and clinicians in a number of countries have established test-retest reliability at 72 hours, 4 and 32 days, with reliability coefficients ranging from  $r = 0.55$  to  $r = 0.88$  for each of the individual symptoms in a geriatric neuropsychiatric sample, and the total score (Spearman's  $r = 0.76$ ) (Cummings, 1997; Choi, et al. 2000; Iverson, et al. 2002; Bidzan and Bidzan, 2005; Boada, et al. 2002). However, when re-testing at three time points over 72 hours with the same staff members, the total score either declined or improved to become more like the average group score (Iverson, 2002). A change in the total score of less than 22 points at re-test should, therefore, be interpreted with caution, as this may relate to the halo effect, whereby the informant's responses are coloured by an unrelated positive or negative experience with the person being assessed.

Inter-rater reliability has been established for the NPI, the NPI-NH, NPI-Q and the Caregiver Distress Scale in English and translated versions (Cummings, 1994; Cummings, et al. 1997; Choi, et al. 2000; Boada, et al. 2002; Politis, 2004; Ikeda, 2004; Bidzan and Bidzan, 2005). Between-rater reliability is reported to vary from 71% to 100% for different items and from 80% to 100% for the total score (Cummings, 1994; Cummings, et al. 1997; Wood, et al. 2000). However, test-retest reliability when conducted with staff of different levels of expertise, Certified Nurses' Aids ratings correlated only moderately well, especially for residents with high levels of neuropsychiatric disturbance (Wood, et al. 2000). It is suggested that the patient's primary nurse is a more reliable source of information for the NPI-NH version (Cummings, 1994; Cummings, et al. 1997; Wood, et al. 2000; Bidzan and Bidzan, 2005).

Clinicians, such as social workers, psychologists, nurses, geriatricians and neurologists agree that the items on the NPI, NPI-NH, NPI-Q and the Caregiver Distress Scale compare favourably with their own clinical assessment processes which employ a combination of health and social history, MMSE test scores (Ikeda, 2004), neurological examination, and rating of behavioural disturbance with other validated measures such as the Brief Psychiatric Rating Scale (BPRS) (Politis, 2004), the Cohen-Mansfield Agitation Inventory (CMAI) (Cummings, et al. 2004, 2006, 2007) and the Caregiver Emotional Distress Scale (EDS) (Lange, et al. 2004; Haloum, et al. 2005). Family caregivers and care staff also agree that the Caregiver Distress Scale items correlate highly with their own perceptions of issues that cause them distress in the caring role (Cummings, 1997; Mega and Cummings, 1996; Fiorello, et al. 1996; Frisoni, et al. 1999; Ikeda, et al. 2004; Politis, 2004).

The NPI subscales correlate well ( $p = 0.001$ ) with the BEHAVE-AD, Hamilton Depression Rating Scale (HAM-D), Clinical Dementia Rating Scale (CDRS) (Lange, 2004) and Brief Psychiatric Rating Scale (BPRS) (Politis, 2004). The Caregiver Distress Scale compares favourably with the Caregiver Emotional Distress Scale (EDS) (Haloum, 2005). All five factors identified in the NPI are similar to the factors identified in other validated measures of behavioural disturbance in dementia, including: the agitation factor in the Cohen-Mansfield Agitation Inventory (CMAI); sleep/aberrant motor activity factor in the Geri-SNAP; elevated behaviour factor in the Mania Rating Scale; aggression and compliance factor in the Geri-SNAP; mood in the Cornell Depression Scale and the Brief Psychiatric Rating Scale; and the psychosis factor in the Geri-SNAP (Ikeda, 2004; Politis, 2004; Lange, 2004; Haloum, 2005; Ikeda, 2004).

The NPI is sensitive to change in dementia severity scores. Only a few of the NPI categories show a minimal response in non-demented controls, indicating that the older person without cognitive impairment/dementia has hardly any of the symptoms identified by the NPI (Cummings, et al. 1994, Mega, et al. 1996). Conversely, elevation of NPI scores is present in persons with dementia, indicating the presence of psychopathology. NPI scores correlate highly with observations for persons with frontotemporal dementias who exhibited significantly more apathy, disinhibition, euphoria, and aberrant motor behaviour than those with Alzheimer's disease, whereas those with supranuclear palsy reveal significantly more apathy and less agitation and anxiety (Levy, et al. 1996). Regional cerebral blood flow to the brain, as measured by single photon emission computed tomography (SPECT) reveal that changes in pre-frontal and anterior temporal perfusion are most highly correlated with NPI apathy scores (Cummings, et al. 1994; Mega, et al. 1996). Scores on the NPI and NPI-NH have been shown to be sensitive to drug treatments with scores improving significantly, e.g. Rivastigmine, Olanzapine, Tacrine (Cummings, et al. 2002; Kaufer, et al. 1998; Hatoum, et al. 2005).

Despite the widespread, international use of the NPI, NPI-NH and the Caregiver Distress Scale Lange (2004) recommends scoring and interpreting the individual items or factors (agitation, mood, psychosis, sleep/motor activity, and elevated behaviour), as opposed to total scores when using the NPI-NH with a heterogeneous population. Lyketsos (2007) also recommends that all raters are systematically trained to use the NPI in a consistent manner, and agrees with Rosenberg, et al. (2005) that NPI ratings are most vulnerable to the effect of caregiver variables,

and reliance on caregiver assessment when they are made without reference to the clinical judgement of experienced clinicians and the person with dementia. These researchers/clinicians recommend developing a revised NPI to include these additional inputs to the assessment. Nevertheless, the reliability and validity of the NPI has been established and has proven applicability for use in community, a range of health care settings, and for research, in a number of different cultures.

### 9.2.3 Behavioural Pathology in Alzheimer's Rating Scale (BEHAVE-AD)

The BEHAVE-AD (Reisberg, Borenstein, et al. 1987; Monteiro, Boksay, et al. 2001) is a popular clinician rated scale developed to measure change in behavioural disturbance in persons with Alzheimer's disease. It was one of the first scales developed for the purpose of measuring behavioural disturbance in persons with dementia and was developed by clinical experts based on retrospective chart review. The original version of the scale rated items on severity, but the scale has been revised to the Behavioural Pathology in Alzheimer's disease Rating Scale Frequency Weighted (BEHAVE-AD-FW) where the items are rated on both severity and frequency. It has been translated into French (see Sclan, 1996), Swedish (Midlov, Bondesson, et al. 2002), German (Auer, Hampel, et al. 2000), Dutch (Engelborghs, Maertens, et al. 2005), Spanish (Boada, Tarraga, et al. 2006), Chinese (Chan, Lam, et al. 2001), and Korean (Suh, Son, et al. 2004). Although originally validated in community-dwelling persons with dementia, the scale has also been extensively used in nursing home residents with dementia (Brodaty, Draper, et al. 2001; Brodaty, Ames, et al. 2005; De Deyn, Katz, et al. 2005). The scale has also been validated for administration by telephone (Monteiro, Boksay, et al. 1998).

The BEHAVE-AD and BEHAVE-AD-FW are rated based on information from a carer and take about 20 minutes to complete. The scale and revised scale are available in the appendix of a published chapter (Reisberg, Borenstein, et al. 1987) and journal article (Monteiro, et al. 2001). The BEHAVE-AD-FW comprises 25 items grouped into 7 major categories: paranoid and delusional ideation, hallucinations, activity disturbance, aggressiveness, diurnal rhythm disturbances, affective disturbances and anxieties and phobias. Items are rated on severity (4-point scale) and frequency (4-point scale). The BEHAVE-AD is scored by adding the severity scores, and the BEHAVE-AD-FW is scored by multiplying the severity and frequency for each item, then summing them. At the end of the scale there is a 4-point global assessment of the overall magnitude of the behavioural symptoms in terms of disturbance to the caregiver and/or dangerousness to the patient.

The BEHAVE-AD and BEHAVE-AD-FW have good to excellent psychometric properties, although no information was located regarding internal consistency and test-retest reliability. Inter-rater reliability is  $>0.70$  for all subscales except anxiety and phobias on the BEHAVE-AD (Sclan, Saillon, et al. 1996), and  $>0.75$  for all subscales except diurnal rhythm disturbance on the BEHAVE-AD-FW (Monteiro, et al. 2001). The scale has good content validity. Factor analyses have revealed differing numbers of factors, which may be because of differences in the sample type or size, or instability in the internal structure of the scale (Harwood, Ownby, et al. 1998; Schreiner, Ballaban, et al. 2005).

Scores on the scale show a curvilinear relationship with severity, increasing with dementia severity and the Mini-Mental State Examination (MMSE) until very severe dementia where scores fall (Sclan, Saillon et al., 1996; Reisberg, Monteiro, et al. 2000; Suh and Kim, 2004). The scale correlates with other measures of behavioural disturbance such as the Cohen-Mansfield Agitation Inventory and the Neuropsychiatric Inventory (Finkel, Lyons, et al. 1992; Finkel, Lyons, et al. 1993; Cummings, McRae, et al. 2006). The scale has been used as the outcome measure in several clinical trials that have evaluated the effects on behavioural disturbance of different models of nursing home care (Brodaty, Draper, et al. 2003), Risperidone (De Deyn, Rabheru, et al. 1999; Chan, Lam, et al. 2001; Brodaty, Ames, et al. 2003; Brodaty, Ames, et al. 2005), Rivastigmine (Burns, Spiegel, et al. 2004), Clozapine (Chacko, Hurley, et al. 1995), Haloperidol (De Deyn,

Rabheru, et al. 1999; Chan, Lam, et al. 2001), and Donepezil (Cummings, McRae, et al. 2006). Data have been published on scores across different levels of dementia severity and MMSE scores (Sclan, Saillon, et al. 1996; Reisberg, Monteiro, et al. 2000) which would assist clinicians in interpreting scores, although there are no published normative data.

The BEHAVE-AD was developed for use in persons with Alzheimer's disease; however, the scale has been used to measure behaviour in persons with Vascular dementia, Lewy Body dementia and Fronto-Temporal dementia (Mendez, Perryman, et al. 1998; Vetter, Krauss, et al. 1999; Engelborghs, Maertens, et al. 2005; Shah, Ellanchenny, et al. 2005; Chiu, Chen, et al. 2006). However the scale may not adequately measure behavioural changes observed in fronto-temporal dementia such as apathy, disinhibition and emotional inappropriateness.

#### **9.2.4 Dementia Behaviour Disturbance Scale (DBDS)**

The DBDS is an instrument designed to measure behavioural disturbance in persons with dementia (Baumgarten, Becker, et al. 1990). The instrument was designed after a literature review and review of symptoms of persons with dementia from the clinical practices of two authors. The scale has been used in outpatient settings (Ott, Tate, et al. 1996; Coen, Swanwick, et al. 1997), residential care facilities (Kurita, Katayama, et al. 1997; Draper and Turner, 2003; Neville and Byrne, 2007) and in the community (Neville and Byrne, 2007). The DBDS has been translated into Japanese (Mizoguchi, Iijima, et al. 1993).

The instrument is completed by a clinician after interviewing the caregiver and takes 15 minutes to complete. The DBDS comprises 28 behavioural items rated for frequency over the past week on a 5 point scale yielding a total score from 0 to 112 (Baumgarten, Becker, et al. 1990).

The DBDS has good psychometric properties. In terms of reliability, internal consistency has been reported as  $>0.80$  (Baumgarten, Becker, et al. 1990), test retest reliability of 0.71 and 0.94 (Baumgarten, Becker, et al. 1990; Neville and Byrne, 2002), and inter-rater reliability of 0.93 (Neville and Byrne, 2002). The DBDS has good construct validity, with persons with dementia scoring more highly than persons without dementia, and correlations in the hypothesized direction with cognition, disability, function and Green's Behaviour and Mood scale (Baumgarten, Becker, et al. 1990; Neville and Byrne, 2002). The scale has not been used as an outcome measure in clinical trials. The scale was used in the population-based Canadian Study of Health and Ageing, however published data on the scale from this study could not be located.

#### **9.2.5 Neurobehavioural Rating Scale (NRS)**

The NRS is an instrument designed to measure psychiatric symptoms in persons with traumatic brain injury which has also been used in persons with dementia (Levin, High, et al. 1987). The NRS has been used extensively in samples of persons with head and brain injury (Levin, High, et al. 1987; Viikki, Ahola, et al. 1994; Cifu, Keyser-Marcus, et al. 1997; Dombovy and Olek, 1997; Mazaux, Masson, et al. 1997; Franulic, Horta, et al. 2000; Vanier, Mazaux, et al. 2000; McCauley, Levin, et al. 2001; Rapoport, McCauley, et al. 2002; Mathias, 2003; Franulic, Carbonell, et al. 2004; de Guise, Feyz, et al. 2005; Nybo, Sainio, et al. 2005; Lippert-Gruner, Kuchta, et al. 2006; van Baalen, Odding, et al. 2006), however it has had more limited use in persons with cognitive impairment and dementia (Sultzer, Levin, et al. 1992; Sultzer, Levin, et al. 1993; Sultzer, Berisford, et al., 1995; Sultzer, Mahler, et al. 1995; Rosen, Bobys, et al. 1999; Kastango, Kim, et al. 2002; Kelly, Todd, et al. 2006).

The instrument is completed by a clinician after a structured interview with the person being rated and clinical observation and takes 15-20 minutes to complete. The original NRS comprises 27 items measuring behaviour, mood and cognitive functioning, each rated on a 7 point severity scale that was summed to provide a total score ranging from 0 to 168 (Rosen, Bobys, et al. 1999). When used in people with dementia, an additional item, 'fluent aphasia' was added (Sultzer, Berisford, et

al. 1995). The NRS items were revised (McCauley, Levin, et al. 2001) and the NRS-R comprises 27 items rated on a 4 point scale which a total score ranging from 0 to 81.

There is evidence from numerous head injury studies on the good psychometric properties of the NRS. Since this review is on the suitability of the NRS for use with persons with dementia, only evidence originating from samples with dementia will be summarized here. There is only limited information on reliability in persons with dementia, where test-retest reliability was > 0.70 in hospitalized persons with dementia (Pollock, Mulsant, et al. 2002). In regards to validity, it is notable that the scale includes measurement of cognitive performance. NRS scores increase with dementia severity (Sultzer, Levin, et al. 1992). Two sets of factor analysis in different samples of persons with dementia revealed 6 and 7 factors respectively (Sultzer, Levin, et al. 1992; Kastango, Kim, et al. 2002). The scale has been evaluated in comparison to clinical determination as suitable for diagnosing behavioural disturbance in nursing homes requiring neuroleptic use with a cut-off of 60 recommended (Rosen, Bobys, et al. 1999).

In persons with dementia, the NRS has been used to demonstrate change on behavioural disturbance after Citalopram treatment, (Foglia, Pollock, et al. 1997) and Citalopram and Perphenazine treatment (Pollock, Mulsant, et al. 2002). The insight item has also been used to measure impaired insight in persons with dementia (Harwood, Sultzer, et al. 2000; Harwood, Sultzer, et al. 2005). The NRS has been used in clinical settings (Sultzer, Levin, et al. 1992; Sultzer, Levin, et al. 1993; Sultzer, Berisford, et al. 1995; Sultzer, Mahler, et al. 1995; Kastango, Kim, et al. 2002), the community (Kelly, Todd, et al. 2006) and in residential care (Rosen, Bobys, et al. 1999).

### **9.2.6 Consortium to Establish a Registry for Alzheimer's Disease – Behavior Rating Scale for Dementia (CERAD-BRSD)**

The CERAD-BRSD (Tariot, 1996; Patterson, Mack, Mackell, Thomas, et al. 1997; Mack, Patterson and Tariot, 1999) is a standardized instrument designed to measure behavioural abnormalities in demented or cognitively impaired persons. It was developed by a team of experts in the field and is one of the assessment instruments that make up the Consortium to Establish a Registry for Alzheimer's Disease battery. Instruments developed by CERAD are used by many researchers worldwide. The CERAD-BRSD has been used in numerous studies in both clinical and research settings. CERAD instruments have also been translated into numerous languages. The CERAD-BRSD has been translated into Spanish.

The instrument is administered by a trained interviewer to an informant who is familiar with the person to be rated and takes about 20-30 minutes to complete. It can be purchased, together with an instruction manual at a cost of \$US85.00. A training video is also available. CERAD-BRSD comprises 46 questions covering six domains: depressive features, inertia, psychotic features, vegetative features, irritability/ aggression, and behavioural dysregulation. Most items are rated on a five point severity scale but some have a yes/no response. Ratings are based on the frequency with which the behaviour occurred during the month prior to the interview. Total scores range from 0 to 164 with higher scores representing greater behavioural disturbance. The original 51 item instrument is still used by some researchers. This version covers eight domains: depressive features; psychotic features; defective self-regulation; irritability/agitation; vegetative features; apathy; aggression; and affective lability. A 17 item abbreviated version is also available which covers depressive symptoms, inertia, vegetative symptoms, irritability/aggression, behavioural dysregulation, and psychotic symptoms. There has been some demand for this version, mainly from clinical practitioners, but the extent to which it has been used is not known and to date no research publications have been found.

Evidence from numerous studies indicates the CERAD-BRSD has very good to excellent psychometric properties. Most studies have provided considerable evidence to ensure the findings can be appropriately interpreted. The internal structure of the instrument has generally

been confirmed through factor analysis confirming the domains proposed by the authors (Mack, Patterson and Tariot, 1999) and inter-item consistency for the subscales was reported to range from 0.48 to 0.80.

Reliability of the instrument has been confirmed through a number of studies (Tariot, Mack, Patterson, Edland, et al. 1995; Patterson, Mack, Mackell, Thomas, et al. 1997; Weiner, Koss, Patterson, Jin, et al. 1998; Mack, Patterson and Tariot, 1999). Test-retest reliability has been shown to be good with correlations and ICC's of 0.70 to 0.89, and inter-rater reliability excellent with kappas ranging from 0.77 to 1.00. Internal consistency for the total scale and for depressive symptoms, irritability/ aggression and psychotic symptoms subscales are reported to be very good with Cronbach's alphas of 0.87, 0.77, 0.75 and 0.80. For the inertia, vegetative symptoms and behavioural dysregulation subscales however, Cronbach's alpha was only 0.48, 0.56 and 0.51.

There is evidence that the instrument has construct validity in terms of the extent to which scores on the CERAD-BRSD relate to other measures in a manner that is consistent with theoretically derived hypotheses concerning the domains measured. Studies (Jacobs, Strauss, Patterson and Mack, 1998; Weiner, Koss, Patterson, Jin, et al. 1998; Weiner, Tractenberg, Teri, Logsdon, et al. 2000; Tractenberg, Weiner, Patterson, Gamst, et al. 2002) have reported expected correlations with several measures of physical and cognitive functioning: Functional Assessment Staging (FAST); Activities of Daily Living (ADL-Functional status); and the Revised Memory and Behavior Problems Checklist (RMBPC). Scores have also been shown to be associated with indicators of depression. Evidence of construct validity in terms of correlations with other well known instruments measuring agitation and/or aggression is also available (Tariot, Mack, Patterson, Edland, et al. 1995; Weiner, Williams and Risser, 1997; Weiner, Koss, Patterson, Jin, et al. 1998; Logsdon, Teri, Weiner, Gibbons, et al. 1999; Weiner, Tractenberg, Teri, Logsdon, et al. 2000). Expected correlations were found with the Cohen-Mansfield Agitation Inventory (CMAI), Agitated Behavior in Dementia (ABID) and the Revised Memory and Behavior Problems Checklist (RMBPC).

Discriminant validity of the CERAD-BRSD has been confirmed through several studies. The instrument has been shown to discriminate between different levels of dementia severity and between demented and non-demented persons (Tariot, Mack, Patterson, Edland, et al. 1995; Whitehouse, Patterson, Strauss, Geldmacher, et al. 1996; Patterson, Mack, Mackell, Thomas, et al. 1997; Weiner, Koss, Patterson, Jin, et al. 1998; Mack, Patterson and Tariot, 1999; Tractenberg, Patterson, Weiner, Teri, et al. 2000; Tractenberg, Weiner, Patterson, Gamst, et al. 2002; Lopez, Becker and Sweet., 2005). CERAD-BRSD scores have also been shown to be associated with white matter changes in the brain (Lee, Choo, Kim, Jhoo, et al. 2006).

Evidence relating to responsiveness is limited and mixed. CERAD-BRSD scores have been shown to be sensitive to the effects of drug treatment (Patterson, Mack, Mackell, Thomas, et al. 1997; Teri, Logsdon, Peskind, Raskind, et al. 2000; Weiner, Martin-Cook, Foster, Saine, et al. 2000). In a study evaluating the effectiveness of a weekly activity based program, scores were sensitive to the intervention, changing significantly, but not in the expected direction, i.e. behaviour did not improve as a result of the intervention (Higgins, Koch, Hynan, Carr, et al. 2005). There is moderate evidence for sensitivity over time with one study (Patterson, Mack, Mackell, Thomas, et al. 1997) reporting significant change over time but only for persons with mild to moderate dementia.

The CERAD-BRSD has been used in both clinical and research settings. Normative data for the instrument is available in the CERAD-BRSD manual. Additionally the CERAD database, available to researchers in the form of a CD-ROM, contains data for 1094 patients with a clinical diagnosis of Alzheimer's disease and 463 control subjects evaluated annually between 1987 and 1996. The data includes clinical findings and neuropsychological test scores, behavioural manifestations of dementia, time to death or admission to a nursing home and neuropathological findings. The CD-ROM is available for purchase at a cost of \$US600.

The CERAD-BRSD has been used in numerous clinical and intervention studies. It has been used as an outcome measure in studies evaluating the effectiveness of drug treatment (Martinon-Torres, Fioravanti and Grimley, 2004; Teri, Logsdon, Peskind, Raskind, et al. 2000; Weiner, Martin-Cook, Foster, Saine, et al, 2000) and in a study assessing an activities based adult dementia care program (Higgins, Koch, Hynan, Carr, et al. 2005). The instrument has also been used in studies investigating comorbidity in community dwelling persons with Alzheimer's disease (AD) (Tractenberg, Weiner, Patterson, Teri, et al. 2003), predicting psychosis onset (Wilkosz, Miyahara, Lopez, Dekosky, et al. 2006) and investigating subtypes of psychosis (Perez-Madrinan, Cook, Saxton, Miyahara, et al. 2004), and in a longitudinal study examining the effects over time of depressive symptoms in persons with AD on depression in their family caregivers (Neundorfer, McClendon, Smyth, Stuckey, et al. 2001). Other clinical studies include, investigating the association between white matter changes and neuropsychiatric symptoms (Lee, Choo, Kim, Jhoo, et al. 2006), and the relationship between nursing home placement and measures of change (Knopman, Berg, Thomas, Grundman, et al. 1999), and a pilot study of a potential new outcome, expected emergence (Tractenberg, Gamst, Thomas, Patterson, et al. 2002).

### 9.2.7 Recommendations Concerning BPSD Instruments

As shown in Table 34, examination of key attributes and psychometric properties of the five final instruments of BPSD Global, measured against the weighting criteria (refer Section 2.5), indicates the NPI and the BEHAVE-AD as the best measures for assessment of BPSD, followed by the DBDS, the NRS and the CERAD-BRSD. All of the five instruments are proxy rated, by interviewing carers/informants who are deemed to have the best knowledge about behavioural and psychological conditions of the persons with dementia.

**Table 34 Summary of Ratings for BPSD Global Instruments**

Criteria	Weight	Instrument				
		NPI	BEHAVE-AD	CERAD-BRSD	DBDS	NRS
Availability of comparison data	3	2	2	2.5	1	2
Length/feasibility of instrument for inclusion in battery	2	3	2	1	2	2
Complexity of administration/cognitive burden	2	3	3	2	2	1
Cultural Appropriateness	1	3	3	2	2	2
Ease of obtaining score	2	3	3	3	3	3
Sensitivity to dementia	3	3	3	3	2	2
Reliability evidence	3	3	3	3	3	3
Validity evidence	3	3	3	3	2	2
Cost of the instrument	2	3	3	1	2	2
Cost of instrument administration	2	2	2	2	2	1
<b>Weighted Total</b>		<b>64</b>	<b>62</b>	<b>54.5</b>	<b>50</b>	<b>49</b>

Both the NPI and the BEHAVE-AD are two of the most popular, widely utilised instruments internationally to assess the presence and severity BPSD in persons with Alzheimer's disease and other types of dementia in pharmacological and non-pharmacological interventions, which are proven to be applicable in various institutional, out-patient and community settings. The NPI has several versions including the NPI with Caregiver Distress Scale, the nursing home version (NPI-NH), and the NPI shorter version for primary care settings (NPI-Q). The BEHAVE-AD in its original form has severity ratings only, but the BEHAVE-AD-FW has frequency weighting where items are rated on both severity and frequency. The NPI takes a shorter time to complete (10-15 minutes) than the BEHAVE-AD (20 minutes) although both can be completed in a shorter time when self administered by carers/informants themselves. Both tools assess impact of BPSD on carers, which can be rated separate from the symptom severity measure of the tool. The NPI comes with a training manual while the BEHAVE-AD does not, which makes it unsuitable for the completion by persons without some clinical training given use of the psychiatric language in the tool. The NPI was originally developed to assess psychopathology in the person with dementia, to help distinguish between different causes of dementia, and includes items pertaining to symptoms known to be rare in Alzheimer's disease, but are characteristic of fronto-temporal dementias. On the other hand, the BEHAVE-AD was developed to measure change in behavioural disturbance in persons with Alzheimer's disease, and may not adequately measure behavioural changes observed in fronto-temporal dementia such as apathy, disinhibition and emotional inappropriateness.

The CERAD-BRSD was designed to measure behavioural abnormalities in demented or cognitively impaired persons. Similar to the NPI and the BEHAVE-AD, the CERAD-BRSD has been used in numerous studies in both clinical and research settings and its psychometric properties are reported to be good to excellent. The main reason for low scores according to the DOMS weighting criteria shown in Table 34 relate to the cost and the lengthy time required for administration.

Unlike the other four tools described above, the NRS was designed to measure psychiatric symptoms in persons with Traumatic Brain Injury (TBI). Whilst the tool has also been used with persons with dementia, the available studies suggest limited use with persons with cognitive impairment and dementia.

The DBDS is an instrument designed to measure behavioural disturbance in persons with dementia. The instrument is completed by a clinician after interviewing the caregiver and takes 15 minutes to complete. The scale was used in the population-based Canadian Study of Health and Ageing and a prevalence study in Australia, however, it has not been used as an outcome measure in clinical trials. Whilst the DBDS has moderate to good psychometric properties further studies are needed to consolidate its sensitivity to dementia.

Based on these reviews of the five final instruments it is recommended the NPI and the BEHAVE-AD be used in both clinical and research settings for assessment of BPSD Global. The CERAD-BRSD is recommended for research in particular given its cost and time required for administration. The 17 item abbreviated version may be considered better for clinical utility, and there has been some demand for it, but limited evidence on this version is currently available.

### **9.3 Differential Diagnosis: Delirium**

Delirium is an acute confusional state with fluctuating course, characterised by disturbance of consciousness, altered attention, impaired cognition, disturbance of thought and perception (delusions and hallucinations), and behaviour. See Table 35 for diagnostic criteria of delirium. Application of different diagnostic criteria produces inconsistent results of delirium prevalence. In a large scale Finnish study comparing different diagnostic classifications of delirium (DSM-III, DSM-III-R, DSM-IV and ICD-10), DSM-IV was found to be most inclusive while ICD-10 was overly

restrictive (Laurila, et al. 2004). Most delirium scales available are based on various DSM versions.

**Table 35 DSM-IV-TR and ICD-10 Diagnostic Criteria of Delirium**

DSM-IV-TR (APA 2000)	ICD-10 (WHO 1992)
<p>a. Disturbance of consciousness (that is, reduced clarity of awareness of the environment, with reduced ability to focus, sustain, or shift attention)</p> <p>b. A change in cognition (such as memory deficit, disorientation, language disturbance) or the development of a perceptual disturbance that is not better accounted for by a pre-existing established or evolving dementia</p> <p>c. The disturbance developed over a short period of time (usually hours to days) and tends to fluctuate during the course of the day</p> <p>d. Where the delirium is due to a <u>general medical condition</u> – there is evidence from the history, physical examination, or laboratory findings that the disturbance is caused by the direct physiological consequences of a general medical condition</p> <p>Where the delirium is due to <u>substance intoxication</u> – there is evidence from the history, physical examination, or laboratory findings of either 1 or 2:</p> <ol style="list-style-type: none"> <li>1. The symptoms in criteria (a) and (b) developed during substance intoxication</li> <li>2. Medication use – aetiologically related to the disturbance</li> </ol> <p>Where the delirium is due to <u>substance withdrawal</u> – there is evidence from the history, physical examination, or laboratory findings that the symptoms in criteria (a) and (b) developed during or shortly after the withdrawal syndrome</p> <p>Where delirium is due to <u>multiple aetiologies</u> – there is evidence from the history, physical examination, or laboratory findings that the delirium has more than one aetiology (for example, more than one aetiological general medical condition, a general medical condition plus substance intoxication, or medication side effects)</p> <p>e. Delirium not otherwise specified – this category should be used to diagnose a delirium that does not meet criteria for any of the specific types of delirium described. Examples include a clinical presentation of delirium that is suspected to be due to a general medical condition or substance use but for which there is insufficient evidence to establish a specific aetiology, or where delirium is due to causes not listed (for example, sensory deprivation)</p>	<p>For a definite diagnosis, symptoms, mild or severe, should be present in each of the following areas:</p> <ol style="list-style-type: none"> <li>a. Impairment of consciousness and attention (ranging from clouding to coma; reduced ability to direct, focus, sustain and shift attention)</li> <li>b. Global disturbance of cognition (perceptual distortions, illusions and hallucinations – most often visual; impairment of abstract thinking and comprehension, with or without transient delusions, but typically with some degree of incoherence; impairment of immediate recall and of recent memory, but with relatively intact remote memory; disorientation for time as well as in more severe cases for place and person)</li> <li>c. Psychomotor disturbances (hypo- or hyperactivity and unpredictable shifts from one to the other; increased reaction time; increased or decreased flow of speech; enhanced startle reaction)</li> <li>d. Disturbance of the sleep/wake cycle (insomnia or, in more severe cases, total sleep loss or reversal of the sleep/wake cycle; daytime drowsiness; nocturnal worsening of symptoms; disturbing dreams or nightmares, which may continue as hallucinations after awakening)</li> <li>e. Emotional disturbances, for example, depression, anxiety or fear, irritability, euphoria, apathy or wandering, perplexity</li> </ol>

As discussed in Section 3.2.4, delirium is one of the reversible/treatable conditions that are often superimposed to the person with dementia which require differential diagnosis. The person with dementia is more susceptible to delirium due to the existing neurological damage (IPA, 2003). Common features to both dementia and delirium include high prevalence in older people; presence of cognitive impairment and behavioural and psychological disturbances; and diurnal variations in symptoms (e.g. sundowning). These make it challenging not only to differentiate between the two syndromes but also to detect delirium among people with existing dementia. Studies have shown rather mixed results as to which symptomatic difference is a better indicator in detecting delirium superimposed to dementia. Laurila, et al. (2004) suggest disturbances in memory, orientation, abstract thinking or motor function have little value for detecting delirium superimposed to dementia. Voyer, et al. (2006) on the other hand report disturbances in attention, thinking, orientation and memory as good indicators of identifying delirium among people with

existing dementia. Both studies agree on that psychomotor symptoms have little value (Laurila, et al. 2004; Voyer, et al. 2006). What is consistent of studies on delirium is that fluctuating symptoms and sudden onset of clouding consciousness are good indicators of detecting delirium superimposed to dementia and differentiating delirium from dementia. In screening and monitoring delirium, therefore it is important to have an instrument that allows repeated measures and monitoring for change of symptom severity.

### 9.3.1 Decision Making Strategies

Of 11 delirium instruments identified through the initial search, the following five tools were listed as leading contenders after additional examination processes:

- Confusion Assessment Method (CAM) (Inouye, et al. 1990)
- Delirium Rating Scale-revised 98 (DRS-R-98) (Trzepacz, et al. 1998)
- Delirium Symptom Interview (DSI) (Albert, et al. 1992)
- Delirium Index (DI) (McCusker, Cole, Bellavance and Primeau, 1998)
- Memorial Delirium Assessment Scale (MDAS) (Breitbart, et al. 1997)

The review team felt it was important to select measures that were sensitive to differentiate delirium symptoms from dementia, well studied, able to assess change of severity of delirium as well as their presence, have good utility and application in a range of settings, and minimal administration time. As shown in Table 36 below, the CAM demonstrated high quality in almost all areas of impact factors, in particular, its high psychometric properties, excellent utility and applicability in a variety of settings. The CAM is a screening tool to detect delirium but does not have power to measure severity. The DRS-R-98 and its earlier version the DRS also showed high quality in most areas of impact factors. In addition they can be used to detect any changes in severity of delirium symptoms over time. A preliminary review suggested the DSI was not sufficiently sensitive to dementia hence it was deemed to be inappropriate for delirium superimposed to dementia. The DI had limited numbers of studies to demonstrate its psychometric properties and did not have the capacity to rate severity. The MDAS initially appeared to be a strong contender however most studies that utilised the MDAS related to persons who were experiencing cancer or sometimes AIDS. Given the scope of the DOMS project and limited resources available it was decided that two reviews for delirium would be sufficient. The review team decided that the CAM and the DRS-R-98 were appropriate for the detailed review given their strong psychometric properties and comprehensive coverage of delirium symptoms.

**Table 36 Decision Summary of the Delirium Leading Contenders****Recommended: CAM and DRS-R-98**

- CAM (Confusion Assessment Method) (Inouye, et al. 1990)
- DRS-R-98 (Delirium Rating Scale-revised 98) (Trzepacz, et al. 1998)
- DSI (Delirium Symptom Interview) (Albert, et al. 1992)
- DI (Delirium Index) (McCusker, Cole, Bellavance and Primeau, 1998)
- MDAS (Memorial Delirium Assessment Scale) (Breitbart, et al. 1997)

	<b>Domains/Sub domains</b>	<b>Applicability/ Stage</b>	<b>Patient</b>	<b>Proxy</b>	<b>Availability/ Cost</b>	<b>Training/Manual</b>	<b>Admin time</b>
<b>CAM</b>	Based on the operational application of DSM-III-R, and consists of 9 features of delirium (acute onset, inattention, disorganised thinking, altered level of consciousness, disorientation, memory impairment, perceptual disturbances, psychomotor agitation, psychomotor retardation and altered sleep-wake cycle)	All stages		✓ Interviewer/ observer rated	Easily available / Free for researchers and clinicians, but fees will be charged for commercial use.	Training required, especially for non-psychiatric clinicians or lay persons / Manual available on the web.	5 minutes (10-15 minutes when combined with other cognitive test).
<b>DRS-R-98</b>	Two components: the 13-item severity section (sleep-wake cycle disturbance, perceptual disturbances and hallucinations, delusions, lability of affect, language, thought process abnormalities, motor agitation, motor retardation, orientation, attention, short-term memory, long-term memory, and visuospatial ability) and the 3-item diagnostic section (temporal onset of symptoms,	All stages		✓ Interviewer/ observer rated	Easily available / Free for use by researchers working in a not-for-profit setting or for research funded by a public/nor-for-profit funding source, but there is charge for use in, for example, pharmaceutical trials.	Training is required for anyone, in particular for those without psychiatric background, for optimal use of the DRS-R-98 / the DRS.	Rated over a 24-hour period, actual time required to implement the tool is not specified.

	fluctuation of symptom severity and physical disorder).						
<b>DSI</b>	Based on DSM–III, 62 items from the seven domains. Assess disorientation, disturbance of sleep, perceptual disturbance, incoherent speech, level of psychomotor activity, general behaviour observations, fluctuating behaviour.	All stages		✓ Interviewer/ Observer rated	Easily accessible. Available in the original paper and through internet. Permission required for its use.	Manual and scoring information available from the second author on request. To receive the detailed the documentation and scoring manual, contact Dr. Sue Levkoff at 617-525-6122 or sue_Levkoff@hms.harvard.edu.	10-15 > minutes.
<b>DI</b>	Adapted from the CAM. Assess attention, disorganised thinking, level of consciousness, disorientation, memory, perceptual disturbance, and motor activity. Rated on the following impairment scale: 0 = absent, 1 = mild, 2 = moderate, 3 = severe. The total score ranges from 0 (no symptoms) to 21 (maximum severity).	All stages		✓ Interviewer/ Observer rated	Easily accessible. Available in the original paper and through internet. No copyright. Free of charge.	Training required. No manual developed, but the manual for the CAM could be utilized.	5-10 minutes, plus time to conduct the MMSE.
<b>MDAS</b>	Based on DSM-IV, 10 item, four point scale (possible range 0-30, 30 worst/most severe), assess disturbance in arousal and level of consciousness, cognitive functioning (memory, attention, orientation, disturbance in thinking) and psychomotor activity.	All stages		✓ Interviewer/ Observer rated	Easily accessible. Available in the original paper and through internet. No fee for its use. Need permission from the author and from the publishers of the Journal of Pain and Symptom Control.	Some training required. Trained lay person can easily implement the tool. The tool was published as part of a validation paper that also describes its use.	10 minutes.

	<b>Domains/Sub domains</b>	<b>Applicability/ Stage</b>	<b>Patient</b>	<b>Proxy</b>	<b>Availability/ Cost</b>	<b>Training/Manual</b>	<b>Admin time</b>
<b>CAM</b>	Based on the operational application of DSM-III-R, and consists of 9 features of delirium (acute onset, inattention, disorganised thinking, altered level of consciousness, disorientation, memory impairment, perceptual disturbances, psychomotor agitation, psychomotor retardation and altered sleep-wake cycle).	All stages		✓ Interviewer/ observer rated	Easily available / Free for researchers and clinicians, but fees will be charged for commercial use.	Training required, especially for non-psychiatric clinicians or lay persons / Manual available on the web.	5 minutes (10-15 minutes when combined with other cognitive test).
<b>DRS-R-98</b>	Two components: the 13-item severity section (sleep-wake cycle disturbance, perceptual disturbances and hallucinations, delusions, lability of affect, language, thought process abnormalities, motor agitation, motor retardation, orientation, attention, short-term memory, long-term memory, and visuospatial ability) and the 3-item diagnostic section (temporal onset of symptoms, fluctuation of symptom severity and physical disorder).	All stages		✓ Interviewer/ observer rated	Easily available / Free for use by researchers working in a not-for-profit setting or for research funded by a public/nor-for-profit funding source, but there is charge for use in, for example, pharmaceutical trials.	Training is required for anyone, in particular for those without psychiatric background, for optimal use of the DRS-R-98 / the DRS.	Rated over a 24-hour period, actual time required to implement the tool is not specified.
<b>DSI</b>	Based on DSM-III, 62 items from the seven domains. Assess disorientation, disturbance of sleep, perceptual disturbance, incoherent speech, level of	All stages		✓ Interviewer/ Observer rated	Easily accessible. Available in the original paper and through internet. Permission required for its use.	Manual and scoring information available from the second author on request. To receive the detailed the documentation and scoring manual, contact	10-15 > minutes.

	psychomotor activity, general behaviour observations, fluctuating behaviour.					Dr. Sue Levkoff at 617-525-6122 or sue_Levkoff@hms.harvard.edu.	
<b>DI</b>	Adapted from the CAM. Assess attention, disorganised thinking, level of consciousness, disorientation, memory, perceptual disturbance, and motor activity. Rated on the following impairment scale: 0 = absent, 1 = mild, 2 = moderate, 3 = severe. The total score ranges from 0 (no symptoms) to 21 (maximum severity).	All stages		✓ Interviewer/ Observer rated	Easily accessible. Available in the original paper and through internet. No copyright. Free of charge.	Training required. No manual developed, but the manual for the CAM could be utilized.	5-10 minutes, plus time to conduct the MMSE.
<b>MDAS</b>	Based on DSM-IV, 10 item, four point scale (possible range 0-30, 30 worst/most severe), assess disturbance in arousal and level of consciousness, cognitive functioning (memory, attention, orientation, disturbance in thinking) and psychomotor activity.	All stages		✓ Interviewer/ Observer rated	Easily accessible. Available in the original paper and through internet. No fee for its use. Need permission from the author and from the publishers of the Journal of Pain and Symptom Control.	Some training required. Trained lay person can easily implement the tool. The tool was published as part of a validation paper that also describes its use.	10 minutes.

Based on information in Table 36 and after a teleconference among the review team members, the final two instruments to be reviewed in detail for delirium are CAM and DRS-R-98.

### 9.3.2 Confusion Assessment Method (CAM)

The Confusion Assessment Method (CAM) (Inouye, et al. 1990) is a brief structured questionnaire that can be implemented by a clinician or a researcher who does not necessarily have formal psychiatric training. It is designed to detect/screen delirium and is for use in older people who are at high risk for the development of delirium (Inouye, et al. 1990). The CAM was originally based on expert opinion and the operational application of the DSM-III-R, however studies have shown that the CAM criteria agree more favourably with the DSM-IV criteria than they did with the previous DSM-III-R criteria or with the ICD-10 (Inouye, 2003; Laurila, Pitkala, Strandberg, and Tilvis, 2002).

The CAM consists of nine features<sup>34</sup> of delirium (acute onset, inattention, disorganised thinking, altered level of consciousness, disorientation, memory impairment, perceptual disturbances, psychomotor agitation, psychomotor retardation and altered sleep-wake cycle), providing a diagnostic algorithm for delirium based on its four cardinal features (referred to as 'Short CAM'). The presence of two cardinal features (acute onset and fluctuating course, and inattention), and at least one of the two secondary features (disorganised thinking and altered level of consciousness) indicate the presence of delirium. The remaining five features have been shown to be non significant to the diagnosis of delirium, however some still use the entire nine questions (referred as 'Long CAM') to fulfil the DSM definition of delirium. An additive score of the four cardinal features, ranging between 0-7, is used to measure the severity of delirium (the higher, the more severe) (Inouye, 2003). This method of rating the severity of delirium is yet to be validated.

Whilst overall results of reliability and validity of the CAM have been excellent (high inter-rater reliability, sensitivity and specificity) some studies showed conflicting outcomes, for example sensitivity ranging between 0.13–0.32 measured by nurses (Inouye, Foreman, Mion, Katz, and Cooney, 2001; Rolfson, McElhaney, Jhangri, and Rockwood, 1999). The review clearly indicates reliability and validity of the CAM algorithm is heavily dependent upon the interviewer/rater's training status on the CAM, and whether or not the CAM is based on clinical observations during a brief, formal cognitive testing (e.g. the MMSE and the Digit Span Test), and thoroughness and time spent during the cognitive testing. Studies have shown these factors produce excellent reliability and validity (Gaudreau, Gagnon, Harel, Tremblay, and Roy, 2005; Lemiengre, et al. 2006; McNicoll, Pisani, Ely, Gifford, and Inouye, 2005). The manual also emphasises on the importance of having sufficient training on the CAM and using information obtained during a formal interview process to score the CAM, rather than relying on the informal observations solely (Inouye, 2003).

The CAM has been used as a screening, diagnostic and/or an outcome instrument in a variety of clinical settings including acute and post acute and nursing home/long-term care facilities. Most were conducted in detecting delirium in acute inpatients, pre-/post-operatively, who had hip fracture and/or were undergoing hip replacement. Clinical studies have examined: occurrence of delirium; clinical features and other predictive factors of delirium (e.g. depression, activities of daily living, dementia, cognitive impairment, co-morbidity, disability, educational level, apolipoprotein E genotypes, melatonin levels, drug metabolism, anaemia/metabolic disorders, changes in plasma and large neutral amino acid concentrations, nutritional status, fluid and electrolyte imbalance, nitrous oxide, serum anticholinergic activity, postoperative pain, and care environment); and relationships between delirium and various outcomes (e.g. mortality, functional outcomes, developing complications, institutionalisation, length of hospital stay, use of restraints, and costs). While most studies examining various aspects of delirium have been devoted to the detection of

---

<sup>34</sup> Sometimes it is known to be ten features as 'acute onset' and 'fluctuating course' are counted as two features.

delirium in older people who are medically ill, the participants in those studies commonly consist of both 'with' or 'without' dementia.

The CAM is used either alone or in conjunction with other delirium assessment instruments such as: the Delirium Symptom Interview (DSI), the Delirium Rating Scale (DRS) and the Memorial Delirium Assessment Scale (MDAS); cognitive test such as the Mini-Mental State Exam (MMSE) and the Digit Span Test; and pathophysiologic examinations. The CAM has been used in both clinical and epidemiological studies conducted. Some studies reported results based on a telephone interview CAM (Marcantonio, Michaels, and Resnick, 1998; Nelson, et al. 2006). To detect delirium for mechanically ventilated patients in ICU, the CAM-ICU version has been developed by adding extra descriptors for the four cardinal features and using pictures and commands. The CAM-ICU showed a high sensitivity and specificity, and excellent inter-rater reliability when used by trained physicians and nurses (Ely, Inouye, et al. 2001; Ely, Margolin, et al. 2001). The CAM has also been adapted to suit emergency care settings (Lewis, Miller, Morley, Nork, and Lasater, 1995) and measure the severity of delirium (Jones, et al. 2006; McCusker, Cole, Bellavance, and Primeau, 1998), however, neither of those two adaptations appear to have been widely utilised nor sufficiently validated, especially for those with delirium superimposed to dementia.

The CAM is the most widely utilised screening/diagnostic tool for delirium, in particular older people with or without dementia, and translated in various languages, for example, in Chilean, Spanish, German, Italian, Portuguese, Dutch, Japanese and French. It is simple, easy to master and implement (5 minutes to complete although it takes about 10-15 minutes when combined with a formal cognitive testing), and easily accessible, with no cost involved when used non-commercially. It is increasingly well recognised amongst clinicians in Australia, as well as internationally. Its successful adaptation in various different languages and countries indicate cultural appropriateness. Whilst there exist some limitations in detecting delirium superimposed to dementia the CAM is by far the most efficient way of screening delirium in both clinical and research contexts.

### **9.3.3 Delirium Rating Scale-Revised-98 (DRS-R-98)**

The Delirium Rating Scale-Revised-98 (DRS-R-98) is a criterion-based symptom rating and observer-rated scale for assessment of both the presence and the severity of delirium symptoms (Trzepacz, et al. 2001). It is rated based on all existing and accessible information from the patient interview, medical status examination, medical history and tests, medical and nursing observations, family reports, etc. over a 24-hour period. The DRS-R-98 was developed through substantial changes to its original form, the Delirium Rating Scale (DRS) (Trzepacz, Baker, and Greenhouse, 1988), to address the shortcomings of the original scale, for instance its limited usefulness in evaluating various aspects of cognitive function, measuring repeated ratings, assessing motor subtypes of delirium, and conducting broad phenomenological and longitudinal intervention research (Trzepacz, et al. 2001). The DRS has been one of the most widely used delirium severity rating scales internationally, evidenced by numerous translated versions of the DRS-R-98 such as French, Italian, Spanish, Dutch, Mandarin Chinese, Korean, Swedish, Japanese, German, and Indian-language translations, which have been successfully applied in a variety of ethnicities and countries (Trzepacz, 1999; Trzepacz, et al. 2001). Given its short history, the range of studies using the DRS-R-98 is small in quantity, however, there is an increasing number of international studies choosing to use the revised version instead of the original scale. However, the developer Dr Trzepacz argues that both tools can be used together in some research settings given the substantial difference between the two versions and distinct usefulness of each tool (e.g., the DRS more useful in patients who are recovering from stupor) (Trzepacz, et al. 2001). The DRS-R-98 has been used successfully in various non-English background countries, including Japan (Takeuchi, et al. 2007); Korea (Pae, et al. 2004); Spain (Fonseca, et al. 2005); and the Netherlands (de Jonghe, Kalisvaart, Timmers, Kat, and Jackson, 2005; Kalisvaart, et al. 2005).

The DRS-R-98 consists of the 13-item severity section (item no.1-13, sleep-wake cycle disturbance, perceptual disturbances and hallucinations, delusions, lability of affect, language, thought process abnormalities, motor agitation, motor retardation, orientation, attention, short-term memory, long-term memory, and visuospatial ability) and the 3-item diagnostic section (item no. 14-16, temporal onset of symptoms, fluctuation of symptom severity and physical disorder). The total summed scores range between 0-46 points (includes the three diagnostic items) and a maximum severity score of 39 points, used separately for repeated measures. Higher scores indicate more severe delirium. The DRS is composed of ten items, with a maximum possible score of 32 points. The items are temporal onset of symptoms, perceptual disturbances, hallucination type, delusions, psychomotor behaviour, cognitive status during formal testing, physical disorder, sleep-wake cycle disturbance, lability of mood, and variability of symptoms (Trzepacz, et al. 1988).

Studies reported moderate to high inter-rater reliability of the DRS, Intraclass correlation coefficient (ICC) ranging between 0.59 and 0.99 (Rockwood, Goodman, Flynn, and Stolee, 1996; Rosen, et al. 1994; Trzepacz, et al. 1988, 2001, 1998), and good sensitivity (ranging from 82% to 95%) and specificity (ranging from 61% to 94%) for the DRS cut-off score of 10 or less (Grassi, et al. 2001; Rockwood, et al. 1996; Rosen, et al. 1994). Sensitivity and specificity change depending on cut-off scores, for example, sensitivity and specificity of 95% and 61% respectively for DRS cut-off of 10; 80% and 76% for DRS cut-off of 12 (Grassi, et al. 2001). Reviews of the literature in particular describing psychometric properties of the DRS-R-98 indicate high inter-rater reliability with ICC ranging from 0.96 to 0.99, and good internal consistency with the Cronbach's alpha coefficient ranging from 0.74 to 0.94 (de Rooij, et al. 2006; Fonseca, et al. 2005; Trzepacz, et al. 2001). Validity of the DRS-R-98 has also been evidenced by its significant and strong correlations with other relevant rating scales such as the DRS, the Cognitive Test for Delirium (CTD) the Clinical Global Impression scale (CGI), the Mini-Mental State Examination (MMSE) and the Delirium-O-Meter (DOM) (de Jonghe, et al. 2005; Fonseca, et al. 2005; Meagher, et al. 2007; Trzepacz, et al. 2001); and its moderate to high sensitivity (Trzepacz, et al. 2001).

The DRS-R-98 can be applied to the phenomenology, pathophysiology and treatment of delirium among people with or without dementia (Trzepacz, et al. 2001). It is recommended the ICD or the DSM criteria along with the DRS-R-98 be used to increase sensitivity when measuring delirium superimposed to dementia. Since its inception, the DRS-R-98 has been largely used in clinical trials to measure pharmacological effectiveness on the severity of delirium symptoms (de Jonghe, et al. 2007; Kalisvaart, et al. 2005; Lee, et al. 2005; Pae, et al. 2004; Straker, Shapiro, and Muskin, 2006; Takeuchi, et al. 2007). The DRS has been widely used to assess patients with medical, surgical or psychiatric illness, with or without dementia, in a variety of clinical settings including geriatric, psychiatric and general hospital (Trzepacz, 1999). The DRS has often been used in post-operative situations (Bohner, et al. 2003; Herrmann, Ebert, Tober, Hann, and Huth, 1999; Karlidag, et al. 2006; Nishikawa, Nakayama, Omote, and Namiki, 2004; Ohki, Matsushima, Shibuya, and Sunamori, 2006; Rothenhausler, et al. 2005; Schneider, et al. 2002).

The DRS-R-98 is a reasonably new tool developed to enable repeated measures of severity of delirium as well as diagnosing the syndrome, and can be applied in both clinical and research settings. Whilst content validity of the DRS-R-98 severity has yet to be established, studies so far have demonstrated its high validity and reliability, including moderate sensitivity to dementia, and applicability in a variety of institutional settings among diverse groups of people with medical and/or psychiatric conditions. Whilst the DRS has been used widely internationally in the last two decades with good reliability and validity demonstrated, the revised version proved to have better reliability and validity so far. Limitations of the DRS-R-98, and the DRS, include that they are time taxing as it is based on detailed clinical observations over a 24 hour period, and require a special training especially for those without psychiatric training background due to some of the psychiatric specific terminologies and descriptors used in the tool.

### 9.3.4 Recommendations Concerning Delirium Instruments

As shown in Table 37, examination of key attributes and psychometric properties of the two delirium instruments shows a considerable difference in their weighted total scores, largely resulting from a limited utility of the DRS-R-98. The weighted total score of 54 for the DRS-R-98 is not significantly low in comparison with other instruments reviewed in the DOMS project.

**Table 37 Summary of Ratings for Delirium Instruments**

Criteria	Weight	Instrument	
		CAM	DRS-R-98
Availability of comparison data	3	3	2
Length/feasibility of instrument for inclusion in battery	2	3	2
Complexity of administration/ cognitive burden	2	3	2
Cultural Appropriateness	1	3	2
Ease of obtaining score	2	3	3
Sensitivity to dementia	3	2	2
Reliability evidence	3	3	3
Validity evidence	3	3	3
Cost of the instrument	2	2	2
Cost of instrument administration	2	2	2
<b>Weighted Total</b>		<b>62</b>	<b>54</b>

The CAM is the most widely utilised screening/diagnostic tool for detecting delirium internationally among older people with or without dementia. It is increasingly well recognised amongst clinicians in Australia. Less well known, however, the DRS-R-98 is also a widely recognised and well validated measure, especially through its original measure, the DRS. Whilst the CAM is superior in its utility to the DRS-R-98, it is a screening measure that does not capture severity of delirium symptoms (just absence or presence) hence is not appropriate for monitoring delirium severity over time. The DRS-R-98 is designed for assessment of both the presence and the severity of delirium symptoms, although most studies have used the severity section of the tool only.

Limitations of the DRS-R-98, and the DRS, include that they are time taxing and require sufficient training. The DRS-R-98 is not appropriate for use in the community setting given its requirement for observation over a 24 hours period. However, it allows for comprehensive assessment of individuals who are at risk or suspected of developing delirium in institutional care settings. Both

the CAM and the DRS-R-98 are reported to have moderate sensitivity when used for people with dementia. In order to ensure reliability and validity it is critical that people who are administering the CAM and the DRS-R-98, especially those who do not have psychiatric background, undertake sufficient training before their use. For the purpose of the DOMS project it is recommended both measures be included as they have two distinct, yet equally important functions.

## **9.4 Individual Symptom Measures for Associated Symptoms**

### **9.4.1 Introduction**

Although many of the Global BPSD instruments reviewed in this chapter have quite substantial coverage of the individual symptoms listed below (for example the NPI), it was decided to include reviews of some individual symptom measures to enable a more detailed assessment of a particular symptom should this be warranted.

### **9.4.2 Aggression**

Aggressive behaviour is one of the six most commonly reported symptoms of dementia (5-22%) (Alzheimer's Disease International, 2003), the incidence rising in the mid to late stage of the illness. Hitting, kicking, resisting care, sexually aggressive behaviour, self-harm and verbal expressions of anger towards others are behaviours that distress family caregivers and also care staff (Wristers, 2007). Yet, care staff and family carers are poor at predicting aggressive behaviour with accuracy (Shah, 1999b). In the 1980s these and other aggressive behaviours were reported to occur in approximately 6.5% of persons with dementia. However, aggression was not well understood by clinicians and researchers and was predominantly studied under the broad category of behavioural disturbance, or included with measures of agitation. Consequently, the specific characteristics of aggression became diluted within these generic measures (Shah, et al. 2000). There was also no definition of aggression specific to dementia (Shah, 1999). This led to the development of clinical instruments focused specifically on identifying and measuring the incidence and severity of aggressive behaviour (Ryden, 1988).

In the process of developing an instrument to measure aggression for the clinical and research settings, Ryden (1988) found an association between the occurrence of aggression pre-morbidly and its presence in dementia, although others do not agree that pre-morbid personality traits are necessarily continued during the course of the illness (e.g. Hamel, et al. 1990). However, aggression was found to be one of the changes in personality that may signal the presence of dementia (Swearer, et al. 1996). Understanding this problematic behaviour is important, as family caregivers find it difficult to provide the level of care required for the person with dementia living in the community, as do care staff in health facilities (Swearer, et al. 1996; Cohen-Mansfield and Werner, 1998; Burke and Morgenlander, 1999). Not only is unmanageable aggression more likely to precipitate placement of the person with dementia in a long-term care facility (Burke and Morgenlander, 1999), it also leads to unnecessary use of chemical and/or physical restraint in care settings (Cohen-Mansfield, et al. 1989; Retsas and Crabbe, 1998; Opie, et al. 2002).

Behaviour-specific measurements are considered by many clinicians to be more useful than the more widely available and broad-based inventories (Teri, et al. 1992). These inventories, like the Cohen-Mansfield Agitation Inventory (Cohen-Mansfield, 1986), the CERAD-BRSD (Tariot, 1996; Patterson, et al. 1999) and the BEHAVE-AD (Reisberg, et al. 1987), subsume aggression under the general rubric of disruptive behaviours. However, aggression like other symptoms of dementia probably has different etiologies and this warrants a more detailed examination if this is one of the most disabling behaviours for the person with dementia (Hope and Fairburn, 1992). While aggression may be related to other symptoms of dementia, and the general inventories make these relationships operational, important differences warrant the use of specific measures when

particular symptoms need to be studied in detail. A closer examination of the nature, incidence, severity and triggers for aggression will assist carers to implement strategies for improvement.

Apart from the larger inventories already mentioned, several other behavioural disturbance inventories also measure aggression and are used in clinical practice. A number of these include aggression with agitation, resistiveness, irritability and other behaviours such as apathy and wandering. In these shorter inventories aggression is only one of three or more sub-scales. Examples include the Caregiver Obstreperous Behaviour Rating Scale (COBRA) (Drachman, et al. 1992), the Disruptive Behaviour Rating Scale (DBRS) (Mungas, et al. 1989) and the Irritability, Aggression and Apathy Scale (Burns, et al. 1990). The Caregiver Obstreperous Behaviour Rating Scale (COBRA) (Drachman, et al. 1992) divides into four main areas with 30 items: aggressive/assaultive; disordered ideas/ personality; mechanical/motor; and vegetative. Studies of the instrument's reliability and validity are limited, but reported by Drachman, et al to be moderate to high, with high inter-rater reliability (0.73-0.99) and significant test-retest reliability coefficients ( $p < 0.01$ ). The Disruptive Behaviour Rating Scales (DBRS) (Mungas, et al. 1989) rates disruptive behaviour with 21 items in four areas: physical aggression, verbal aggression, agitation, and wandering. Reliability and validity reports of this measure are sparse, however Mungas, et al. found clear evidence of convergent validity for all four sub-scales, good psychometric properties for most items, and inter-rater reliability of greater than 0.83. The Irritability, Aggression and Apathy Scale (Burns, et al. 1990) incorporates the Yudovsky Aggression Scale (1 item each for showing anger, getting into an argument, pouting/sulking and raising the voice in anger), with 1 item relating to irritability and 5 items to apathy. When developing the scale, inter-rater reliability was over 0.85, and scores correlated closely with those on the Dependency Rating Scale.

None of these scales are reported more favourably in measuring aggression than the larger inventories such as the Cohen-Mansfield Agitation Inventory (Cohen-Mansfield, 1986), the CERAD-BRSD (Tariot, 1996; Patterson, et al. 1999) and the BEHAVE-AD (Reisberg, et al. 1987). A more recently developed inventory measuring aggression as well as other symptoms of dementia is the Challenging Behaviour Scale (CBS) (Moniz-Cook, Woods, Gardiner, Silver, et al. 2001). This scale was developed to measure challenging behaviour in residential settings. It comprises 25 items measuring physical aggression, verbal aggression and noise making, wandering, urinating in public, stripping, inappropriate sexual behaviour and deviant behaviour. The CBS is completed by a member of staff, usually a key worker familiar with the resident. It takes about five to seven minutes to complete. There are four rating measures, three rated by the staff member and the fourth, a computed score. First, a rating is given to indicate if the behaviour has been displayed in the past eight weeks. If yes, the frequency and severity of that behaviour is also rated. Scores for the total number (0-25), frequency (0-100) and difficulty (0-100) are calculated. The fourth measure, total level of 'challenge' is calculated as the sum of the products of frequency and difficulty ratings for each behavioural item on the scale.

The authors (Moniz-Cook, Woods, Gardiner, Silver, et al. 2001) reported the CBS to have good reliability and adequate validity. Internal consistency was excellent with Cronbach's alpha for the four measures ranging from 0.82 to 0.87, and test-retest reliability was excellent with ICC's ranging from 0.72 to 0.99. Inter-rater reliability was better when staff received training or when staff of mixed qualification completed the scale (ICC's ranged from 0.72 to 0.96), than when untrained raters completed it (ICC's ranged from 0.27 to 0.67). There was some support for construct validity. CBS scores were significantly correlated with the Social Disturbance Scale of the Clifton Assessment Procedures for the Elderly (CAPE-BRS), and there was a moderate to strong relationship between the 'challenges' score and time sampled direct observations of challenging behaviour. The CBS was also shown to discriminate between dementia and no dementia suggesting support for discriminative validity.

The Chinese version of the CBS has also been found to have very good internal consistency (Cronbach's alpha of 0.86) and inter-rater reliability (ICC = 0.79), and excellent test-retest reliability (ICC = 0.98) (Lam, Chan, Mok, Li, et al. 2006).

The CBS has been used as an outcome measure in two clinical studies. One assessed the effectiveness of staff training (Asthill, 2004) and another evaluated the effect of an antiepileptic drug on challenging behaviour (Hurtado, Koepp, Sander and Thompson, 2006). In both studies CBS scores improved significantly as a result of the intervention.

This measure shows some promise, however, at this time there is insufficient evidence to recommend its use and it was difficult to obtain a full copy of the measure for assessment. The eight week timeframe for rating may be problematic for some residential settings and may present issues of staff burden. The CBS also is not really an individual symptom measure for aggression as it contains other behaviours, such as wandering, which would not be classed as aggression. In view of these issues the RAGE seems a more appropriate instrument to recommend at this stage.

The most widely reported dementia-specific aggression-instrument with good reliability and validity, and used widely in different clinical and research settings is the Rating Scale for Aggressive Behaviour in the Elderly (RAGE) (Patel and Hope, 1992a). Other reliable measures include the Social Dysfunction Aggression Scale (SDAS) (Wistedt, et al. 1990), the Staff Observation Aggression Scale (SOAS) (Palmstierna and Winstead, 1988), the Overt Aggression Scale (OAS) (Yudovsky, et al. 1986), and the Ryden Aggression Scale (Ryden, 1988). In these dementia-specific scales, aggressive behaviour is divided into four domains, which may overlap: verbal aggression; aggression directed at others; aggression directed at self; aggression directed at objects, or using objects. When these different facets of aggression are measured with these scales and compared with the inventories that cover aggression as well as other related behaviours, the degree of overlap is comparatively small (Shah and Allen, 1999).

Table 38 below identifies the relative merits of each of the leading aggression-specific measures.

**Table 38 Decision Summary of the BPSD Global Leading Contenders**

- Rating Scale for Aggression in the Elderly (RAGE)
- Overt Aggression Scale (OAS)
- Social Dysfunction and Aggression Scale (SDAS)
- Staff Observation Aggression Scale (SOAS)
- Ryden Aggression Scale (RYDEN)
- Challenging Behaviour Scale (CBS)

	Domains/Sub domains	Applicability/ Stage	Patient	Proxy	Availability/ Cost	Training/Manual	Admin time
<b>RAGE</b>	19 aggressive behaviours assessed over preceding 3-5 days: 1 item assesses the effect of aggressive behaviours on caregiver's use of restraint 1 item assesses caregiver's overall assessment of level of aggression.	Mild to severe		✓	Free to all users	Brief training recommended and written instructions available.	5 minutes.
<b>OAS</b>	4 categories of aggression assessed and quantified for time of occurrence and duration: verbal, physical against objects, physical against others and self, physical against self Assesses 11 possible interventions for aggression.	Moderate to severe		✓	Free to all users	No training required.	5-10 minutes.
<b>SDAS</b>	9 items cover incidents and severity of outward aggression to objects and others. Two items cover incidents and severity of inward aggression (suicidal and self injurious behaviour) observed at time of incident and over a medium period of time.	Mild to severe		✓	Free for all users	Limited training required.	5-10 minutes.

<b>SOAS</b>	2 factors cover the type of verbal and physical aggression, and the factors provoking aggression, the target of aggression, the consequences and the measures taken to stop or control aggression. Nurse or care staff assessment of nature, number and severity of aggression occurs through direct observation, immediately following each incident. Staff rate the severity of aggression on a visual analogue scale.	No aggression to extremely severe.		✓	Free for all users	No training required.	5-10 minutes.
<b>RYDEN</b>	17 items of physical aggression, 4 items of verbal aggression and 5 items of sexual aggression. Frequency rated for all items from one or more times daily to less than once a year or never.	No aggression incidents to one or more times daily.		✓	Free for all users	No training required.	10 minutes.
<b>CBS</b>	25 items measuring physical aggression, verbal aggression and noise making, wandering, urinating in public, stripping, inappropriate sexual behaviour and deviant behaviour.	Moderate to severe.		✓	Difficult to obtain copy of the instrument.	Training recommended.	5 to 7 minutes.

	Citations	Psycho-metrics	Use in Practice (to date)	Judgments/Comments
<b>RAGE</b>	112 journal articles 2 books	Good to very good	<p>Assesses: Informal and formal caregiver's observation and evaluation of the incidence, frequency and severity of aggressive behaviours in dementia in all care settings over the previous 3-5 days, and the caregiver's use of restraint for aggression.</p> <p>Able to be employed in community and all health care settings.</p> <p>Employed in research for baseline behaviours, intervention studies and for care planning.</p>	<p>Able to be used in the community and all health care settings and in research as a specific measure of aggression in dementia.</p> <p>Has very good to excellent psychometric properties in comparison to other dementia-specific aggression measures in English and translated versions.</p> <p>Is rated higher for assessing aggression than other measures which assess aggression as only one symptom of behavioural disturbance in dementia.</p> <p>Covers the full range of aggressive behaviours that occur in Alzheimer's disease and other dementias, e.g. fronto-lobal.</p>
<b>OAS</b>	280 journal articles 2 books	Good	<p>Assesses: Formal caregiver's direct observation and rating of a limited number of aggressive behaviours over a short period of time, and the caregiver's use of a range of interventions, some of which are more suited to mental health programs.</p> <p>Able to be employed in all health care settings.</p> <p>Employed in research in intervention studies.</p>	<p>Is a good indicator of type and severity of aggression in person with dementia, although not applicable only to this population. Some interventions identified not suitable for dementia care.</p> <p>Short instrument and easy to use by care-staff requiring direct observation of aggression by care staff/clinicians.</p> <p>OAS-Revised scale increases the range of possible interventions to reflect current dementia and neuro-rehabilitation care practices.</p> <p>Does not cover the full range of aggressive behaviours occurring in all dementias, rather those that occur in mental illness.</p>
<b>SDAS</b>	159 journal articles 1 book	Adequate to good for the outward aggression	<p>Assesses: Formal caregiver's direct observations of aggression in all health care settings.</p>	<p>Compares favourably with other aggression scales measuring nature, incidence and severity of aggression in dementia.</p> <p>Scores do not significantly change between different cultures.</p>

		items and the inward aggression items	Employed in research for intervention studies.	Does not cover the full range of aggressive behaviours occurring in all dementias.
<b>SOAS</b>	101 journal articles 1 book	Adequate to good	Assesses: Specifically designed for nurses and assessed by direct observation at the time of incident and over a short period of time.  Able to be employed in all health care settings.  Employed in research for care planning, to monitor care approaches/interventions for aggression.	Has good to very good psychometric properties. SOAS-Revised has similar results in English speaking subjects. SOAS scores decline significantly during observation periods (Hawthorn effect) and are subject to a halo effect, whereby the effect of staffs' experience with the aggressive behaviours will influence their judgment of aggressiveness severity.  Covers a wide range of aggressive behaviours in dementia.
<b>RYDEN</b>	56 journal articles 2 books	Good to very good	Assesses: Can be used by caregivers in the community as well as care staff in all health care settings.  Employed also in intervention studies.	Has good psychometric properties compared with other measures of aggressive behaviour.  Covers a wide range of aggressive behaviours in dementia.  Does not assess aggression severity or the consequences of aggression for the person or others, or management approaches.
<b>CBS</b>	10 journal articles 1 book	Good to very good	Assesses: Challenging behaviour of patients in long term care facilities.  Used as an outcome measure in intervention studies.	Has good psychometric properties.  Contains items addressing behaviours other than aggression.  Difficult to obtain copy of the instrument.  Very few citations and the instrument has not been extensively used.

None of the aggression-specific instruments have been considered as a “gold standard”. The retrospective nature of data collection from caregivers can lead to recall bias, however, an instrument that focuses specifically on aggression moderates this potential (Shah and Allen, 1999). At the same time, while direct observation of each aggressive incident within a short time frame does not necessarily capture the overall pattern of aggression, when data on each incident is combined it may provide a more accurate picture of the types and severity of aggression, as well as the triggers and modifiers. Behavioural mapping techniques, such as the Staff Observation Aggression Scale (SOAS) (Palmstierna and Wistedt, 1987), do capture the immediate detail required; however, they also may influence the behaviour being observed. While the Staff Observation Aggression Scale (SOAS) (Palmstierna and Wistedt, 1987) has been cited frequently in the literature as the measure of choice by care staff, like some other measures of aggression in dementia, scores tend to spontaneously decline during a period of serial measurement. This is likely to occur when raters systematize their observations, which leads to identification of an individual’s aggression pattern and avoidance of precipitating and provocation factors in the clinical setting (Shah, 1999). If the measure employed aims to not only measure aggression but also to influence its incidence and severity through the constant presence of the observer, then it is understandable that the SOAS is the measure of choice for care staff. However, the SOAS necessarily involves an educational process for the raters, which may potentially obscure the effect of any alternative intervention for the amelioration of aggression (Shah and Allen, 1999).

The Rating Scale for Aggression in the Elderly (RAGE) is a more objective measure and specifically excludes analysis of the reasons for aggression, therefore, is less likely to influence the ward/unit culture and care practices (Shah and Allen, 1999). The scores are recorded by a senior nurse who observes incidences of aggression every two days for some weeks. The RAGE does not show the tendency for aggression to decline during staff observation. This feature suggests the RAGE is preferable for use in the clinical and community setting. Similarly, unlike RAGE, other scales measuring aggression in dementia have good reliability, but lack sufficient evidence of validity, for example the Ryden scale (Ryden, 1988), Social Dysfunction and Aggression Scale (SDAS) (Wistedt, et al. 1990) and the Overt Aggression Scale (Yudovsky, et al. 1986). Because of the many structural problems in care settings for persons with dementia, such as staff shortages, poor skill mix, staff burnout and care strategies/routines that are task-focused, it is important that the instrument selected to measure aggression is both reliable and valid so as to minimise observer-rated biases that might occur in staff who are affected by these problems. So while each of the previously developed contender measures of aggression in dementia have many favourable features, they do not rate as highly as the RAGE on a number of criteria such as validity and observer bias. It is important that raters are able to discriminate between a genuine change in the level of aggressive behaviour and a change in the perception of aggressive behaviour by staff or carers (Shah and Allen, 1999). The RAGE is designed as an objective measure of aggression in dementia, making it the measurement of choice for assessing the presence, nature, incidence and severity of aggression.

#### **9.4.2.1 Rating Scale for Aggressive Behaviour in the Elderly (RAGE)**

Aggression was defined by Patel and Hope (1992) as “an overt act involving the delivery of noxious stimuli to (but not necessarily aimed at) another organism, object or self which is clearly not accidental.” This definition progressed the thinking about measuring aggression specific to dementia which, at the time, was routinely included in the broader concept of behavioural disturbance and linked to agitation and aggression in the same measure, such as in the Cohen-Mansfield Agitation Inventory (Cohen-Mansfield and Billings, 1986). Given the relatively high incidence of aggression they were encountering in people with dementia in the hospital setting, and their dissatisfaction with previously developed measures of aggression in dementia, Patel and Hope developed the Rating Scale for Aggressive Behaviour in the Elderly (RAGE). The initial purpose of doing so was to assist nurses in measuring aggressive behaviour within in-patient psychogeriatric populations. Patel and Hope also aimed to employ RAGE for research on both the

effects of potential treatment of aggressive behaviour and on the relationships between aggressive behaviours and other factors, such as the care context.

The RAGE has 21 items and provides a composite measure of the quantity and severity of aggressive behaviour and gives a score on a 4-point scale (0-3) for each of the 21 items and the total score. Each of the sub-scales scores and total score agree with checklists of aggressive behaviours commonly occurring in persons with dementia and mental illness that have been constructed by expert clinicians, regular care staff, family caregivers and researchers when observing these behaviours over various time periods. Item 20 identifies the consequences of staff response to the aggressive behaviour in relation to using restraint. Item 21 asks the rater to make an overall assessment of aggressive behaviour using a 4-point scale (0 – not at all; 1 – mildly; 2 – moderately; 3 – severely). The RAGE items group into three factors: verbal aggression, physical aggression and anti-social behaviour, reflecting the most commonly occurring aggressive behaviours observed in the person with dementia (Patel and Hope, 1992a; Patel and Hope, 1992b; Shah, et al. 1997; Shah, et al. 1998).

### **Clinical applications**

The main use of RAGE is for caregivers and/or care staff to identify: specific frequencies and types of aggressive behaviours; aggression patterns occurring over the previous three days; and staff's use of restraint (physical and chemical) to manage aggressive behaviour; and an overall assessment of the frequency of aggressive behaviour over the past three days. Regular caregivers are chosen to convey this information because the person with dementia may not exhibit aggression during direct assessment by a clinician in a shorter time period. RAGE is also suitable for use in research to identify baseline behaviours and also in drug and treatment/therapy trials (Patel and Hope, 1992a, 1992b; Shah and De, 1998; Shah, Evans, Parkash, 1998; Patel, Hope, Hall, Fairburn, 1995; Shah, et al. 2000). Translations of the English version of the RAGE are reliable and valid measures of aggression in dementia from non-English speakers, including the Chinese (Lam, Chui and Ng, 1997) and Scandinavian populations (Patel and Hope, 1992b).

### **Reliability and validity**

As there was no "gold standard" aggression instrument available when RAGE was being developed, reference to experienced expert medical staff, nurses and caregivers occurred to develop the instrument, initially by observing aggressive behaviours in 125 patients with cognitive impairment and/or mental illness, and then tested with 90 psychogeriatric patients. The RAGE was developed according to the basic steps outlined by Hall (1997, 1980) for ensuring content validity in observer completed rating scales.

Internal consistency was high, with a Cronbach's alpha of 0.89 when used on 13 people with dementia displaying aggressive behaviours (Shah, Evans and Parkash, 1998). Internal consistency of the Chinese translation of RAGE (C-RAGE) is high with an alpha coefficient of 0.74. The split half reliability of the C-RAGE is 0.79 (Lam, Chui and Ng, 1997). The questions can be answered reliably by anyone who provides the day-to-day care with the person being assessed, such as informal and formal caregivers and nurses, to assist in care planning and treatment regimens. The inter-rater reliability of the total score and individual scores are high in all these respondent groups (Patel and Hope, 1992a, 1992b; Patel, Hope, Hall, Fairburn, 1995; Shah, Evans, Parkash, 1998). In the original studies (Patel and Hope, 1992a, 1992b) total score IRR was estimated using the correlation co-efficient. Fifty residents with dementia were rated by nurses with access to a checklist of usual aggressive behaviours occurring in the wards where they worked, and 40 residents were assessed by nurses who had no checklist. The correlation was 0.94 ( $p < 0.001$ ) when using a checklist and 0.54 without a checklist. The values for the individual items of weighted Kappa for inter-rater reliability ranged between 0.61 and 0.92. These correlations have been confirmed in subsequent studies (Patel, Hope, Hall, Fairburn, 1995; Shah, Evans, Parkash, 1998).

The original RAGE was designed to measure aggressive behaviour for the preceding three days. However, this was adapted for use over the preceding week because the original reliability study conducted by Patel and Hope (1992a, 1992b) reported little change over 7 days and 14 days in early studies in nursing homes (Shah, et al. 1997), and in acute and continuing care psychogeriatric wards (Shah, et al. 1998; Shah and De, 1998). Test-retest and inter-rater reliability studies show high agreement when RAGE and C-RAGE are used by clinicians, care staff and researchers (Lam, Chui and Ng, 1997). There is no significant decrease in reliability between 6 hour test-retest interval (0.75) and the 7 day (0.76) and 14 day (0.84) interval using stable patient groups (Patel, Hope, Hall, Fairburn, 1995; Shah, Evans, Parkash, 1998).

RAGE has been extensively tested and found suitable as an outcome measure in treatment studies and is sensitive to change. For 14 subjects independently rated to have shown a decrease in aggressive behaviours, mean RAGE scores for all items fell from 17.8 to 6.5 ( $p < .05$ ), while for 7 subjects independently rated to have shown an increase in aggressive behaviours, mean RAGE scores for all items rose from 6.7 to 16 ( $p > .05$ ) (Patel and Hope, 1992b). It is a reliable and valid measure of aggression not only in dementia, but also in mental illness and intellectual disability, in different care contexts and in the community (Patel and Hope, 1992a, 1992b; Shah, and De, 1998; Shah, Evans, Parkash, 1998; Patel, Hope, Hall, Fairburn, 1995; Shah, et al. 2000). The Chinese translated version (C-RAGE) has been validated in the Chinese population (Lam, Chui and Ng, 1997). None of the studies reviewed identified any flaws in the psychometric properties of the instrument and it compares favourably with other well-known and widely used validated measures of aggression in dementia, such as the Cohen-Mansfield Agitation Inventory (CMAI), the Brief Agitation Rating Scale (BARS), and the Behavioural Pathology in Alzheimer's Disease Scale (BEHAVE-AD) (Shah, Evans, Parkash, 1998; Gormley, Rizwan, Lovestone, 1998; Lam, Chui and Ng, 1997; Shah and De, 1997; Oye, Loke, Chan Kwok, 2005; Bathareethan and Shah, 2000).

While O'Malley, et al. (2002) recommended against relying on any one measure for aggression in the population with dementia, in order to avoid the possibility of missing data, the RAGE does capture the nature, incidence and severity of aggression routinely occurring in this population, as reported by care staff and family carers over different periods of time. Research shows this to be preferable to relying on reports of aggression by caregivers at one point of time, or over only one day, as this may be triggered by particular stimuli occurring either in the person, or in the care context, at that point in time. The aggression may subsequently abate once the noxious stimuli is reduced or removed (Patel and Hall, 1995; Shah, 1999; Shah, Chiu and Ames, 2000).

In reviewing measures of aggression, particular attention was paid to the number of citations, the coverage of aggression specifically, application in a range of settings and administration time, as well as the psychometric properties. The RAGE addressed these factors more than any other measure of aggression in dementia. The most salient features of RAGE as a measure of aggression in dementia are that it is valid and reliable in English and some other languages, compares favourably with other validated measures of aggression; focuses specifically on aggression incidence and severity, is accessible to caregivers and care staff and can be relied on in clinical intervention studies. RAGE scores do not decline over time with repeated observations by the same raters as a result of perceived patterns of aggressive behaviours. Importantly, it captures information about use of restraint associated with aggression, which is very useful in care planning and care monitoring. The summary rating sheet for RAGE can be found below:

**Table 39 Summary of Ratings for Aggression Instruments**

Criteria	Instrument	
	Weight	RAGE
Availability of comparison data	3	1
Length/feasibility of instrument for inclusion in battery	2	2
Complexity of administration/ cognitive burden	2	3
Cultural Appropriateness	1	2
Ease of obtaining score	2	3
Sensitivity to dementia	3	3
Reliability evidence	3	3
Validity evidence	3	2
Cost of the instrument	2	3
Cost of instrument administration	2	2
<b>Weighted Total</b>		<b>55</b>

### 9.4.3 Agitation

The Neuropsychiatric Inventory, which has been reviewed under Global BPSD measures in an earlier section, is also often used to assess agitation. Further literature searches identified two other commonly used measures to assess agitation. These were the Cohen-Mansfield Agitation Inventory and the Pittsburgh Agitation Scale. Summary details are provided in the table and sections below:

**Table 40 Decision Summary Table for Agitation Instruments**

- CMAI (Cohen-Mansfield Agitation Inventory)
- PAS (Pittsburgh Agitation Scale)

	Domains/Sub domains	Applicability/ Stage	Patient	Proxy	Availability/ Cost	Training/Manual	Admin time
<b>CMAI</b>	29 items describing agitated behaviours, summarised into three domains or factors: aggressive behaviour, physically non-aggressive behaviour, and verbally aggressive behaviour. Each item is rated on a 7 point scale.			✓	Free with authors permission.	Training recommended. Instruction manual and training provided by authors.	10-15 minutes.
<b>PAS</b>	4 general categories of severity of agitation: aberrant vocalisation, motor agitation, aggressiveness and resisting care. It comprises one item for each of these categories, with each item rated on a 4 point scale. 1 indicator of interventions used during rating period.	Not present to severe		✓	Free for all users.	Minimal training.	1 minute.

	Citations	Psycho-metrics	Use in Practice (to date)	Judgments/Comments
<b>CMAI</b>	242 journals 2 books	Good to very good	Assesses: Incidence and severity of agitated behaviours that can be summarised into three domains or factors: aggressive behaviour, physically non-aggressive behaviour, and verbally aggressive behaviour.  Different versions of the instrument are also available. Long (normal 29	Frequently used in the clinical setting and been adapted for use in the nursing home. Can be used by the family carer with training.  Is widely used as a baseline measure in research and in intervention studies.  Reliability and validity is good to very good.  Translated in several languages without compromising validity and reliability.

			<p>items), short and community for use in different settings by different raters, such as medical staff, nurses and family carers.</p> <p>Frequently used in a variety of research studies worldwide.</p>	
<b>PAS</b>	<p>152 journals 2 books</p>	<p>Good</p>	<p>Assesses: Presence and severity of 4 types of agitated behaviours, as well as type of intervention used during the observation period.</p> <p>Easily and frequently used by care staff.</p> <p>Quickly administered in the clinical setting and suitable for use in the community.</p> <p>Employed in intervention studies and research as a baseline measure of agitated behaviour.</p>	<p>Readily accessible for clinical and community settings, and used in research.</p> <p>Combination of categories and intensity makes it comprehensive.</p> <p>Reliability is limited, however, good validity.</p> <p>No translated versions available as yet.</p>

### 9.4.3.1 Cohen Mansfield Agitation Inventory (CMAI)

The Cohen-Mansfield Agitation Inventory (Cohen-Mansfield, Marx and Rosenthal, 1989) is a caregiver rating scale developed by a team of experts, and based on information obtained from interviews with nursing home staff members and extensive literature reviews. It is designed to assess the frequency of agitated behaviours in elderly persons with cognitive impairment and has been widely used in studies evaluating the effectiveness of drug treatments or other interventions, and research studies investigating agitated behaviour. The instrument is free and comes with an instruction manual. Training is recommended and is provided either through the instruction manual or a training video which is also available from the authors. The CMAI has been translated into Dutch, Danish, French, German, Greek, Norwegian, Spanish, Chinese, Korean, Japanese and Hebrew. Information about how to obtain these translations is available in the Instruction manual.

The CMAI comprises 29 items describing agitated behaviours that can be summarised into three domains or factors: aggressive behaviour; physically non-aggressive behaviour; and verbally aggressive behaviour. It is administered by a staff member who rates each item on a seven point scale based on the frequency with which the patient has engaged in the behaviour in the previous two weeks. Completion time is 10-15 minutes.

Several other versions of the instrument are also available. The 14 item version (CMAI-Short) (Werner, Cohen-Mansfield, Koroknay and Braun, 1994), also used in the nursing home population, has the same domains as the CMAI and is rated on a five point scale. The 37 item CMAI-Community (CMAI-C) (Cohen-Mansfield, Werner, Watson and Pasis, 1995) is available for use in the community and can be used by both professional and family caregivers. This instrument can be summarised into four domains or factors: physically non-aggressive behaviour; physically aggressive behaviour; verbally non-aggressive behaviour; and verbally aggressive behaviour. It consists of 36 frequency items (rated on a 7 point scale) and one item to determine the time of day the behaviour occurred. In addition to these versions there is also the CMAI with expanded definitions (which provide additional examples of each behaviour) and the disruptiveness form in which the disruptiveness of the behaviour is rated along with the frequency. This is available in the Long (i.e. the normal 29 item version), Short and Community forms.

Ratings for the CMAI and CMAI-C range from a score of 1 (never) to 7 (several times an hour). Ratings for the CMAI-Short range from 1 (never) to 5 (a few times an hour or continuous for half an hour or more). Item scores can be summed to give a total score of 29-203 for the CMAI, 14-56 for the CMAI-Short, and 36-216 for the CMAI-C. Alternatively, or in addition, items relating to specific behaviours of interest only, or items relating to each of the domains can be summed. Regardless of the scoring used, higher scores indicate greater agitation or behavioural disruption.

There is considerable evidence indicating the CMAI has good to excellent psychometric properties. Most studies have provided considerable information to ensure the findings can be appropriately interpreted. The internal structure of the instrument has generally been confirmed through factor analysis confirming the domains proposed by the authors (Cohen-Mansfield, 1986; Cohen-Mansfield, Marx and Rosenthal, 1989; Cohen-Mansfield, 1991; Cohen-Mansfield, Werner, Watson and Pasis, 1995; Miller, Snowdon and Vaughan, 1995; de Jonghe and Kat, 1996; Schreiner, Yamamoto and Shiotani, 2000; Choy, Lam, Chan, Li, et al. 2001; Vespa, Gori, Bonaiuto, Cruciani, et al. 2002; Weiner, Tractenberg, Jin, Gamst, et al. 2002; Suh, 2004; O'Leary, Jyringi and Sedler, 2005; Rabinowitz, Davidson, De Deyn, Katz, et al. 2005; Zuidema, de Jonghe, Verhey and Koopmans, 2007). In the main, studies have generally confirmed the three factor structure of aggressive behaviour, physically non-aggressive behaviour, and verbally agitated behaviour for the CMAI and CMAI-Short. Some studies however, have reported a four factor solution that included hiding and hoarding (Schreiner, Yamamoto and Shiotani, 2000; Suh, 2004; Rabinowitz, Davidson, De Deyn, Katz, et al. 2005). One Dutch study (Zuidema, de Jonghe, Verhey and Koopmans, 2007) reports both a restricted three factor solution (as reported in other studies) and an unrestricted six factor solution (aggressive behaviour, physically non-aggressive

behaviour, verbally agitated behaviour, hiding and hoarding, vocal agitation and miscellaneous items factor (repetitious mannerisms, spitting)).

Findings for the CMAI-C are mixed. One study (Cohen-Mansfield, 1991) reports a four factor solution, (physically non-aggressive behaviour, physically aggressive behaviour, verbally non-aggressive behaviour, and verbally aggressive behaviour), one (Cohen-Mansfield, Werner, Watson and Pasis, 1995) suggests both a three (as for the CMAI) and a four factor solution may be appropriate and another (Weiner, Tractenberg, Jin, Gamst, et al. 2002) stated that the four factor solution was not appropriate and that a total score was best suited to describe behaviour.

The reliability and internal consistency of the CMAI is excellent (Cohen-Mansfield, 1986; Finkel, Lyons and Anderson, 1992; Miller, Snowdon and Vaughan, 1995; de Jonghe and Kat, 1996; Shah, Evans and Parkash, 1998; Choy, Lam, Chan, Li, et al. 2001; Vespa, Gori, Bonaiuto, Cruciani, et al. 2002; Suh, 2004; Rabinowitz, Davidson, De Deyn, Katz, et al. 2005). Studies report Cronbach's alpha as ranging from 0.75 to 0.91. Findings for the subscales have also been generally very good with most studies reporting Cronbach's alpha above 0.70 except for two (Choy, Lam, Chan, Li, et al. 2001; Rabinowitz, Davidson, De Deyn, Katz, et al. 2005) that reported alphas ranging from 0.62 to 0.78 for physically non aggressive behaviour, and 0.59 to 0.78 for verbally agitated behaviour. Internal consistency for the CMAI-C and CMAI-Short has not been reported.

The instrument has excellent test-retest and inter-rater reliability (Cohen-Mansfield, 1986; Cohen-Mansfield, Marx and Rosenthal, 1989; Cohen-Mansfield, 1991; Finkel, Lyons and Anderson, 1992; Werner, Cohen-Mansfield, Koroknay and Braun, 1994; Cohen-Mansfield, Werner, Watson and Pasis, 1995; Miller, Snowdon and Vaughan, 1995; Buettner, Lundegren, Lago, Farrell, et al. 1996; de Jonghe and Kat, 1996; Koss, Weiner, Ernesto, Cohen-Mansfield, et al. 1997; McGee, Orengo, Kunik, Molinari, et al. 1997; Kunik, Graham, Snow-Turek, Molinari, et al. 1998; Ponce, Molinari, Kunik, Orengo, et al. 1998; Shah, Evans and Parkash, 1998; Weiner, Koss, Patterson, Jin, et al. 1998; Whall, Black, Yankou, Groh, et al. 1999; Choy, Lam, Chan, Li, et al. 2001; Suh, 2004). Studies investigating test-retest reliability report correlations ranging from 0.79 to 0.97 for the CMAI and a correlation of 0.83 for the CMAI-C. Test-retest reliability for the CMAI-Short has not been reported. Studies investigating inter-rater reliability cite correlations ranging from 0.76 to 0.96 for CMAI, 0.71 to 0.92 for CMAI-C, and 0.82 to 0.92 for CMAI-Short.

The validity of the CMAI has also been confirmed. Evidence for construct validity in terms of the extent to which scores on the CMAI relate to other measures in a manner consistent with theoretically derived hypotheses comes from several studies (Cohen-Mansfield, 1986; Cohen-Mansfield, Taylor and Werner, 1998; Weiner, Tractenberg, Teri, Logsdon, et al. 2000; Villanueva, Smith, Erickson, Lee, et al. 2003; O'Leary, Jyringi and Sedler, 2005). The CMAI shows expected correlations with the Rapid Disability Rating Scale (RDRS) and the Pain Assessment for the Dementing Elderly (PADE). It is also associated with the presence of psychotic symptoms such as delusions, and paranoia as determined by nurses' ratings. The CMAI-C is associated with the presence of delusions and hallucinations.

There is considerable evidence for construct validity in terms of correlations with other well known instruments that measure behavioural disturbance (Finkel, Lyons and Anderson, 1992; Miller, Snowdon and Vaughan, 1995; de Jonghe and Kat, 1996; Weiner, Williams and Risser, 1997; Shah, Evans and Parkash, 1998; Weiner, Koss, Patterson, Jin, et al. 1998; Logsdon, Teri, Weiner, Gibbons, et al. 1999; Ramadan and Naughton, 1999; Volicer, Camberg, Hurley, Ashley, et al. 1999; Weiner, Tractenberg, Teri, Logsdon, et al. 2000; Choy, Lam, Chan, Li, et al. 2001; Cohen-Mansfield and Libin, 2004; Skjerve, Holsten, Aarsland, Bjorvatn, et al. 2004; Suh, 2004; Nagels, Engelborghs, Vloeberghs, Van Dam, et al. 2006). The CMAI has been found to have significant correlations with the following well known measures of agitation and aggression: Agitated Behaviors Mapping Instrument (ABMI); Brief Agitation Rating Scale (BARS); and the Rating Scale for Aggressive Behaviour in the Elderly (RAGE). It also correlated with other indicators of agitation such as: Daily Verbalisation Scores (VS) and Actigraphic Recordings made using an octagonal motion-logger. Significant correlations with the following global measures of behavioural

disturbance are also reported: Behavioral and Emotional Activities Manifested in Dementia (BEAM-D); Behavioral Pathology in Alzheimer's Disease Scale (BEHAVE-AD); Behavioral Syndrome Scale for Dementia (BSSD); Dutch Behaviour Rating Scale for Psychogeriatric Inpatients (GIP); Nursing Home Problem Behaviour Scale (NHPBS); and the Revised Memory and Behavior Problems Checklist (RMBPC). The CMAI-C has been shown to significantly correlate with Agitated Behaviour in Dementia (ABID) and the global scale, Consortium to establish a Registry for Alzheimer's Disease-Behavior Rating Scale dementia (CERAD-BRSD). The CMAI-Short correlated with an agitation indicator, the Visual Analogue Scale (VAS)-agitation.

There is also evidence for the discriminant validity of the CMAI (Koss, Weiner, Ernesto, Cohen-Mansfield, et al. 1997; Ponce, Molinari, Kunik, Orengo, et al. 1998; Weiner, Koss, Patterson, Jin, et al. 1998; O'Leary, Jyringi and Sedler, 2005; Schreiner, Ballaban, Brannath, Lang, et al. 2005). Both the CMAI and CMAI-C have been shown to discriminate between different levels of dementia severity.

There is considerable evidence that the CMAI meets the criteria for responsiveness. The CMAI and CMAI-C have been shown to be sensitive to drug treatment with scores improving significantly as a result of treatment (Calkin, Kunik, Orengo, Molinari, et al. 1997; De Deyn, Rabheru, Rasmussen, Bocksberger, et al. 1999; Ramadan and Naughton, 1999; Tariot, Schneider, Mintzer, Cutler, et al. 2001; Brodaty, Ames, Snowdon, Woodward, et al. 2003; Fujikawa, Takahashi, Kinoshita, Kajiyama, et al. 2004; Rabinowitz, Katz, De Deyn, Brodaty, et al. 2004; Rainer, Mucke, Kruger-Rainer, Haushofer, et al. 2004; Suh, Son, Ju, Jcho, et al. 2004; De Deyn, Katz, Brodaty, Lyons, et al. 2005; Suh, Greenspan and Choi, 2006). Sensitivity to non-pharmacological interventions has also been reported for the CMAI. These include music therapy (Gerdner and Swanson, 1993; Goddaer and Abraham, 1994; Richeson and Neill, 2004; Hicks-Moore, 2005), air mat therapy (Buettner, Lundegren, Lago, Farrell, et al. 1996; Shalek, Richeson and Buettner, 2004) and a motor and occupational intervention that included music and other social and occupational activities (Vespa, Gori and Spazzafumo, 2002). Richeson (2003) and Skjerve, Holsten, Aarsland, Bjorvatn, et al. (2004) reported sensitivity of the instrument to the effects of animal assisted therapy and bright light therapy, however these studies only had a sample size of 10 so results should be viewed with caution. The CMAI-Short showed sensitivity to air mat therapy (Shalek, Richeson and Buettner, 2004), and snoezelen rooms (van Diepen, Baillon, Redman, Rooke, et al. 2006).

Findings regarding sensitivity to change over time are not strong. One study (Koss, Weiner, Ernesto, Cohen-Mansfield, et al. 1997) reported a non-significant change in CMAI-C scores in a sample of moderate to severely demented persons. Another (Weiner, Koss, Patterson, Jin, et al. 1998) found no significant difference in CMAI-C scores over time in a sample of persons with mild dementia.

#### **9.4.3.2 Pittsburgh Agitation Scale (PAS)**

The Pittsburgh Agitation Scale (PAS) (Rosen, Burgio, Kollar, Cain, et al. 1994) is a user friendly rating scale developed by experts in the field to assess the level of agitation in persons with dementia both in the nursing home and inpatient setting. It has also been used as an outcome measure in intervention studies. The PAS is completed by staff members during the course of their direct observation and documentation and takes less than one minute to complete. It can be obtained at no cost and training is minimal. No translations are available as yet.

The instrument measures the severity of agitation in four general categories: aberrant vocalisation, motor agitation, aggressiveness and resisting care and comprises one item for each of these categories. Each item is rated from 0 (not present) to 4 (most disruptive or unsafe behaviour). Ratings are based on behaviours observed during a rating period of typically 4 to 8 hours. Scores for 'vocalisation' and 'motor agitation' are determined by the intensity and disruptiveness within the environment, and the ease with which the behaviour can be redirected. Scores for the 'aggressiveness' dimension are based on a general description of aggressive behaviour. Scores

for 'resisting care' are based on behaviour associated with specific identified activities such as washing, dressing etc. Scores of 3 or 4 reflect behaviours that are not responsive to redirection, distraction, or other behavioural interventions. Total scores range from 0 to 16.

Evidence regarding the psychometric properties of the PAS is limited but available data suggests it has very good reliability and moderate validity. Internal consistency and inter-rater reliability have been found to be very good with Cronbach's alpha ranging from 0.75 to 0.84, and ICC's or Pearson's correlations ranging from 0.70 to 0.93 (Rosen, Burgio, Kollar, Cain, et al. 1994; Wells, Dawson, Sidani, Craig, et al. 2000; Zieber, Hagen, Armstrong-Esther and Aho, 2005).

Support for construct validity is mixed. Results from one study (Zieber, Hagen, Armstrong-Esther and Aho, 2005) show that PAS scores significantly correlated with two pain measures, the Discomfort Scale for Dementia of the Alzheimer's type (DS-DAT) and pain ratings made by palliative care and facility nurses. PAS scores have also been found to correlate significantly with scores on other measures of agitation and aggression (Rosen, Burgio, Kollar, Cain, et al. 1994). In the acute psychiatric setting, scores correlated with ratings for direct observation for vocalisation, motor agitation and aggressiveness. In the nursing home setting scores correlated with the need for restraints (either chemical or physical). PAS scores did not however show significant correlations with the Neurobehavioral Rating Scale (NBRS) (Rosen, Bobys, Mazumdar, Mulsant, et al. 1999).

The diagnostic accuracy of the PAS attests to its discriminant validity. Area under the curve (AUC) analysis indicates the instrument has a sensitivity and specificity of 78% and 95% respectively (Rosen, Bobys, Mazumdar, Mulsant, et al. 1999).

Evidence for responsiveness is limited. Findings from some studies suggest the PAS is sensitive to the effects of interventions, others do not. Results from all the studies however, need to be interpreted with caution, as none of them provide any information about the magnitude of score differences which would be clinically meaningful. Two studies reported a change in PAS scores as a result of non-pharmacological interventions. One study reported significant improvement in agitation levels as a result of stage and age based activities programs (Mahoney, 2003). In the other study, levels of agitation significantly decreased for residents receiving morning care from nurses trained in an abilities focussed program (Wells, Dawson, Sidani, Craig, et al. 2000). Evidence of sensitivity to other behavioural interventions is much weaker. One small study (Holmes, Hopkins, Hensford, MacLaughlin, et al. 2002) reported a change in score as a result of aromatherapy treatment but the sample size in this study was only nine. Another small study (Perivolaris, LeClerc, Wilkinson and Buchanan, 2006) showed no significant effect resulting from an enhanced dining program.

#### **9.4.4 Conclusions Concerning Measures of Agitation**

The summary of the ratings for the measures of agitation are provided below. It can be seen that the CMAI has higher ratings than the PAS and this is partly because the CMAI has been used more extensively in the field and thus there is more data concerning its psychometric properties. The CMAI is recommended for the assessment of agitation. However, the PAS is a new and promising instrument and its ease of implementation makes it appealing for use in clinical settings.

**Table 41 Summary of Ratings for Measures of Agitation**

Criteria	Weight	Instrument	
		CMAI	PAS
Availability of comparison data	3	2	1
Length/feasibility of instrument for inclusion in battery	2	2	3
Complexity of administration/ cognitive burden	2	2	3
Cultural Appropriateness	1	3	2
Ease of obtaining score	2	3	2
Sensitivity to dementia	3	3	2
Reliability evidence	3	3	3
Validity evidence	3	3	3
Cost of the instrument	2	3	3
Cost of instrument administration	2	2	2
<b>Weighted Total</b>		<b>60</b>	<b>54</b>

### 9.4.5 Anxiety

“Compared with the amount of research concerning depression and dementia, little research concerning anxiety in older adults has been done” (page 139) (Kane and Kane, 2000). This lack of research has also affected the development of anxiety measures for elderly populations, let alone their application to people with dementia. This is despite the fact that anxiety is very common in people with dementia, occurring in half of all patients (McKeith, et al. 1999) and it co-exists with depression in older adults (Lang and Stein, 2001).

This situation is now starting to change with a number of reviews being available for anxiety instruments in the elderly (Lang and Stein, 2001; Alwahhabi, 2003; Lauderdale and Sheikh, 2003; Kane and Kane, 2000). The present state of the art is characterised by using general anxiety measurement instruments designed for younger age groups. As Lang and Stein (2001) state, “Information about the specific assessment needs of older adults is lacking, so clinicians must largely rely on instruments that have been developed for use with younger populations.” However, there are some problems with this approach as Lang and Stein (2001) summarize:

- Instruments have not been tested on representative samples of older adults
- There is a lack of validity as instruments rely heavily on somatic symptoms which may be endorsed by the elderly because of increased incidence of physical problem
- They may tap into different underlying constructs when compared to younger age groups
- In terms of interpretation of results, the following issues also apply according to Lang and Stein (2001)
- Higher scores may be due to cardiac or respiratory problems

- It is hard to distinguish agitation from anxiety in people with dementia
- Impaired memory can be interpreted as a feature of dementia or anxiety conditions
- Fears about the future in relation to ones own life circumstances may be realistic rather than excessive

These points reflect the lack of research data about anxiety measurement in older people and suggest the need for more condition specific measurement approaches. Unfortunately, no reviews examining anxiety measurement in people with dementia were found. But a recent paper by Gibbons, et al. (2006) highlights many of the key issues, especially in terms of symptom overlap and refinement.

In light of this examination of the scientific literature, the key issues in defining the present state of anxiety measurement in dementia are:

- A lack of research data and condition specific measurement instruments
- Need to examine the key anxiety symptoms that emerge in relation to dementia
- Need to examine the relationship between anxiety and depression (as well as agitation) in this population
- Need to assess the impact of cognitive impairment on anxiety measurement

Finally, as with all associated symptoms of dementia mentioned in this chapter, it is important to determine how detailed the assessment is required to be. Is a full work-up required using specialised instruments or do clinicians wish to flag the issue as part of a wider assessment process? If clinicians wish to flag the issue, it is suggested that the relevant items from the Neuropsychiatric Inventory (NPI) (Cummings, et al. 1994) rating scale can do the job. A recent paper by Cummings, et al. (2006) shows how this instrument can be used in practice.

#### **9.4.5.1 Selection of Instruments**

Instruments were selected for comprehensive review based on literature search and instrument selection methodology outlined previously (refer Section 2). For the selection of the anxiety instruments these steps are summarised below.

#### **9.4.5.2 Literature Search**

19 anxiety instruments were initially identified on the Master Database as a result of the academic literature and practice surveys. However the vast majority of these instruments were not dementia specific, but were used in studies with older adults. In fact, as mentioned above, it is only until recently that research activity has applied these general anxiety measures to older age groups.

#### **9.4.5.3 Short-listed Instruments**

The list of 19 instruments was reduced to a short list of contender instruments based on the application of the additional selection criteria (availability of instrument, number of citations, psychometric evidence, used in clinical practice, availability of normative and clinical reference data, administration time, able to be used with the various severity levels of dementia, cost considerations, and applicability for routine care). The short list of anxiety instruments is presented below:

**Table 42 Short-listed Anxiety Instruments**

<b>Instrument</b>	<b>Original Article Cite Author(s) + Publication Date</b>
<b>Rating Anxiety in Dementia</b>	Shankar, et al. (1999)
<b>Hamilton Anxiety Rating Scale</b>	Hamilton (1959)
<b>Padua Inventory</b>	Sanavio (1988)
<b>Participant Anxiety Scale</b>	Westhuis and Thyer (1989)
<b>Beck Anxiety Inventory</b>	Beck, et al. (1988)
<b>Fear Questionnaire</b>	Marks and Matthews (1979)
<b>State-Trait Anxiety Inventory</b>	Spielberger, et al. (1970)
<b>Manifest Anxiety Scale – Elderly version</b>	Lowe and Reynolds (2000)
<b>Hospital Anxiety and Depression Scale</b>	Zigmond and Snaith (1983)
<b>Short Anxiety Screening Test</b>	Sinoff, et al. (1999)
<b>Worry Scale</b>	Wisocki, et al. (1988)
<b>Penn State Worry Questionnaire</b>	Meyer, et al. (1990)

This list includes the major anxiety measures; however few studies have applied these instruments to older adults.

Other scales which were found to be noteworthy in this area were the self-report instruments: Short Anxiety Screening Test (SAST) (Sinoff, et al. 1999); Participant Anxiety Scale (PAS) (Westhuis and Thyer, 1989) (used in Gibbons, et al. 2006), the Worry Scale (Wisocki, et al. 1988). and the new Geriatric Anxiety Inventory (GAI) (Pachana, et al. (2006). However, these instruments had few citations and less available psychometric evidence and were therefore not included in the short list.

**9.4.5.4 Reviewed Instruments**

The short-listed instruments were now classified into the following four categories: Accept; Reject; Grey Area; Grey Area – but promising new instrument. The specific criteria used to select the instruments for comprehensive review were: (1) Whether there is a copy of the instrument and original article available for review; (2) The number of citations found; (3) The amount of published psychometric evidence; and (4) whether the instrument is used in clinical practice.

Owing to the limited application of most of these measures to older populations, it was decided during the course of this analysis, to focus on the one anxiety measure that was designed for people with dementia – the Rating Anxiety in Dementia (RAID) (Shankar, et al. 1999).

**9.4.5.5 Rating Anxiety in Dementia (RAID)**

The Rating Anxiety in Dementia (RAID) (Shankar, et al. 1999) is an 18 item rating scale to assess the symptoms of anxiety in dementia (administration time: 5 - 10 minutes) which was modelled on the Cornell Scale for Depression in Dementia (CSDD). According to Gibbons, et al. (2006), the RAID has good psychometric properties, especially in regard to inter-rater agreement (82-100% for individual items) and test-retest reliability (kappa’s in the range of 0.53-1.00 for individual items, over a 7-10 day timeframe). The RAID also has high internal consistency (Cronbach’s alpha = 0.83). However, due to the co-morbidity of depression and anxiety, and symptom overlap issues (plus factor structure limitations), further validity work is required to refine and better understand individual items in the RAID. As Gibbons, et al. (2006) said, “Further refinement of the definition and measurement of anxiety in dementia will result in a better understanding of anxiety and its complex relationship with depression and other aspects of dementia” (page 207).

In summary, the RAID is a promising new rating scale for anxiety in people with dementia. However, it requires: further validity information in relation to item loadings and factor structure; and in its relationship to the construct of depression. It also needs wider application especially in the areas of cultural and language adaptation; as well as in the measurement of sensitivity to change. Finally, the RAID requires better training resources.

#### 9.4.5.6 Instrument Rankings – Summary Rating

Considering the attributes of each test as reviewed in the appendix, a scoring judgement was made according to the criteria outlined, and this information is summarised in the table below.

**Table 43 Summary of Ratings for Anxiety Instruments**

Criteria	Instrument	
	Weight	RAID
Availability of comparison data	3	2
Length/feasibility of instrument for inclusion in battery	2	2
Complexity of administration/ cognitive burden	2	3
Cultural Appropriateness	1	3
Ease of obtaining score	2	3
Sensitivity to dementia	3	3
Reliability evidence	3	3
Validity evidence	3	2
Cost of the instrument	2	3
Cost of instrument administration	2	1
<b>Weighted Total</b>		<b>57</b>

This analysis finds mostly high ratings for the RAID instrument in terms of its administration, scoring and psychometric properties.

#### 9.4.5.7 Recommendations

Based on the high summary rankings of the RAID it is recommended for the specialist assessment of anxiety symptoms in people with dementia.

#### 9.4.6 Apathy

“Apathy is the most common behavioural change evidenced by patients with AD. Early in the clinical course, in concert with the onset of memory abnormalities, patients manifest progressive emotional distancing, loss of interests, disengagement, reduced motivation, and decreased initiation. These symptoms are disproportionate to the accompanying cognitive impairment and may not be accompanied by reduced physical activity. Apathy is present in approximately half of

the patients in the early phases of the illness and is evident in nearly all of them in the final stages of AD” (McKeith, et al. 1999).

While clinically important, the emergence of interest in examining apathy as a construct in its own right is relatively recent. It can be traced to a landmark paper by Marin in 1990 into the differential diagnosis and classification of apathy. Since this time a small but growing body of literature on this topic has emerged. A number of key works are highlighted here.

Another landmark paper is by Stuss, et al. (2000) which proposes a different model of apathy to Marin (1990) which is based on sub-types reflecting neuroanatomical / neuropsychological structures. These two seminal works shape this emerging field and highlight the symptom boundary issues concerning the apathy construct, in particular its relationship to depression and withdrawal.

Other significant papers include: Verkaik, et al. (2005) which reviews the effectiveness of treatments for apathetic behaviours in dementia; the work of Aalten, et al. (2006) which demonstrates a relationship between unawareness and apathy (and psychosis), while a higher level of awareness is associated with depression and anxiety symptoms. (This work also has important implications with regard to using self-report measures with people with dementia); and Steffens, et al. (2006) in outlining a collaborative research agenda for cognitive impairment and late-life depression researchers, highlight the importance of measuring clinically relevant behaviour manifestations like apathy.

In tandem with the emergence of clinical / scientific interest in apathy as a construct, a number of review papers comment on apathy measurement issues (see Malloy and Grace, 2005; Malloy and Boyle, 2005; van Reekum, et al. 2005; and Williams, 2005). van Reekum, et al. (2005) quite rightly state there is no consensus on the appropriate clinical “gold standard” for apathy.

In light of this examination of the scientific literature, the key issues in defining the present state of apathy measurement in dementia are:

- The lack of a clinical “gold standard”
- The need for clarification of the construct - boundary issues with regard to depression and withdrawal symptoms
- That further examination is required into the relationship with cognitive functioning
- That current measures have been developed for a range of neurological conditions (brain injury, dementia, stroke and Parkinson’s Disease) and as such they may not be dementia specific

Finally, as with all associated symptoms of dementia mentioned in this chapter, it is important to determine how detailed an assessment is required. Is a full work-up required using specialised instruments or do clinicians wish to flag the issue as part of a wider assessment process? If clinicians just wish to flag the issue, it is suggested that the relevant items from the Neuropsychiatric Inventory (NPI) (Cummings, et al. 1994) rating scale can do the job.

#### **9.4.6.1 Selection of Instruments**

Instruments were selected for comprehensive review based on literature search and instrument selection methodology outlined previously (see the section on dementia staging and descriptive instruments or cognitive instruments). For the selection of the depression instruments these steps are summarised below.

### 9.4.6.2 Literature Search

Only 2 apathy instruments were initially identified on the Master Database as a result of the academic literature and practice surveys. This was supplemented by further research of the literature to find a total of 14 instruments. This makes sense as the measurement of apathy is a newly emerging field of study, that can be traced back to the work of Robert S. Marin in 1990 (Marin, 1990).

### 9.4.6.3 Short-listed Instruments

The list of 14 instruments was reduced to a short list of contender instruments based on the application of general culling criteria (availability of instrument, number of citations, psychometric evidence, used in clinical practice, availability of normative and clinical reference data, administration time, able to be used with the various severity levels of dementia, cost considerations, and applicability for routine care). The short list of apathy instruments is presented below:

**Table 44 Short-listed Apathy Instruments**

<b>Instrument</b>	<b>Original Article Cite Author(s) + Publication Date</b>
<b>Frontal Systems Behavior Scale</b>	Grace and Molloy (2001)
<b>Apathy Evaluation Scale</b>	Marin, et al. (1991)
<b>Apathy Inventory</b>	Robert, et al. (2002)
<b>Irritability Apathy Scale</b>	Burns, et al. (1990)
<b>Apathy Scale for Parkinson's Disease</b>	Starkstein, et al. (1992)
<b>Frontal Behavior Inventory</b>	Kertesz, et al. (1997)
<b>Lille Apathy Rating Scale</b>	Sockeel, et al. (2006)

Other scales which were found to be noteworthy in this area were: two structured interviews - the Dementia Apathy Interview and Rating (DAIR) (Strauss and Sperry, 2002) and the Structured Interview for Apathy (SIA) (Starkstein, et al. 2005); the withdrawal subscale of the Multidimensional Observation Scale for Elderly Subjects (MOSES) (Helmes, et al. 1987); as well as a recent paper using the Apathy items from the Hamilton Rating Scale for Depression (ApHRSD) (Marin, et al. 2003). These types of measures (i.e. structured interviews and sub-scales of other related instruments) because of their length they were regarded as out of scope for the project.

### 9.4.6.4 Reviewed Instruments

The short-listed instruments were now classified into the following four categories: Accept; Reject; Grey Area; Grey Area – but promising new instrument. The specific criteria used to select the instruments for comprehensive review were: (1) Whether there is a copy of the instrument and original article available for review; (2) The number of citations found; (3) The amount of published psychometric evidence; and (4) whether the instrument is used in clinical practice.

Following a thorough going analysis of these measures it was decided to examine the most influential, namely the Apathy Evaluation Scale (AES). In reviewing the literature, it was found that the AES seems to have contributed to the development of a number of measurement tools in this area, e.g. the Apathy Inventory and the new Lille Apathy Rating Scale (LARS) (which at this stage has published data for Parkinson's disease). The AES is also related to the Apathy Scale for Parkinson's disease (Starkstein, et al. 1992), also known as Starkstein's 14 item scale, as they have the same source from Marin (1990).

In terms of the other short-listed instruments, the Frontal Systems Behavior Scale (FrSBe) (Grace and Molloy, 2001) has 46 behavioural rating items and is a proprietary measure. The FrSBe and Frontal behavioural Inventory (FBI) (Kertesz, et al. 1997) examine “frontal symptoms” (e.g. apathy, dis-inhibition, lack of insight, executive functioning) not just apathy. The Irritability Apathy Scale (Burns, et al. 1990) is a forgotten instrument in the literature, including both irritability and apathy scales, but it may be that the apathy subscale could form a useful short measure of apathy. However, this would need to be assessed in further research.

Based on the above analysis, the one instrument selected for comprehensive review was the Apathy Evaluation Scale. (This review will also examine the relationship and commonalities between the AES and other apathy instruments).

#### **9.4.6.5 Apathy Evaluation Scale (AES)**

The Apathy Scale (AES) (Marin, et al. 1991) is an 18 item rating scale to assess symptoms of apathy (response categories: 4 point, Likert-type) with clinician, informant and self-rated versions. The rating is based on an interview with the patient about their interests, activities and daily routine. The AES is a scale with reasonable to good psychometric properties (see Glenn, 2005) (For the clinician’s version: Cronbach’s alpha = 0.90; test-retest reliability = 0.88 for a mean timeframe of 25 days; inter-rater reliability = 0.94 for two raters). However Glenn (2002) did find poor sensitivity and specificity in respect to clinician’s judgement of apathy in a Traumatic Brain Injury (TBI) sample (for the informant and self-rated versions of the scale).

The AES is a highly influential instrument in the emerging area of apathy syndrome measurement. However, there is some concern noted in the literature about the lack of adequate training and guidance for the clinical and informant ratings (Malloy and Boyle, 2005; Clarke et al. 2007). (Here the abridged version of the AES by Starkstein et al. (1992) may be a useful alternative) Also following the work of Clarke et al. (2007), the informant and self-report versions of the AES require further examination of their factor structure. There is also a need for normative data.

#### **9.4.6.6 Instrument Rankings – Summary Rating**

Considering the attributes of each test as reviewed in the appendix, a scoring judgement was made according to the criteria outlined, and this information is summarised in the table below.

**Table 45 Summary of Ratings for Apathy Instruments**

Criteria	Instrument	
	Weight	AES
Availability of comparison data	3	3
Length/feasibility of instrument for inclusion in battery	2	2
Complexity of administration/ cognitive burden	2	3
Cultural Appropriateness	1	3
Ease of obtaining score	2	2
Sensitivity to dementia	3	3
Reliability evidence	3	3
Validity evidence	3	2
Cost of the instrument	2	3
Cost of instrument administration	2	1
<b>Weighted Total</b>		<b>55</b>

This analysis finds mostly high ratings for the AES instrument in terms of its administration, scoring and psychometric properties.

#### 9.4.6.7 Recommendations

Based on the high summary rankings for the Apathy Evaluation Scale it is recommended for the specialist assessment of apathy symptoms in people with dementia.

#### 9.4.7 Depression

According to a recent review by Onega (2006), approximately 50% of older adults with dementia have a minor depressive disorder / depressive symptoms and 15-20% have a major depressive disorder. Thus depression is one of the most common symptoms associated with dementia. Likewise a number of reviews have been written and this is a major area of psychometric development work (Burns, et al. 2004).

In brief, the key issues in determining depression measurement in this population are:

- 1) Whether to use patient self-report or a clinician rating scale
- 2) Whether the instrument been specially designed for people with dementia or has the measure been developed for older adults more generally
- 3) How to measure depression in people who are also cognitively impaired

Finally, as with all associated symptoms of dementia mentioned in this chapter, it is important to determine how detailed the assessment is required. Is a full work-up required using specialised instruments or do clinicians wish to flag the issue as part of a wider assessment process?

#### 9.4.7.1 Selection of Instruments

Instruments were selected for comprehensive review based on literature search and instrument selection methodology outlined previously (see Section 2). For the selection of the depression instruments these steps are summarised below.

#### 9.4.7.2 Literature Search

30 depression instruments were identified on the Master Database as a result of the academic literature searches and practice surveys. However the majority of these instruments were not dementia specific, but had been used in studies with older adults.

#### 9.4.7.3 Short-listed Instruments

The list of 30 instruments was reduced to a short list of contender instruments based on the application of the additional selection criteria (availability of instrument, number of citations, psychometric evidence, used in clinical practice, availability of normative and clinical reference data, administration time, able to be used with the various severity levels of dementia, cost considerations, and applicability for routine care). The short list of depression instruments is presented below:

**Table 46 Short-listed Depression Instruments**

Instrument	Original Article Cite Author(s) + Publication Date
<b>Geriatric Depression Scale</b>	Yesavage, et al. (1983)
<b>Cornell Scale for Depression in Dementia</b>	Alexopoulos, et al. (1988)
<b>Hamilton Rating Scale for Depression</b>	Hamilton (1960)
<b>Centre for Epidemiological Studies Depression Scale</b>	Radloff and Teri (1986)
<b>Montgomery and Asberg Depression Rating Scale</b>	Montgomery and Asberg (1979)
<b>Hospital Anxiety and Depression Scale</b>	Zigmond and Snaith (1983)
<b>Beck Depression Inventory I, II</b>	Beck, et al. (1961)

Other scales which were found to be noteworthy in this area were the: Depressive Symptom Assessment for Older Adults (Onega, 2006), Apparent Emotion Rating Instrument (Snyder, et al. 1998), Dementia Mood Assessment Scale (DMAS) (Sunderland, et al. 1988), Depressive Signs Scale (DSS) (Katona and Aldridge, 1985), Even Briefer Assessment Scale for Depression (EBAS-DEP) (Weyerer, et al. 1999), Brief Assessment Schedule Depression Cards (BASDEC) (Adshead, et al. 1992). However, these instruments had few citations and less available psychometric evidence and were therefore not included in the short list.

#### 9.4.7.4 Reviewed Instruments

The short-listed instruments were now classified as previously outlined in the sections on anxiety and apathy. For instance, regarding the Beck Depression Inventory and the Hamilton Rating Scale for Depression, although they are widely used, it was noted that they were not designed for older people. The Beck Depression Inventory is also a proprietary instrument.

Using the impact measure data and the above criteria it was decided to examine the two most popular instruments for this population: the Geriatric Depression Scale (GDS Yesavage) and Cornell Scale for Depression in Dementia (CSDD). In making this decision we opted to choose one self-report instrument and one clinical rating scale. This selection was further confirmed by the work of Ramirez Diaz, et al. (2005) which found that the GDS (Yesavage) and CSDD were in the top ten of assessment tools used across Europe. Both measures are included in the “Silver Book” of the Royal Australian College of General Practitioners (RACGP) and in a recent major study known as the Challenge Depression Project conducted by the Hammond Care Group (web-site: [http://www.health.gov.au/internet/wcms/Publishing.nsf/Content/ageing-chall-depress.htm/\\$FILE/challenge04.pdf](http://www.health.gov.au/internet/wcms/Publishing.nsf/Content/ageing-chall-depress.htm/$FILE/challenge04.pdf)). The CSDD is also used in the new Aged Care Funding Instrument (ACFI) (web-site: <http://www.health.gov.au/internet/wcms/publishing.nsf/content/ageing-acfi-outcome.htm>).

Based on the above analysis, the two instruments elected for comprehensive review were the Geriatric Depression Scale and the Cornell Scale for Depression in Dementia.

#### **9.4.7.5 Geriatric Depression Scale (GDS)**

The Geriatric Depression Scale (GDS Yesavage) is a short self-report screening and assessment instrument (Administration time: 5 – 10 minutes) for depression in elderly people. It has 30, 15, 10 and 4 item versions. A notable feature of the instrument is that it uses dichotomous (Yes / No) response items to ease administration and completion by elderly people.

Widely used and researched, the GDS (Yesavage) compares favourably with other rating scales and self report measures of depression, for example, the Hamilton Rating Scale for Depression (HRSD) and the Center for Epidemiological Studies Depression Scale (CES-D) (McDowell, 2006). The GDS (Yesavage) has been used in hospital, community / primary care and residential settings (Bowling, 2005), and has good psychometric properties (Cronbach’s alpha is in the range of 0.69 – 0.94; Test-retest reliability is in the range of 0.85 – 0.98 for a 10 to 12 day timeframe in residential care). However, care is needed when interpreting data from the GDS-15 obtained from community and hospital samples, as there is some evidence of lower reliability for this version of the scale outside of residential care settings.

Also further research work is needed in the following areas: (1) the detection of minor depression (Watson and Pigone, 2003); (2) the use of the GDS (Yesavage) for those that are 75 years and older (McDowell, 2006); and (3) the applicability and suitability of the GDS (Yesavage) for those with dementia / cognitive impairment. Here the evidence is mixed at best, and restricts the applicability of this instrument to those with milder forms of dementia - though it must be remembered that this scale was not specifically designed for people with dementia. Finally, recent normative data for the United Kingdom has been provided by Osborn, et al. (2002).

#### **9.4.7.6 Cornell Scale for Depression in Dementia (CSDD)**

The Cornell Scale for Depression in Dementia (CSDD) is a 19 item clinical rating scale based on semi-structured interview questions with the informant (nursing staff or relative) and interview questions and signs from the patient (administration time: 20 minutes). It focuses on the identification of depressive symptoms and signs in people with dementia (McKeith, et al. 1999).

The CSDD is a widely used and highly respected measure, as Burns, et al. (2004) indicate, the CSDD “sets the standard” in the area of depression measurement in severe dementia when measurement by an informant is required. The CSDD’s internal consistency reliability is in the range of 0.84 – 0.98 (Cronbach’s alpha) and inter-rater reliability is in the range of 0.67 – 0.74. The comprehensive review found that the CSDD compares well to established diagnostic criteria (DSM-IV, ICD-10, RDC) i.e. it has criterion validity. However, since the original publications of the CSDD, little work has been published on scale’s inter-rater reliability. (Up to date, reliability

information is of vital importance if a clinical rating scale is going to be used in routine assessments by different practitioners, across different practice settings.)

Finally, it should be noted that for the effective use of clinical rating scales in mental health settings adequate training is required in order to ensure consistent ratings. The need for adequate training is especially relevant for the illness of dementia where depression and dementia symptoms (as well as other associated symptoms) are intertwined.

#### 9.4.7.7 Instrument Rankings – Summary Rating

Considering the attributes of each test as reviewed in the appendix, a scoring judgement was made according to the criteria outlined, and this information is summarised in Table 47 below.

**Table 47 Summary of Ratings for Depression Instruments**

Criteria	Weight	Instrument	
		CSDD	GDS
Availability of comparison data	3	3	3
Length/feasibility of instrument for inclusion in battery	2	2	2
Complexity of administration/ cognitive burden	2	3	3
Cultural Appropriateness	1	3	2
Ease of obtaining score	2	3	3
Sensitivity to dementia	3	3	1
Reliability evidence available	3	2	3
Validity evidence available	3	3	2
Cost of the instrument	2	3	3
Cost of instrument administration	2	2	3
<b>Weighted Total</b>		<b>61</b>	<b>57</b>

This analysis finds very high ratings for both instruments in terms of their administration, scoring and psychometric properties. However, the self-report nature of the GDS (Yesavage) limits its applicability to people with less severe dementia.

#### 9.4.7.8 Recommendations

Based on the summary rankings both the Geriatric Depression Scale and the Cornell Scale for Depression in Dementia instruments are recommended. The GDS (Yesavage) is able to be used with those people with less severe dementia, while the Cornell rating scale can be used with people with dementia across the range of severity, if required. However, for both measures, their application to Australian clinical settings for dementia patients needs further research and investigation.

## 9.5 Other Omnibus Measures: HoNOS 65+

Some members of the National Expert Panel indicated that the HoNOS 65+ (Burns, et al. 1999a) was being used for the assessment of elderly patients in mental health settings throughout Australia and thus it should also be examined with regard to its use in assessing dementia patients across a broader range of service settings. The HoNOS is a comprehensive measure which covers most ICF domains, for example, environmental aspects such as living conditions are included.

The HoNOS 65+ involves rating patients (from 0-4) on twelve scales which ostensibly assess their behaviour on four factors (behaviour, impairment, symptoms, social function) as is indicated in Table 48 below:

**Table 48 HoNOS 65+ Scales and Factors**

12 Items/Scales	Factor
1. Behavioural disturbance (e.g. overactive, agitated, aggressive, resistant)	Behaviour
2. Non accidental self injury	Behaviour
3. Problem drinking or drug	Behaviour
4. Cognitive problems	Impairment
5. Physical illness or disability problems	Impairment
6. Hallucinations and delusions	Symptoms
7 Depressive symptoms	Symptoms
8 Other mental and behavioural symptoms (phobia, panic, compulsion etc)	Symptoms
9. Problems with relationships	Social
10 Problems with ADL	Social
11 Problems with living conditions	Social
12 Problems with occupation and activities	Social

A total score can also be derived from the HoNOS +65 as well as the item subscale scores. However, there is limited published information on the internal consistency of the overall scale. Shergill, et al. (1999) indicated an internal consistency of 0.61 for the scale total but this was based on using the original HoNOS with elderly people and not the HoNOS 65+.

The four factor structure (refer above) proposed by Burns, et al. (1999a) has not been replicated in other studies (Turner, 2004). Burns, et al. (1999) found the 4 factors identified above explained 57.4% of the variance but Turner (2004) even queries whether Burn's own data is actually supporting the proposed factor structure that is outlined (e.g. item 1 does not load with items 2 and 3 as indicated above but with hallucinations for initial assessment, and with hallucinations, depression and 'other mental' items on follow up). Turner (2004) in reviewing the few available studies on this issue concludes that the finding of different factor structures in different studies throws doubt on the validity of the original 4 factors that were outlined.

There is some support for correlation with other measures (Spear, et al. 2002) – both for the total score and item/subscale scores. The total score has moderate correlations with MMSE, the Crichton Royal Behavioural Rating Scale and the Barthel Index of ADL. The highest correlations were generally between subscales/items and other measures of that symptom – and the correlations tend to be moderate to good and in expected direction e.g. cognition item with MMSE; items 1, 4, 9 with the Brief Agitation Rating Scale (BARS); and items 6-9 with Brief Psychiatric Rating Scale.

However, correlations between the depression item and other depression scales are variable. Spear, et al. (2002) found the depression item correlated quite well (0.61) with the Geriatric Depression Inventory (GDS Yesavage), however, Turner (2004) notes that Burns, et al. (1999a) report a correlation of -0.20 with the GDS (Yesavage). Turner (2004) suggests some sub-dimensions of HoNOS +65 may need additional measures. The HoNOS 65+ also does not appear to assess apathy as distinct from depression and this is an important symptom in dementia.

There is only a moderate correlation of the total score with the Clinical Dementia Rating Scale and a moderate correlation for change scores with CIBIC+/ clinical impression of change, however, these correlations were significant. Spear, et al. (2002) concludes there are some indications that the HoNOS 65+ is sensitive to change over time.

Spear, et al. (2002) found the inter-rater reliability (IRR) for most items was generally adequate (>.70) but IRR was problematic (<.70) for items 4, 5, 9 and 10 and the IRR for item 9 (problems with relationships) was particularly poor. Shergill, et al. (1999) reports poor IRR for items 10-12. It was also found that Item 3 (problem drinking etc) was hard to rate for elderly patients and there were too few instances of drug/alcohol problems for this data to be included in the analysis (Shergill, et al. 1999). One might also question the relevance of items 11 and 12 (living conditions, occupation and activities) for those in residential care.

The HoNOS 65+ was designed to assess the elderly (65+) in mental health service settings. There is some limited evidence that the pattern of scores and the total score differs between people with dementia and those with mood disorders and thus it can differentiate people with dementia from those with mood disorders in the settings studied (Spear, et al. 2002; Turner, 2004). It is not widely used in other aged care settings. There appear to be problems with inter-rater reliability for some of the items. Pirkis, et al. (2005) have reported that the predictive validity and the test-retest reliability of the HoNOS 65+ have not been reported in the published literature.

Although it is used widely within the mental health sector within Australia, and the data collected to date must be considerable, as yet there are few published papers referring to its psychometric properties. Given these considerations other reviewed measures of the omnibus type may be currently preferred, such as the NPI, as there is far more published evidence available concerning its' psychometric properties.

## **9.6 Conclusions and Recommendations**

A number of global measures of behavioural and psychological disturbance (Global BPSD) have been reviewed. Based on these reviews it is recommended the NPI and the BEHAVE-AD be used in both clinical and research settings for assessment of Global BPSD. These instruments both have well established psychometric properties.

The CERAD-BRSD is recommended for research rather than routine practice given its cost and the time required for its administration. A 17 item abbreviated version may be considered better for clinical utility, but limited evidence on this version is currently available.

A number of delirium measures were also assessed in order to aid in the differential diagnosis of dementia and delirium. The Confusion Assessment Methodology (CAM) is the most widely utilised screening/diagnostic tool for detecting delirium internationally among older people with or without dementia. Less well known, however, the Delirium Rating Scale (DRS-R-98) is also a widely recognised and well validated measure. Whilst the CAM is superior in its utility to the DRS-R-98, it does not capture severity of delirium symptoms hence is not appropriate for repeated measures of delirium severity. The DRS-R-98 is designed for assessment of both the presence and the severity of delirium symptoms. Limitations of the DRS-R-98, and the DRS, include that they are time taxing and require sufficient training, especially for those who do not have psychiatric background. The DRS-R-98 is not appropriate for use in the community setting given its requirement for observation

over a 24 hours period. However, it allows for comprehensive assessment of individuals who are at risk or suspected of developing delirium in institutional care settings.

For the purpose of the DOMS project it is recommended both measures be included as they have two distinct, yet equally important functions.

In many cases the use of Global BPSD measures such as the NPI may suffice for the assessment of the associated symptoms of dementia. However, if a more detailed assessment of a particular symptom is required the following recommendations are made:

Aggression: Rating Scale for Aggressive Behaviour in the Elderly (RAGE)  
 Agitation: Cohen Mansfield Agitation Inventory (CMAI)  
 Anxiety: Rating Anxiety in Dementia (RAID)  
 Apathy: Apathy Evaluation Scale (AES)  
 Depression: Cornell Scale for Depression in Dementia (CSDD)  
 Geriatric Depression Scale (GDS Yesavage) - less severe cases and in community settings)

## References

Aalten P, van Valen E, de Vugt ME, et al. (2006) *Awareness and behavioral problems in dementia patients: a prospective study*. International Psychogeriatrics. Vol. 18, No.1, pp.3-17.

Adshead F, Cody DD and Pitt B (1992) *BASDEC: a novel screening instrument for depression in elderly medical inpatients.[see comment]*. BMJ. Vol. 305, No.6850, pp.397.

Albert MS, Levkoff SE, Reilly C, Liptzin B, Pilgrim D, Cleary PD, et al. (1992) *The delirium symptom interview: an interview for the detection of delirium symptoms in hospitalized patients*. Journal of Geriatric Psychiatry and Neurology. Vol. 5, No.1, pp.14-21.

Alexopoulos GS, Abrams RC, Young RC, et al. (1988) *Cornell scale for depression in dementia*. Biological Psychiatry. Vol. 23, No.3, pp.271-284.

Allen NH, Gordon S, Hope T and Burns A (1996) *Manchester and Oxford Universities Scale for the Psychopathological Assessment of Dementia (MOUSEPAD)*. British Journal of Psychiatry. Vol. 169, No.3, pp.293-307.

Alwahhabi F (2003) *Anxiety symptoms and generalized anxiety disorder in the elderly: a review*. Harvard Review of Psychiatry. Vol. 11, No.4, pp.180-193.

Alzheimer's Disease International (2003) *Behavioural and psychological symptoms of dementia*. 10/66 Pilot Studies. Findings. p.1, available at [www.alz.co.uk/1066/research/pilot/bpsd1.html](http://www.alz.co.uk/1066/research/pilot/bpsd1.html), accessed 14 March 2007.

American Psychiatric Association (1994) *Diagnostic and Statistical Manual of Mental Disorders DSM-IV (4th ed.)* American Psychiatric Association, Washington DC.

American Psychiatric Association (2000) *Diagnostic and statistical manual of mental disorders, fourth edition, text revision*. American Psychiatric Association, Washington DC.

Asthill LF (2004) *Staff training and challenging behaviour in a day hospital*. Dementia: The International Journal of Social Research and Practice. Vol. 3, No.3, pp.384-392.

Auer S, Hampel H, et al. (2000) *Translations of measurements and scales: Opportunities and diversities*. International Psychogeriatrics. Vol. 12, Suppl.1, pp. 391-394.

Bathareethan M and Shah A (2000) *Aggressive behaviour, staff attitude and staff perception of patients on two continuing care psychogeriatric wards*. Ageing and Mental Health. Vol. 4, No.1, pp. 66-71.

- Baumgarten M, Becker R, et al. (1990) *Validity and reliability of the dementia behavior disturbance scale*. Journal of the American Geriatrics Society. Vol. 38, No.3, pp.221-6.
- Beck AT, Epstein N, Brown G, et al. (1988) *An inventory for measuring clinical anxiety: psychometric properties*. J Consult Clin Psychol. Vol. 56, No.6, pp.893-7.
- Beck AT, Ward CH, Mendelson M, et al. (1961) *An inventory for measuring depression*. Archives of General Psychiatry. Vol. 4, pp.561-71.
- Boada M, Cejudo JC, Tarraga L, et al. (2002) *Neuropsychiatric Inventory Questionnaire (NPI-Q): Spanish validation of an abridged form of the Neuropsychiatric Inventory (NPI)*. Neurologica. Vol. 17, No.6, pp.317-323.
- Boada M, Tarraga L, et al. (2006) *Behavioral pathology in Alzheimer's Disease Rating Scale (BEHAVE-AD): Spanish validation*. Neurologia. Vol. 21, No.1, pp.19-25.
- Bohner H, Friedrichs R, Habel U, Muller EE, Sandmann W and Schneider F (2003) *Delirium increases morbidity and length of stay after vascular surgery operations. Results of a prospective study*. Chirurg. Vol. 74, No.10, pp. 931-936.
- Bowling A (2005) *Measuring health: a review of quality of life measurement scales (3rd ed.)* Open University Press, Berkshire, England.
- Breitbart W, Rosenfeld B, Roth A, Smith MJ, Cohen K and Passik S (1997) *The Memorial Delirium Assessment Scale*. Journal of Pain and Symptom Management. Vol. 13, No.3, pp.128-137.
- Brodaty H, Ames D, et al. (2003) *A randomized placebo-controlled trial of risperidone for the treatment of aggression, agitation, and psychosis of dementia*. Journal of Clinical Psychiatry. Vol. 64, No.2, pp.134-43.
- Brodaty H, Ames D, et al. (2005) *Risperidone for psychosis of Alzheimer's disease and mixed dementia: results of a double-blind, placebo-controlled trial*. International Journal of Geriatric Psychiatry. Vol. 20, No.12, pp.1153-7.
- Brodaty H, Ames D, Snowden J, Woodward M, et al. (2003) *A randomized placebo-controlled trial of risperidone for the treatment of aggression, agitation, and psychosis of dementia*. Journal of Clinical Psychiatry. Vol. 64, No.2, pp.134-43.
- Brodaty H, Draper B, et al. (2001) *Psychosis, depression and behavioural disturbances in Sydney nursing home residents: prevalence and predictors*. International Journal of Geriatric Psychiatry. Vol. 16, No.5, pp.504-12.
- Brodaty H, Draper BM, et al. (2003) *Randomized controlled trial of different models of care for nursing home residents with dementia complicated by depression or psychosis*. Journal of Clinical Psychiatry. Vol. 64, No.1, pp.63-72.
- Buettner LL, Lundegren H, Lago D, Farrell P, et al. (1996) *Therapeutic recreation as an intervention for persons with dementia and agitation: an efficacy study*. American Journal of Alzheimer's Disease and Other Dementias. Vol. 11, No.5, pp.4-12.
- Burke JR and Morgenlander JC (1999) *Managing common behavioural problems in dementia*. Post-graduate Medicine. Vol. 106, No.5, pp.131-140.
- Burns A, Beevor A, Lelliott P, Wing J, et al. (1999a) *Health of the Nation Outcome Scales for elderly people (HoNOS 65+)*. British Journal of Psychiatry. Vol. 174, pp.424-427.
- Burns A, Folstein S, Brandt J and Folstein M (1996) *Clinical assessment of irritability, aggression and apathy in Huntington and Alzheimer's Disease*. Journal of Nervous and Mental Disease. Vol. 178, pp.20-26.
- Burns A, Folstein S, Brandt J, et al. (1990) *Clinical assessment of irritability, aggression and apathy in Huntington and Alzheimer's Disease*. Journal of Nervous and Mental Disease. Vol. 178, pp.20-26.

- Burns A, Lawlor B and Craig S (2004) *Assessment scales in old age psychiatry (2nd ed.)* Taylor and Francis, London.
- Burns A, Spiegel R, et al. (2004) *Efficacy of rivastigmine in subjects with moderately severe Alzheimer's disease*. International Journal of Geriatric Psychiatry. Vol. 19, No.3, pp.243-9.
- Caine E (1996) *Behavioural disturbances of dementia in our current nomenclature system: Diagnostic classification of neuropsychiatric signs and symptoms in patients with dementia*. International Psychogeriatrics. Vol. 8, Suppl.3, pp.273-279.
- Calkin PA, Kunik ME, Orengo CA, Molinari V et al. (1997) *Tolerability of clonazepam in demented and non-demented geropsychiatric patients*. International Journal of Geriatric Psychiatry. Vol. 12, No.7, pp.745-9.
- Chacko RC, Hurley RA, et al. (1995) *Clozapine for acute and maintenance treatment of psychosis in Parkinson's disease*. Journal of Neuropsychiatry and Clinical Neurosciences. Vol. 7, No.4, pp.471-5.
- Chan WC, Lam LC, et al. (2001) *A double-blind randomised comparison of risperidone and haloperidol in the treatment of behavioural and psychological symptoms in Chinese dementia patients*. International Journal of Geriatric Psychiatry. Vol. 16, No.12, pp.1156-62.
- Chiu MJ, Chen TF, et al. (2006) *Behavioral and psychologic symptoms in different types of dementia*. Journal of the Formosan Medical Association. Vol. 105, No.7, pp.556-62.
- Choi SH, Na DL, Kwon HM, Yoon SJ, Jeong JH and Ha CK. (2000) *The Korean Version of the Neuropsychiatric Inventory: a scoring tool for neuropsychiatric disturbance in dementia patients*. Journal of Korean Medical Science. Vol. 15, No.6, pp.609-15.
- Choy CN, Lam LC, Chan WC, Li SW, et al. (2001) *Agitation in Chinese elderly: validation of the Chinese version of the Cohen-Mansfield Agitation Inventory*. International Psychogeriatrics. Vol. 13, No.3, pp.325-35.
- Cifu DX, Keyser-Marcus L, et al. (1997). *Acute predictors of successful return to work 1 year after traumatic brain injury: a multicenter analysis*. Archives of Physical Medicine and Rehabilitation. Vol. 78, No.2, pp.125-31.
- Clarke DE, Reekum R, Simard M, et al. (2007) *Apathy in dementia: an examination of the psychometric properties of the apathy evaluation scale*. Journal of Neuropsychiatry and Clinical Neurosciences Vol.19, No.1, pp.57-64.
- Coen RF, Swanwick GR, et al. (1997) *Behaviour disturbance and other predictors of carer burden in Alzheimer's disease*. International Journal of Geriatric Psychiatry. Vol. 12, No.3, pp.331-6.
- Cohen-Mansfield J (1986) *Agitated behaviors in the elderly: II Preliminary results in the cognitively deteriorated*. Journal of the American Geriatrics Society. Vol. 34, pp.722-27.
- Cohen-Mansfield J (1991) *Instruction manual for the Cohen-Mansfield Agitation Inventory (CMAI)*. The Research Institute of the Hebrew Home of Greater Washington, Rockville.
- Cohen-Mansfield J (1999) *Measurement of inappropriate behaviour associated with dementia*. Journal of Gerontological Nursing. Vol. 25, No.2, pp.42-51.
- Cohen-Mansfield J and Billing N (1986) *Agitated behaviours in the elderly. A conceptual review*. Journal of the American Geriatrics Society. Vol. 34, pp. 711-721.
- Cohen-Mansfield J and Libin A (2004) *Assessment of agitation in elderly patients with dementia: correlations between informant rating and direct observation*. International Journal of Geriatric Psychiatry. Vol. 19, No.9, pp.881-91.
- Cohen-Mansfield J and Werner P (1998) *Predictors of aggressive behaviour: a longitudinal study of senior day care centres*. Journal of Gerontology. Vol. 53B, pp.300-310.
- Cohen-Mansfield J, Marx M and Rosenthal A (1989) *A description of agitation in a nursing home*. Journal of Gerontology. Vol. 44, No.3, pp.77-84.

- Cohen-Mansfield J, Taylor L and Werner P (1998) *Delusions and hallucinations in an adult day care population. A longitudinal study.* American Journal of Geriatric Psychiatry. Vol. 6, No.2, pp.104-21.
- Cohen-Mansfield J, Werner P, Watson V and Pasis S (1995) *Agitation among elderly persons at adult day-care centers: the experiences of relatives and staff members.* International Psychogeriatrics. Vol. 7, No.3, pp.447-58.
- Craig AH, Cummings JL, Rairbanks L, et al. (1996) *Cerebral blood flow correlates of apathy in Alzheimer's disease.* Arch Neurol. Vol. 53, pp.1116-1120.
- Cummings JL (1996) *Theories behind scales and measurements: Theories behind existing scales for rating behaviour in dementia.* International Psychogeriatrics. Vol. 8, suppl.3, pp.293-300.
- Cummings JL (1997) *The Neuropsychiatric Inventory: assessing psychopathology in dementia patients.* Neurology. Vol. 48, suppl.6, pp.S10-16.
- Cummings JL, Diaz C, Levy ML, et al. (1996) *Behavioural syndromes in neurodegenerative diseases: frequency and significance.* Semin. Neuropsychiatry. Vol. 1, pp.241-247.
- Cummings JL, McRae T, et al. (2006) *Effects of donepezil on neuropsychiatric symptoms in patients with dementia and severe behavioral disorders.* American Journal of Geriatric Psychiatry. Vol. 14, No.7, pp.605-12.
- Cummings JL, Mega M, Gray K, et al. (1994) *The Neuropsychiatric Inventory: comprehensive assessment of psychopathology in dementia.* Neurology. Vol. 44, No.12, pp.2308-2314.
- Cummings JL, Mega M, Gray K, Rosenberg-Thompson S, Carusi DA and Gornbein J (1994) *The Neuropsychiatric Inventory: comprehensive assessment of psychopathology in dementia.* Neurology. Vol. 44, pp.2308-2314.
- Cummings JL, Raman R, Ernstrom K, et al. (2006) *ADCS Prevention Instrument Project: Behavioral Measures in Primary Prevention Trials.* Alzheimer Disease and Associated Disorders Vol. 20, Suppl.3, pp.S147-S151.
- Cummings JL, Street J, Masterman D and Clarke WS (2002) *Efficacy of olanzipine in the treatment of psychosis in dementia with lewy bodies.* Dementia and Geriatric Cognitive Disorders. Vol. 13, No.2, pp.67-73.
- de Deyn PP, Katz IR, Brodaty H, Lyons B, et al (2005) *Management of agitation, aggression, and psychosis associated with dementia: a pooled analysis including three randomized, placebo-controlled double-blind trials in nursing home residents treated with risperidone.* Clinical Neurology and Neurosurgery. Vol. 107, No.6, pp.497-508.
- de Deyn PP, Rabheru K, Rasmussen A, Bocksberger JP, et al. (1999) *A randomized trial of risperidone, placebo, and haloperidol for behavioral symptoms of dementia.[see comment].* Neurology. Vol. 53, No.5, pp.946-55.
- de Guise E, Feyz M, et al. (2005) *Overview of traumatic brain injury patients at a tertiary trauma centre.* Canadian Journal of Neurological Sciences. Vol. 32, No.2, pp.186-93.
- de Jonghe JF and Kat MG (1996) *Factor structure and validity of the Dutch version of the Cohen-Mansfield Agitation Inventory (CMAI-D).* Journal of the American Geriatrics Society. Vol. 44, No.7, pp.888-9.
- de Jonghe JF, Kalisvaart KJ, Dijkstra M, van Dis H, Vreeswijk R, Kat MG, et al. (2007) *Early symptoms in the prodromal phase of delirium: a prospective cohort study in elderly patients undergoing hip surgery.* Am J Geriatr Psychiatry. Vol. 15, No.2, pp.112-121.
- de Jonghe JFM, Kalisvaart KJ, Timmers JFM, Kat MG and Jackson JC (2005) *Delirium-O-Meter: a nurses' rating scale for monitoring delirium severity in geriatric patients.* International Journal of Geriatric Psychiatry. Vol. 20, No.12, pp.1158-1166.

- de Rooij SE, van Munster BC, Korevaar JC, Casteelen G, Schuurmans MJ, van der Mast RC, et al. (2006) *Delirium subtype identification and the validation of the Delirium Rating Scale--Revised-98 (Dutch version) in hospitalized elderly patients*. Int J Geriatr Psychiatry. Vol. 21, No.9, pp.876-882.
- Devanand DP, Miller L, Richards M, Marder K, Bell K, Mayeux R and Stern Y (1992) *Columbia University Scale for Psychopathology in Alzheimer's Disease*. Archives of Neurology. Vol. 49, pp.371-6.
- Dombovy M L and A C Olek (1997) *Recovery and rehabilitation following traumatic brain injury*. Brain Injury. Vol. 11, No.5, pp.305-18.
- Drachman DA, Swearer JM, O'Donnell BF, Mitchell AL and Maloon A (1992) *The caretaker obstreperous-behaviour rating assessment (COBRA) scale*. Journal of the American Geriatric Society. Vol. 40, pp.463-470.
- Draper B and Turner J (2003) *Outcome of treatment of vocally disruptive behaviour in nursing home residents*. Vol. 2, pp.80-4.
- Ely EW, Inouye SK, Bernard GR, Gordon S, Francis J, May L, et al. (2001) *Delirium in mechanically ventilated patients: validity and reliability of the confusion assessment method for the intensive care unit (CAM-ICU)*. JAMA. Vol. 286, No.21, pp.2703-2710.
- Ely EW, Margolin R, Francis J, May L, Truman B, Dittus R, et al. (2001) *Evaluation of delirium in critically ill patients: validation of the Confusion Assessment Method for the Intensive Care Unit (CAM-ICU)*. Critical Care Medicine. Vol. 29, No.7, pp.1370-1379.
- Engelborghs S, Maertens K, et al. (2005) *Neuropsychiatric symptoms of dementia: cross-sectional analysis from a prospective, longitudinal Belgian study*. International Journal of Geriatric Psychiatry. Vol. 20, No.11, pp.1028-37.
- Finkel S, Lyons J and Anderson R (1992) *Reliability and validity of the Cohen-Mansfield Agitation Inventory in institutionalized elderly*. International Journal of Geriatric Psychiatry. Vol. 7, pp.487-90.
- Finkel SI, Lyons JS, et al. (1993) *A brief agitation rating scale (BARS) for nursing home elderly*. Journal of the American Geriatrics Society. Vol. 41, No.1, pp.50-2.
- Foglia JP, Pollock BG, et al. (1997) *Plasma levels of citalopram enantiomers and metabolites in elderly patients*. Psychopharmacology Bulletin. Vol. 33, No.1, pp.109-12.
- Fonseca F, Bulbena A, Navarrete R, Aragay N, Capo M, Lobo A, et al. (2005) *Spanish version of the Delirium Rating Scale-Revised-98: reliability and validity*. Journal of Psychosomatic Research. Vol. 59, No.3, pp.147-151.
- Franulic A, Carbonell CG, et al. (2004) *Psychosocial adjustment and employment outcome 2, 5 and 10 years after TBI*. Brain Injury. Vol. 18, No.2, pp.119-29.
- Franulic A, Horta E, et al. (2000) *Organic personality disorder after traumatic brain injury: cognitive, anatomic and psychosocial factors. A 6 month follow-up*. Brain Injury. Vol. 14, No.5, pp.431-9.
- Frisoni GB, Rozzini L, Gozzetti A, Binetti G, Zanetti O, Bianchetti A and Trabucchi M (1999) *Behavioural syndromes in Alzheimer's Disease: description and correlates*. Dementia Geriatr Cogn Discord. Vol. 10, pp.130-138.
- Fujikawa T, Takahashi T, Kinoshita A, Kajiyama H, et al. (2004) *Quetiapine treatment for behavioral and psychological symptoms in patients with senile dementia of Alzheimer type*. Neuropsychobiology. Vol. 49, No.4, pp.201-4.
- Gaudreau JD, Gagnon P, Harel F, Tremblay A and Roy MA (2005) *Fast, systematic, and continuous delirium assessment in hospitalized patients: the nursing delirium screening scale*. Journal of Pain and Symptom Management. Vol. 29, No.4, pp.368-375.
- Gerdner LA and Swanson EA (1993) *Effects of individualized music on confused and agitated elderly patients*. Archives of Psychiatric Nursing. Vol. 7, No.5, pp.284-91.

Gibbons LE, Teri L, Logsdon RG, et al. (2006) *Assessment of Anxiety in Dementia: An Investigation into the Association of Different Methods of Measurement*. Journal of Geriatric Psychiatry and Neurology. Vol.19, No.4, pp.202-208.

Glenn M (2005) *The Apathy Evaluation Scale*. The Center for Outcome Measurement in Brain Injury, available at <http://www.tbims.org/combi/aes>, accessed May 26, 2007.

Glenn MB, Burke DT, O'Neil-Pirozzi T, et al. (2002) *Cutoff score on the apathy evaluation scale in subjects with traumatic brain injury*. Brain Injury. Vol. 16, No.6, pp.509-516.

Goddaer J and Abraham IL (1994) *Effects of relaxing music on agitation during meals among nursing home residents with severe cognitive impairment*. Archives of Psychiatric Nursing. Vol. 8, No.3, pp.150-8.

Gormley N, Lyons D and Howard R (2001) *Behavioural management of aggression in dementia: a randomised controlled trial*. Age and Ageing. Vol. 30, pp.141-145.

Gormley N, Rizwan M and Lovestone S (1998) *Clinical predictors of aggressive behaviours in Alzheimer's disease*. Int J Geriatric Psychiatry. Vol. 15, pp.109-115.

Grace J and Malloy PF (2001) *Frontal Systems Behavior Scale (FrSBe): Professional Manual*. Psychological Assessment Resources.

Grassi L, Caraceni A, Beltrami E, Borreani C, Zamorani M, Maltoni M, et al. (2001) *Assessing delirium in cancer patients: the Italian versions of the Delirium Rating Scale and the Memorial Delirium Assessment Scale*. Journal of Pain and Symptom Management. Vol. 21, No.1, pp.59-68.

Grube M (2004) *Which types of aggressive behaviour are associated with suicidal and self-injurious behaviour at the time of admission?* Psychopathology. Vol. 37, No.1, pp.41-49.

Hall JN (1977) *The content of ward rating scales for long-stay patients*. Br. J Psychiatr. Vol. 130, pp.287-293.

Hall JN (1980) *Ward rating scales for long-stay patients*. Psychol.Med. Vol. 10, pp.277-288.

Hamilton M (1959) *The assessment of anxiety states by rating*. British Journal of Medical Psychology. Vol. 32, No.1, pp.50-55.

Hamilton M (1960) *A rating scale for depression*. Journal of Neurology, Neurosurgery and Psychiatry. Vol. 23, pp.56-62.

Harwood DG, Ownby RL, et al. (1998) *The behavioral pathology in Alzheimer's Disease Scale (BEHAVE-AD): factor structure among community-dwelling Alzheimer's disease patients*. International Journal of Geriatric Psychiatry. Vol. 13, No.11, pp.793-800.

Harwood DG, Sultzer DL, et al. (2000) *Impaired insight in Alzheimer disease: association with cognitive deficits, psychiatric symptoms, and behavioral disturbances*. Neuropsychiatry, Neuropsychology, and Behavioral Neurology. Vol. 13, No.2, pp.83-8.

Harwood DG, Sultzer DL, et al. (2005) *Frontal lobe hypometabolism and impaired insight in Alzheimer disease*. Am J Geriatr Psychiatry. Vol. 13, No.11, pp.934-41.

Hatoum HT, Lin S-J, Arcona S, Thomas SK, Koumaris B and Mirski D (2005) *The use of the occupational disruptiveness scale of the neuropsychiatric inventory-nursing home version to measure the impact of rivastigmine on the disruptive behaviour of nursing home residents with Alzheimer's disease*. Journal of the American Medical Directors Association. Vol. 6, No.4, pp.238-45.

Helmes E, Csapo KG and Short J-A (1987) *Standardization and validation of the Multidimensional Observation Scale for Elderly Subjects (MOSES)*. Journal of Gerontology. Vol. 42, pp.395-405.

Herrmann M, Ebert AD, Tober D, Hann J and Huth C (1999) *A contrastive analysis of release patterns of biochemical markers of brain damage after coronary artery bypass grafting and valve replacement and their*

association with the neurobehavioral outcome after cardiac surgery. European Journal of Cardio-Thoracic Surgery. Vol. 16, No.5, pp.513-518.

Hicks-Moore SL (2005) *Relaxing music at mealtime in nursing homes: effects on agitated patients with dementia*. Journal of Gerontological Nursing. Vol. 31, No.12, pp.26-32.

Higgins M, Koch K, Hynan LS, Carr S, et al. (2005) *Impact of an activities-based adult dementia care program*. Neuropsychiatric Disease And Treatment. Vol. 1, No.2, pp.165-169.

Hirono M, Mori E, Ikejiri Y, et al. (1997) *Japanese version of the Neuropsychiatric Inventory- a scoring system for neuropsychiatric disturbance in dementia patients*. No To Shinkei. Vol. 49, No.3, pp.266-271.

Holmes C, Hopkins V, Hensford C, MacLaughlin V, et al. (2002) *Lavender oil as a treatment for agitated behaviour in severe dementia: a placebo controlled study*. International Journal of Geriatric Psychiatry. Vol. 17, No.4, pp.305-8.

Hurtado B, Koepf MJ, Sander JW and Thompson PJ (2006) *The impact of levetiracetam on challenging behavior*. Epilepsy and Behavior. Vol. 8, No.3, pp.588-92.

Ikeda M, Fukuhara R, Shigendobu K, Hokoishi NM, Nebu A, Komori K and Tanabe H (2007) *Dementia associated mental and behavioural disturbances in elderly people in the community: findings from the first Nakayama study*. Journal Neral. Neurosurg. Psychiatry. Vol. 200, No.75, pp.146-148, available at [www.jnmp.com](http://www.jnmp.com), accessed 8 March, 2007.

Inouye S (2003) *The Confusion Assessment Method (CAM): Training Manual and Coding Guide*, University School of Medicine, Yale.

Inouye SK, Foreman MD, Mion LC, Katz KH and Cooney LM Jr (2001) *Nurses' recognition of delirium and its symptoms: comparison of nurse and researcher ratings*. Archives of Internal Medicine. Vol. 161, No.20, pp.2467-2473.

Inouye SK, van Dyck CH, Alessi CA, Balkin S, Siegel AP and Horwitz RI (1990) *Clarifying confusion: the confusion assessment method. A new method for detection of delirium*. Annals of Internal Medicine. Vol. 113, No.12, pp.941-948.

IPA (International Psychogeriatric Association) (2003) *The IPA Behavioural and Psychological Symptoms of Dementia*. Available at <http://www.ipa-online.net/pdfs/bpsd.asp>, accessed 1 June 2006.

Iverson GL, Hopp GA, de Wolfe K and Solomons K (2002) *Measuring change in psychiatric symptoms using the Neuropsychiatric Inventory: nursing home version*. Intern J Geriatric Psychiatry. Vol. 17, pp.438-443.

Jacobs MR, Strauss ME, Patterson MB and Mack JL (1998) *Characterization of depression in Alzheimer's disease by the CERAD Behavior Rating Scale for Dementia (BRSD)*. American Journal of Geriatric Psychiatry. Vol. 6, No.1, pp.53-8.

Jones RN, Yang FM, Zhang Y, Kiely DK, Marcantonio ER and Inouye SK (2006) *Does educational attainment contribute to risk for delirium? A potential role for cognitive reserve*. Journals of Gerontology Series A-Biological Sciences and Medical Sciences. Vol. 61, No.12, pp.1307-1311.

Kalisvaart KJ, de Jonghe JFM, Bogaards MJ, Vreeswijk R, Egberts TCG, Burger BJ, et al. (2005) *Haloperidol prophylaxis for elderly hip-surgery patients at risk for delirium: a randomized placebo-controlled study*. Journal of the American Geriatrics Society. Vol. 53, No.10, pp.1658-1666.

Kane R and Kane R (2000) *Assessing older persons: measures, meaning and practical applications*. Oxford University Press, USA.

Karlidag R, Unal S, Sezer OH, Bay Karabulut A, Battalolu B, But A, et al. (2006) *The role of oxidative stress in postoperative delirium*. General Hospital Psychiatry. Vol. 28, No.5, pp.418-423.

Kastango KB, Kim Y, et al. (2002) *Verification of scale sub-domains in elderly patients with dementia: a confirmatory factor-analytic approach*. American Journal of Geriatric Psychiatry. Vol. 10, No.6, pp.706-14.

- Kat MG and de Jonghe IF (2002) *Neuropsychiatric symptoms of dementia: psychometric properties of the Dutch Neuropsychiatric Inventory (NPI)*. Tijdschrift Gerontologie Geriatrie. Vol. 33, No.4, pp.150-155.
- Katona C, Livingston G, Cooper C, Ames D, Brodaty H and Chui E (2007) *International Psychogeriatric Association statement on defining and measuring treatment benefits in dementia*. International Psychogeriatrics. Doi: 10.1017/S1041610207005145.
- Katona CL and Aldridge CR (1985) *The dexamethasone suppression test and depressive signs in dementia*. Journal of Affective Disorders. Vol. 8, No.1, pp.83-9.
- Kaufers DL, Cummings JL and Christine D (1998) *Differential neuropsychiatric symptom response in Tacrine in Alzheimer's disease: relationship to symptom severity*. J Neuropsychiatry Clin Neurosci. Vol. 10, pp.55-63.
- Kaufers DL, Cummings JL, Christine D, Bray T, Castellon S, Masterman D, MacMillan A, Ketchel P and Dekovsky ST (1998) *Assessing the impact of neuropsychiatric symptoms in Alzheimer's disease; the Neuropsychiatric Inventory Caregiver Distress Scale*. Journal of the American Geriatric Society. Vol. 46, pp.210-215.
- Kelly G, Todd J, et al. (2006) *The Overt Behaviour Scale (OBS): a tool for measuring challenging behaviours following ABI in community settings*. Brain Injury. Vol. 20, No.3, pp.307-19.
- Kertesz A, Davidson W and Fox H (1997) *Frontal behavioral inventory: diagnostic criteria for frontal lobe dementia*. Can J Neurol Sci. Vol. 24, pp.29-36.
- Knopman DS, Berg JD, Thomas R, Grundman M, et al. (1999) *Nursing home placement is related to dementia progression: experience from a clinical trial*. Alzheimer's Disease Cooperative Study. Neurology. Vol. 52, No.4, pp.714-8.
- Koss E, Weiner M, Ernesto C, Cohen-Mansfield J, et al. (1997) *Assessing patterns of agitation in Alzheimer's disease patients with the Cohen-Mansfield Agitation Inventory*. The Alzheimer's Disease Cooperative Study. Alzheimer Disease and Associated Disorders. Vol. 11 Suppl.2, pp.S45-50.
- Kunik ME, Graham DP, Snow-Turek AL, Molinari VA, et al. (1998) *Contribution of cognitive impairment, depression, and psychosis to the outcome of agitated geropsychiatric inpatients with dementia*. Journal of Nervous and Mental Disease. Vol. 186, No.5, pp.299-303.
- Kurita A, Katayama K, et al. (1997) *Relationship between cognitive deficits, behavioral disturbances and falls in patients with dementia*. Nippon Ronen Igakkai Zasshi. Vol. 34, No.8, pp.662-7.
- Lam CL, Chan WC, Mok CC, Li SW, et al. (2006) *Validation of the Chinese Challenging Behaviour Scale: clinical correlates of challenging behaviours in nursing home residents with dementia*. International journal of geriatric psychiatry. Vol. 21, No.8, pp.792-9.
- Lam LC, Chui HF and Ng J (1997) *Aggressive behaviour in the Chinese elderly-validation of the Chinese version of the Rating Scale for Aggressive Behavior in the Elderly (RAGE) in hospital and nursing home settings*. International Journal of Geriatric Psychiatry. Vol. 12, No.6, pp.678-681.
- Lang AJ and Stein MB (2001) *Anxiety Disorders: How to recognize and treat the medical symptoms of emotional illness*. Geriatrics. Vol. 56, No.5, pp.24-34.
- Lange RT, Hopp G and Kang N (2004) *Psychometric properties and factors structure of the Neuropsychiatric Inventory Nursing Home version in an elderly neuropsychiatric population*. International Journal of Geriatric Psychiatry. Vol. 19, pp.440-448.
- Lauderdale SA and Sheikh JI (2003) *Anxiety disorders in older adults*. Clinics in Geriatric Medicine. Vol. 19, No.4, pp.721-741.
- Laurila JV, Pitkala KH, Strandberg TE and Tilvis RS (2002) *Confusion assessment method in the diagnostics of delirium among aged hospital patients: would it serve better in screening than as a diagnostic instrument?* International Journal of Geriatric Psychiatry. Vol. 17, No.12, pp.1112-1119.

- Laurila JV, Pitkala KH, Strandberg TE and Tilvis RS (2004) *Delirium among patients with and without dementia: does the diagnosis according to the DSM-IV differ from previous classifications?* International Journal of Geriatric Psychiatry. Vol. 19, pp.271-277.
- Lee DY, Choo IH, Kim KW, Jhoo JH, et al. (2006) *White matter changes associated with psychotic symptoms in Alzheimer's disease patients*. Journal of Neuropsychiatry and Clinical Neurosciences. Vol. 18, No.2, pp.191-8.
- Lee KU, Won WY, Lee HK, Kweon YS, Lee CT, Pae CU, et al. (2005) *Amisulpride versus quetiapine for the treatment of delirium: a randomized, open prospective study*. International Clinical Psychopharmacology. Vol. 20, No. 6, pp.311-314.
- Lemiengre J, Nelis T, Joosten E, Braes T, Foreman M, Gastmans C, et al. (2006) *Detection of delirium by bedside nurses using the confusion assessment method*. Journal of the American Geriatrics Society. Vol. 54, No.4, pp.685-689.
- Levin HS, High WM, et al. (1987) *The neurobehavioural rating scale: assessment of the behavioural sequelae of head injury by the clinician*. Journal of Neurology, Neurosurgery and Psychiatry. Vol. 50, No.2, pp. 183-93.
- Levy MH, Miller BL, Cummings JL, Fairbanks LAS and Craig A (1996) *Alzheimer's Disease and frontotemporal dementias: behavioural distinctions*. Arch. Neurol. Vol. 53, pp.687-690.
- Lewis LM, Miller DK, Morley JE, Nork MJ and Lasater LC (1995) *Unrecognized delirium in ED geriatric patients*. American Journal of Emergency Medicine. Vol. 13, No.2, pp.142-145.
- Lippert-Gruner M, Kuchta K, et al. (2006) *Neurobehavioural deficits after severe traumatic brain injury (TBI)*. Brain Injury. Vol. 20, No.6, pp.569-74.
- Litvan I, Mega MS, Cummings JL and Fairbanks L (1996) *Neuropsychiatric aspects of progressive supranuclear palsy*. Neurology. Vol. 47, pp.1184-1189.
- Logsdon RG, Teri L, Weiner MF, Gibbons LE, et al. (1999) *Assessment of agitation in Alzheimer's disease: the agitated behavior in dementia scale*. Alzheimer's Disease Cooperative Study. Journal of the American Geriatrics Society. Vol. 47, No.11, pp.1354-8.
- Lopez OL, Becker JT and Sweet RA (2005) *Non-cognitive symptoms in mild cognitive impairment subjects*. Neurocase. Vol. 11, No.1, pp.65-71.
- Lowe PA and Reynolds CR (2000) *Exploratory analyses of the latent structure of anxiety among older adults*. Educational and Psychological Measurement. Vol. 60, pp.100-116.
- Lyketsos CG (2007) *Neuropsychiatric symptoms (behavioural and psychological symptoms of dementia) and the development of dementia treatments*. Consensus Conference Paper. International Psychogeriatric Association. International Psychogeriatrics. pp.1-12.
- Lyketsos CG, Breitner JCS and Rabins PV (2001) *An evidence based proposal for the classification of neuropsychiatric disturbance in Alzheimer's disease*. International Journal of Geriatric Psychiatry. Vol. 16, pp.1037-1042.
- Lyketsos CG, Steinberg M, Tschanz JY, Norton MC, Steffens DC and Breitner JC (2000) *Mental and behavioural disturbances in dementia: findings for the Cache County study on memory in aging*. Am J Psychiatry. Vol. 157, pp.708-714.
- Mack JL, Patterson MB and Tariot PN (1999) *Behavior Rating Scale for Dementia: development of test scales and presentation of data for 555 individuals with Alzheimer's disease*. Journal of Geriatric Psychiatry and Neurology. Vol. 12, No.4, pp.211-23.
- Mahoney AE (2003) *Age- or stage-appropriate? Recreation and the relevance of Piaget's theory in dementia care*. American Journal of Alzheimer's Disease and Other Dementias. Vol. 18, No.1, pp.24-30.

- Malloy PF and Boyle PA (2005) *Apathy and Its Treatment in Alzheimer's Disease and Other Dementias*. Psychiatric Times. Vol. 22, No.13.
- Marcantonio ER, Michaels M and Resnick NM (1998) *Diagnosing delirium by telephone*. Journal of General Internal Medicine. Vol. 13, No.9, pp.621-623.
- Marin RS (1990) *Differential diagnosis and classification of apathy*. American Journal of Psychiatry. Vol. 147, pp.22-30.
- Marin RS, Biedrzycki RC and Firinciogullari S (1991) *Reliability and validity of the Apathy Evaluation Scale*. Psychiatry Research. Vol. 38, No.2, pp.143-162.
- Marin RS, Butters MA, Mulsant BH, et al. (2003) *Apathy and executive function in depressed elderly*. Journal of Geriatric Psychiatry and Neurology. Vol.16, No.2, pp.112-116.
- Marks IM and Mathews AM (1979) *Brief standard self-rating for phobic patients*. Behaviour Research and Therapy. Vol. 17, pp.263-267.
- Martin C, McKenzie S and Ames D (1994) *Disturbed behaviour in dementia sufferers: A comparison of three nursing home settings*. International Journal of Geriatric Psychiatry. Vol. 9, No.5, pp.393-398.
- Martinon-Torres G, Fioravanti M and Grimley EJ (2004) *Trazodone for agitation in dementia*. Cochrane Database of Systematic Reviews. No.4, pp.CD004990.
- Mathias JL (2003) *Neurobehavioral functioning of persons with Parkinson's disease*. Applied Neuropsychology. Vol. 10, No.2, pp.57-68.
- Mazaux JM, Masson F, et al. (1997) *Long-term neuropsychological outcome and loss of social autonomy after traumatic brain injury*. Archives of Physical Medicine and Rehabilitation. Vol. 78, No.12, pp.1316-20.
- McCauley SR, Levin HS, et al. (2001) *The neurobehavioural rating scale-revised: sensitivity and validity in closed head injury assessment*. Journal of Neurology, Neurosurgery and Psychiatry. Vol. 71, No.5, pp.643-51.
- McCusker J, Cole M, Bellavance F and Primeau F (1998) *Reliability and validity of a new measure of severity of delirium*. International Psychogeriatrics. Vol. 10, No.4, pp.421-433.
- McDowell I (2006) *Measuring Health: A guide to Rating Scales and Questionnaires (3rd ed)*. Oxford University Press.
- McGee S, Orenge C, Kunik ME, Molinari V, et al. (1997) *Delirium in geropsychiatric patients: Patient characteristics and treatment outcomes*. Journal of Geriatric Psychiatry and Neurology. Vol. 10, pp.7-10.
- McKeith I, Cummings J, Lovestone S, et al. (1999) *Outcome measures in Alzheimer's Disease*. Martin Dunitz, London.
- McNicoll L, Pisani MA, Ely EW, Gifford D and Inouye SK (2005) *Detection of delirium in the intensive care unit: comparison of confusion assessment method for the intensive care unit with confusion assessment method ratings*. Journal of the American Geriatrics Society. Vol. 53, No.3, pp.495-500.
- Meagher DJ, Moran M, Raju B, Gibbons D, Donnelly S, Saunders J, et al. (2007) *Phenomenology of delirium. Assessment of 100 adult cases using standardised measures*. Br J Psychiatry. Vol. 190, pp.135-141.
- Mega MS, Cummings JL, Fiorello R and Gornbein J (1996) *The spectrum of behavioural change in Alzheimer's Disease*. Neurology. Vol. 46, pp.130-135.
- Mendez MF, Perryman KM, et al. (1998) *Behavioral differences between frontotemporal dementia and Alzheimer's disease: a comparison on the BEHAVE-AD rating scale*. International Psychogeriatrics. Vol. 10, No.2, pp.155-62.

- Meyer TJ, Miller ML, Metzger RL, et al. (1990) *Development and validation of the Penn State Worry Questionnaire*. Behaviour Research and Therapy. Vol. 28, pp.487-495.
- Midlov P, Bondesson A, et al. (2002) *Descriptive study and pharmacotherapeutic intervention in patients with epilepsy or Parkinson's disease at nursing homes in southern Sweden*. European Journal of Clinical Pharmacology. Vol. 57, No.12, pp.903-10.
- Miller RJ, Snowdon J and Vaughan R (1995) *The use of the Cohen-Mansfield Agitation Inventory in the assessment of behavioral disorders in nursing homes*. Journal of the American Geriatrics Society. Vol. 43, No.5, pp.546-9.
- Mizoguchi T, Iijima S, et al. (1993) *Reliability and validity of a Japanese version of the Dementia Behavior Disturbance Scale*. Nippon Ronen Igakkai Zasshi. Vol. 30, No.10, pp.835-40.
- Moniz-Cook E, Woods R, Gardiner E, Silver M, et al. (2001) *The Challenging Behaviour Scale (CBS): development of a scale for staff caring for older people in residential and nursing homes*. British Journal of Clinical Psychology. Vol. 40, No.3, pp.309-22.
- Monteiro IM, Boksay I, et al. (2001) *Addition of a frequency-weighted score to the Behavioral Pathology in Alzheimer's Disease Rating Scale: the BEHAVE-AD-FW: methodology and reliability*. European Psychiatry: The Journal of the Association of European Psychiatrists. Vol. 16, Suppl.1, pp.5s-24s.
- Montgomery SA and Asberg M (1979) *A new depression scale designed to be sensitive to change*. British Journal of Psychiatry. Vol. 134, pp.382-9.
- Mungas D, Weller P, Franz C and Henry R (1989) *Assessment of disruptive behaviour associated with dementia: the Disruptive Behaviour Rating Scale*. Journal of Geriatric Psychiatry. Vol. 2, pp.196-202.
- Nagels G, Engelborghs S, Vloeberghs E, Van Dam D, et al. (2006) *Actigraphic measurement of agitated behaviour in dementia*. International Journal of Geriatric Psychiatry. Vol. 21, No.4, pp.388-93.
- Nelson JE, Tandon N, Mercado AF, Camhi SL, Ely EW and Morrison RS (2006) *Brain dysfunction: another burden for the chronically critically ill*. Archives of Internal Medicine. Vol. 166, No.18, pp.1993-1999.
- Neundorfer MM, McClendon MJ, Smyth KA, Stuckey JC, et al. (2001) *A longitudinal study of the relationship between levels of depression among persons with Alzheimer's disease and levels of depression among their family caregivers*. Journals of Gerontology Series B-Psychological Sciences and Social Sciences. Vol. 56, No.5, pp.P301-13.
- Neville CC and Byrne GJ (2002) *Behaviour of older people admitted for residential respite care*. Aust J Adv Nurs. Vol. 20, No.1, pp.8-12.
- Neville CC and Byrne GJ (2007) *Prevalence of disruptive behaviour displayed by older people in community and residential respite care settings*. International Journal of Mental Health Nursing. Vol. 16, No.2, pp.81-5.
- Nishikawa K, Nakayama M, Omote K and Namiki A (2004) *Recovery characteristics and post-operative delirium after long-duration laparoscope-assisted surgery in elderly patients: propofol-based vs. sevoflurane-based anesthesia*. Acta Anaesthesiologica Scandinavica. Vol. 48, No.2, pp.162-168.
- Nybo T, Sainio M, et al. (2005) *Middle age cognition and vocational outcome of childhood brain injury*. Acta Neurologica Scandinavica. Vol. 112, No.5, pp.338-42.
- O'Malley KJ, Orengo CA, Kunik ME, Snow L and Molinari V (2002) *Measuring aggression in older adults: a latent variable modeling approach*. Ageing Mental Health. Vol. 6, pp.231-238.
- Ohki T, Matsushima E, Shibuya M and Sunamori M (2006) *An evaluation strategy for the early detection of postoperative delirium*. Psychiatry and Clinical Neurosciences. Vol. 60, No.3, pp.277-282.
- O'Leary D, Jyringi D and Sedler M (2005) *Childhood conduct problems, stages of Alzheimer's disease, and physical aggression against caregivers*. International Journal of Geriatric Psychiatry. Vol. 20, No.5, pp.401-5.

- Onega LL (2006) *Assessment of psychoemotional and behavioral status in patients with dementia*. Nursing Clinics of North America. Vol. 41, No.1, pp.23-41.
- Opie J, Doyle C and O'Connor DW (2002) *Challenging behaviours in nursing home residents with dementia: a randomized controlled trial of multi-disciplinary interventions*. International Journal of Geriatric Psychiatry. Vol. 17, No.1, pp.6-13.
- Osborn DPJ, Fletcher AE, Smeeth L, et al. (2002) *Geriatric Depression Scale Scores in a representative sample of 14 545 people aged 75 and over in the United Kingdom: results from the MRC Trial of Assessment and Management of Older People in the Community*. International Journal of Geriatric Psychiatry. Vol.17, No.4, pp.375-382.
- Ott BR, Tate CA, et al. (1996) *Gender differences in the behavioral manifestations of Alzheimer's disease*. J Am Geriatr Soc. Vol. 44, No.5, pp.583-7.
- Oye T, Loke YA, Chan MF and Kwok T (2005) *Psychometric properties of the Chinese version of the Cost of Care Index to measure caregiver burden among Chinese informal caregivers*. Asian Journal of Nursing Studies. Vol. 8, No.2, pp.4-13.
- Pachana NA, Byrne GJ, Siddle H, et al. (2007) *Development and validation of the Geriatric Anxiety Inventory*. International Psychogeriatrics. Vol.19, No.1, pp.103-114.
- Pae CU, Lee SJ, Lee CU, Lee C and Paik IH (2004) *A pilot trial of quetiapine for the treatment of patients with delirium*. Human Psychopharmacology. Vol. 19, No.2, pp.125-127.
- Palmstierna T and Wistedt B (1987) *Staff Observation Aggression Scale. Presentation and observation*. Acta Psychiatrica Scandinavia. Vol. 76, pp.657-663.
- Patel V and Hope RA (1992a) *A rating scale for aggressive behaviour in the elderly: RAGE*. Psychological Medicine. Vol. 22, pp.211-221.
- Patel V and Hope RA (1992b) *Aggressive behaviour in elderly psychogeriatric inpatients*. Acta. Psychiatr Scand. Vol. 85, pp.131-135.
- Patel V and Hope RA (1992c) *The assessment of behavioural phenomena in dementia*. In A Burns (ed.). Ageing and Dementia: a Methodological Approach. Edward Arnold, UK.
- Patel V and Hope RA (1993a) *Aggressive behaviour in elderly people with dementia: A review*. International Journal of Geriatric Psychiatry. Vol. 8, pp.457-472.
- Patel V, Hope T, Hall J and Fairburn CG (1995) *Three methodological issues in the development of observer-rated behaviour rating scales: experience from the development of the Rating Scale for Aggressive Behaviour in the Elderly*. International Journal of Methods in Psychiatric Research. Vol. 5, pp.21-27.
- Patterson MB, Mack JL, Mackell JA, Thomas R, et al. (1997) *A longitudinal study of behavioral pathology across five levels of dementia severity in Alzheimer's disease: the CERAD Behavior Rating Scale for Dementia*. The Alzheimer's Disease Cooperative Study. Alzheimer Disease and Associated Disorders. Vol. 11 Suppl.2, pp.S40-44.
- Perez-Madrinan G, Cook SE, Saxton JA, Miyahara S, et al. (2004) *Alzheimer disease with psychosis: excess cognitive impairment is restricted to the misidentification subtype*. American Journal of Geriatric Psychiatry. Vol. 12, No.5, pp.449-456.
- Perivolaris A, LeClerc CM, Wilkinson K and Buchanan S (2006) *An enhanced dining program for persons with dementia*. Alzheimer's Care Quarterly. Vol. 7, No.4, pp.258-267.
- Pirkis JE, Burgess PM, Kirk PK, Dodson S, et al. (2005) *A review of the psychometric properties of the Health of the Nation Outcome Scales (HoNOS) family of measures*. Health and Quality of Life Outcomes. Vol. 3.
- Politis AM, Mayer LS, Passa M and Maillis A (2004) *Validity and reliability of the newly translated Hellenic Neuropsychiatric Inventory (H-NPI) applied to Greek outpatients with Alzheimer's Disease: a study of the*

- disturbing behaviours among referrals to a memory clinic.* International Journal of Geriatric Psychiatry. Vol. 19, pp.203-208.
- Pollock BG, Mulsant BH, et al. (2002) *Comparison of citalopram, perphenazine, and placebo for the acute treatment of psychosis and behavioral disturbances in hospitalized, demented patients.* American Journal of Psychiatry. Vol. 159, No.3, pp.460-5.
- Ponce H, Molinari V, Kunik ME, Orengo C, et al. (1998) *Placement predictors of geropsychiatric inpatients: Home versus nursing home.* Journal of Gerontological Social Work. Vol. 29, No.4, pp.3-12.
- Rabinowitz J, Davidson M, De Deyn PP, Katz I, et al. (2005) *Factor analysis of the Cohen-Mansfield Agitation Inventory in three large samples of nursing home patients with dementia and behavioral disturbance.* American Journal of Geriatric Psychiatry. Vol. 13, No.11, pp.991-8.
- Rabinowitz J, Katz IR, De Deyn PP, Brodaty H, et al. (2004) *Behavioral and psychological symptoms in patients with dementia as a target for pharmacotherapy with risperidone.* Journal of Clinical Psychiatry. Vol. 65, No.10, pp.1329-34.
- Radloff LS and Teri L (1986) *Use of the Center for Epidemiological Studies-Depression Scale with older adults.* Clinical Gerontologist. Vol. 5, No.1-2, pp.119-136.
- Rainer MK, Mucke HA, Kruger-Rainer C, Haushofer M, et al. (2004) *Zotepine for behavioural and psychological symptoms in dementia: an open-label study.[see comment].* CNS Drugs. Vol. 18, No.1, pp.49-55.
- Ramadan FH and Naughton BJ (1999) *Paroxetine for verbal agitation.* Annals of Long Term Care. Vol. 7, No.7, pp.282-286.
- Ramirez Diaz SP, Gregorio PG, Ribera Casado JM, et al. (2005) *The need for a consensus in the use of assessment tools for Alzheimer's disease: the Feasibility Study (assessment tools for dementia in Alzheimer Centres across Europe), a European Alzheimer's Disease Consortium's (EADC) survey.* International Journal of Geriatric Psychiatry. Vol. 20, No.8, pp.744-8.
- Rapoport M, McCauley S, et al. (2002) *The role of injury severity in neurobehavioral outcome 3 months after traumatic brain injury.* Neuropsychiatry, Neuropsychology, and Behavioral Neurology. Vol. 15, No.2, pp.123-32.
- Ray WA, Taylor JA, Lichtenstein MJ and Meador KG (1992) *The Nursing Home Behavior Problem Scale.* Journal of Gerontol. Vol. 47, No.1, pp.M9-16.
- Reisberg B, Borenstein J, et al. (1987) *BEHAVE-AD: A clinical rating scale for the assessment of pharmacologically remediable behavioral symptomatology in Alzheimer's disease.* In Altman, H.J (ed.) Alzheimer's disease: Problems, prospects, and perspectives. Plenum Press, New York. pp.1-16.
- Reisberg B, Monteiro I, et al. (2000) *Do many of the behavioral and psychological symptoms of dementia constitute a distinct clinical syndrome? Current evidence using the BEHAVE-AD.* International Psychogeriatrics. Vol. 12, Suppl.1, pp.155-164.
- Retsas AP and Crabbe H (1998) *Use of physical restraints in aged care facilities in New South Wales.* International Journal of Nursing Studies. Vol. 35, pp.177-183.
- Richeson NE (2003) *Effects of animal-assisted therapy on agitated behaviors and social interactions of older adults with dementia: an evidence-based therapeutic recreation intervention.* American Journal of Recreation Therapy. Vol. 2, No.4, pp.9-16.
- Richeson NE and Neill DJ (2004) *Therapeutic recreation music intervention to decrease mealtime agitation and increase food intake in older adults with dementia.* American Journal of Recreation Therapy. Vol. 3, No.1, pp.37-41.
- Robert PH, Clairet S, Benoit M, et al. (2002) *The apathy inventory: assessment of apathy and awareness in Alzheimer's disease, Parkinson's disease and mild cognitive impairment.* International Journal of Geriatric Psychiatry. Vol. 17, No.12, pp.1099-105.

- Rockwood K, Goodman J, Flynn M and Stolee P (1996) *Cross-validation of the Delirium Rating Scale in older patients*. Journal of the American Geriatrics Society. Vol. 44, No.7, pp.839-842.
- Rolfson DB, McElhaney JE, Jhangri GS and Rockwood K (1999) *Validity of the confusion assessment method in detecting postoperative delirium in the elderly*. International Psychogeriatrics. Vol. 11, No.4, pp.431-438.
- Rosen J, Bobys PD, Mazumdar S, Mulsant BH, et al. (1999) *OBRA regulations and neuroleptic use: defining agitation using the Pittsburgh Agitation Scale and the Neurobehavioral Rating Scale*. Annals of Long-Term Care. Vol. 7, No.12, pp.429-36.
- Rosen J, Burgio L, Kollar M, Cain M, et al. (1994) *The Pittsburgh Agitation Scale: a user friendly instrument for rating agitation in dementia patients*. American Journal of Geriatric Psychiatry. Vol. 2, pp.52-9.
- Rosen J, Sweet RA, Mulsant BH, Rifai AH, Pasternak R and Zubenko GS (1994) *The Delirium Rating Scale in a psychogeriatric inpatient setting*. Journal of Neuropsychiatry and Clinical Neurosciences. Vol. 6, No.1, pp.30-35.
- Rosenberg PR, Meikle MM and Lyketsos CG (2005) *Caregiver assessment of patient's depression in Alzheimer's Disease: longitudinal analysis in a drug treatment study*. American Journal of Geriatric Psychiatry. Vol. 13, pp.822-826.
- Rothenhausler HB, Grieser B, Nollert G, Reichart B, Schelling G and Kapfhammer HP (2005) *Psychiatric and psychosocial outcome of cardiac surgery with cardiopulmonary bypass: a prospective 12-month follow-up study*. General Hospital Psychiatry. Vol. 27, No.1, pp.18-28.
- Royal Australian College of General Practitioners (RACGP) (2005) *Medical care of older persons in residential aged care facilities (Silver book)* (4th Ed) Royal Australian College of General Practitioners.
- Ryden M (1988) *Aggressive behaviour in persons with dementia who live in the community*. Alzheimer Disease and Associated Disorders. Vol. 2, pp.342-355.
- Sanavio E (1988) *Obsessions and compulsions: the Padua Inventory*. Behaviour Research and Therapy. Vol. 26, pp.167-177.
- Schneider F, Bohner H, Habel U, Salloum JB, Stierstorfer A, Hummel TC, et al. (2002) *Risk factors for postoperative delirium in vascular surgery*. General Hospital Psychiatry. Vol. 24, No.1, pp.28-34.
- Schreiner AS, Yamamoto E and Shiotani H (2000) *Agitated behavior in elderly nursing home residents with dementia in Japan*. Journals of Gerontology Series B-Psychological Sciences and Social Sciences. Vol. 55, No.3, pp.P180-6.
- Schreiner D, Ballaban T, Brannath W, Lang T, et al. (2005) *Components of behavioral pathology in dementia*. International Journal of Geriatric Psychiatry. Vol. 20, No.2, pp.137-45.
- Sclan SG, Saillon A, et al. (1996) *The Behavior Pathology in Alzheimer's Disease Rating Scale (BEHAVE-AD): Reliability and analysis of symptom category scores*. International Journal of Geriatric Psychiatry. Vol. 11, No.9, pp.819-830.
- Shah A and Allen H (1999) *Editorial: Is Improvement possible in the measurement of behaviour disturbance in dementia?* International Journal of Geriatric Psychiatry. Vol. 14, pp.512-519.
- Shah A (1999a) *Some methodological issues in using aggression rating scales in intervention studies among institutionalised elderly*. Int. Psychogeriatric. Vol. 11, No.4, pp.439-444.
- Shah A (1999b) *Aggressive behaviour in the elderly*. Int J. Psychiat.Clin. Pract. Vol. 3, pp.85-103.
- Shah A and De T (1997) *The relationship between two scales measuring aggressive behaviour among continuing-care psychogeriatric inpatients*. International Psychogeriatrics. Vol. 9, pp.471-477.

- Shah A, Chiu E and Ames D (2000) *Aggressive behaviour and environmental characteristics in Australian nursing homes*. Ageing and Mental Health. Vol. 4, No.1, pp.43-47.
- Shah A, Chiu E, Ames D, Harrigan S and McKenzie D (2000) *Characteristics of aggressive subjects in Australian (Melbourne) nursing homes*. International Psychogeriatrics. Vol. 12, No.2, pp.145-161.
- Shah A, Ellanchenny N, et al. (2005) *A comparative study of behavioral and psychological symptoms of dementia in patients with Alzheimer's disease and vascular dementia referred to psychogeriatric services in Korea and the U.K.* Int Psychogeriatr. Vol. 17, No.2, pp.207-19.
- Shah A, Evans H and Parkash N (1998) *Evaluation of three aggression/agitation behaviour rating scales for use on an acute admission and assessment psychogeriatric ward*. International Journal of Geriatric Psychiatry. Vol. 13, No.6, pp.415-420.
- Shalek M, Richeson NE and Buettner LL (2004) *Air mat therapy for treatment of agitated wandering: an evidence-based recreational therapy intervention*. American Journal of Recreation Therapy. Vol. 3, No.3, pp.18-26.
- Shankar KK, Walker M, Frost D, et al. (1999) *The development of a valid and reliable scale for rating anxiety in dementia (RAID)*. Ageing and Mental Health Vol.3, No.1, pp.39-49.
- Shergill S, Shankar K, Seneviratna K and Orrell M (1999) *The validity and reliability of the Health of the Nation Outcome Scales (HoNOS) in the elderly*. Journal of Mental Health (UK). Vol. 8, pp.511-521.
- Sinoff G, Ore L, Zlotogorsky D, et al. (1999) *Short Anxiety Screening Test - a brief instrument for detecting anxiety in the elderly*. International Journal of Geriatric Psychiatry. Vol. 14, pp.1062-1071.
- Skjerve A, Holsten F, Aarsland D, Bjorvatn B, et al. (2004) *Improvement in behavioral symptoms and advance of activity acrophase after short-term bright light treatment in severe dementia*. Psychiatry and Clinical Neurosciences. Vol. 58, No.4, pp.343-7.
- Snyder M, Ryden MB, Shaver P, et al. (1998) *The Apparent Emotion Rating Instrument: assessing affect in cognitively impaired elders*. Clinical Gerontologist. Vol.18, No.4, pp.17-29.
- Sockeel P, Dujardin K, Devos D, et al. (2006) *The Lille Apathy Rating Scale (LARS): a new instrument for detecting and quantifying apathy: Validation in Parkinson's disease*. Journal of Neurology, Neurosurgery and Psychiatry. Vol.77, No.5, pp.579-584.
- Spear J, Chawla S, O'Reilly M and Rock D (2002) *Does the HoNOS 65+ meet the criteria for a clinical outcome indicator for mental health services for older people?* International Journal of Geriatric Psychiatry. Vol. 17, pp.226-230.
- Spielberger CD, Gorsuch R and Lushene RE (1970) *Manual for the State-Trait Anxiety Inventory*. Consulting Psychologists Press, Palo Alto, CA.
- Starkstein SE, Ingram L, Garau ML, et al. (2005) *On the overlap between apathy and depression in dementia*. Journal of Neurology, Neurosurgery and Psychiatry. Vol. 76, No.8, pp.1070-1074.
- Starkstein SE, Mayberg HS, Preziosi TJ, et al. (1992) *Reliability, validity, and clinical correlates of apathy in Parkinson's disease*. Journal of Neuropsychiatry and Clinical Neurosciences. Vol. 4, No.2, pp.134-9.
- Steffens DC, Otey E, Alexopoulos GS, et al. (2006) *Perspectives on depression, mild cognitive impairment, and cognitive decline*. Archives of General Psychiatry. Vol. 63, No.2, pp.130-8.
- Steinberg, M., Tschanz, JT., Norton, M., Breitner, JCS. and Liketsos, CG. (2006) *Point and five-year period prevalence of neuropsychiatric symptoms in dementia: The Catch County Study*. Poster, International Conference on Alzheimer's Disease, Madrid, July 2006.
- Straker DA, Shapiro PA and Muskin PR (2006) *Aripiprazole in the treatment of delirium*. Psychosomatics. Vol. 47, No.5, pp.385-391.

- Strauss ME and Sperry SD (2002) *An informant-based assessment of apathy in Alzheimer disease*. Neuropsychiatry, Neuropsychology, and Behavioral Neurology. Vol.15, No.3, pp.176-183.
- Stuss DT, van Reekum R, Murphy KJ (2000) *Differentiation of states and causes of apathy*. In Borod J (ed.) The Neuropsychology of Emotion. Oxford University Press, New York. pp 340–363.
- Suh GH (2004) *Agitated behaviours among the institutionalized elderly with dementia: validation of the Korean version of the Cohen-Mansfield Agitation Inventory*. International Journal of Geriatric Psychiatry. Vol. 19, No.4, pp.378-85.
- Suh GH and Kim SK (2004) *Behavioral and Psychological Signs and Symptoms of Dementia (BPSD) in antipsychotic-naïve Alzheimer's Disease patients*. International Psychogeriatrics. Vol. 16, No.3, pp.337-50.
- Suh GH, Greenspan AJ and Choi SK (2006) *Comparative efficacy of risperidone versus haloperidol on behavioural and psychological symptoms of dementia*. International journal of geriatric psychiatry. Vol. 21, No.7, pp.654-60.
- Suh GH, Son HG, Ju YS, Jcho KH, et al. (2004) *A randomized, double-blind, crossover comparison of risperidone and haloperidol in Korean dementia patients with behavioral disturbances*. American Journal of Geriatric Psychiatry. Vol. 12, No.5, pp.509-16.
- Sultzer DL, Berisford MA, et al. (1995) *The Neurobehavioral Rating Scale: reliability in patients with dementia*. Journal of Psychiatric Research. Vol. 29, No 3, pp.185-91.
- Sultzer DL, Levin HS, et al. (1992) *Assessment of cognitive, psychiatric, and behavioral disturbances in patients with dementia: the Neurobehavioral Rating Scale*. Journal of the American Geriatrics Society. Vol. 40, No.6, pp.549-55.
- Sultzer DL, Levin HS, et al. (1993) *A comparison of psychiatric symptoms in vascular dementia and Alzheimer's disease*. American Journal of Psychiatry. Vol. 150, No.12, pp.1806-12.
- Sultzer DL, Mahler ME, et al. (1995) *The relationship between psychiatric symptoms and regional cortical metabolism in Alzheimer's disease*. Journal of Neuropsychiatry and Clinical Neurosciences. Vol. 7, No.4, pp.476-84.
- Sunderland T, Hill JL, Lawlor BA, et al. (1988) *NIMH Dementia Mood Assessment Scale (DMAS)*. Psychopharmacology Bulletin. Vol. 24, No.4, pp.747-53.
- Swearer JM, Hoople NE, Kane KJ and Drachman DA (1996) *Predicting aberrant behaviour in Alzheimer's Disease*. Neuropsychiatry, Neuropsychology and Behavioural Neurology. Vol. 9, No.3, pp.162-170.
- Takeuchi T, Furuta K, Hirasawa T, Masaki H, Yukizane T, Atsuta H, et al. (2007) *Perospirone in the treatment of patients with delirium*. Psychiatry and Clinical Neurosciences. Vol. 61, No.1, pp.67-70.
- Tariot PN (1996) *CERAD Behavior Rating Scale for Dementia*. International Psychogeriatrics. Vol. 8, Suppl.3, pp.514-515.
- Tariot PN, Mack JL, Patterson MB, Edland SD, et al. (1995) *The Behavior Rating Scale for Dementia of the Consortium to Establish a Registry for Alzheimer's Disease*. The Behavioral Pathology Committee of the Consortium to Establish a Registry for Alzheimer's Disease. American Journal of Psychiatry. Vol. 152, No.9, pp.1349-57.
- Tariot PN, Schneider LS, Mintzer JE, Cutler AJ, et al. (2001) *Safety and tolerability of divalproex sodium in the treatment of signs and symptoms of mania in elderly patients with dementia: Results of a double-blind, placebo-controlled trial*. Current Therapeutic Research, Clinical and Experimental. Vol. 62, No.1, pp.51-67.
- Teri L, Logsdon RG, Peskind E, Raskind M, et al. (2000) *Treatment of agitation in AD: a randomized, placebo-controlled clinical trial.[see comment][erratum appears in Neurology 2001 Feb 13;56(3):426]*. Neurology. Vol. 55, No.9, pp.1271-8.

- Teri L, Rabins P, Whitehouse P, Berg L, Reisberg B, Sunderland T, Eicheleman B and Phelps C (1992) *Management of behaviours disturbance in Alzheimer's disease: Current knowledge and future directions.* Alzheimer's Disease and Associated Disorders. Vol. 6, No.2, pp.77-88.
- Teri L, Truax P, Logsdon R, Uomoto J, Zarit S and Vitaliano PP (1992) *Assessment of behavioral problems in dementia: the revised memory and behavior problems checklist.* Psychology and Ageing. Vol. 7, No.4, pp.622-631.
- Tractenberg RE, Gamst A, Thomas RG, Patterson M, et al. (2002) *Investigating emergent symptomatology as an outcome measure in a behavioral study of Alzheimer's disease.* Journal of Neuropsychiatry and Clinical Neurosciences. Vol. 14, No.3, pp.303-10.
- Tractenberg RE, Patterson M, Weiner MF, Teri L, et al. (2000) *Prevalence of symptoms on the CERAD behavior rating scale for dementia in normal elderly subjects and Alzheimer's disease patients.*[erratum appears in *J Neuropsychiatry Clin Neurosci* 2001 Winter;13(1):95]. Journal of Neuropsychiatry and Clinical Neurosciences. Vol. 12, No.4, pp.472-9.
- Tractenberg RE, Weiner MF, Patterson MB, Gamst A, et al. (2002) *Emergent psychopathology in Alzheimer's disease patients over 12 months associated with functional, not cognitive, changes.* Journal of Geriatric Psychiatry and Neurology. Vol. 15, No.2, pp.110-7.
- Tractenberg RE, Weiner MF, Patterson MB, Teri L, et al. (2003) *Comorbidity of psychopathological domains in community-dwelling persons with Alzheimer's disease.* Journal of Geriatric Psychiatry and Neurology. Vol. 16, No.2, pp.94-9.
- Trzepacz PT (1999) *The Delirium Rating Scale. Its use in consultation-liaison research.* Psychosomatics. Vol. 40, No.3, pp.193-204.
- Trzepacz PT, Baker RW and Greenhouse J (1988) *A symptom rating scale for delirium.* Psychiatry Research. Vol. 23, No.1, pp.89-97.
- Trzepacz PT, Mittal D, Torres R, Canary K, Norton J and Jimerson N (2001) *Validation of the Delirium Rating Scale-revised-98: comparison with the delirium rating scale and the cognitive test for delirium.*[erratum appears in *J Neuropsychiatry Clin Neurosci* 2001 Summer;13(3):433]. Journal of Neuropsychiatry and Clinical Neurosciences. Vol. 13, No.2, pp.229-242.
- Trzepacz PT, Mulsant BH, Dew MA, Pasternak R, Sweet RA and Zubenko GS (1998) *Is delirium different when it occurs in dementia? A study using the delirium rating scale.* J Neuropsychiatry Clin Neurosci. Vol. 10, No.2, pp.199-204.
- Turner S (2004) *Are the Health of the Nation Outcome Scales (HoNOS) useful for measuring outcomes in older people's mental health services?* Ageing and Mental Health. Vol. 8, No. 5, pp. 387-396.
- van Baalen B, Odding E, et al. (2006) *Reliability and sensitivity to change of measurement instruments used in a traumatic brain injury population.* Clinical Rehabilitation. Vol. 20, No.8, pp.686-700.
- van Diepen E, Baillon SF, Redman J, Rooke N, et al. (2006) *A pilot study of the physiological and behavioural effects of Snoezelen in dementia.* British Journal of Occupational Therapy. Vol. 65, No.2, pp.61-66.
- van Reekum R, Stuss DT and Ostrander L (2005) *Apathy: Why care?* J Neuropsychiatry Clin Neurosci. Vol. 17, No.1, pp.7-19.
- Vanier M, Mazaux JM, et al. (2000) *Assessment of neuropsychologic impairments after head injury: interrater reliability and factorial and criterion validity of the Neurobehavioral Rating Scale-Revised.* Archives of Physical Medicine and Rehabilitation. Vol. 81, No.6, pp.796-806.
- Verkaik R, van Weert JC and Francke AL (2005) *The effects of psychosocial methods on depressed, aggressive and apathetic behaviors of people with dementia: a systematic review.* International Journal of Geriatric Psychiatry. Vol.20, No.4, pp.301-314.

- Vespa A, Gori G and Spazzafumo L (2002) *Evaluation of non-pharmacological intervention on antisocial behavior in patients suffering from Alzheimer's disease in a day care center.* Archives of gerontology and geriatrics. Vol. 34, No.1, pp.1-8.
- Vespa A, Gori S, Bonaiuto S, Cruciani G, et al. (2002) *Validation of the Cohen-Mansfield agitation inventory (CMAI-long form) on a sample of Italian demented patients.* Archives of gerontology and geriatrics. Vol. 8, pp.377-83.
- Vetter PH, Krauss S, et al. (1999) *Vascular dementia versus dementia of Alzheimer's type: do they have differential effects on caregivers' burden?* Journals of Gerontology Series B-Psychological Sciences and Social Sciences. Vol. 54, No.2, pp.S93-8.
- Vilkki J, Ahola K, et al. (1994) *Prediction of psychosocial recovery after head injury with cognitive tests and neurobehavioral ratings.* Journal of Clinical and Experimental Neuropsychology: Official Journal of the International Neuropsychological Society. Vol. 16, No.3, pp.325-38.
- Villanueva MR, Smith TL, Erickson JS, Lee AC, et al. (2003) *Pain Assessment for the Dementing Elderly (PADE): reliability and validity of a new measure.[see comment].* Journal of the American Medical Directors Association. Vol. 4, No.1, pp.1-8.
- Volicer L, Camberg L, Hurley AC, Ashley J, et al. (1999) *Dimensions of decreased psychological well-being in advanced dementia.* Alzheimer Disease and Associated Disorders. Vol. 13, No.4, pp.192-201.
- Voyer P, McCusker J, Cole M and Khomenko L (2006) *Influence of prior cognitive impairment on the severity of delirium symptoms among older patients.* Journal of Neuroscience Nursing. Vol. 38, No.2, pp.90-101.
- Watson LC and Pignone MP (2003) *Screening accuracy for late-life depression in primary care: a systematic review.* Journal of Family Practice. Vol. 52, No.12, pp.956-964.
- Weiner MF, Koss E, Patterson M, Jin S, et al. (1998) *A comparison of the Cohen-Mansfield agitation inventory with the CERAD behavioral rating scale for dementia in community-dwelling persons with Alzheimer's disease.* Journal of Psychiatric Research. Vol. 32, No.6, pp.347-51.
- Weiner MF, Martin-Cook K, Foster BM, Saine K, et al. (2000) *Effects of donepezil on emotional/behavioral symptoms in Alzheimer's disease patients.* Journal of Clinical Psychiatry. Vol. 61, No.7, pp.487-92.
- Weiner MF, Tractenberg R, Teri L, Logsdon R, et al. (2000) *Quantifying behavioral disturbance in Alzheimer's disease patients.* Journal of Psychiatric Research. Vol. 34, No.2, pp.163-7.
- Weiner MF, Tractenberg RE, Jin S, Gamst A, et al. (2002) *Assessing Alzheimer's disease patients with the Cohen-Mansfield Agitation Inventory: scoring and clinical implications.* Journal of Psychiatric Research. Vol. 36, No.1, pp.19-25.
- Weiner MF, Williams B and Risser RC (1997) *Assessment of behavioral symptoms in community-dwelling dementia patients.* American Journal of Geriatric Psychiatry. Vol. 5, No.1, pp.26-30.
- Wells DL, Dawson P, Sidani S, Craig D, et al. (2000) *Effects of an abilities-focused program of morning care on residents who have dementia and on caregivers.* Journal of the American Geriatrics Society. Vol. 48, No.4, pp.442-9.
- Werner P, Cohen-Mansfield J, Braun J and Marx MS (1989) *Physical restraint and agitation in nursing home residents.* Journal of the American Geriatric Society. Vol. 37, pp.1122-1126.
- Werner P, Cohen-Mansfield J, Koroknay V and Braun J (1994) *The impact of a restraint-reduction program on nursing home residents.* Geriatric Nursing. Vol. 15, No.3, pp.142-46.
- Westhuis D and Thyer BA (1989) *Development and validation of the clinical anxiety scale: a rapid assessment instrument for empirical practice.* Educational and Psychological Measurement. Vol. 49, pp.153-163.

- Weyerer S, Killmann U, Ames D, et al. (1999) *The Even Briefer Assessment Scale for Depression (EBAS DEP): its suitability for the elderly in geriatric care in English- and German-speaking countries*. International Journal of Geriatric Psychiatry. Vol. 14, No.6, pp.473-80.
- Whall AL, Black MEA, Yankou DJ, Groh CJ, et al (1999) *Nurse aides' identification of onset and level of agitation in late stage dementia patients*. American Journal of Alzheimer's Disease. Vol. 14, No.4, pp.202-206.
- White HK, McConnell ES, Bales CW and Kuchibhatia M (2004) *The 6<sup>th</sup> month study of the relationship between weight loss and behavioural symptoms in institutionalised Alzheimer's disease subjects*. Journal of the American Medical Directors Association. Vol. 5, No.2, pp. 89-97.
- Whitehouse PJ, Patterson MB, Strauss ME, Geldmacher DS, et al. (1996) *Hallucinations*. International Psychogeriatrics. Vol. 8, No.Suppl 3, pp.387-392.
- Wilkosz PA, Miyahara S, Lopez OL, Dekosky ST, et al. (2006) *Prediction of psychosis onset in Alzheimer disease: The role of cognitive impairment, depressive symptoms, and further evidence for psychosis subtypes*. American Journal of Geriatric Psychiatry. Vol. 14, No.4, pp.352-60.
- Williams AK (2005) *Motivation and dementia*. Topics in Geriatric Rehabilitation. Vol.21, No.2, pp.123-126.
- Wisocki PA (1988) *Worry as a phenomenon relevant to the elderly*. Behavior Therapy. Vol. 19, No.3, pp.369-379.
- Wistedt B, Rasmussen A, Pedersen L, Malm U, Tracksman-Bendtz L, Wakelin J 7 Bech P (1990) *The development of an observer-scale for measuring social dysfunction and aggression*. Pharmacopsychiatry. Vol. 23, No.6, pp.249-52.
- Wood S, Cummings JL, Hsu MA, et al. (2000) *The use of the Neuropsychiatric Inventory in nursing home residents: characteristics and measurement*. American Journal of Geriatric Psychiatry. Vol. 8, pp.75-83.
- World Health Organisation (1992) *The ICD-10 Classification of Mental and Behavioural Disorders: Clinical descriptions and diagnostic guidelines*. WHO, Geneva.
- Yesavage JA, Brink TL, Rose TL, et al. (1983) *Development and validation of a geriatric depression screening scale: a preliminary report*. Journal of Psychiatric Research. Vol. 17, No.1, pp.37-49.
- Yudovsky SC, Silver JM, Jackson W, Endicott J and Williams D (1986) *The Overt Aggression Scale for objective rating of verbal and physical aggression*. American Journal of Psychiatry. Vol. 143, pp.35-39.
- Zieber CG, Hagen B, Armstrong-Esther C and Aho M (2005) *Pain and agitation in long-term care residents with dementia: use of the Pittsburgh Agitation Scale*. International Journal of Palliative Nursing. Vol. 11, No.2, pp.71-8.
- Zigmond AS and Snaith RP (1983) *The Hospital Anxiety and Depression Scale*. Acta Psychiatrica Scandinavica. Vol. 67, No.6, pp.361-370.
- Zuidema SU, de Jonghe JF, Verhey FR and Koopmans RT (2007) *Agitation in Dutch institutionalized patients with dementia: factor analysis of the Dutch version of the Cohen-Mansfield Agitation Inventory*. Dementia and Geriatric Cognitive Disorders. Vol. 23, No.1, pp.35-41.

## 10 Measures of Function for Dementia

### 10.1 Introduction

This chapter summarizes a number of key geriatric and dementia review papers (Pearson, 2000 in Kane and Kane, 2000; Burns, et al. 2004; Spector, 1997; McKeith, et al. 1999) in the area of the measurement of function. Both Spector (1997) and Pearson (2000) also provide useful accounts of the historical development of important measures and concepts in this area. This chapter highlights a number of key issues, as well as describing the current state of the research literature. These issues include:

- A summary of the measurement literature for the assessment of function for people with dementia
- Challenges for generic functional assessment instruments when used with people with dementia
- Challenges for dementia specific functional assessment instruments
- Some recent research highlights

Following this the process of instrument selection and recommendations will be discussed.

#### 10.1.1 Importance of the Measurement of Function for People with Dementia

“Measurement of function is an essential part of clinical practice and is one of the major outcomes used in the assessment of interventions in dementia. It is also one of the main determinates of people being admitted to long-term care” (Burns, et al. 2004, page 185).

These views are reaffirmed by major reviewers in the dementia field. McKeith, et al. (1999) points out that a change in function status is an independent criterion for the diagnosis of AD, representing the outward face of global decline and dictating care needs. Spector (1997) highlights the importance of ongoing functional assessment in tracking the relationship between cognitive decline and the performance on basic activities for clinical practice. This is especially in terms of identifying milestones, access to home and institutional care and preventing unnecessary behaviour problems.

Ideally assessment should be undertaken by a multidisciplinary team with the occupational therapist having a leading role, using a combination of data from direct observation, caregiver interview and validated scales (McKeith, et al. 1999). In terms of clinical practice, it naturally follows that this approach toward the formal assessment of functional status, highlights the importance of examining the performance of the individual in their own environment, rather than to just view them as an older person, in terms of their chronological age. As Pearson (2000) in Kane and Kane (2000) comments, the tendency toward ageism in assessment must be avoided.

#### 10.1.2 A Simple Working Definition of Function

Pearson (2000) in Kane and Kane (2000) provides a useful working definition for this area of functional measurement:

“Over the years, physical functioning has come to mean a person’s ability to perform those activities deemed necessary to survive adequately in modern society. Functional assessment includes three domains: activities of daily living or self-care activities, instrumental activities of daily living (IADLs), and mobility.” (page 17)

Activities of daily living (ADLs) or self-care activities are also known as personal activities of daily living (PADL) or basic activities of daily living (BADLs). IADL is also known as extended activities of daily living (EADL). Chong (1995) indicates the phrase “instrumental activities of daily living” (IADL) was originally introduced in 1969 by Lawton and Brody in their seminal work *Assessment of Older People: Self Maintaining and Instrumental Activities of Daily Living*. Rather than provide a definition of IADL, they describe the schema of competence into which behaviours would fit, taking life maintenance at the lowest level. They measured this with the Physical Self- Maintenance Scale, which corresponds to the present general understanding of an ADL scale. Behaviours that indicated successively more complex levels of function were ascribed to the IADL scale.

However, Pearson (2000) in Kane and Kane (2000) summarises two key issues from the literature. Firstly, that there is no universally recognised, “gold standard” functional assessment tool; and secondly, while there are many IADL instruments and the concept of IADL is generally understood, there is no universally accepted operational definition.

These two definitional issues colour much of the research activity being carried out today in the measurement of function across all disease groups. A further complication is the lack of psychometric development work at the other end of the functional continuum, namely the measurement of participation in society. Following on from the work in developing the International Classification of Functioning, Disability and Health (ICF), the challenge here is to move from the traditional measurement of functional activities and skills toward broader measures of functional outcome in home, community and social participation (for more discussion see Jette, et al. 2005 and Heinemann, 2005). As an exercise the recommended functional measures were compared against the ICF framework. This analysis showed that each of the recommended measures only covered one or two elements of each of the ICF descriptors (these descriptors may contain 3-5 or even more elements) and none of the instruments covered environmental factors. The psychometric examination of an individual’s performance of activities (or function status level) in the local environment or their participation in society in general is a new area which requires further developmental work.

### **10.1.3 A Possible Analysis Framework for Functional Assessment Instruments**

Pearson (2000) in Kane and Kane (2000) also provides a useful starting point for the development of an analysis framework or typology for the examination of different types of functional assessment instruments. Pearson (2000) in Kane and Kane (2000) outlines a number of categories which include:

- Single item or two item questions (or very short measures e.g. 4 items, see Li, et al. 2006)
- Self-report measures
- Proxy / informant report measures
- Direct observation measures (i.e. clinical rating)
- Performance based measures

These categories are then applied across the domains of ADL, or IADL or some combination of ADL/IADL items. These instruments can then be further categorized into either generic measures of function for older people or dementia specific measures of function.

### **10.1.4 Challenges for Functional Assessment Instruments**

Pearson (2000) highlights the main challenges for functional assessments:

- Gender and culture bias in instruments
- The relationship with cognitive functioning has not been fully explored
- A need for greater alignment with patient goals and expectations

- The relationship to chronic health problems needs to be examined
- The focus on disability and dependence; the instruments do not consider strengths
- The need to examine the impact of the home design and the local environment on functioning
- The ADL scales used in the community have ceiling effects
- How to deal with the competency vs. actual performance problem in ADL and IADL measurement, otherwise known as the “can-do” vs. the “do-do” problem. For instance in a household there may be the sharing of duties e.g. with cooking or finances, meaning that one partner while being able to do a task but does not undertake it on a regular basis. This has important implications for measurement and in establishing prevalence rates (see Bootsma-van der Wiel, et al. 2001 for an example)
- How to score the use of assistance (provided by a person or equipment)

Additional issues from the scientific literature were:

- The bandwidth fidelity problem – the trade-off between range (and detail) vs. practicality of the scale (see Lindeboom, et al. 2003)
- Whether there is one underlying dimension for ADL and IADL or whether they are separate and have multiple dimensions
- And following from this, do ADL and IADL items have a particular structure or order e.g. in terms of functional incapacity does one lose IADLs before ADLs, or lose bathing skills before feeding skills (as in the Katz Index)? Bootsma-van der Wiel, et al. (2001) and Thomas, et al. (1998) provide data which does not support the traditional conceptualisation of hierarchical structure)

Finally, there is the issue of the sheer number of instruments available to users. As Pearson (2000) suggests “The future of functional assessment might well be served by calling a moratorium on developing new instruments and by weeding out instruments that have poor psychometric characteristics.” (page 46)

These challenges outline the research that needs to be undertaken and represent the state of the art in the functional assessment area.

### **10.1.5 Summary of the Measurement Literature for the Assessment of Function for People with Dementia**

Spector (1997) indicates the most popular generic ADL scales for use with people with dementia are the Katz Activities of Daily Living scale (Katz, et al. 1963), the Physical Self-Maintenance Scale (Lawton and Brody, 1969), and the Functional Independence Measure (Hamilton, et al. 1987) and its precursor the Barthel Index (Mahoney and Barthel, 1965). However, Spector (1997) also states that there has been little activity in constructing or adjusting generic scales so that they are not biased for demented populations.

He notes that validation of dementia specific scales can be characterised as preliminary and inconclusive and that none of these scales have gained general acceptance in the clinical or research arenas, except perhaps the functional questions on the Blessed scale, which is the oldest scale. This would suggest that further research is needed in this area.

### 10.1.6 Challenges for Generic Functional Assessment Instruments when used with People with Dementia

While they have good psychometric properties and have been used in many studies, the major criticism of generic functional assessment instruments for people with dementia is that they are insensitive to the functional losses from cognitive impairment associated with dementia and are thus biased against them (Spector, 1997). Burns, et al. (2004) also comments on this issue of generic instruments focusing on functional problems resulting from physical impairments, rather than those resulting from and secondary to cognitive deficits.

Spector (1997) goes on to outline other related problems with generic instruments:

- They do not address impairments in the ability to plan, organise, sequence and remember for people with dementia.
- They do not address a person with dementia's need for prompting to cue for imitation, to reduce confusion and to aid concentration.
- They do not address the person with dementia's need for supervision for safety reasons.
- They do not reduce the complexity / difficulty of the task which may be necessary for people with dementia, e.g. setting out clothing in the morning to help them get dressed.
- They do not contain enough cognitive items.
- They measure the person's ability to do the whole task, rather than the separate aspects of the task that may be problematic for people with dementia.

In selecting a generic measure of function for people with dementia these issues need to be borne in mind. It suggests that these established instruments may require adaptation for persons with dementia.

### 10.1.7 Challenges for Dementia Specific Functional Assessment Instruments

A major challenge for dementia specific instruments is the need to span the whole range of dementia severity types, where in milder forms of dementia only IADLs are affected, while in severe forms both ADLs and IADLs are affected (Burns, et al. 2004). Burns, et al (2004) suggests that depending on the setting it may be important to know which functions are intact in terms of basic daily living tasks or more complex instrumental tasks.

Other measurement challenges noted by Spector (1997) include the need to:

- Further examine the relationship between dementia specific functional measures with generic measures and cognitive measures. At present it is difficult to determine whether these instruments are measuring functional disability, cognitive impairment or some combination of these dimensions.
- Further examine the relationship between dementia specific observation measures of function with proxy / informant measures and performance based measures.
- Examine how best to measure help / assistance (especially with performance measures – use of prompts and gestures).
- Confirm that when informants and self-reporters are making ratings they understand the concepts like 'initiates', 'appropriateness' and 'supervision'.
- Clarify whether the instrument is measuring function or cognitive impairment. For example the functional items of the Blessed ask about a tendency to dwell in the past, the inability to recall recent events, and the inability to remember a short list of items.
- Target item selection. What is the right mix of ADL and IADL items for mild, moderate and severe dementia?

- Study the properties of these instruments with different types of dementia.
- Identify whether there is one dimension or multiple dimensions underlying the ADL / IADL model? (This issue applies to all functional instruments, see above) (Spector, 1997 makes the analogy to cognition where multiple underlying dimensions can be used to construct a model).
- Assess whether the scales can be streamlined. (This issue applies to all functional instruments, see above).

These issues (as well as those raised in Section 10.5.1) suggest that there is no one measurement solution at this stage.

### 10.1.8 Recent Research Highlights

Some highlights of the recent research literature include:

- A recent paper by Ayis, et al. (2007) found that asking people using various definitions of functional decline (deterioration in functioning, occurrence of difficulty in performing and change from independence to dependency) produces different patterns or results according to health condition, gender and environmental factors; as well as suggesting that ADL items should not be summed together.
- A paper by Neugebauer, et al. (2003) in the area of rheumatoid arthritis, examined changes in functional status and explored “the loss of valued activities” and its relationship to psychological well-being.
- The work of Coster, et al. (2004) used Rasch analysis in the development of new instruments from the existing instruments and examined the underlying structure between items. They challenge the traditional conceptualisation of basic and instrumental activities, to suggest a structure of functional activities which require either whole body movement or skilled upper limb use or cognitive skills.
- The work of Lindeboom, et al. (2003) in the Amsterdam Liner Disability Score Project, which used IRT to calibrate functional items to create new scales and item banks.

These papers highlight the need for further work in this area and the potential statistical methods for new research.

## 10.2 Selection of Instruments

Functional instruments were selected for comprehensive review based on literature search and instrument selection methodology outlined previously (refer Section 2). For the selection of the measurement of function instruments these steps are summarised below.

### 10.2.1 Literature Search

98 instruments in the area of function were identified on the Master Database by the literature search and instrument selection methodology. This list was cut in half by focusing on those instruments with high citations from the academic literature and / or their use in practice survey. This had the effect of creating a more manageable list, leaving behind those instruments with one or two mentions only. This revised list was further examined to produce a list of contender instruments for comprehensive review. The steps in the selection process are summarised below.

For this analysis a number of different types of tests were excluded. These were:

- Comprehensive assessments for older people where functional items are included but do not form a recognised separate component assessment tool, like the InterRAI (Hawes, et al. 1997), Revised Elderly Person’s Disability Scale (REPDS) (Fleming and Kramer, 1995),

OASIS-B (Shaughnessy, et al. 1997) or EASY-Care (Philp, 2000). For a detailed review of some of these instruments see the recent report by the Lincoln Centre for Ageing and Community Care Research (2004).

- Mental health instruments which rate social, behavioural and psychological functioning in general. These included the Life Skills Profile (LSP) (Rosen, et al. 1989) and the Global Assessment of Functioning (GAF) Scale (American Psychiatric Association, 1994).
- Wider disability measures examining social functioning, like the London Handicap Scale (Harwood, et al. 1994).
- Disease specific measures in areas other than dementia. For example, the Functional Assessment Measure (FAM) (Hall, et al. 1996) for traumatic brain injury, or the Nottingham Extended Activities of Living Scale (EADL) (Nouri and Lincoln, 1987) for stroke, or the Functional Status Index (FSI) (Jette, 1980) for Arthritis.
- Dementia staging instruments which have functional elements, like the Clinical Dementia Rating Scale (CDR) or the Functional Assessment Staging (FAST) (Reisberg, 1998) (the FAST is related to the Global Deterioration Scale (GDS – Reisberg). These dementia staging and descriptive instruments have already been reviewed in Section 4 of this report.
- Global ratings like the Clinicians' Global Impression of Change (CGIC) (Guy, 1976) which included functional elements in the clinician's judgement. Burns, et al. (2004) provides a useful summary of the available measures. (Also excluded here were single item or two item measures of function).
- Performance based or timed instruments (both generic and dementia specific). These type of measures, while important both clinically (occupational therapy assessment) and for research purposes (an attempt to examine real world performance - ecological validity - of clinical or proxy ratings), were deemed too long and labour intensive (Burns, et al. 2004). They may require special equipment and thus these measures fell out of scope for this report. They are probably best regarded as specialist measures. Examples of dementia specific performance measures include: the Direct Assessment of Functional Status (Loewenstein, et al. 1989); the Structured Assessment of Individual Living Skills (SALES) (Mahurin, et al. 1991) and the Texas Functional Living Scale (TFLS) (Cullum, et al 2001). These also include generic measures such as: the Direct Assessment of Functional Abilities (DAFA) (Karagiois, et al. 1998); the Performance Test of Activities of Daily Living (PADL) (Kuriansky and Gurland, 1976); the Physical Performance Test (PPT) (Reuben and Siu, 1990); the Timed Manual Performance Test (TMP) (Williams, et al. 1994); and the Assessment of Motor and Process Skills (AMPS) (Merritt and Fisher, 2003). Other examples include the 6 minute walk distance test (Lord and Menz, 2002). An extensive review of these performance based measures is provided by Moore, et al. (2007).

As can be seen from this list of exclusions, this analysis has taken a traditional, as well as practical approach to the assessment of function by focusing on two key assessment types - direct observation measures (i.e. clinical rating) and proxy / informant report measures. This is in line with the framework outlined in Section 10.1.3.

Readers are referred to the measurement issues section of this report on (Refer Section 12) which addresses the challenges and issues concerning cognitive impairment and self-report, and proxy measurement. For example a paper by Talbert, et al. (2002) found that a subset of MCI patients overestimated their higher-level functional abilities and this apparent lack of awareness predicted future diagnosis of AD. A paper by Doble, et al. (1999) provides evidence of where family informants of AD patients over-estimated ADL functioning even when cognitive impairment was mild.

For performance based measures there are also a number of advantages and disadvantages. Performance based measures are seen as an alternative to self-report or proxy informant or clinical observations of function. Their chief characteristic and advantage is that they can break up

common activities or tasks into their objective parts and sequences, allowing for standardised assessment (e.g. cooking skills), and for the assessor to intervene with assistance or prompting. Gross and fine motor skills as well cognitive functioning are important aspects of these measures (Burns, et al. 2004). Pearson (2000) in Kane and Kane (2000) also points out that speed of functioning is usually a major component.

However, Spector (1997) has also outlined the some of the main problems with performance measures of function:

- The performance task is somewhat artificial and may not be the same as performance in the real world.
- Training is required to use the task and it may require special equipment.
- Performance tasks require a longer assessment time.
- Patients need to be able to respond to simple commands.
- The issue of motivation during the performance of the task.

### **10.2.2 Short-listed Instruments**

For ease of analysis, the remaining instruments were sorted into three tables: ADL only, IADL only and Combination instruments (as per section 10.1.3). These tables were also sorted into generic and dementia specific measures, and the basic features of each measure were also obtained e.g. number of items, estimated time to complete.

**Table 49 ADL Instruments**

Type	Name	Alternate names	Abbrev.	Source	Direct obs.	Self-report	Proxy / informant report	Number of items	Time to complete	Other elements
Generic	Katz Index of ADL	Katz Index of Independence in Activities of Daily Living, Index of Activities of Daily Living	ADL	Katz et al. (1963)	✓	✓		6	5 - 10 min	
Generic	Physical Self Maintenance Scale		PSMS	Lawton and Brody (1969)	✓	✓		6	5 - 10 min	
Generic	Rapid Disability Rating Scale-2		RDRS-2	Linn & Linn (1982)			✓	18	10 min	1 IADL item, plus items on disability and special problems
Generic	Resource Utilisation Groups – Activities of Daily Living		RUG-ADL	Buckingham et al. 1998	✓			4	5 min	measure of nursing dependency
Generic	Barthel Index + Modifications		BI	Mahoney & Barthel (1958)	✓	✓	✓	10	5 - 15 min	* <i>check version</i>
Generic	Functional Independence Measure		FIM	Granger & Hamilton (1987)	✓	✓	✓	18	20 - 30 min (in conference)	Includes communication and cognition items
Dementia Specific	Functional Dementia Scale		FDS	Moore et al. (1983)			✓	20	15 min	Includes items on orientation and behaviour

**Table 50 IADL Instruments**

Type	Name	Alternate names	Abbrev.		Direct observation	Self-report	Proxy / informant report	Number of items	Time to complete	Other elements
Generic	Instrumental Activities of Daily Living Scale		IADL	Lawton & Brody (1969)	✓	✓		8	5 min	
Generic	Older Americans Resources and Services Multi-Dimensional Functional Assessment Questionnaire - IADL		OARS, OMFAQ, OARS-IADL	Fillenbaum et al. (1988)	✓	✓		8	5 min	IADL component only
Generic	Functional Activities Questionnaire (Pfeffer)		FAQ, PFAQ	Pfeffer et al. (1982)			✓	10	10 min	Later versions include 4 ADL items + one item on initiation
Dementia Specific	Bayer Activities of Daily Living Scale		B-ADL	Hindmarch et al. (1998)			✓	25	20 min	Includes 4 ADL items - very general / overall level, Plus cognitive items

**Table 51 Combination Instruments**

Type	Name	Alternate names	Abbrev.		Direct observation	Self-report	Proxy / informant report	Number of items	Time to complete	Other elements
Generic	Nurses Observational Scale for Geriatric Patients		NOSGER	Spiegel et al. (1991)	✓			30	20 min	Plus memory, mood and behaviour items
Dementia Specific	Blessed Dementia Scale		Blessed	Blessed et al. (1968)	✓			22	15 min	Includes questions on changes in personality, interests, drive. Plus cognition items
Dementia Specific	Present Functioning Questionnaire		PFQ	Crockett et al. (1989)			✓	65	20 min	Includes cognitive items and personality items
Dementia Specific	Progressive Deterioration Scale		PDS	DeJong et al. (1989)			✓	27	90 min	Includes QoL items
Dementia Specific	Interview for Deterioration in Daily Living Activities in Dementia	Interview for Deterioration in Daily Functioning Activities	IDDD	Teunisse & Derix (1991)			✓	33	20 min	Includes items on initiation of activity eg. brushing hair or teeth
Dementia Specific	Disability Assessment for Dementia Scale		DAD	Gelinas et al. (1999)			✓	40	15 min	
Dementia Specific	Cleveland Scale for ADL		CSADL	Patterson et al. (1992)	✓			66	25 min	
Dementia Specific	Bristol Activities of Daily Living Scale			Bucks et al. (1996)			✓	20	15 min	Plus orientation and communication items
Dementia Specific	Alzheimer's Disease Functional Assessment and Change Scale		ADFACS	Mohs et al. 2001			✓	16	10 - 15 min	
Dementia Specific	Alzheimer's Disease Co-operative Study - Activities of Daily Living		ADCS-ADL	Galasko et al. (1997)			✓	24	20 min	

*NB: A proxy / informant measure can be based on either an interview or questionnaire format*

The 19 instruments in these three tables were then reduced to a short list of contender instruments based on the application of additional selection criteria (availability of instrument, number of citations, psychometric evidence, used in clinical practice, availability of normative and clinical reference data, administration time, able to be used with the various severity levels of dementia, cost considerations, and applicability for routine care). For instance, the Progressive Deterioration Scale (PDS) (De Jong, et al. 1989) was removed because of its length and the fact that it also contains quality of life items.

The short list of functional assessment instruments is presented below:

**Table 52 Short-listed Measurement of Function Instruments**

<b>Instrument</b>	<b>Original Article Cite Author(s) + Publication Date</b>
<b>Katz Index of ADL</b>	Katz, et al. (1963)
<b>Physical Self Maintenance Scale (PSMS)</b>	Lawton and Brody (1969)
<b>Resource Utilisation Groups – Activities of Daily Living (RUG-ADL)</b>	Buckingham et al. (1998)
<b>Barthel Index</b>	Mahoney and Barthel (1965)
<b>Functional Independence Measure (FIM)</b>	Hamilton, Granger, et al. (1987)
<b>IADL (Lawton and Brody)</b>	Lawton and Brody (1969)
<b>Older Americans Resources and Services – IADL (OARS – IADL)</b>	Fillenbaum (1988)
<b>Blessed Dementia Scale</b>	Blessed, et al. (1968)
<b>Interview for Deterioration in Daily Living Activities in Dementia (IDDD)</b>	Teunisse, Derix and Crevel (1991)
<b>Disability Assessment for Dementia Scale (DAD)</b>	Gelinas, et al. (1999)
<b>Cleveland Scale for ADL (CSADL)</b>	Patterson, et al. (1992)

Other scales which were found to be noteworthy in this area were newer instruments like the Activities of Daily Living Questionnaire (ADLQ) (Johnson, et al. 2004) a combination, informant / proxy based measure for the outpatient population with dementia and the Alzheimer's Disease Activities of Daily Living International Scale (ADL-IS) (Reisberg, et al. 2001) which is a 40 item combination measure which is relatively free from gender and culture bias (Burns, et al. 2004). As these newer instruments had few citations and less psychometric evidence available they were not included in the short list.

Two other short-listed instruments that were not selected for comprehensive review are worthy of further discussion. These are the Resource Utilisation Groups – Activities of Daily Living (RUG-ADL) scale and the Bayer Activities of Daily Living Scale (B-ADL).

The RUG-ADL (Buckingham et al. 1998) is part of NSW Health's Mental Health Outcomes and Assessment Tools (MH-OAT) initiative. It is used if an inpatient is 65 years or older or has a diagnosis of chronic organic brain syndrome. The RUG-ADL consists of four items: eating, bed mobility, transferring and toileting; and was designed to measure nursing dependency or the need for assistance in ADLs in skilled nursing facilities (total score range 4 – 18). The RUG-ADL is useful for casemix classification, and because of its ease of use it could also be considered for application in examining ADL dependency in residential care facilities. However, it does not address IADLs.

The Bayer Activities of Daily Living Scale (B-ADL) (Erzigkeit, et al. 2001) is another new informant based measure, with relatively few citations. Used in international studies, the scale was designed to measure the performance of everyday activities of living for community dwelling elders with mild cognitive impairment or mild to moderate dementia. It consists of 25 questions activities of everyday living (mainly IADL) and cognitive functioning (Burns, et al. 2004), using a numerical 10 point response scale, labelled from 'never' to 'always'. The B-ADL has high internal consistency (Cronbach's alpha = 0.98), and a uni-dimensional factor structure, as well as the ability to distinguish between Global Deterioration Scale stages 1 to 5. It was also found to have comparable or better discrimination than the MMSE for Global Deterioration Scale stages 4 and 5. However, further information is required on its construct validity in relation to other measures of function (including direct performance measures), and in relation to its undefined and ungraded numerical response format.

### 10.2.3 Reviewed Instruments

The short-listed instruments were now classified into the following four categories: Accept; Reject; Grey Area; Grey Area – but promising new instrument. The specific criteria used to select the instruments for comprehensive review are provided in Section 2.

In order to adequately cover developments in the field of functional measurement in dementia specific and elderly populations in general, it was decided to make recommendations incorporating generic and dementia specific measures. Of the dementia specific measures combination measures (ADL + IADL) covering a range of activities were preferred, as they had undergone the most development in the literature (see the number of dementia specific combination instruments vs. the dementia specific ADL and IADL only instruments in Table 52). It was also decided to include direct observation / clinical rating instruments and proxy / information report instruments to cover the range of practice settings.

In terms of the generic measures, the Functional Independence Measure (FIM) and the Barthel Index (using Collin, et al. 1988 scoring) for residential and community applications are the stand out measures for ADL. These have been comprehensively reviewed in the Continence Outcomes Measurement Suite (COMS) project (see Thomas, et al. 2006). In general ADL instruments will be used in acute care or nursing home settings, while IADL instruments are more commonly used in the community settings. For IADL, the Lawton and Brody IADL instrument and its adaptation the Older Americans Resources and Services (OARS-IADL) were chosen. They have been reviewed by (see Eagar, et al. 2001, 2006) and adapted for Australian practice settings (Green, et al. 2006).

A brief description of these instruments follows including their psychometric properties and application with people with dementia. These descriptions highlight the limited use of generic functional instruments in dementia or cognitively impaired older populations, as discussed in Section 10.1.5 above.

The FIM includes 18 items including independence in self-care, sphincter control, mobility, locomotion, communication and cognition, scored on a 7 point scale (from total assistance to complete independence). It has excellent psychometric properties, median test-retest reliability = 0.95 (Thomas, et al. 2006), and has been used in a number of patient groups and benchmarking studies. The FIM also has the advantage in dementia populations of incorporating cognitive as well as motor functioning items.

In terms of the dementia literature, only a few papers have used the FIM measures with dementia or elderly cognitively impaired people. Examining validity, Cotter, et al. (2002) provided information on the criterion validity of the FIM by comparing caregiver reports with performance data (correlation range for items = 0.62 – 0.91) in people with dementia. While Ruchinskas, et al. (2001) provides data comparing FIM scores with the clock drawing test, showing a relationship between cognitive ability and poorer physical ability at discharge from a geriatric rehabilitation unit. MacNeil and Litchenberg (1997) also used the FIM to predict rehabilitation outcomes using logistic regression, suggesting that a high FIM score on the social cognition item is related to independent living outcomes post discharge. In terms of clinical applications, Goldstein, et al. (1997) found improvements in FIM motor score improvements (moving patients from dependence to independence) for both cognitively impaired and unimpaired individuals, as a result of a geriatric rehabilitation program for hip fracture. While a study by Petracca, et al. (1996) using the drug Clomipramine with depressed and probable AD patients found that the drug improved mood levels but did not affect FIM scores. These papers suggest that the FIM can be applied to people with dementia to examine overall functional status.

The Barthel Index was the forerunner of the FIM, and includes items on feeding, transfers, toileting, grooming, bathing, dressing, bowel and bladder continence and mobility, scored on a 3 point scale. Like the FIM, the Barthel Index is used widely and Burns, et al. (2004) reports that the Barthel Index has excellent validity, reliability, sensitivity and clinical utility.

In terms of validity, a study by Wolstenholme, et al. (2002) into health care cost of persons with dementia shows that a one-point fall in the Barthel Index is associated with a dramatic increase in the costs of care. While Minicuci, et al. (2003) in Italy found that people with moderate to severe total dependency in ADL at admission were three times more likely to have died at discharge than those who were independent. This also applied at the one-year follow-up point. In terms of comparison studies, Silver, et al. (2001) compared the Barthel Index and CDR scores in the community finding to find a high correlation ( $r = -0.73$ ). While Ballard, et al. (2001) found significant correlations between low performance on activities of daily living and Dementia Care Mapping QoL indices in nursing homes. In terms of clinical application, Stone, et al. (1994) demonstrated the usefulness of the Barthel Index as an outcome measure, finding a median 6 point change from admission to discharge for a large group of patients ( $n = 102$ ) on an acute geriatric rehabilitation ward. Along with cognitive performance the Barthel Index was related to discharge destination for this study. Challis, et al. (2000) also showed using the Barthel Index that many low dependency patients were entering UK nursing homes rather than being supported at home, while Quartararo, et al. (1991) report using the Barthel in a geriatric assessment program for planning residential care needs. These papers highlight the application of the Barthel functioning instrument to the potential dementia population, especially in community and nursing home settings.

**Table 53 Summary of Ratings for the Generic Measurement of Function Instruments**

Criteria	Weight	ADL Instruments			IADL Instruments	
		FIM	Barthel Index	Katz	OARS-IADL	Lawton & Brody IADL
Availability of comparison data	3	3	2	1	3	2
Length/feasibility of instrument for inclusion in battery	2	2	3	3	3	3
Complexity of administration/cognitive burden	2	3	3	2	3	2
Cultural Appropriateness	1	3	3	3	3	3
Ease of obtaining score	2	3	3	3	3	3
Sensitivity to dementia	3	1	1	2	2	2
Reliability evidence	3	3	2	1	3	2
Validity evidence	3	3	2	1	2	2
Cost of the instrument	2	2	3	3	3	3
Cost of instrument administration	2	1	1	1	1	1
<b>Weighted Total</b>		<b>55</b>	<b>50</b>	<b>42</b>	<b>59</b>	<b>51</b>

The Lawton and Brody IADL instrument consists of 8 items reflecting higher order activities (ability to use telephone, shopping, food preparation, housekeeping, laundry, mode of transportation, responsibility for own medications, ability to handle finances). Each item is scored on a 3, 4, or 5 point scale, with higher scores indicating greater severity. Pearson (2000), Burns, et al. (2004) and McDowell (2006) report high inter-rater reliability in range of 0.85 to 0.94. Expected significant validity coefficients were also reported for ADLs, mental status tests, behaviour and physical health measures. McDowell (2006) and Pearson (2000) report that the instrument items form a Guttman type scale for women and three items need to be eliminated for men (food preparation, laundry and housekeeping). The authors acknowledge that the scale is associated with gender role stereotypes (Burns, et al. 2004). It has been recommended for use with community dwelling older persons (Pearson, 2000).

Some studies have used the Lawton and Brody IADL instrument for persons with dementia. In terms of comparative validity, Lechowski, et al. (2005) in France used the IADL instrument to examine the nature of functional decline for AD patients living in the community, showing a relationship with cognition (ADAS-Cog, MMSE) and functional staging (CDR). While Farias, et al. (2003) found that neuropsychological test performance accounted for 25% of the variance in IADL scores in an AD sample. A study by Green, et al. (1993) investigated reliability and functional decline using the IADL instrument in a longitudinal study (31 months). They reported adequate reliability and suggested the need for better IADL items for men. Green, et al. (1993) also found that IADL scores decline on average about 2 points per year (with smaller changes for those people with severe dementia).

Older Americans Resources and Services (OARS-IADL) is an adaptation of the Lawton and Brody IADL instrument (Pearson, 2000), containing seven items (telephone, transportation, shopping, meal preparation, housework, medication management, money management) with a core three point response format: without help, with help or unable. (Note: Fillenbaum, 1985 has also created a 5 item screening instrument - using the items on transportation, shopping, meal preparation, housework, money management). Pearson (2000) reports adequate test-retest reliability ( $r = 0.71$ ) and validity correlations with physical and mental health (SF-20). While McDowell (2006) also reports adequate internal consistency (Cronbach's  $\alpha = 0.68$ ) and a correlation of 0.33 with the Katz ADL scale. Normative data has also been provided for US populations (McDowell, 2006). Pearson (2000) recommends this measure as a screening tool for determining need for services by community dwelling adults. Eagar, et al. (2006) and Green, et al. (2006) have adapted this measure as a screening assessment for Australian conditions.

Some relevant studies using the OARS-IADL in the literature include, comparative papers by Doble, et al. (1997) and Rogers, et al. (1994) that show that the OARS-IADL correlates significantly with performance based measures (e.g. AMPS) in AD patients. While Reuben, et al. (1995) reports on the correlation between the OARS-IADL with the modified Katz ( $r = 0.33$ ) and the self-administered SF-36 Physical Functioning scale (PF-10) ( $r = 0.36$ ) for a group of community based older persons. In terms of validity, the relationship between functional decline as measured by OARS items and cognitive impairment (3MS) has been investigated in a very large study ( $n = 5874$ ) over 5 years by Njegovan, et al. (2001), demonstrating that IADLs were lost on the whole, before ADLs at higher cognitive performance levels. In terms of clinical application, the OARS- IADL has been used in studies examining APOE epsilon4 allele (see Blazer, et al. 2001).

In summary, the Functional Independence Measure (FIM) and the Barthel Index (for community applications), the Lawton and Brody IADL instrument and its adaptation the Older Americans Resources and Services (OARS-IADL) are the recommended measures for ADL and IADL respectively. The above analysis shows that these instruments have superior psychometric properties in general and some limited application in dementia populations. However a thorough work-up of their psychometric properties with people with dementia is recommended.

\*\*\*

In terms of dementia specific instruments, the Blessed Dementia Scale and the Cleveland Scale for Activities of Daily Living (CSADL) were chosen for review as they were the only two direct observation, and ADL + IADL combination, instruments in the short list. While the Alzheimer's Disease Co-operative Study – ADL (ADCS-ADL) and Disability Assessment for Dementia Scale (DAD) were chosen for review as they were the two dementia specific proxy report instruments and because of their comparatively large number of citations in the literature. These instruments are described below.

#### **10.2.3.1 Blessed Dementia Scale (BDS)**

The Blessed Dementia Scale (BDS) is a direct observation measure containing 22 ADL + IADL items. It is a highly influential scale, if perhaps a little dated. For instance, the Blessed functional section tends to combine functioning with cognitive impairment asking about a tendency to dwell in the past, the inability to recall recent events, and the inability to remember a short list of items (Spector, 1997). Also the item asking to discriminate between people in a hospital (e.g. patients, doctors and nurses) (Burns, et al. 2004) needs refinement when applied to different practice settings. Further information on the Blessed Dementia Scale can be found in Section 4 of this report. It has adequate internal consistency reliability (Cronbach's  $\alpha = 0.66$ ) and test-retest reliability was reported to be in the range of 0.79 – 0.88 over a four week period. Finally, in terms of validity, Stern, et al. (1990) caution against relying on the BDS to detect functional change in persons with dementia since disparate functional domains are assessed. For these reasons it was not considered further for the assessment of function.

### 10.2.3.2 Cleveland Scale for ADL (CSADL)

The Cleveland Scale for ADL (CSADL) is a direct observation measure containing 66 brief items across 16 ADL + IADL domains. It includes items from the OARS and the Functional Activities Questionnaire of Pfeffer (Burns, et al. 2004). The response codes range from 0 – 3 (fully independent to completely independent). The CSADL is a promising new scale for the proxy measurement of function in people with dementia. The three papers, Patterson et al. (1992), Patterson and Mack (2001), and Mack and Patterson (2006) chart the rigorous psychometric work-up of the scale, including information on internal consistency (Cronbach's alpha = 0.97) and inter-rater reliability (0.84 to 0.99 for items), correlation with other well-known dementia measures (MMSE and Blessed Dementia Scale), factor structure, item difficulty and discriminant validity. However, further work is required concerning the areas of test – retest reliability and its' sensitivity to change. The CSADL also needs to be compared to other dementia specific measures of function (including performance and timed tests) in order to better gauge validity (especially criterion related validity). Finally, further studies are required to replicate the findings of the CSADL development team.

### 10.2.3.3 Disability Assessment for Dementia Scale (DAD)

The Disability Assessment for Dementia Scale (DAD) is a proxy / informant report instrument containing 17 ADL and 23 IADL items. The instrument can be administered via questionnaire or interview. Burns, et al. (2004) reports a high degree of reliability (test-retest = 0.96 [ICC for one week], inter-rater [0.95 with two raters] and internal consistency [Cronbach's alpha = 0.96]). While Schneider (2001) reports that the DAD has been used in a number of clinical trials of cholinesterase inhibitors and has been shown to measure functional decline at 12 months. Similar findings are reported by recent papers by Feldman, et al. (2001, 2003) and by Behl, et al. (2006). An advantage of the DAD is that it looks, not only at impaired activities, but at specific aspects of performance (e.g. initiation, planning and organisation, and effective performance).

In summary, the DAD is a logically developed and reliable measure for assessing functional disability in persons with dementia and has been used in a number of clinical trials. While shown to be sensitive to change in a number of studies, further development work is needed in the areas of internal structure (i.e. factor analysis of the whole scale, and the value of the cognitive component sub-domains) and construct / criterion validity. In particular, this informant / proxy rating inventory needs to be compared to other dementia specific measures of function, clinical rating scales and the direct assessments of function (e.g. performance and timed tests).

### 10.2.3.4 Alzheimer's Disease Co-operative Study – Activities of Daily Living (ADCS-ADL)

The Alzheimer's Disease Co-operative Study – Activities of Daily Living (ADCS-ADL) is another proxy /informant report instrument. It consists of 24 ADL + IADL items attempted during the past 4 weeks (depending on the version used), derived from clinician derived pool of 45 items. The ADCS-ADL has good test-retest reliability (0.91 over a four week period) and has been shown to distinguish between the stages of dementia severity (Schneider, 2001) and change at 12 months (Burns, et al. 2004).

The ADCS-ADL is a psychometrically well designed instrument for measuring decline in functional performance in clinical trials, where it has been used extensively (see for example, Galatamine – Galasko, et al. 2004, Brodaty, et al. 2005, Loy and Schneider, 2006 [Cochrane Review]; Memantine – Feldman, et al. 2006, Peskind, et al. 2006; APOE epsilon4 allele – Farlow, et al. 2004). While shown to be sensitive to change in a number of studies, further development work is needed on the ADCS-ADL in the areas of inter-rater reliability and construct / criterion validity. Like other measures reviewed in this section, the ADCS-ADL needs to be compared to other dementia specific measures of function, clinical rating scales and the direct assessments of function (e.g.

performance and timed tests). A wider application beyond clinical drug trials to other settings like hospitals and nursing homes is also required. These studies should also examine the performance of the ADCS-ADL in people with dementia with other co-morbid conditions.

### 10.2.4 Instrument Rankings – Summary Rating

Considering the attributes of each test as reviewed in the appendix, a scoring judgement was made according to the criteria outlined, and this information is summarised in the table below.

**Table 54 Summary of Ratings for the Measurement of Dementia Specific Function Instruments**

Criteria	Weight	Instruments		
		DAD	ADCS-ADL	CS-ADL
Availability of comparison data	3	3	3	2
Length/feasibility of instrument for inclusion in battery	2	1	2	2
Complexity of administration/ cognitive burden	2	3	3	3
Cultural Appropriateness	1	3	3	3
Ease of obtaining score	2	3	2	3
Sensitivity to dementia	3	3	3	3
Reliability evidence available	3	3	2	2
Validity evidence available	3	2	2	2
Cost of the instrument	2	3	3	3
Cost of instrument administration	2	2	2	2
<b>Weighted Total</b>		<b>60</b>	<b>57</b>	<b>56</b>

This analysis finds mostly high ratings for the selected dementia specific and combination (ADL + IADL) functional instruments in terms of their administration, scoring and psychometric properties.

It should be noted that, in terms of clinical practice, it is important to monitor functional status as people with dementia progress from the mild stages of the disease through to the severe stages of the disease. This has a bearing on drug treatment strategies and residential or care planning options; and in these situations informant measures of people with dementia living in the community, like the DAD and ADCS-ADL, may be the most suitable instruments to pick up these changes.

### 10.2.5 Recommendations

As indicated earlier the Functional Independence Measure (FIM), the Barthel Index and the Lawton and Brody IADL and the Older Americans Resources and Services (OARS-IADL)

instruments were chosen as generic measures of ADL and IADL respectively. These instruments have been reviewed recently (Eagar, et al. 2001, 2006; Thomas, et al. 2006) and have been shown to have good psychometric properties and have been used in geriatric settings. They represent the “industry standard” for ADL and IADL measurement in Australia. However, further research into their application with dementia populations is required.

With regard to the activities of daily living the FIM is probably more appropriate for acute care and high level residential care settings but it is noted that accredited training is required for its use. However, it is already widely used in acute care rehabilitation settings within Australia. The Barthel Index is an easier to use measure and may be more appropriate for use in primary and community care settings.

Although the Katz ADL instrument has been quite widely used in dementia settings the review of this instrument by Thomas, et al. (2006) indicated it has weak psychometric properties and thus it is not recommended for use. Although it is not reviewed again here a comparative assessment with other instruments is included in Table 53.

The Lawton and Brody IADL and the Older Americans Resources and Services (OARS-IADL) are recommended as generic instruments for the assessment of instrumental activities of daily living (IADL). The OARS-ADL is preferred as it is an advance on the Lawton and Brody IADL scale with improved psychometric properties and less reliance on gender role stereotypes; and it has been adapted for use in primary and community care settings in Australia (see Green, et al. 2006).

The recommended dementia specific instruments for the assessment of function (ADL and IADL) for people with dementia include both proxy measures and measures of direct observation. The Alzheimer’s Disease Co-operative Study – ADL (ADCS-ADL) and Disability Assessment for Dementia Scale (DAD) are the two proxy report instruments that are recommended. For the direct observation of functioning the Cleveland Scale for Activities of Daily Living (CSADL) is recommended.

The recommended dementia specific instruments for the assessment of function (ADL and IADL) for people with dementia include both proxy measures and clinical rating scales. While it is acknowledged that proxy reports have their limitations (refer Section 12), they will generally be used where assessment by interview or self rating is no longer possible due to the degree of cognitive impairment of the person with dementia. Proxy measures are also extremely useful in primary and community care settings in order to monitor the maintenance of functional status or its decline, in conjunction with drug therapy or in terms of care management as the disease progresses. The direct observation rating scale may be more appropriate for acute care and residential care settings. By recommending both proxy and direct observation rating scales different practice settings and clinical situations (e.g. a person with dementia may not have a carer) can be addressed.

From this discussion, which has highlighted a number of measurement problems with regard to the assessment of function of people with dementia it is clear there is an urgent need for a program of research and development in this area. In the absence of a research consensus for the measurement of function in dementia, and given a high degree of overlap in items, there is a clear need for a streamlining the various functional instruments and items (Spector, 1997) across each of the practice settings (Spector, 1997). The work of Lindeboom, et al. (2003) in the Amsterdam Liner Disability Score Project using IRT to calibrate ADL instruments in neurology could be used as a guide. A similar study with a large group of dementia patients could examine and calibrate functional items from the short-listed instruments (both generic and dementia specific) to create a comprehensive item bank. This dementia item bank could then be used to examine item redundancy and coverage across the range of severity levels and could be used to develop new tools or provide cross-calibration between the existing instruments. This project would also need to

examine the relationship of these items with recommended cognitive and functional assessment staging instruments.

## References

- American Psychiatric Association (1994) *Diagnostic and Statistical Manual of Mental Disorders*. American Psychiatric Association, Washington DC.
- Ayis SA, Bowling A, Gooberman-Hill R, et al. (2007) *The effect of definitions of activities of daily living on estimates of changing ability among older people*. International Journal of Rehabilitation Research. Vol. 30, No.1, pp.39-46.
- Ballard C, O'Brien J, James I, et al. (2001) *Quality of life for people with dementia living in residential and nursing home care: the impact of performance on activities of daily living, behavioral and psychological symptoms, language skills, and psychotropic drugs* International Psychogeriatrics. Vol. 13, No.1, pp.93-106.
- Behl P, Lanctot KL, Streiner DL, et al. (2006) *Cholinesterase inhibitors slow decline in executive functions, rather than memory, in Alzheimer's disease: a 1 year observational study in the Sunnybrook dementia cohort*. Current Alzheimer Research. Vol. 3, No.2, pp.147-156.
- Blazer DG, Fillenbaum GG and Burchett B (2001) *The APOE-E4 allele and the risk of functional decline in a community sample of African American and white older adults*. Journals of Gerontology Series A - Biological Sciences and Medical Sciences. Vol. 56, No.12, pp.M785-M789.
- Blessed G, Tomlinson BE and Roth M (1968) *The association between quantitative measures of dementia and of senile change in the cerebral grey matter of elderly subjects*. British Journal of Psychiatry. Vol. 114, pp.797-811.
- Bootsma-van der Wiel A, Gussekloo J, de Craen AJM, et al. (2001) *Disability in the oldest old: "Can do" or "Do do"*. JAGS. Vol. 49, No.7, pp.909-914.
- Brodaty H, Corey-Bloom J, Potocnik FCV, et al. (2005) *Galantamine prolonged-release formulation in the treatment of mild to moderate Alzheimer's disease*. Dementia and Geriatric Cognitive Disorders. Vol. 20, No.2-3, pp.120-132.
- Buckingham W, Burgess P, Solomon S, Pirkis J and Eagar K. (1998) *Developing a Casemix Classification for Mental Health Services*. Volume 1: Main Report. Canberra: Commonwealth Department of Health and Family Services.
- Burns A, Lawlor B and Craig S (2004) *Assessment scales in old age psychiatry (2nd ed.)* Taylor and Francis, London.
- Challis D, Mozley CG, Sutcliffe C, et al. (2000) *Dependency in older people recently admitted to care homes*. Age and Ageing. Vol. 29, No.3, pp.255-260.
- Chong D (1995) *Measurement of Instrumental Activities of Daily Living in Stroke*. Stroke. Vol. 26, pp. 1119 - 1122.
- Collin C, Wade DT, Davis S, et al. (1988) *The Barthel ADL Index: A reliability study*. International Disability Studies. Vol. 10, pp.61-63.
- Coster WJ, Haley SM, Andres PL, et al. (2004) *Refining the conceptual basis for rehabilitation outcome measurement: Personal care and instrumental activities domain*. Medical Care. Vol. 42, No.1, pp.162-172.
- Cotter EM, Burgio LD, Stevens AB, et al. (2002) *Correspondence of the functional independence measure (FIM) self-care subscale with real-time observations of dementia patients' ADL performance in the home*. Clinical Rehabilitation. Vol. 16, No.1, pp.36-45.

- Cullum C, Saine K, Chan LD, et al. (2001) *Performance-based instrument to assess functional capacity in dementia: The Texas Functional Living Scale*. Neuropsychiatry, Neuropsychology and Behavioral Neurology. Vol. 14, No.2, pp.103-108.
- deJong R, Osterlund OW and Roy GW (1989) *Measurement of quality-of-life changes in patients with Alzheimer's disease*. Clinical Therapeutics. Vol. 11, pp.545-554.
- Doble SE, Fisk JD, MacPherson KM, et al. (1997) *Measuring functional competence in older persons with Alzheimer's disease*. International Psychogeriatrics. Vol. 9, No.1, pp.25-38.
- Doble SE, Fisk JD and Rockwood K (1999) *Assessing the ADL functioning of persons with Alzheimer's Disease: comparison of family informants' ratings and performance-based assessment findings*. International Psychogeriatrics. Vol. 11, No.4, pp.399-409.
- Eagar K, Owen A, Cromwell D, et al. (2001) *Towards a national measure of functional dependency for home and community care services in Australia - Stage 1 report of the HACCC Dependency Data Items Project*. Centre for Health Service Development, University of Wollongong.
- Eagar K, Owen A, Marosszeky N, et al. (2006) *Towards a measure of function for home and community care services in Australia: Part 1 - Development of a Standard National Approach*. Australian Journal of Primary Health. Vol. 12, No.1, pp.73-81.
- Erzigkeit H, Lehfeld H, Pena-Casanova J, Bieber F, Yekrangi-Hartman C, Rupp M, Rappard F, Arnold K, and Hindmarch I (2001) *The Bayer-Activities of Daily Living Scale (B-ADL): Results from a validation study in three European countries*. Dementia and Geriatric Cognitive Disorders. Vol. 12, pp. 348-358.
- Farias ST, Harrell E, Neumann C, et al. (2003) *The relationship between neuropsychological performance and daily functioning in individuals with Alzheimer's disease: ecological validity of neuropsychological tests*. Archives of Clinical Neuropsychology. Vol. 18, No.6, pp.655-672.
- Farlow MR, He Y, Tekin S, et al. (2004) *Impact of APOE in mild cognitive impairment*. Neurology. Vol. 63, No.10, pp.1898-1901.
- Feldman HH, Gauthier S, Hecker J, et al. (2003) *Efficacy of donepezil on maintenance of activities of daily living in patients with moderate to severe Alzheimer's disease and the effect on caregiver burden*. JAGS. Vol. 51, No.6, pp.737-744.
- Feldman HH, Sauter A, Donald A, et al. (2001) *The disability assessment for dementia scale: a 12-month study of functional ability in mild to moderate severity Alzheimer disease*. Alzheimer Disease and Associated Disorders. Vol. 15, No.2, pp.89-95.
- Feldman HH, Schmitt FA, Olin JT, et al. (2006) *Activities of daily living in moderate-to-severe Alzheimer disease: An analysis of the treatment effects of memantine in patients receiving stable donepezil treatment*. Alzheimer Disease and Associated Disorders. Vol. 20, No.4, pp.263-268.
- Fillenbaum GG (1985) *Screening the elderly: A brief instrumental activities of daily living measure*. JAGS. Vol. 33, No.10, pp.698-706.
- Fillenbaum GG (1988) *Multidimensional functional assessment of older adults: the Duke Older Americans Resources and Services procedures*. Lawrence Erlbaum Associates, Hillsdale, NJ.
- Fleming R and Kramer T (1995) *Dementia Care Planning: individual, integrated and effective. A national action plan for dementia project*. Hammond Care Group.
- Galasko D, Bennett D, Sano M, et al. (1997) *An inventory to assess activities of daily living for clinical trials in Alzheimer's Disease*. Alzheimer Disease and Associated Disorders. Vol. 11, No.Suppl 2, pp.S33-39.
- Galasko D, Kershaw PR, Schneider LS, et al. (2004) *Galantamine maintains ability to perform activities of daily living in patients with Alzheimer's disease*. JAGS. Vol. 52, No.7, pp.1070-1076.

- Gelinas I, Gauthier L, McIntyre M, et al. (1999) *Development of a functional measure for persons with Alzheimer's Disease: The Disability Assessment for Dementia*. American Journal of Occupational Therapy. Vol. 53, pp.471-481.
- Gill T, Williams CS, Richardson E, et al. (1997) *A predictive model for ADL dependence in community-living older adults based on a reduced set of cognitive status items*. JAGS. Vol. 45, No.4, pp.441-445.
- Goldstein FC, Strasser DC, Woodard JL, et al. (1997) *Functional outcome of cognitively impaired hip fracture patients on a geriatric rehabilitation unit*. JAGS. Vol. 45, No.1, pp.35-42.
- Green CR, Mohs RC, Schmeidler J, et al. (1993) *Functional decline in Alzheimer's disease: a longitudinal study*. JAGS. Vol. 41, No.6, pp.654-661.
- Green J, Eagar K, Owen A, et al. (2006) *Towards a measure of function for home and community care services in Australia: Part 2 - Evaluation of the Screening Tool and Assessment Instruments*. Australian Journal of Primary Health. Vol. 12, No.1, pp.82-90.
- Guy W (ed) (1976) *Clinical Global Impressions (CGI)*. ECDEU Assessment Manual for Psychopharmacology, revised (DHEW Publ No ADM 76-338.) National Institute for Mental Health, Rockville, MD, pp. 218-222.
- Hall KM, Mann N, High WM, et al. (1996) *Functional measures after traumatic brain injury: ceiling effects of FIM, FIM+FAM, DRS, and CIQ*. J Head Trauma Rehabil. Vol. 11, No.5, pp.27-39.
- Hamilton BB, Granger CV, Sherwin FS, et al. (1987) *A uniform national data system for medical rehabilitation*. In Fuhrer M (ed.) Rehabilitation outcomes: analysis and measurement. P.H. Brookes, Baltimore. pp.137-147.
- Harwood RH, Gompertz P and Ebrahim S (1994) *Handicap one year after stroke: validity of a new scale*. Journal of Neurology, Neurosurgery and Psychiatry. Vol. 57, No.7, pp.825-829.
- Hawes C, Morris JN, Phillips CD, et al. (1997) *Development of the Nursing Home Resident Assessment Instrument in the USA*. Age and Ageing. Vol. 26, No.Suppl 2, pp.19-25.
- Heinemann AW (2005) *Putting outcome measurement in context: A rehabilitation psychology perspective*. Rehabilitation Psychology. Vol. 50, No.1, pp.6-14.
- Hindmarch I, Leffeld H, de Jongh P, et al. (1998) *The Bayer Activities of Daily Living Scale (B-ADL)*. Dementia and Geriatric Cognitive Disorders. Vol. 9, No.Suppl 2, pp.20-26.
- Jette AM (1980) *Functional capacity evaluation: an empirical approach*. Archives of Physical Medicine and Rehabilitation. Vol. 61, pp.85-89.
- Jette AM, Keysor J, Coster WJ, et al. (2005) *Beyond function: predicting participation in a rehabilitation cohort*. Archives of Physical Medicine and Rehabilitation. Vol. 86, pp.2087-2094.
- Johnson N, Barion A, Rademaker A, et al. (2004) *The Activities of Daily Living Questionnaire: A validation study in patients with dementia*. Alzheimer Disease and Associated Disorders. Vol. 18, No.4, pp.223-230.
- Karagiozis H, Gray S, Sacco J, et al. (1998) *The Direct Assessment of Functional Abilities (DAFA): A comparison to an indirect measure of instrumental activities of daily living*. Gerontologist. Vol. 38, No.1, pp.113-121.
- Katz S, Ford AB, Moskowitz RW, et al. (1963) *Studies of illness in the aged. The index of ADL: A standardized measure of biological and psychological function*. JAMA. Vol. 185, No.12, pp.914-919.
- Kuriansky JB and Gurland B (1976) *Performance test of activities of daily living*. International Journal of Ageing and Human Development. Vol. 7, pp.343-352.
- Lawton M and Brody E (1969) *Assessment of older people: Self-maintaining and instrumental activities in daily living*. Gerontologist. Vol. 9, pp.179-186.

- Lechowski L, De Stampa M, Tortrat D, et al. (2005) *Predictive factors of rate of loss of autonomy in Alzheimer's disease patients. A prospective study of the REAL.FR cohort.* Journal of Nutrition, Health and Ageing. Vol. 9, No.2, pp.100-104.
- Li M, Ng TP, Kua EH, et al. (2006) *Brief Informant Screening Test for Mild Cognitive Impairment and Early Alzheimer's Disease.* Dementia and Geriatric Cognitive Disorders. Vol. 21, No.5-6, pp.392-402.
- Lincoln Centre for Ageing and Community Care Research (2004) *The review and identification of an existing, validated, comprehensive assessment tool.* Australian Institute for Primary Care, La Trobe University.
- Lindeboom R, Vermeulen M, Holman R, et al. (2003) *Activities of daily living instruments: Optimizing scales for neurologic assessments.* Neurology. Vol. 60, pp.738-742.
- Linn M and Linn B (1982) *The Rapid Disability Rating Scale - 2.* JAGS. Vol. 30, No.6, pp.378-382.
- Loewenstein DA, Amigo E, Duara R, et al. (1989) *A new scale for the assessment of functional status in Alzheimer's disease and related disorders.* Journal of Gerontology. Vol. 44, No.4, pp.114-121.
- Lord SR and Menz HB (2002) *Physiologic, psychologic, and health predictors of 6-minute walk performance in older people.* Archives of Physical Medicine and Rehabilitation. Vol. 83, No.7, pp.907-911.
- Loy C and Schneider LS (2006) *Galantamine for Alzheimer's disease and mild cognitive impairment.* Cochrane Database of Systematic Reviews. Vol. 1, CD001747.
- Mack JL and Patterson MB (2006) *An empirical basis for domains in the analysis of dependency in the activities of daily living (ADL): results of a confirmatory factor analysis of the Cleveland Scale for Activities of Daily Living (CSADL).* Clinical Neuropsychologist. Vol. 20, No.4, pp.662-677.
- MacNeil SE and Lichtenberg PA (1997) *Home alone: the role of cognition in return to independent living.* Archives of Physical Medicine and Rehabilitation. Vol. 78, No.7, pp.755-758.
- Mahoney FI and Barthel DW (1965) *Functional evaluation: the Barthel Index.* Maryland State Medical Journal. Vol. 14, pp.61-65.
- Mahurin RK, DeBettignies BH and Pirozzolo FJ (1991) *Structured assessment of independent living skills: preliminary report of a performance measure of functional abilities in dementia.* Journal of Gerontology. Vol. 46, No.2, pp.P58-P66.
- McDowell I (2006) *Measuring health: A guide to rating scales and questionnaires.* Oxford University Press, Oxford.
- McKeith I, Cummings J, Lovestone S, Harvey R and Wilkinson D (1999) *Outcome measures in Alzheimer's Disease.* Martin Dunitz, London.
- Merritt BK and Fisher AG (2003) *Differences between men and women in ADL performance.* Archives of Physical Medicine and Rehabilitation. Vol. 84, pp.1872-1877.
- Minicuci N, Maggi S, Noale M, et al. (2003) *Predicting mortality in older patients: The Velca Study.* Ageing-Clinical and Experimental Research. Vol. 15, No.4, pp.328-335.
- Mohs RC, Doody RS, Morris JC, et al (2001) *A 1-year, placebo-controlled preservation of function survival study of donepezil in AD patients* Neurology. Vol. 57, No.3, pp.481-488.
- Moore DJ, Palmer BW, Patterson TL, et al. (2007) *A review of performance-based measures of functional living skills.* Journal of Psychiatric Research. Vol. 41, pp.97-118.
- Moore JT, Bobula JA, Short TB, et al. (1983) *A functional dementia scale.* The Journal of Family Practice. Vol. 16, pp.499-503.

- 
- Neugebauer A, Katz PP and Pasch LA (2003) *Effect of valued activity disability, social comparisons, and satisfaction with ability on depressive symptoms in rheumatoid arthritis*. Health Psychology. Vol. 22, No.3, pp.253-262.
- Njegovan V, Hing MM, Mitchell SL, et al. (2001) *The hierarchy of functional loss associated with cognitive decline in older persons*. Journals of Gerontology Series A - Biological Sciences and Medical Sciences. Vol. 56, No.10, pp.M638-M643.
- Patterson MB, Mack JL, Neundorfer MM, et al. (1992) *Assessment of functional ability in Alzheimer disease: A review and a preliminary report on the Cleveland Scale for Activities of Daily Living* Alzheimer Disease and Associated Disorders. Vol. 6, pp.145-163.
- Patterson MB and Mack JL (2001) *The Cleveland Scale for Activities of Daily Living (CSADL): Its reliability and validity*. Journal of Clinical Geropsychology. Vol. 7, No.1, pp.15-28.
- Pearson VI (2000) *Assessment of function in older adults*. In Kane R, Kane R, (eds.) Assessing older persons: measures, meaning and practical applications. Oxford University Press, USA.
- Peskind ER, Potkin SG, Pomara N, et al. (2006) *Memantine treatment in mild to moderate Alzheimer disease: a 24-week randomized, controlled trial*. American Journal of Geriatric Psychiatry. Vol. 14, No.8, pp.704-715.
- Petracca G, Teson A, Chemerinski E, et al. (1996) *A double-blind placebo-controlled study of clomipramine in depressed patients with Alzheimer's disease*. Journal of Neuropsychiatry and Clinical Neurosciences. Vol. 8, No.3, pp.270-275.
- Pfeffer RI, Kurosaki TT, Harrah CH, et al. (1982) *Measurement of functional activities in older adults in the community*. Journal of Gerontology. Vol. 37, pp.323-329.
- Philp I (2000) *EASY-Care: A systematic approach to the assessment of older people*. Geriatric Medicine. Vol. 30, No.5, pp.15-19.
- Quartararo M, O'Neill TJ, Tang C, et al. (1991) *Assessing the residential care needs of nursing home applicants*. Australian Journal of Public Health. Vol. 15, No.3, pp.222-227.
- Reisberg B (1988) *Functional assessment staging (FAST)*. Psychopharmacology Bulletin. Vol. 24, pp.653-659.
- Reisberg B, Finkel S, Overall J, et al. (2001) *The Alzheimer's Disease Activities of Daily Living International Scale (ADL-IS)*. International Psychogeriatrics. Vol. 13, No.2, pp.163-181.
- Reuben DB and Siu AL (1990) *An objective measure of physical function of elderly outpatients: The Physical Performance Test*. JAGS. Vol. 38, No.10, pp.1105-1112.
- Reuben DB, Valle LA, Hays RD, et al. (1995) *Measuring physical function in community dwelling older persons: A comparison of self-administered, interview-administered, and performance-based measures*. JAGS. Vol. 43, No.1, pp.17-23.
- Rogers JC, Holm MB, Goldstein G, et al. (1994) *Stability and change in functional assessment of patients with geropsychiatric disorders*. American Journal of Occupational Therapy. Vol. 48, No.10, pp.914-918.
- Rockwood K (2007) *The measuring, meaning and importance of activities of daily living (ADLs) as an outcome*. International Psychogeriatrics. Vol. 19, No.3, pp.467-482.
- Ruchinkas RA, Singer HK and Repetz NK (2001) *Clock drawing, clock copying, and physical abilities in geriatric rehabilitation*. Archives of Physical Medicine and Rehabilitation. Vol. 82, No.7, pp.920-924.
- Schneider LS (2001) *Assessing outcomes in Alzheimer Disease*. Alzheimer Disease and Associated Disorders. Vol. 15, No.Suppl 1, pp.S8-S18.
-

- 
- Shaughnessy PW, Crisler KS and Schlenker RE (1997) *Medicare's OASIS: Standardized Outcome and Assessment Information Set for Home Health Care, OASIS-B*. Center for Health Services and Policy Research, University of Colorado Health Sciences Center.
- Silver MH, Jilinskaia E and Perls TT (2001) *Cognitive functional status of age-confirmed centenarians in a population-based study*. Journals of Gerontology Series B - Psychological Sciences and Social Sciences. Vol. 56, No.3, pp.P134-P140.
- Spector WD (1997) *Measuring functioning in daily activities for persons with dementia*. Alzheimer Disease and Associated Disorders. Vol. 11, No.Suppl 6, pp.81-90.
- Spiegel R, Brunner C, Ermini-Funfschilling D, et al. (1991) *A new behavioural assessment scale for geriatric out- and in-patients: the NOSGER (Nurses' Observation Scale for Geriatric Patients)*. JAGS. Vol. 39, pp.339-347.
- Stern Y, Hesdorffer D, Sano M and Mayeux R (1990) *Measurement and prediction of functional capacity in Alzheimer's disease*. Neurology. Vol. 40, No. 1, pp.8-14
- Stone SP, Ali B, Auberleek I, et al. (1994) *The Barthel Index in clinical practice: use on a rehabilitation ward for elderly people*. Journal of the Royal College of Physicians of London. Vol. 28, No.5, pp.419-423.
- Tabert MH, Albert SM, Borukhova-Milov L, et al. (2002) *Functional deficits in patients with mild cognitive impairment: Prediction of AD*. Neurology. Vol. 58, pp.758-764.
- Teunisse S, Derix MMA and van Crevel H (1991) *Assessing the severity of dementia: patient and caregiver*. Archive of Neurology. Vol. 48, pp.274-277.
- Thomas S, Nay R, Moore K, et al. (2006) *Continence Outcomes Measurement Suite Project*. Department of Health and Ageing, Canberra.
- Thomas VS, Rockwood K and McDowell I (1998) *Multidimensionality in Instrumental and Basic Activities of Daily Living*. J Clin Epidemiol. Vol. 51, No.4, pp.315-321.
- Williams M, Gaylord S and Gerrity M (1994) *The Timed Manual Performance test as a predictor of hospitalization and death in a community-based elderly population*. JAGS. Vol. 42, No.1, pp.21-27.
- Wolstenholme J, Fenn P, Gray A, et al. (2002) *Estimating the relationship between disease progression and cost of care in dementia*. British Journal of Psychiatry. Vol. 181, pp.36-42.

## 11 Measures of Patient and Carer Satisfaction

### 11.1 Patient Satisfaction

Patient satisfaction has increased in popularity due to three changes in health care. First, the role of clinicians has changed from one of helping patients through their illness to where the clinician is expected to either cure the patient or alleviate chronic symptoms. Second, the rise of the patients' rights movement, which presents patients as consumers of health care, has led to patient views being taken into account during medical decision-making. Third, patient perspectives are increasingly sought for inclusion in the monitoring of health care and the legitimizing of health policy. This paper takes the position that health care recipients are patients rather than consumers because (a) most patients in Australia are not fully informed consumers and (b) this review is concerned with their personal health care satisfaction.

#### 11.1.1 Defining Patient Satisfaction

This brief section is based on the fuller treatment of patient satisfaction theories presented in Hawthorne (2006).

Patient satisfaction was initially perceived as being related to issues around access to medical infrastructure (Bashshur, et al. 1967; Caplan and Sussman, 1966; James, 1967; Rouse, 1967; Rowbotham, 1953) and nursing care (Abdellah and Levine, 1957). Donabedian, with a focus on quality assessment, also saw it as arising from the medical infrastructure (the quality of amenities), but extended it to include the technical health outcomes from the treatment process and process quality which focussed on personal relationships within health care systems (Donabedian, 1980; Donabedian, 1988). There are thus three general dimensions to patient satisfaction: the amenities where care is provided, the interpersonal relationships between the patient and health care providers, and the technical competence of the care provider.

This position was reiterated by Wilson and Goldschmidt (1995) who separated patient satisfaction from patient outcomes on the basis that patient satisfaction was concerned only with the interpersonal aspects of health care. The corollary was that patient satisfaction could be used as a surrogate indicator enabling the incorporation of the patient perspective into the process of monitoring and improving health care services or, even more broadly, as an evaluative tool in the formulation of social and health policy (Sitzia and Wood, 1997; Williams, 1994) — even though patient satisfaction per se is not an evaluation of medical care (Locker and Dunt, 1978).

The major patient satisfaction theories were all published during the 1980s; almost all research since then is based on these. Ware, et al. (1983) argued that patient satisfaction was a function of patients' subjective responses to experienced care mediated by personal preferences and expectations. Linder-Pelz (1982) postulated it was mediated by personal beliefs and values about care as well as prior expectations of the care. Fox and Storms (1981) advocated that a person's orientation determined satisfaction; dissatisfaction, therefore, occurred where there was transgression of the relationship between expectation and experience (Jackson and Kroenke, 2001; Ross, et al. 1987; Thompson and Sunol, 1995; Zebiene, et al. 2004). Fitzpatrick and Hopkins (1983) argued that expectations were socially mediated, reflecting the health goals of the patient and the extent to which illness and health care violated the patient's personal sense of self. As mentioned above, Donabedian (1980, 1988) postulated it was based on personal relationships within health care systems and health care outcomes from treatment, where these were mediated by the values of the patient.

Patient satisfaction is defined here as the patient's judgement on the quality of care, particularly the interpersonal relationships with clinicians and other care providers (Donabedian, 1988, p1746):

*Patient satisfaction may be considered to be one of the desired outcomes of care, even an element in health status itself. An expression of satisfaction or dissatisfaction is also the patient's judgement on the quality of care in all its aspects, but particularly as concerns the interpersonal process.*

The implication is that the construct of patient satisfaction covers all aspects of care quality, particularly the interpersonal processes. Patient dissatisfaction will occur where there are a cluster of small transgressions of these dimensions or a major failure in service provision. Hawthorne (2006) argued that the key dimensions of this construct were:

1. Appropriate access to health services, including the environment within which treatment takes place and the level of care coordination (Fox and Storms, 1981; Hardy, et al. 1996; Ware, et al. 1983);
2. The provision of health information which helps to set patient expectations (Donabedian, 1980; Fox and Storms, 1981; Hardy, et al. 1996; Hawthorne and Harmer, 2000; Newsome and Wright, 1999; Suchman, 1965; Thompson and Sunol, 1995);
3. The relationship between the patient and health care staff, specifically empathy with the patient (Ben-Sira, 1976; Donabedian, 1988; Fitzpatrick, 1990; Hardy, et al. 1996; Kane, et al. 1997; Sitzia and Wood, 1997; Ware, et al. 1983);
4. Participation in making choices regarding health treatment, including the associated fears and sense of loss of control as well as the appropriate use of treatment therapies and medications (Hardy, et al. 1996);
5. Satisfaction with the treatment provided, i.e. the technical quality of the care provided (Fox and Storms, 1981; Hardy, et al. 1996; Hawthorne and Harmer, 2000; Kane, et al. 1997; Linder-Pelz and Struening, 1985; Ware, et al. 1983);
6. The effectiveness of treatment, including the extent to which treatment helps the patient in his/her daily life (Donabedian, 1988; Hardy, et al. 1996; Hawthorne and Harmer, 2000; Ware, et al. 1983); and
7. A general satisfaction with the experience of health care.

This model of patient satisfaction postulates that in a comprehensive assessment of patient satisfaction all seven dimensions will contribute and should be measured.

The rise in the use of patient satisfaction measures has also been justified on health outcome grounds, despite the lack of evidence showing that these measures are widely used in routine clinical practice or that they actually influence day-to-day clinical practice (Greenhalgh and Meadows, 1999). Reasons for this rise include that interventions with higher patient satisfaction outcomes are to be preferred, that satisfied patients are more likely to seek medical care, have greater compliance with treatment, continuing relationships with the clinician and have better health outcomes (Baker, 1990; Chung, et al. 2002; Fitzpatrick, 1991; Hara, et al. 2003; Hardy, et al. 1996; Larsen and Rootman, 1976; Lindsey, et al. 2002).

Additionally, when reading this report it should be kept in mind that much of the research on patient satisfaction is from the USA. Given that the structure, operation and financing of the health care system in the USA is systematically different to that of many other countries, caution needs to be exercised in assuming generalisability of theories, instruments and studies. At least six recent studies have shown that it cannot be assumed that a patient satisfaction scale developed in one country is directly transferable to another country without modification (Firestone, et al. 2004; Henderson, et al. 2003; Labarere, et al. 2004; Meakin and Weinman, 2002; Zebiene, et al. 2004), and Baker, et al. (2003) reported that patient satisfaction was lower among US patients when compared with UK patients. This suggests there are cultural nuances in how satisfaction is understood, assessed and reported.

### 11.1.2 Method

This report is based on Hawthorne, et al.'s patient satisfaction reviews (Hawthorne, 2006; Hawthorne, et al. 2006). These comprehensive reviews were carried out as part of the National Continence Management Strategy. The literature searches for those reviews were carried out in 2004; for this report these searches have been updated in two ways. First, the original searches were replicated for the period 2005/6. Second, these searches were supplemented with dementia-specific searches as described below.

The original searches of papers discussing or reporting patient satisfaction were carried out in MEDLINE /Pub Med and the internet using the terms *patient, client, or consumer* and crossing these with any of the terms *satisfaction, questionnaires, instrument, measurement or theory*. In addition separate searches were made of the terms *patient/client/consumer, satisfaction* and *theory/instrument*.

Using the term 'patient satisfaction', 38,193 articles were identified through MEDLINE /Pub Med and over 10,000 websites in January 2004. Refinement of the search terms as described above led to the identification of 858 unique articles or reports and a further 126 websites. Abstracts (or, in the case of internet sites, first paragraphs) were reviewed. Based on an assessment of the contribution of the paper to the literature in a way not made elsewhere or providing a particularly good illustration of an issue of interest, unique articles and reports were obtained for close reading or critique. The reference lists were scanned for additional papers of interest. These were then extracted in turn. Altogether 130 unique articles were extracted and reviewed in the Hawthorne, et al. reports.

These searches have been supplemented for the present report. Replication of the original search identified an additional 3885 articles published between 2005 and July 2007 of which 223 were concerned with patient satisfaction theory or new measures. Following abstract review, 8 were extracted for detailed review. Altogether, 87 articles were identified by crossing the terms *patient satisfaction* with *dementia, Alzheimer's disease* and *cognitive impairment*. Following review, 12 were extracted for review. No dementia-specific instruments were identified.

In selecting instruments and papers for review, the following criteria were used.

1. Patient satisfaction theories had to be original and developed as a generic theory;
2. Patient satisfaction instruments had to be formal self-report instruments;
3. Patient satisfaction instruments had to be generic, that is designed for use in all health conditions, with all patients and across studies or research settings;
4. Instruments had to be concerned with patient satisfaction assessment at the intervention level rather than the health care system level;
5. Papers must have been published in English;
6. They must have been accessible through the academic press or internet; and
7. Basic psychometric data must have been reported.

Theories, instruments and papers were excluded on the following grounds:

- Qualitative approaches to the assessment of patient satisfaction were excluded since this review was concerned with formal measurement. For example, Aggarwal's, et al. (2003) exploration of personal experiences in residential care was excluded on these grounds.
- Theories were excluded if they were elaborations of earlier theories, or if they had been developed for use with a specific condition (other than dementia or Alzheimer's disease). The review revealed that most of the modern theories or models of patient satisfaction are either restatements of earlier theories or have been developed for specific health conditions. An important observation is that there are multiple competing theories, many of which have not been fully explicated (Aspinal, et al. 2003). For example, Hudak's (2004) embodiment theory of patient satisfaction for those with hand surgery was excluded, as was Aragon's (Aragon and

Gesell, 2003) primary provider theory of patient satisfaction for use in hospital emergency departments.

- Instruments measuring constructs other than patient satisfaction or that measured just one aspect of patient satisfaction were excluded. For example, the EUROPEP (Grol and Wensing, 2000; Grol, et al. 1999) was excluded because it was designed to assess patients' views of their medical care, not satisfaction with that care. This exclusion draws the distinction between patients' cognitive awareness of care and their satisfaction with that care. Likewise the Patient Perception of Hospital Experience with Nursing (Dozier, et al. 2001) was excluded because it specifically focussed on nursing care within hospital settings. Also excluded were Elzinga and Barlow's (1991) adaptation of MacDonald, et al's (1988) patient satisfaction questionnaire providing assessment of the physical and social conditions of long-stay wards in psychiatric hospitals because many of the items are about constraints on the patients (e.g. being able to make a cup of tea when wanting to).
- Instruments designed for use in single studies or for specific medical procedures were excluded. Thus the Surgery Satisfaction Questionnaire (Baker, 1991) was excluded. Likewise, Kane, et al's (1997) patient satisfaction measure was excluded because it was developed for a single study and a single type of medical procedure.
- Instruments designed for use with specific groups of patients were excluded. For example, MacDonald, et al's (1988) patient satisfaction questionnaire for long-stay psychiatric patients.
- Instruments designed for use with patients suffering specific conditions were excluded. Thus the Verona Service Satisfaction Scale (Ruggeri and Dall'Agnola, 1993), while enjoying widespread international support (particularly across Europe), was excluded because it is concerned with the assessment of health care just for those with mental health conditions. Similarly the Psychiatry Outpatient Consult Clinic Patient Care Survey was also excluded (Camara, 1991). Margolis, et al's (2003) patient satisfaction scale was excluded because it was thought to be culturally limited (e.g. that males and females are separated in a medical clinic). Similarly, Westaway, et al's (2003) diabetes patient satisfaction scale was also excluded.
- Scales that were embedded within instruments were also excluded. For example, the Treatment Outcome Profile (TOP; Holcomb, et al. 1998) is a 27-item measure designed for use with psychiatric patients to assess changes in quality of life, symptomatology and functioning level. A fourth scale of 9 items, embedded within the instrument, measures patient satisfaction with services (the effectiveness of treatment, perceived competence of staff and the treatment environment). Another example was the HIV/AIDS satisfaction questionnaire developed by Beck, et al. (1999).
- Finally, scales that were culturally inappropriate for the Australian health care system were excluded. For example, the 10-item Older Patients Satisfaction Scale (OPSS) was excluded on this ground. This American scale was developed in the early 1980s to assess satisfaction with the (then) new American health maintenance organizations (HMOs) because older adults at the time were underrepresented in HMO subscribers (Cryns, et al. 1989). Although well constructed, the item content reflects this background; thus there are items assessing getting more health care from the enrolled HMO than from any other health plan, that the HMO covered more services than other health plans, and that hospital admission was pre-arranged.

The exception to these exclusion rules was in relation to the SAPS (the Short Assessment of Patient Satisfaction instrument) (Hawthorne, et al. 2006). The SAPS is a generic patient satisfaction instrument that was developed for inclusion in the Australian National Continence Management Strategy because of identified issues with other leading patient satisfaction instruments. The inclusion of the SAPS in this review is in the interests of national uniformity.

After applying the exclusion criteria above, the instruments for review were:

- Single item assessments

- 
- The Client Satisfaction Questionnaire (CSQ-18 and CSQ-8);
  - The Consultation Satisfaction Questionnaire (CSQ, described here as the ConsultSQ);
  - The La Monica-Oberst patient satisfaction scale (LOPSS);
  - The Linder-Pelz satisfaction scales;
  - The Medical Interview Satisfaction Scale (MISS);
  - The Patient Satisfaction Index (PSI);
  - The Patient Satisfaction Questionnaire (PSQ);
  - The Patient Visit Rating Questionnaire (PVRQ);
  - Gonzalez, et al's patient satisfaction questionnaire;
  - Inpatient Evaluation of Service Questionnaire (IESO); and
  - The Short Assessment of Patient Satisfaction instrument (SAPS).

A common feature of most of the articles reviewed was that although researchers reported patient satisfaction estimates, the actual measures used are very poorly reported, if at all. It is usually reported in a single sentence, where it is offered as evidence complementing treatment success. Few papers report either the instruments used, their psychometric properties in the study samples, or the actual results. This situation applied regardless of whether the paper was written in the 1980s or since 2000; it applied to single-item measures as well as to formal instruments.

The psychometric properties of instruments used were particularly poorly reported — a finding consistent with that of Sitzia (1999) who found, based on a review of 195 patient satisfaction papers published in 1994, that less than half reported any psychometric data, yet that 81% reported using a new patient satisfaction instrument and a further 10% reported modifying a previously existing instrument. He reported that most of the study instruments reviewed had little evidence of reliability or validity, and that of those papers reporting a new instrument, 60% reported no psychometric data whatsoever. For example, Johannsson, et al. (2002) in a long-term follow-up study of haemorrhoidectomy reported patient satisfaction and examined the relationship between continuing faecal symptoms and patient satisfaction. Yet nowhere in their paper was the patient satisfaction measure described, nor was there any reference to it so that it could be tracked down. As reported by Sitzia (1999), this situation is unacceptable research practice.

A second consistent finding across the literature reviewed is that most people are satisfied with their health care. Typically, between 70-90% of patients report satisfaction, even when there is evidence of continuing health problems. The reasons for this are primarily to do with health literacy, the unequal relationship between patients and their clinicians, instrument administration and bias, and that most people are satisfied with their lives generally. It may also be that patients' initial expectations of the health care system are lower than their actual experience, thus causing a high level of satisfaction (the expectancy disconfirmation theory) (Newsome and Wright, 1999; Thompson and Sunol, 1995).

The third key finding is that most studies report patient satisfaction based on a single item, such as How satisfied are you with your health care? These kinds of items are short, quick and easy to administer. They are widely used in clinical settings because they are easy to understand and interpret immediately, and they are frequently used by clinicians as discussion starters with patients. Almost no research, however, has been undertaken regarding their psychometric properties, and, since patients are usually in a dependent relationship with their clinician when responding to such questions, the value of the responses is extremely suspect.



The number of points provided on the continuum may vary from none to 10 or more, or use a visual analogue scale (VAS). Fox and Storms (1981, p.560) assessed patient satisfaction on a scale of 10-points with a single question:

*If a score of ten represents the best possible medical care available and one represents a very poor quality of medical care, how would you rate the medical care you have received in the past year?*

Alternately, the patient may be asked to respond to a question with categorical answers, like this:<sup>35</sup>

*How satisfied are you with your health care?*

- Very satisfied*
- Satisfied*
- Not sure*
- Dissatisfied*
- Very dissatisfied*

Two recent examples of a very simple categorical patient satisfaction measure is that reported by Grumbach, et al. (1999) who used a global satisfaction question with the five response categories listed above in a study of patient satisfaction with their primary clinician (the GP). The results showed that 82% were satisfied.

Similarly, Hawthorne, et al. (2006) reported an almost identical item (How satisfied are you with the outcome of your treatment?) with the same response set. Unlike other researchers they reported the psychometric properties of their item, viz., variation by treatment type, differences between known groups, an absence of differential item functioning by known group, and sensitivity to a then-test of health status.

A third form of single question is where patients are asked would they either (a) have the procedure again, or (b) would they recommend the procedure to their friends. Positive responses are interpreted as indicators of satisfaction. However, invariably most patients state that they would have the procedure again even when the medical procedure involves considerable health losses including complications (Everaert, et al. 2000).

#### **11.1.4.2 Client Satisfaction Questionnaire (CSQ-18 and CSQ-8)**

The American CSQ-18 (18 items) developed by Larsen, et al. (1979) was intended to measure satisfaction with services. From the literature, 9 service dimensions were identified and 9 items written for each dimension. These were then assessed by health professionals and the resulting pool of 31 items administered to 248 mental health patients. Factor analysis revealed a single factor accounting for 75% of the common variance (Nguyen, et al. 1983). A shorter version, the CSQ-8 (8 items), is also available; it was developed through removal of items from the CSQ-18 where the criteria for removal was those items with the lowest internal consistency properties; thus the CSQ-8 is a more homogenous scale. The CSQ-18 (CSQ-8 items are marked below with an asterisk) consists of 18 items measuring the promptness of being seen, the comfort and attractiveness of the facility and building, the amount of help received (\*), the appropriateness of the help given, the helpfulness of the services (\*), how well the patient was listened to, whether the patient received the service(s) he/she wanted (\*), whether there were other services the patient

<sup>35</sup> Where an item has a number of categories from which the respondent endorses the one that best describes him/her, like those shown, these categories are referred to as a 'response set'.

wanted but did not receive, how clearly the patient was understood, the competence of the clinician, rating the quality of service received (\*), overall satisfaction with services received (\*), recommending the service to a friend (\*), being understood, having needs met (\*), having rights respected and returning to the service (\*). Each item is scored on a 4-point scale, where the responses cover a poor service through to an excellent service. Scoring of the CSQ-18 and CSQ-8 is by simple summation. For the CSQ-18 the score range is 18-72 and for the CSQ-8 it is 8-32.

The psychometric properties of the CSQ-18 and CSQ-8 were reported by Attkisson and Zwick (1982) and Nguyen, et al. (1983). The Attkisson and Zwick sample was a sub-set of 45 cases who completed the CSQ-18 as part of a larger trial (n=62) at an urban community mental health centre where participants were randomly assigned to treatment or control group. The treatment group viewed a videotape on pre-therapy orientation whereas the controls were admitted normally. Follow-up was at one month. Data for the CSQ-8 was extracted and properties of both the CSQ-18 and CSQ-8 reported. Cronbach  $\alpha$  for the CSQ-18 was 0.91 and 0.93 for the CSQ-8. Regarding predictive validity, this was assessed through correlation with three service use variables during the month following administration. For remaining in treatment (Yes/No) the CSQ-18 correlation was 0.61 (CSQ-8: 0.57), for the number of sessions it was 0.54 (0.56) and for the proportion of sessions missed it was 0.06 (0.01). Correlations between various outcome measures (various symptom measures as assessed by both the patient and the clinician) ranged from -0.01 to -0.35 (0.01 to -0.40). In general, the findings showed that greater CSQ satisfaction ratings were associated with more sessions attended and with greater symptom reduction, but not with current symptomatology.

The Nguyen, et al. split-half reliability study sample equivalent forms analysis was based on 34 cases, where the CSQ-18 was randomly split into two scales; the correlation was  $r = 0.82$ . In a further analysis, based on 44 cases, where the CSQ-18 was administered in written and oral modes, the oral mode produced scores that were 10% higher (more satisfied) (Nguyen, et al. 1983). Also reported in the Nguyen, et al. paper was a study of the CSQ-8 involving 49 cases with 4-week follow-up; the Cronbach  $\alpha$  was 0.92 and scores were correlated with self-reported clinical improvement scores ( $r = 0.53$ ), and that in a further study (n=3,120) the CSQ-8 mean score was 27.09 (sd = 4.01) with an  $\alpha = 0.87$ . Although these reports suggest good psychometric properties, they also suggest that the CSQ is subject to differences in administration. The implication is that in a study where, say, the CSQ was administered, post-treatment, orally to the treatment group and self-completed by the control group, any differences favouring the treatment group could be due to the difference in administration mode and not the new treatment.

Pang, et al. (2003) in a study of concomitant tension-free vaginal tape insertion during pelvic floor reconstruction surgery follow-up at 1-year post-operation reported on the Chinese version of the CSQ-8. Regarding data distribution across the items, only 10% of all responses involved the third and fourth level (poor service) of the response scale, across all items. The implication is that half the response scale was redundant and the responses on the CSQ-8 items were essentially dichotomous. In a Costa Rican study of diabetes, Firestone, et al. (2004) reported psychometric limitations (the original article was unable to be extracted, so no further details are available). Hilton, et al. (2001) used a shortened version of the CSQ-18 (through removal of 9 items, leaving just 9 items in the version used) and reported a Cronbach  $\alpha = 0.78$ . They also commented that there was positive skew on item responses and recommended a different method of assessing dissatisfaction was needed.

Roberts, et al. (1983) directly compared the CSQ-18 with the Patient Satisfaction Questionnaire (PSQ, see below) in a study of 148 public health patients. The two measures were shown to assess different aspects of patient satisfaction. While the CSQ-18 provided information that was orientated towards service planning and monitoring, the data from the PSQ was more highly correlated with global life satisfaction and well-being rather than with the specific health care services used. Generally, the CSQ-18 scores were significantly higher than those obtained on the PSQ.

#### 11.1.4.3 Consultation Satisfaction Questionnaire (CSQ, described here as the ConsultSQ)

Based on a literature review and iterative consultation with clinicians and patients, the British ConsultSQ assesses patient's satisfaction with a consultation with a general practitioner (Baker, 1990). From the review and consultation an item bank was developed and administered to patients in a surgery following a consultation. After further modification, it was re-issued. This procedure was iteratively followed and the bank progressively modified as more data about the performance of items within the bank became available. Following iteration, factor and correlation analyses were used to discard further items and refine the final form of the ConsultSQ.

The ConsultSQ comprises 18 items located in four scales: general satisfaction (3 items); professional care (7 items describing the patient's concerns, the provision of information, treatment by the doctor, agreement with the doctor's advice, and the doctor treating the person as a whole); depth of relationship (5 items measuring the doctor's intimate knowledge of the patient and the transmission of personal information to the doctor); and perceived time (3 items measuring the length of the consultation in relation to the patient's perceived needs). A limitation of the ConsultSQ is that there are no items assessing treatment effects.

Items were written as attitude statements, such as *I am totally satisfied with my visit to this doctor*, and comprised both positive and negative statements. Responses were 5-point Likert scales. Scoring is by simple summation following reversal of negative items. Following administration to 239 patients, the psychometric properties were examined. Internal reliability of the ConsultSQ was Cronbach  $\alpha = 0.91$ , and for the scale professional care it was 0.87, for depth of relationship 0.83, for perceived time 0.82 and for general satisfaction 0.67. Spearman correlations between the general satisfaction scale and other scales were 0.50 for depth of relationship and 0.64 for professional care, suggesting these were measuring different, but related, constructs.

The ConsultSQ was assessed in a trial comparing those who changed doctor ( $n=272$ ) with those who did not ( $n=711$ ) (Baker and Whitfield, 1992). The results showed ConsultSQ scores systematically varied as predicted, from which the authors concluded that the ConsultSQ possessed validity because changing doctors is a strong statement of dissatisfaction.

Kinnersley, et al. (1996) compared the ConsultSQ with the Medical Interview Satisfaction Scale (MISS, see below) in a sample of 198 patients attending GP surgeries. The findings were that there was very little difference in the psychometric properties of the two measures. The correlation between the ConsultSQ and MISS was 0.82, suggesting they were measuring the same latent concept. For the ConsultSQ the mean score was 72% (sd = 12.6) of the scale range. Correlations between sub-scales ranged from 0.40 to 0.79 for the ConsultSQ. The reliability of the ConsultSQ scales were Cronbach  $\alpha = 0.73$  to 0.94. In a study examining the competence of medical students, McKinley, et al. (2004) reported that the correlation between consultation assessment and the ConsultSQ scales ranged from 0.16 to 0.44; they suggested it should not be used for assessing medical students. The mean scores on the ConsultSQ scales ranged from 37-69% of the scale ranges.

#### 11.1.4.4 La Monica-Oberst Patient Satisfaction Scale (LOPSS)

Developed using factor analysis by La Monica, et al. (1986) for measuring satisfaction in oncology in the USA, the LOPSS originally consisted of 48 items. It was revised by Munro, et al. (1994) through the removal of redundant items, making it more suitable for general health care satisfaction assessment. The standard version has 28 items, although Vahey, et al. (2004) reported using a 21-item version and O'Connell, et al. (1999) an 18-item version.

The LOPSS measures interpersonal support (9 items), good impressions (5 items) and dissatisfaction with nursing care (14 items). A typical item is *In general, the nurse seems more*

*interested in completing tasks than in listening to concerns.* Responses are on a 5-point Likert scale, from strongly agree to strongly disagree, although Vahey, et al. (2004) used a forced choice 4-point response scale. Scoring is by summation after reversal, giving a range of 28 through 140 where the highest scores reflect the greatest satisfaction with nursing.

Regarding the LOPSS's psychometric properties, the factor structure (3 sub-scales) reported by La Monica, et al. was confirmed by Munro, et al. (1994). Likewise, reliability was reported by La Monica, et al. (1986) at Cronbach  $\alpha = 0.98$  and the revised version at 0.97 (Munro, et al. 1994). O'Connell, et al. (1999) reported that Cronbach  $\alpha = 0.96$  and Vahey, et al. (2004) reported 0.93. Munro, et al. reported that the mean score was 118.7 (sd = 17.3).

O'Connell, et al. (1999) investigated the psychometric properties of the LOPSS in a sample of 105 surgical patients who were questioned about their nursing care. The mean score was 115.7 (sd = 17.4). When LOPSS scores were examined by presumed correlates (age, gender, length of stay) of satisfaction, no significant differences were observed. Telephone interviews revealed that dissatisfaction with several aspects of care did not appear to be reflected in instrument scores, leading to the conclusion that the LOPSS items were too insensitive and that the measure may be prone to acquiescent response bias.

#### **11.1.4.5 Linder-Pelz Satisfaction Scales**

Linder-Pelz (1982) developed scales to test the expectancy hypothesis arising from her work on the theory of patient satisfaction in Israel. Three scales were developed: the Doctor Conduct (DCS), General Satisfaction (GSS) and Convenience scales (CS). The 10 items in DCS were all negative in tone, e.g. *Doctor should have told me how to care for condition*, and measured condition care, being thorough, showing interest, doing the patient a favour, explaining the medical problem, having better clinical equipment, ordering tests, making the patient feel foolish, ignoring previous medical problems and the patient liking more time with the doctor. Six of the 7 items in the GSS were all positive, e.g. *My questions were answered to my complete satisfaction*. It measured answering questions, the doctor understanding the patient, the patient being satisfied with the visit, understanding the medical condition better, receiving better medical care than most people, the doctor being one of the best and not wanting to see the same doctor again. The Convenience Scale had 4 items measuring easy getting to the clinic, the waiting area being comfortable, how the staff treated the patient and having to wait too long. Scoring of the scales was through simple summation.

Validation of the scales was through administration to all first-time patients in a medical centre in Upper Manhattan (n=125) following a session with the doctor. Cronbach  $\alpha$  for each scale was 0.81 (DCS), 0.77 (GSS) and 0.49 (CS). No other psychometric properties were reported. Other than Linder-Pelz's own work, there appears to have been no further psychometric work on her scales.

#### **11.1.4.6 Medical Interview Satisfaction Scale (MISS)**

The MISS was developed in the USA to measure patient satisfaction with a clinical consultation (Wolf, et al. 1978; Wolf and Stiles, 1981). Originally the MISS consisted of 29 items, however recent work in the UK (Meakin and Weinman, 2002) has suggested a more coherent structure with 21 items nominally organised in four scales (the MISS-21). The scales measure distress relief (told what the trouble is, how serious the illness is, how long before getting better, worries relieved, and that the clinician knew what to do); communication comfort (uncertain, embarrassed, not allowed to say what the patient wanted, and clinician did not understand); rapport (clinician interested in patient, clinician warm and friendly, clinician treated problems seriously, patient felt free to talk about private matters, patient given chance to say what was on his/her mind, being understood by the clinician, feeling trust in the clinician, and the clinician knew what he/she was doing); and compliance intent (easy to follow clinician's advice, difficult to follow clinician's advice, and not sure

if worth the trouble of following clinician's advice). A typical item is *the physician told me the name of my disease in words I could understand*.

In addition to revision of the instrument length, a further difficulty is that different research teams have used different response scales. The original item response scales were 7-point Likert scales, but 5-point Likert scales have been used (Zebiene, et al. 2004). Scoring is through summation.

Several studies have reported its psychometric properties. Wolf, et al. (1978; Wolf and Stiles, 1981) reported that the internal consistency of the construction sample was  $\alpha = 0.93$  and that the four original scales explained 40% of the variance. In the Kinnersley, et al. (1996) study reported above the MISS mean score was 76.6% (sd = 11.4) of the scale range; and the reliability of the MISS scales was Cronbach  $\alpha = 0.78$  to 0.96.

Zebiene, et al. (2004) administered the MISS to 460 cases, and examined the internal properties. They reported that factor analysis revealed eleven factors explaining 62% of the variance. Of these there were four substantive factors explaining 42% of the variance. They labelled these emotional support, understanding and explanation, information and diagnosis, and treatment. The Cronbach  $\alpha$  for each of the four substantive factors was 0.88 (emotion support), 0.85 (understanding), 0.70 (information) and 0.66 (diagnosis and treatment). Although Zebiene, et al. reported similar psychometric properties to those obtained by Wolf, et al. (1978; Wolf and Stiles, 1981), they also reported that 5 items did not contribute to the scale, perhaps because of differences in expectations in Lithuania.

Meakin and Weinman (2002) administered the MISS to 150 patients and examined its internal structure. They obtained a five factor solution, which they then constrained to four factors through the removal of 8 items which failed to load on any factor  $>0.40$ , or which cross-loaded. The correlations between the four scales were from 0.46 to 0.65, perhaps suggesting the presence of an underlying single construct. Internal consistency was examined for each of the four scales, and the Cronbach  $\alpha$  range was 0.67 to 0.92.

Finally, several studies have suggested that the MISS is culturally specific and that cultural adaptation is needed prior to it being used in other countries than the US (Meakin and Weinman, 2002; Zebiene, et al. 2004).

#### **11.1.4.7 Patient Satisfaction Index (PSI)**

The Canadian PSI was designed to discriminate between patients with a life-threatening illness who were satisfied with their medical care and those who were not (Guyatt, et al. 1995). An initial item bank was assembled from the literature, patient interviews, and interviews with family members and health care providers. Following item review, three parallel questionnaires were constructed and interviewer-administered to 102 patients and 153 relatives. Preliminary scales were developed based on the most frequent and important items. These were then re-administered at two week interval along with other items measuring satisfaction. Logical criteria were used to sort items into dimensions, and 8 domains were identified. Health care providers were asked to verify the domains and sort items into the domains. The pattern of responses suggested that three different scales were needed: one for patients (the PSI); one for relatives of competent patients; and one for relatives of incompetent patients. Only the PSI is reviewed here.

The PSI comprises 23 items measuring gone through a lot, decisions made without involving patients, went through more than expected, felt helpless in decision-making, felt out of control of the situation, wanted decisions made by clinicians, feeling overwhelmed, involved in decisions too late, didn't understand what was happening, problems not clearly explained, not firm enough about wishes, options explained, co-operation from clinician, understood by clinician, understood clinician, family involved, respected by clinician, received appropriate care level, decision choices available, comfortable with decision-making, sharing same goals as clinician, clinicians clarify

wishes and feel clinicians care. Item response scales are Guttman-type 7-point scales where 7 indicates the highest level of satisfaction and 1 the lowest. Scoring is by summation, providing a range of scores from 23 to 161. The PSI is designed for interviewer administration, and takes 20-30 minutes to administer.

The PSI properties were assessed through administration to a sample of 105 patients, and re-test was carried out with 97 patients. All data were collected through interview. The intraclass correlation coefficient for the test-retest was 0.86. Correlations with other patient satisfaction measures were in the range 0.67 to 0.75, correlations with health care provider (mainly nurses) were 0.19, and with relatives' estimates were 0.28.

#### **11.1.4.8 Patient Satisfaction Questionnaire (PSQ)**

The American PSQ (Ware, et al. 1983) was derived from a review of the literature and the responses of convenience samples of patients. All the statements written for the PSQ were based on classic attitude measurement theory such that they expressed an opinion, e.g. *I'm very satisfied with the medical care I receive*. Altogether 2,300 statements were prepared and submitted to a panel of judges, who reduced the item pool to about 500 items. The item pool was then administered to patients in four different patient groups and various psychometric tests, including factor analysis, used to eliminate most items. A reduced item pool of 87 items was then administered to fresh samples and, the PSQ-II constructed based on both logical and statistical criteria. The PSQ-II had 68 items; further revision led to the PSQ-III with 51 items. The items are presented 7 dimensions of satisfaction covering: Access to care (emergency care (3 items), convenience of services (2) and access (2 items)); Financial aspects (cost of care (4 items), payment mechanisms (4) and insurance coverage (3)); Availability of resources (family doctors (2 items), specialists (2) and hospitals (2)); Continuity of care (family (2 items) and self (2)); Technical quality (quality competence (9 items), prudence-risks (2) and doctor's facilities (2)); Interpersonal manner (explanations (3 items), consideration (5), prudence-expenses (2)); and Overall satisfaction (4 items). Scoring of the PSQ-III scales is by simple addition of items within scales.

Tests of reliability (Cronbach  $\alpha$  and test-retest) on the original construction sample ranged from 0.77 to 0.88 (Hays, et al. 1987). However, subsequent reliability tests produced more varied results in that the reliability of PSQ sub-scales ranged from 0.23 (Prudence – risks) to 0.93 (Availability – hospitals) (Ware, et al. 1983). Overall, 78% of the PSQ sub-scales obtained reliability estimates  $>0.50$ , which was the standard adopted from Helmstadter (Helmstadter, 1964; Ware, et al. 1983). An abbreviated version of the PSQ was used by Ross, et al. (1993) in a study of patient preferences; they reported that the internal consistency of six scales (Access to care, Availability of services, Technical quality of care, Inter-personal care, Communication and Financing of care) ranged from 0.79 to 0.91.

Regarding the validity of the PSQ, Ware, et al. (1983) argued that because the internal structure of the PSQ was replicated across their many field trials, validity was implied. They also assessed PSQ scales against respondents' concerns, and reported that the scales performed as expected (e.g. those who complained of technical deficiency obtained lower scores on the Technical Quality sub-scale). When assessed against single-item measures of satisfaction, the PSQ behaved as expected and discriminated between groups; although Ware, et al. did report that the PSQ provided lower scores than did single-item measures.

#### **11.1.4.9 Patient Visit Rating Questionnaire (PVRQ)**

The American PVRQ is also referred to as the Medical Outcomes Trust patient satisfaction scale and the RAND 9-item patient satisfaction survey. Developed as part of the Medical Outcomes Study (MOS), the PVRQ was designed to provide comprehensive measurement of all aspects of patient satisfaction with a medical consultation for the purpose of comparing patients' views of the quality of care in different systems of care (Rubin, et al. 1993). It consists of 9-items measuring the

visit overall, the technical skills of the clinician, the personal manner of the clinician, the time to get an appointment, the convenience of the medical rooms, contacting the medical rooms by telephone, the time spent waiting at the medical rooms, the time spent with the clinician and the explanation of the treatment. The response categories are *poor/fair/good/very good/excellent*. Two scoring methods were reported by Rubin, et al. (1993). Scores were summed and transformed into percentile scores (scale 0 to 100, where 0 = poor and 100 = excellent). The second scoring system was where item responses were dichotomized at excellent/not excellent.

Regarding the psychometric properties of the PVRQ, Rubin, et al. (1993) reported data from the MOS survey (n=17,671) which was carried out in three US cities (Boston, Chicago and Los Angeles) where within each city three different types of medical practice were sampled (a prepaid medical insurance practice (health maintenance organization, HMO), a multi-speciality practice (MSG) and with solo clinician practices (SOLO). The PVRQ was administered after consultation, but before the patient left the office; thus there may well have been a Hawthorne effect. In terms of data distribution, on average less than 5% of respondents endorsed the lowest two categories (reporting that the service was either fair or poor), implying that the response scales were essentially 3-point scales. Based on dichotomization, 55% of respondents reported excellent satisfaction. Responsiveness was assessed by comparing across the different types of medical practice; this revealed that the PVRQ showed SOLO practices obtained higher satisfaction levels, followed by HMO and then MSG practices. Furthermore, differences were also obtained by prepaid and fee-for-service practices and whether patients changed clinicians.

These findings stand in marked contrast with those of Kikano, et al. (1998), who used the PVRQ in a study comparing self-employed clinicians (n=2,185 patients) compared with those who were employed (n=1,351 patients). The findings showed there was no significant difference in PVRQ satisfaction scores, even though there were significant differences on other aspects of health care (e.g. history taking, planning treatment, doing physical examinations).

#### 11.1.4.10 Gonzalez, et al. Patient Satisfaction Questionnaire

This Spanish patient satisfaction questionnaire was developed to provide an assessment of acute hospital care following patient discharge (Gonzalez, et al. 2005). It was developed following a review of patient satisfaction measures published in the 30 years before 2005 – a review which concluded that most instruments had poor psychometric properties and were, perhaps, culturally bound. First, a literature review of patient satisfaction measures was undertaken and eight focus groups were conducted to elicit the views of patients. An item bank was developed based on interleaving the review and focus group findings. The draft items were then reviewed by both patients and clinicians, leading to item revision. The final item bank was administered to 650 patients at 2-weeks after hospital discharge; the response rate was 74%. Factor analysis was used to explore the extent to which the data supported the hypothesized model.

The questionnaire consists of 34 items sorted into six dimensions: information and medical care (12 items), nursing care (8 items), comfort (6 items), visiting (4 items), privacy (2 items), and cleanliness (2 items). Different response sets were used with different items, ranging from three to six categories. Gonzalez, et al. (2005) reported that 77% of patients obtained ceiling scores on privacy, 64% on cleanliness, 22% on nursing care and 20% on information and medical care. The reliability of each of the scales was Cronbach  $\alpha = 0.85$  for information, 0.82 for nursing care, 0.71 for comfort, 0.77 for visiting, 0.60 for privacy, and 0.74 for cleanliness. All six scales were monotonically<sup>36</sup> sensitive to a global patient satisfaction question which had 4 response levels;

---

36 Monotonicity describes where mean scores on an instrument vary in order on a known response set from a criterion. For example, if the criterion is good health and the response set is Excellent/Very good/Good/Fair/Poor, and the scores of interest are walking rate, a monotonic relationship would be where the mean walking rates, for each response set level, were Excellent: 140cm/sec; Very Good: 135 cm/sec; Good: 128 cm/sec; Fair: 113 cm/sec; Poor: 102 cm/sec.

however the comfort, visiting and cleanliness scales were non-monotonic when tested against perceived health improvement. There were statistically significant differences in scores by age group, except for the visiting scale.

A re-analysis of the same dataset (Quintana, et al. 2006) showed scores systematically varied by education level, marital status, gender, work status, length of stay, the number of previous hospital admissions, and whether the patient had assistance completing the questionnaire.

#### **11.1.4.11 Inpatient Evaluation of Service Questionnaire (IESQ)**

The Australian Inpatient Evaluation of Service Questionnaire (Meehan, et al. 2002) was developed as a psychiatric patient satisfaction measure. Eight discussion groups with 66 psychiatric care inpatients were conducted to draft items related to patient satisfaction with hospital stay. Based on the literature, several more items were added to the item pool. The second step was to administer all 51 items to a sample of 72 patients and for each item to be assessed for its importance. The least important items were discarded, leaving 20 items in the pool. An overall satisfaction item was added, as were two items probing recommend the hospital and intention to return to the hospital. Two open-ended items were added for general feedback. There are thus 20 core items in the instrument. The response scales for the core items are poor/ fair/ good/ very good/ excellent. Scores are simply the summed endorsements across all items.

Factor analysis revealed three factors, labelled the staff-patient alliance (10 items), satisfaction with environment (6 items) and satisfaction with treatment (4 items). The reliability of the three scales was reported to be Cronbach  $\alpha = 0.93, 0.78$  and  $0.86$ , respectively. Cronbach  $\alpha$  for the full 20 items was  $0.95$ .

The researchers excluded from their sample any patient who was in the hospital for <7 days on the grounds that these patients would not have had sufficient experience of the hospital environment to be able to make meaningful evaluations. This requirement would obviously prevent the instrument being used generally.

#### **11.1.4.12 The Short Assessment of Patient Satisfaction (SAPS)**

The 7-item Australian Short Assessment of Patient Satisfaction (SAPS) (Hawthorne, et al. 2006) was developed as a generic patient satisfaction instrument. The underlying construct was Donabedian's and it was constructed to ensure measurement of each of the 7 dimensions of patient satisfaction defined in Section 11.1.1 (Hawthorne, 2006). A sample of women (N=178) who had treatment for urinary incontinence completed a patient satisfaction questionnaire consisting of the Client Satisfaction Questionnaire (CSQ-18) (Larsen, et al. 1979), the Consultation Satisfaction Questionnaire (Consult SQ) (Baker, 1990), the Genito-Urinary Treatment Satisfaction Scale (GUTSS) (Hawthorne and Harmer, 2000) and the Patient Satisfaction Index (PSI) (Guyatt, et al. 1995). Iterative item response theory analysis was used to construct the scale (Loevinger  $H = 0.55$ , indicating a strong unidimensional scale). The 7 items are: happy with the effect of your treatment, satisfied with explanations of treatment results, the clinician was careful to check everything, satisfied with the health care choices, being respected, having sufficient time with the clinician, and being generally happy with the care received. The item response sets are 5-point scales with the anchorpoints descriptors varying to match the item stems. For example, *How satisfied were you with the choices you had in decisions affecting your health care?* Very dissatisfied/ Dissatisfied/ Neither satisfied nor dissatisfied/ Satisfied/ Very satisfied. Scoring is by summation. Cronbach  $\alpha = 0.86$ .

Regarding the psychometric properties of the SAPS, 20% of respondents obtained scores at the ceiling, scores were sensitive to health status, treatment type, treatment success and to information given. SAPS scores correlated with the Consult SQ scores  $r_s = 0.73$ , the CSQ-18  $r_s = 0.78$ , the GUTSS  $r_s = 0.83$  and the PSI  $r_s = 0.83$  ( $p < 0.01$  for all). These Spearman correlations

are suggestive that the SAPS measures the same construct as the other measures. When compared with the other four instruments included in the study, the relative efficiency of the SAPS was 2.18 (when compared with the Consult-SQ), 1.89 (PSI), 1.35 (GUTSS) and 1.30 (CSQ-18). These findings suggested it was more sensitive than any of the other measures.

## 11.1.5 Discussion

### 11.1.5.1 Discussion of Single-item Patient Satisfaction Measures

Often single item assessments are asked in the context of health care, perhaps at the end of a consultation when the patient is in a dependent relationship with the clinician. Many clinicians use this as lead-in material for more detailed discussions of issues arising from treatment. Not surprisingly, given this relationship, most patients report very high levels of satisfaction.

Quite apart from this administrative issue, there are substantive psychometric grounds for rejecting this model of patient satisfaction measurement.

It assumes that patient satisfaction is a single holistic dimension, which is adequately captured by a single item. As shown in Section 11.1.1, however, patient satisfaction is a construct with at least 6 substantive dimensions (other than general satisfaction) implying that the level of satisfaction will vary depending upon which aspect of medical care is being assessed by the patient. Where different dimensions of care are assessed globally, there may be no way of determining which aspects are in need of improvement (Locker and Dunt, 1978).

It is also assumed that a given response to a single item is reliable. However, none of the studies reviewed reported on the reliability of a single item. There are two exceptions. Ware, et al. (1983) reported that, based on Helmstadter's recommendation of a reliability estimate of 0.50 or greater for group comparisons (Helmstadter, 1964), 75% of single items from the PSQ (Patient Satisfaction Questionnaire) failed this criterion, compared with 18% of the PSQ sub-scales. Hawthorne, et al. (2006) examined the psychometric properties of 5 single global items before endorsing the use of a single item. He reported variation by treatment type, differences between known groups, an absence of differential item functioning by known group, and sensitivity to a then-test of health status.

Where single item reliability has been systematically investigated elsewhere, the results suggest that single items are of doubtful reliability. Wyrwich (2002), for example, reported 1-4 day test-retest kappa agreement of 0.64 to 0.73 for single item patient change scores; estimates which fall outside the normally accepted psychometric standard for reliability (generally reported to be >0.80).

### 11.1.5.2 Discussion of Patient Satisfaction Instruments

Over 20 years ago Nguyen, et al. (1983) noted that it is almost impossible to make any meaningful comparisons between different patient satisfaction scale scores for two key reasons: first, that satisfaction scores across studies are so high that comparative interpretation is almost impossible, and second that because there are almost no standard instruments that are widely used or reported it is difficult to equate scores from one study to another. They pointed out that most patient satisfaction questionnaires have been developed based on the researchers' views i.e. that at best most patient satisfaction measures have face validity only. Hardy, et al. (1996) observed that most patient satisfaction measures were developed in hospitals in the USA, the implication being that they may have little applicability elsewhere.

Furthermore, Sitzia (1999) in a review paper found that 81% of studies reported using a new patient satisfaction instrument and a further 10% reported modifying a previously existing instrument, yet 60% of studies examined failed to report any psychometric data.

The findings from this review are consistent with these earlier reports, although some improvements are evident, such as the reporting of basic psychometric tests (e.g. scale means, standard deviations and estimates of internal reliability). In general, however, the literature is still characterized by measures developed for particular studies where almost no psychometric data are available. Surprisingly, very few of the recognized patient satisfaction measures were identified as being used in studies other than by the original authors, and in many cases instruments were modified without appropriate psychometric testing. An additional difficulty uncovered in this review is that even where patient satisfaction measures have been available over time and are widely cited, there is almost no further psychometric work reported in the literature beyond that of the instrument developers. As a consequence, very few patient satisfaction instruments met the study criteria.

Finally, a key finding was that there are almost no head-to-head studies comparing the measurement properties of different patient satisfaction instruments. This is a truly extraordinary situation suggesting that different research teams have simply failed to sufficiently acknowledge previous work in the field.

Concerning the instruments reviewed, when these were examined against the psychometric criteria outlined in Sections 2 and 11.1.3, none met all the criteria.

### Evidence of a latent construct of patient satisfaction

The definition of patient satisfaction varies by the purpose of the researchers. Regarding the instruments reviewed, the stated purposes are shown in Table 55. This reveals that five instruments were primarily developed to assess satisfaction with a clinical or medical consultation (ConsultSQ, Linder-Pelz, MISS, PSQ-III and PVRQ).

**Table 55 Purpose of Instruments Reviewed**

Purpose of instruments reviewed												
	<i>CSQ -18</i>	<i>CSQ -8</i>	<i>ConsultS Q</i>	<i>LOPPS -18</i>	<i>Linder -Pelz</i>	<i>MISS -21</i>	<i>PSI</i>	<i>PSQ -III</i>	<i>PVRQ</i>	<i>Gonzalez</i>	<i>IESQ</i>	<i>SAPS</i>
Clinical consultation	—	—	Yes	—	Yes	Yes	—	Yes	Yes	—	—	—
Health care generally	Yes	Yes	—	—	—	—	—	—	—	—	—	Yes
Life-threat care	—	—	—	—	—	—	Yes	—	—	—	—	—
Nursing care	—	—	—	Yes	—	—	—	—	—	Yes	Yes	—
Self-completed	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Length (a)	M	S	M	M	M	M	M	L	S	L	M	S

Notes:

a = S: short instrument (<15 items); M: medium length (15-30 items); L: long instrument. (>30 items).

Other than the ConsultSQ, all instrument developers referred to a theory of patient satisfaction in general. However, none invoked a theoretical model and then tested their manifest instrument against the model, with the exception of the Linder-Pelz scales, the PSQ-III and the SAPS. All instrument developers stated that the model of patient satisfaction used was created by reading the literature and consulting with clinicians. The more thorough instrument developers also consulted with patients; these instruments were the ConsultSQ, LOPSS, PSI, PSQ-III, Gonzalez and IESQ. Generally though, the impression was that instrument developers defined a theoretical model in accordance with their particular concerns, and then created the instrument around those

concerns. This judgement rests on the fact that of the latent dimensions contributing to patient satisfaction presented in Section 11.1.1, only one instrument covered all these; i.e. instrument developers chose to measure some of the theoretical parts of the patient satisfaction construct. The exception was the SAPS which was constructed to cover these dimensions (Hawthorne, et al. 2006). Under these circumstances, the accepted interpretation of what is being measured in psychometric terms is to examine the content of the instrument.

- The preferred instruments are those where the stated purpose is consistent with the theories of satisfaction given in section 11.1.1, i.e. that patient satisfaction is primarily around the interaction relationship between the patient and the clinician. Based on this criterion the better instruments are ConsultSQ, Linder-Pelz, MISS, PSQ-III, PVRQ and SAPS.

**Adequate coverage of the latent construct**

Table 56 shows the coverage of the reviewed instruments, where the instrument items, based on the item content, have been mapped against the dimensions of patient satisfaction outlined in Section 11.1.1 Other than the SAPS, no instrument provided complete coverage; the best instruments were the Linder-Pelz scales, PSQ-III and Gonzalez. The PSQ-III, however, had very strong measurement of access to health care, other areas (most of which were measuring the patient’s ability to pay for health care), and the technical skill of the clinician. The Gonzalez instrument had very strong measurement of hospital facilities. Other instruments with particular emphases were the PSI for measuring patient participation, the PVRQ for measuring access, the ConsultSQ for technical skill, the CSQ-18/8 for general satisfaction, and the IESQ for exploring patient-clinician relationships. The LOPSS primarily measures information and patient-clinician relationships.

**Table 56 Content Validity (Coverage)**

Content validity (coverage)												
	<i>CSQ -18</i>	<i>CSQ -8</i>	<i>ConsultS Q</i>	<i>LOPPS -18</i>	<i>Linder -Pelz</i>	<i>MISS -21</i>	<i>PSI</i>	<i>PSQ -III</i>	<i>PVRQ</i>	<i>Gonzalez</i>	<i>IESQ</i>	<i>SAPS</i>
Access & facilities	3	—	3	—	4	—	—	12	5	9	4	1
Information	—	—	2	5	4	3	3	3	1	5	4	1
Relationship	2	—	4	5	5	6	6	8	1	4	6	1
Participation	2	—	1	5	1	6	9	3	—	1	1	1
Technical skill	2	1	5	2	5	3	3	9	1	5	3	1
Effectiveness	1	1	—	—	—	3	—	—	—	—	—	1
Satisfaction general	5	6	3	1	1	—	—	6	1	1	—	1
Other	3	—	—	—	1	—	2	10	—	6	2	—
Total items	18	8	18	18	20	21	23	51	9	34	20	7
Length (a)	M	S	M	M	M	M	M	L	S	L	M	S

Notes:

a = S: short instrument (<15 items); M: medium length (15-30 items); L: long instrument. (>30 items).

Areas that were poorly measured were treatment effectiveness or outcome (this was measured by the SAPS, MISS, and CSQ-18), and satisfaction generally (not measured by the MISS, PSI, and IESQ).

The PSQ-III, although offering the most comprehensive coverage of any instrument, is excessively weighted towards issues around access and payment — indeed these constitute 20/51 of its items, and this emphasis is likely to be misplaced in an Australian context (see Section 11.1.5.3).

Although the CSQ-18 has good coverage, the CSQ-8's coverage is poor because 6/8 items are about satisfaction in general (the content is primarily about help being given, and needs being met).

Although the coverage of the ConsultSQ is very good in that it has items about 6 of the 7 satisfaction dimensions, this coverage is subject to issues around item repetition. For example, *I am totally satisfied with my visit to this doctor* and *I am not completely satisfied with my visit to the doctor*. Essentially these are the same item, one expressed positively and the other negatively. Altogether there are 6 pairs of such items in the instrument; thus 6/18 items are repetitive.

Although the coverage of the MISS is good, examination of the actual items reveals that most are about the doctor being fully in charge of the health of the patient, particularly with respect to decision-making. For example, *The doctor told me just what my trouble is*, or *The doctor gave me a chance to say what was really on my mind*. As such the tone of the MISS is out of step with one of the key reasons for the rise of patient satisfaction measurement: to give patients a voice. The fundamental issue this tone raises is whether the MISS actually measures patient satisfaction at all rather than measuring the behaviour and attitude of the clinician towards the patient.

The items in the PSI pose a different problem because many are concerned with the inner feeling of the patient in coping with the life-threatening condition. For example, *Gone through a lot*, or *Felt out of control of situation*. Because these kinds of items comprise most items in the PSI it is difficult to know whether the PSI is measuring patient satisfaction or patient internalization of their experiences with their health care.

Although the Gonzalez instrument is the longest instrument reviewed, a quarter of its items are concerned with the hospital facilities, such as toilet and room cleanliness. Compared with this emphasis, other areas contributing to patient satisfaction are lightly covered (e.g. participation). The IESQ has good coverage of access and facilities, information, technical skill and relationships. In contrast, the SAPS has just 1 item from each of the 7 dimensions.

- Based on the coverage of content criteria, the better instruments are the CSQ-18, Linder-Pelz, and SAPS.

Regarding the number of items, i.e. instrument length, in clinical work and epidemiological studies parsimony is important. Clinicians do not have the time to administer long instruments or the resources to score them, and in most research studies instrument batteries are administered where there are competing demands for the available space. Additionally, there are major concerns regarding self-report among those with mild dementia, Alzheimer's disease or cognitive impairment in relation to current affective state, adaptation, lack of insight, neuroticism and emotional adjustment, and possible effects of neuroleptic therapy (Awad, et al. 1995; Coucill, et al. 2001; Diener, et al. 1999; Jenkins, 1992; Kring, et al. 1993; Magaziner, 1997; Wood, et al. 1985).

- Based on the need for parsimony, the shorter instruments are the CSQ-8, PVRQ and SAPS.

### **Data distribution and ceiling effects**

Regarding the response scale used in the instruments, Likert scales are used by the ConsultSQ, LOPPS, MISS, PSQ-III and SAPS. The other instruments use Guttman scales. The details are given in Table 57.

**Table 57 Scoring of the Instruments**

Scoring of the instruments		CSQ -18	CSQ -8	ConsultSQ	LOPPS -28	Linder -Pelz	MISS -21	PSI	PSQ -III	PVRQ	Gonzalez	IESQ	SAPS
Responses	Likert-type	—	—	Yes	Yes	—	Yes	—	Yes	—	—	—	Yes
	Guttman-type	Yes	Yes	—	—	—	—	Yes	—	Yes	Yes	Yes	—
	N. points	4	4	5	5	—	7	7	5	5	(a)	5	5
Scale range	Minimum	18	8	18	28	—	—	23	—	—	0	0	0
	Maximum	72	32	90	140	—	—	161	—	—	100	100	35
	Mean scores (range)	49-55	24-27	52-77	115-120	—	—	122	—	—	67-90	56-61	29
Cummins %SM range (b)		56-68%	65-78%	46-82%	77- 82%	—	77% (c)	72%	—	—	67- 90%	56- 61%	82%

Notes:

a = Varies 3-6

b = mean score as percentage of potential scale range

c = not computed from data, but reported in the papers

All instruments had different score ranges; but all suffered from assigning high levels of satisfaction. To examine whether these scores were so high that ceiling effects were likely Cummins' %SM (Cummins, 1995) was computed for all instruments, where the standard was 75%SM of the theoretical score range. Instruments with %SM scores above this standard are more likely to suffer ceiling effects, whereas instruments below this standard are less likely to. As shown in Table 56, score ranges were not reported for the Linder-Pelz, MISS, PSQ-III or PVRQ instruments.

The Gonzalez instrument was that with the highest %SM, suggesting ceiling effects were more likely. The other instruments with %SM above 75%SM were the ConsultSQ, the LOPPS-28 and the SAPS. The LOPPS-28 observation is consistent with O'Connell, et al. who reported high levels of acquiescent response bias through item insensitivity (O'Connell, et al. 1999). The same situation has been reported for the PVRQ where less than 5% of respondents endorsed the lower two response categories.

An important point to note is that the %SM values presented in Table 57 show that there is a divide in %SM scores by response scale type: those instruments using Likert scales (ConsultSQ, LOPPS, MISS and SAPS) all obtained higher %SM scores at the upper end of the scale than did the instruments using Guttman scales (the exception was the Gonzalez). This finding is consistent with the literature (Hendriks, et al. 2001; Ware and Hays, 1988). Since high patient satisfaction scores are both endemic and problematic in the measurement of patient satisfaction, this finding would, prima facie, suggest that instruments with Guttman-type scales might be preferred.

- Based on scoring ranges and the likelihood of ceiling effects, the better instruments are the CSQ-8, CSQ-18 and IESQ.

**Validity evidence**

As the individual instrument reviews show, there is very little sustained evidence of validity for any of the instruments reviewed. Generally, the available evidence is from the instrument developers and perhaps one or two other research teams. This evidence is summarized in Table 58. When reading the table, it should be remembered that 'Yes' means that some evidence is available, and a null entry (—) that no evidence was reported in the papers reviewed (this does not mean that the

instruments have not been assessed against the validation criteria, just that no evidence was uncovered).

**Table 58 Validity Evidence**

Validity evidence	<i>CSQ -18</i>	<i>CSQ -8</i>	<i>ConsultS Q</i>	<i>LOPPS -18</i>	<i>Linder -Pelz</i>	<i>MISS -21</i>	<i>PSI</i>	<i>PSQ -III</i>	<i>PVRQ</i>	<i>Gonzalez</i>	<i>IESQ</i>	<i>SAPS</i>
Ecological (a)	—	—	Yes	Yes	—	—	Yes	Yes	—	Yes	Yes	—
Factor analysis (b)	Yes	Yes	Yes	Yes	—	Yes	—	Yes	—	Yes	Yes	Yes
IRT analysis (c)	—	—	—	—	—	—	—	—	—	—	—	Yes
Concurrent (d)	Yes	Yes	Yes	Yes	—	—	Yes	Yes	Yes	Yes	—	Yes
Convergent (e)	Yes	Yes	Yes	No	—	—	Yes	Yes	Yes	—	—	Yes
Predictive (f)	—	—	—	—	—	—	—	—	Yes	—	—	—

**Notes:**

a = involving patients during instrument construction.

b = a measure of construct validity because it examines the relationships between items

c = item-response theory analysis is a measure of construct validity because it examines the relationships between items

d = correlation with other measures of patient satisfaction

e = assessed against patients groups known to be satisfied/dissatisfied or treatment success/failure

f = assessed by power to predict future outcomes

Regarding the validity evidence for the LOPPS, it must be recognised that although there is some evidence available, this does not fully support the LOPPS in that the LOPPS did not show significant variation in scores by groups known to differ in satisfaction (based on age), and that it did not register patient dissatisfactions.

- The instruments that have the most validity evidence are the ConsultSQ and PSQ-III. Although there is validity evidence for the SAPS, it comes from the one study.

### Reliability and responsiveness evidence

Three forms of reliability were reported, Cronbach  $\alpha$ , split-half and test-retest. The data are given in Table 59. The conventional range for internal consistency as assessed by Cronbach  $\alpha$  is from 0.70 to 0.90 where cohorts are to be compared (Nunnally and Bernstein, 1994). However,  $\alpha$  is a function of both the correlation between items within a scale and of the scale length (Cortina, 1993). The implication is that where  $\alpha$  exceeds 0.90 there is likely to be redundancy in the scale because the same concept is being asked twice or more often. High  $\alpha$ -values can also be brought about where there is little variance in responses; typically where the range of responses is truncated. In this study this was the case with many instruments because respondents mainly utilized the first two or three categories (i.e. where most respondents ticked a 1 or 2 in a scale of 5). Equally where  $\alpha < 0.70$  the scale is likely to be made up of items that are too disparate and that do not form a homogenous scale.

**Table 59 Reliability and Responsiveness Evidence**

Reliability and responsiveness evidence												
	<i>CSQ -18</i>	<i>CSQ -8</i>	<i>ConsultS Q</i>	<i>LOPPS -18</i>	<i>Linder -Pelz</i>	<i>MISS -21</i>	<i>PSI</i>	<i>PSQ -III</i>	<i>PVRQ</i>	<i>Gonzalez</i>	<i>IESQ</i>	<i>SAPS</i>
Cronbach $\alpha$ (a)	0.78 - 0.91	0.92- 0.93	0.91	0.93- 0.98	—	0.93	—	0.88	—	—	0.95	0.86
Cronbach $\alpha$ (b)	N/A	N/A	0.67- 0.94	—	0.49- 0.81	0.66- 0.96	N/A	0.23- 0.93	N/A	0.60- 0.85	0.78- 0.93	N/A
Split –half correlation	0.82	—	—	—	—	—	—	—	—	—	—	—
Test-retest (c)	—	—	—	—	—	—	0.86	—	—	—	—	—
Responsivene ss (d)	Yes	Yes	Yes	?	—	Yes	Yes	Yes	?	Yes	—	Yes

**Notes:**

a = for summated instrument scores. Where different estimates are available, the lower and upper estimates are given.

b = for scales within the instrument, where applicable. Where different estimates are available, the lower and upper estimates are given.

c = intra-class correlation

d = ability to detect differences between groups of patients

Table 59 shows that most of the instruments were within the acceptable range, albeit at the upper end. No instrument was reported as being unreliable, although several instruments' scales fell outside the conventional range for reliability (scales on the ConsultSQ, Gonzalez, Linder-Pelz, MISS and PSQ). It is likely, however, that the LOPPS contained redundant items; and there may be redundancy in the CSQ-18/8, Gonzalez and MISS.

As reported in Table 56, in the case of the CSQ-18/8 and the Gonzalez this may be a function of the lack of breadth of measurement because a high proportion of the items measure the same dimension (e.g. in the case of the CSQ-8 6/8 items are concerned with overall or general satisfaction, thus there is likely to be repetitive measurement). In the case of the LOPPS where the items are spread out over different dimensions this situation is almost certainly caused by insensitive items leading to the endorsement of high end categories indicating satisfaction (as shown in Table 57).

Regarding responsiveness, the evidence suggests that all instruments were responsive, although there was mixed evidence for the LOPPS and for the PVRQ, and insufficient for the Linder-Pelz scales. No evidence was sighted for the IESQ.

- Based on reliability criteria, the better instruments are the CSQ-18, PSI, and the SAPS.

### 11.1.5.3 Additional Criteria

In addition to the psychometric criteria discussed in the previous section, there are two contextual issues that are relevant in assessing patient satisfaction instruments for use in Australian settings.

#### Relevance to the Australian health care system

Patient satisfaction measures may be culturally specific (Baker, et al. 2003; Firestone, et al. 2004; Henderson, et al. 2003; Labarere, et al. 2004; Meakin and Weinman, 2002; Zebiene, et al. 2004). This implies that different instruments are not equally relevant in different settings.

The Australian health care system is characterised by multi-level funding: the Commonwealth and State Governments provide about 70% of all health costs, primarily through Medicare and the Pharmaceutical Benefits Scheme (both of which provide subsidized services) and the funding of public hospitals where emergency and outpatient services are provided free of charge (AIHW, 2004). Safety nets for high consumption users who have limited resources also apply (e.g. for

those on unemployment or pension benefits). For most Australians the first contact for health care is a general practitioner (GP). Access to GP services, from the patient perspective, is uncapped. Although GPs are located across Australia, thus ensuring ready access to primary health care, there are some distributional issues that affect access to and quality of services, mostly in country areas (Birrell and Hawthorne, 2001; Hawthorne, et al. 2003). Additionally, about 49% of the population has private health insurance for hospital and ancillary health care (AIHW, 2004).

The implication is that patient satisfaction measures which focus on costs borne by the patient, access to health care or emphasis on the buildings within which care is provided are likely to be less relevant in the Australian context. Additionally, the literature suggests that these issues play little part in determining patient satisfaction.

The instruments which have scales or several items measuring these aspects of care are the CSQ-18, Gonzalez, IESQ, Linder-Pelz, PSQ-III and the PVRQ. The CSQ-18 has two items measuring the promptness of being seen, and the comfort and attractiveness of the facility and building. The Linder-Pelz scales have 3 items assessing a patient's entitlements (rights to see a clinician immediately, not wait for an appointment and the right to tell the clinician everything) and 3 questions on access (ease of getting to the clinic, the comfort of the waiting room, and having to wait). The PSQ-III has 7 items assessing the effect of health costs (e.g. going without services due to the cost), and 9 items assessing access to care (e.g. that health care facilities should have longer opening hours). The PVRQ has items covering the time to get an appointment, the convenience of the medical rooms, contacting the medical rooms by telephone, the time spent waiting at the medical rooms. The Gonzalez has items probing room conditions, physical description of the hospital, food quality, room comfort, toilet cleanliness, room cleanliness and privacy. Similarly, the IESQ has items covering activities for the patients (e.g. videos, games, outings), ward cleanliness, group activities, food quality and privacy. Additionally, the LOPSS has an item on responsiveness to the patient ringing the nurse call bell; which is a reflection of its primary use inside hospitals.

Although some of these items (e.g. the convenience of the medical rooms) may be appropriate for Australians living in locations where access to services is compromised, it is doubtful these items are relevant beyond this. Table 60 summarizes these issues.

This criterion would suggest that the better instruments are the CSQ-8, the ConsultSQ, the MISS-21, the PSI and the SAPS.

**Table 60 Additional Criteria**

Additional criteria													
	<i>Single item</i>	<i>CSQ -18</i>	<i>CSQ -8</i>	<i>Consult SQ</i>	<i>LOPPS -18</i>	<i>Linder -Pelz</i>	<i>MISS -21</i>	<i>PSI</i>	<i>PSQ -III</i>	<i>PVRQ</i>	<i>Gonzalez</i>	<i>IESQ</i>	<i>SAPS</i>
Australian relevance	Yes	—	Yes	Yes	—	—	Yes	Yes	—	—	—	?	Yes
<i>Best for:</i>													
Clinicians	Yes	—	—	—	—	?	—	—	—	—	—	—	Yes
Specialists	?	Yes	—	?	—	?	—	?	—	Yes	—	—	Yes
Researchers	No	Yes	—	?	—	?	—	?	—	—	—	—	Yes

### Instrument users

There are three main users of patient satisfaction instruments: (a) clinicians, (b) specialists, and (c) researchers or program evaluators.

At the clinical level, patient satisfaction is likely to be related to the clinical management of individual patients. Assessment may be near the end of a consultation, and the patient's response may be used by the clinician as a discussion starter. At the clinician level there are time and data collection issues: busy clinicians may not have the time and expertise required to use long, multi-scaled instruments. Additionally, they may need a measure where the scoring is instant so that they can discuss the results with the patient immediately. Data collection should be as brief as possible and there should be few data analysis demands upon clinicians.

Specialists working with patients may have somewhat different needs. Although when working with individual patients their needs may be similar to those of clinicians, many specialists also need more information and are often involved in research or evaluation. Instruments used at this level may need to possess sufficient nomological evidence to be used at the case level; i.e. for individual patient assessment.

Researchers and program evaluators' needs centre round data that are useful for answering research questions where analyses are group-based; where data collection procedures may be remote; and where findings are aimed at demonstrating the effect of new treatments or at influencing policy decisions.

Finally, in a health care system committed to evidence-based practice, basic data which is collected and held at the clinician level should be suitable for transfer to local level analysis and also to research settings (e.g. for inclusion in monitoring or surveillance).

These different needs imply that at each level different patient satisfaction measures may be needed. Based on the reviews of instruments, when assessed against this criterion, the rankings of instruments presented in Table 61 were made. This suggests that:

- For clinicians working with individual patients a single global question may be sufficient. Where more information is sought, the SAPS may be appropriate because its brevity, coverage and ease of use.
- For specialists involved in research studies, in addition to a single global question, short instruments assessing satisfaction with incontinence care and treatment outcomes may be needed. The preferred instruments would be the CSQ-18, ConsultSQ and SAPS.
- For researchers, in addition to a single global question and short patient satisfaction instruments, generic patient satisfaction measures may be needed which cover the full range of patient satisfaction dimensions and that have excellent reliability, validity and responsiveness properties. The preferred instruments would be the CSQ-18 or the SAPS.

### **11.1.6 Recommendations**

Based on the criteria for measuring patient satisfaction (Section 11.1.3) and the reviews of instruments in sections 11.1.4 and 11.1.5, it was possible to compare the multi-item instruments reviewed. This is shown in Table 61.

For each of the study criteria described in Section 2 and 11.1.3, each instrument was rated on a scale of 1, 2 or 3, where 1 indicated the instrument did not meet the criterion, 2 indicated there was some evidence the instrument partly met the criterion, and 3 indicated the instrument met the criterion.

**Table 61 Summary Assessing Patient Satisfaction Instruments Against the Study Criteria**

Criteria	Weight	Instruments												
		SAPS	Consult-SQ	PVRQ	LOPPS-18	Single item	CSQ-8	CSQ-18	PSI	MISS-21	IESQ	Linder-Pelz	PSQ-III	Gonzalez
Availability of comparison data	3	3	2	1	1	1	2	2	2	1	1	1	2	1
Length/feasibility of instrument for inclusion in battery	2	3	2	3	2	3	1	2	2	2	2	2	1	1
Complexity of administration/cognitive burden	2	3	2	3	2	3	3	2	2	2	2	2	1	1
Cultural Appropriateness	1	3	2	3	2	3	2	1	2	1	1	2	1	1
Ease of obtaining score	2	3	2	3	2	3	3	2	2	2	2	2	1	1
Sensitivity to dementia	3	1	1	1	1	1	1	1	1	1	1	1	1	1
Reliability evidence available	3	2	3	1	3	1	2	2	2	2	2	1	2	2
Validity evidence available	3	2	3	2	2	1	2	2	1	1	1	1	2	1
Cost of the instrument	2	3	3	3	3	3	1	1	3	3	3	3	3	3
Cost of instrument administration	2	3	3	3	3	3	3	3	2	2	2	2	1	1
<b>Weighted Total</b>		<b>57</b>	<b>53</b>	<b>48</b>	<b>47</b>	<b>45</b>	<b>45</b>	<b>42</b>	<b>42</b>	<b>38</b>	<b>38</b>	<b>36</b>	<b>36</b>	<b>30</b>

Scores on each criterion were weighted and then summed. The two standout instruments were the SAPS and the ConsultSQ. None of the other instruments reviewed could be considered truly satisfactory.

\*\*\*

The key finding from this review of patient satisfaction instruments is that no instrument has been sufficiently validated for its use in Australia to be automatically recommended. There are three key reasons for this finding:

- A. There is evidence throughout the literature that patient satisfaction is culturally specific. It cannot be assumed that an instrument that is relevant, valid and reliable in one culture retains those properties in another culture. Thus instruments developed overseas may not be appropriate in Australian settings.
- B. There is no agreed theoretical model of patient satisfaction or of its constituent parts. As this review has shown, the consequence is that instrument designers have proceeded on an ad hoc basis with the result that there are thousands of patient satisfaction measures available.
- C. Among recognised generic patient satisfaction instruments there is insufficient evidence of their psychometric properties for any instrument to be fully accepted as possessing a nomological net of validity evidence.

The recommendations below should be read with these caveats in mind. They are:

1. That a single item patient satisfaction measure should be adopted for use in Australian settings by clinicians wishing to assess the satisfaction of their patients 'on the spot'. Strategies should be put in place to encourage clinicians to adopt this measure as a common metric across Australia. Encouragement should be given to specialists and researchers to also include this common metric in their work. In this way a bank of shared understanding will be progressively established. It may be possible that a single item measure could be drawn from the generic instruments recommended in #2 or #3, or that the Hawthorne, et al. recommended single item for the National Continence Management Strategy be re-examined for this role (Hawthorne, et al. 2006).
2. That the SAPS and ConsultSQ are validated in dementia-populations. These were the two better generic patient satisfaction instruments identified in this report. For the reasons outlined in the report, however, neither can be recommended outright because there is no evidence of their reliability, validity or responsiveness in dementia populations. It is recommended that a head-to-head validation study be undertaken in dementia-populations.
3. Until recommendation #2 is implemented and the results published, it is recommended that the SAPS be used.

## **11.2 Carer Satisfaction with Services: A Review**

### **11.2.1 Introduction**

Informal or non-professional carer satisfaction with health services (hereafter carer satisfaction) provided to the care recipient can be viewed from two related perspectives.

On the one hand it is related to carer burden, where it is concerned with the carer's perceptions of how well the care recipient (and the carer) is being supported or cared for by health care providers. Simply, this is satisfaction with the care provided to the care recipient and/or carer (Kristjanson, 1993). Its origins can be traced to two concerns: the provision of appropriate care for

those with chronic incurable or terminal illnesses living in hospices (which was perceived to frequently be of poor quality or troublesome to family carers (Bowers, 1988; Nolan, et al. 1990) and the need for support or respite care from the burden of caring for families providing palliative care at home for these people (Blake and Lincoln, 2000; Hasselkus, 1988; Lawton, et al. 1989; Mason, et al. 2007; Nolan and Grant, 1992; Seale, 1989). The core dimensions, based on family or carer concerns, relate to satisfaction with quality care availability, physical and psychosocial care, and information giving (Hare, et al. 2006; Kristjanson, 1989; Kristjanson, 1993). Carer satisfaction from this perspective has been primarily studied in respite care, in community-based services which are a response to home caregiving, the British consumer audit requirements, casemix funding, and the increasing number of hospital in the home schemes (Abbott, et al. 2005; Bauld, et al. 2000; Bindman, et al. 1996; Clare and Hofmeyer, 1998; Fulford and Farhall, 2001; Gaddini, et al. 2005; Lawton, et al. 1989; Mason, et al. 2007; Montalto, 1996; Nicoll, et al. 2002; Pritchard and Dewing, 2001; Simpson, et al. 1995; Stephenson, et al. 1995; Wilson, et al. 2002). Service satisfaction surveys have also been conducted in several Australian states (e.g. Victoria: Gill and Maas, 2000).

On the other hand carer satisfaction has been defined as an evaluative procedure for quality assurance, marketing and health care planning (Buttle, 1996; Parasuraman, et al. 1988). Its origins can be found in the early 1980s experiences of the American health maintenance organizations (HMOs). Their explicit role was to control health care costs through restricting services covered by the health care plan and reducing hospital admissions – roles that did not fit well with the need to provide long-term palliative health care for those with chronic and/or terminal health conditions (Bates and Brown, 1988; Manning, et al. 1984). Bates and Brown (1988) provide an interesting example of this mismatch noting that 12/24 benefit packages offered by HMOs provided dental health care whereas just 5/24 provided denture cover). The need to provide better matching of services with geriatric patient need led directly to an interest in assessing satisfaction with health care services. For example, the content of early patient satisfaction instruments included items such as getting more health care from the enrolled HMO than from other health plans, that the HMO covered more services than other health plans, and that hospital admissions were pre-arranged (Cryns, et al. 1989). More recently a questionnaire included items covering cleanliness, privacy, and the quantity/quality of institutional meals (Lubart, et al. 2004).

The underlying theory is that satisfaction with a specific transaction is a function of disconfirmation which reflects the extent to which a person's initial expectations are met. Parasuraman, et al. (1985) and Parasuraman, et al. (1988) argued that the accumulation of incidents of satisfaction over time results in service quality perceptions. They identified five dimensions, viz., tangibles (the physical facilities, equipment and the appearance of personnel), reliability (the ability to perform the service dependably and accurately), responsiveness (a willingness to help customers and provide prompt service), assurance (the knowledge and courtesy of employees and their ability to gain the trust and confidence of customers), and empathy (the caring, individualized attention paid to customers).

In general, care quality assurance is discussed in the literature in negative terms, viz., poor facilities or infrastructure, physical abuse of the patient, his/her psychological abuse, physical and psychological neglect and exploitation (Schulz and Williamson, 1997), whereas care satisfaction is usually asked in more neutral terms, focussing on the extent to which the carer is satisfied with the care of the care recipient. This difference in perspective may well explain differences in reported assessment levels between quality assurance and carer satisfaction (Soliman, 1992). These two perspectives imply that although the assessment of the quality of caring provided by a health service provider and carer satisfaction are different constructs which should not be confused or conflated, quality of caring cannot be adequately assessed without some consideration of both – especially where a care recipient moves from being cared for at home to being cared for in an institution, or where studies compare home care with institutional care (Kessler, et al. 2005). It is a matter of emphasis as to which perspective is of greater interest to carers, clinicians, researchers and policy makers.

This review is concerned with the first of these two perspectives (carer satisfaction) for three reasons. First, in dementia care the primary concern of a carer is that his/her care recipient is well taken care of by community-based health care clinicians, service personnel or teams where necessary, or within institutional care. Second, there is gross market failure in the Australian health care system generally, and particularly in the dementia care sector: most Australians are not fully informed consumers and most Australians do not have the opportunity to make meaningful choices regarding available services for the care of their loved ones. Third, assessments of quality assurance are a function of service provider characteristics and carer expectations and information; areas that most carers have little experience of when they begin caregiving with the implication that immature or uninformed assessments regarding quality assurance can be easily made (Buttle, 1996; Chesterman, et al. 2001; Soliman, 1992).

\*\*\*

Turning to carer satisfaction, the literature reviewed for this study reports high levels of satisfaction (typically >70% (Bekelman, et al. 2005; Clare and Hofmeyer, 1998; Grunfeld, et al. 2004; Hwang, et al. 2003; Kealey and McIntyre, 2005; Lubart, et al. 2004; Ringdal, et al. 2002; Teno, et al. 1997), although this does appear to vary by inpatient status (Shepperd and Iliffe, 1998; Shepperd and Iliffe, 2005) and care coordination (Walker, et al. 2001). Other than this, there appears to be little variation in the level of carer satisfaction. A repeated cross-sectional study of carer satisfaction in Victoria showed no significant change over time with health care services (Gill and Maas, 2000) and another study comparing stroke family support with standard care reported significant differences in carer satisfaction only for practical help and emotional support (Lincoln, et al. 2003).

The literature, however, suggests that results like these may be confounded or may lack generalisability due to the personal characteristics of carers, the structure of health care systems and health care insurance arrangements, the carer's relationship with the health care provider, poor participation rates, and high attrition rates (usually due to death) confounding longitudinal studies (Addington-Hall and McPherson, 2001; Chesterman, et al. 2001; Hasselkus, 1988; Ingleton, et al. 2004; Jaglal, et al. 2007; Nicoll, et al. 2002; Wilkinson, et al. 1999). For example, in a study of the perceptions of family caregivers of nursing home care Bowers (1988) reported that although the participants attributed responsibility for care tasks to nursing home staff, they held themselves responsible for teaching the staff how to provide quality care. The clear implication is that there may be a difference in the expectations between family carers and professional nursing staff. This judgement, however, is subject to the researcher's agenda: the questions used to probe care quality were: Are you involved in your relative's care?/ Have you had any problems with the staff?/ What would happen if you stopped providing that care?/ and Is the care you provide different than the care provided by the staff?. Clearly these were leading questions designed to elicit any issues with the quality of nursing care. Similar findings were also reported by Hasselkus (1988). In light of the literature, illustrated by this example, it seems reasonable to draw the general conclusion that perceived carer satisfaction will be a complex interplay between carer expectations, professional standards of care and the researcher's perspective.

A third key finding is that many researchers have simply asked carers to describe the care received on a single item (e.g. excellent, good, fair, poor, no care) and reported this as satisfaction with health services. The item may be administered for different health care services, such as general practitioners, counselling and/or nursing (Butters, et al. 1993; Kelleher and Mannix, 2001). Others have used qualitative research methods where the questions asked of participants are not standardized (Gessert, et al. 2000). Because of the epistemological limitations of these approaches, neither can be reliably used in large-scale studies or for the routine research assessment of carer satisfaction. In addition, although many researchers have reported that they asked about carer satisfaction no details regarding what they asked are provided. For example, Harris, et al. (2005) in an economic evaluation of a hospital at home program reported that

acceptability of the program was assessed by asking both the patient and his/her caregiver 30 questions relating to service satisfaction. No further details were given in the paper.

### 11.2.2 Method

To identify published instruments assessing carer satisfaction with health services providing caring for people with dementia or cognitive impairment in clinical, epidemiological and research situations a search of MEDLINE, CINAHL and PsycINFO was undertaken using the terms *carer* and *caregiver*, with *satisfaction*, crossed with the keywords *health services*, *dementia*, *Alzheimer's disease* and *mild cognitive impairment*. The results were, in turn, crossed with *instrument*, *questionnaire*, *measure*, *measurement*, *scale* and *tool*. No studies were identified.

Relaxation of the search terms through exclusion of *dementia*, *Alzheimer's disease* and *mild cognitive impairment* identified 23 unique articles. All titles and abstracts were searched to identify instruments. Where papers reported using a measure, the bibliography was scanned to identify the original source.

Fourteen carer satisfaction instruments or scales were identified. Of these 8 were not reviewed following scrutiny of the relevant papers (in chronological order):

- The Caregiving Activities Scale (White, 1972). This scale of 50 items, derived from the literature, covering physical care, psychological care, medical care and preparation for hospital discharge. Rather than measuring carer satisfaction with health care, it measures the quality of nursing (Johnson, 1987). For example, Hancock, et al. (2003) used it to assess the satisfaction of the carer with various aspects of nursing care.
- McCusker's scales to measure satisfaction (McCusker, 1984). McCusker developed 12 scales, with either 3 or 4 items per scale, to assess carer satisfaction with health care. Altogether there are 42 items assessing general satisfaction, availability of care, continuity of care, physician availability, physician competence, personal qualities of the physician, communication with the physician, preference for home care, preference for physician decisions, involvement in care decisions, freedom from pain and pain control. Many of the items do not measure 'satisfaction' (e.g. when the time comes, my relative would prefer to die in his/her own home) and the internal consistency of the scale was poor (Cronbach  $\alpha = 0.53-0.85$ ; only 3 scales achieved a carer  $\alpha > 0.70$ ).
- The SERVQUAL (Buttle, 1996; Parasuraman, et al. 1985; Parasuraman, et al. 1988), which is a 22-item scale for assessing consumer perceptions of service quality, where quality was defined as the consumer's judgement of the excellence or superiority of the entity being assessed. The purpose of the SERVQUAL was to track service quality trends, identify areas of concern, to identify and target market segments, and to assess service performance relative to that offered by competitors.
- Dennis, et al's (1997) modified version of Pound, et al's (1994) 20-item Patient Satisfaction with Stroke Services scale. The modifications were designed to make the scale appropriate for carers rather than patients. Although the modified items were listed in their paper (e.g. I think the ambulance service is reliable) no psychometric properties were reported and each item was treated as a separate dichotomous item for scoring.
- The Carer Assessment Scale (Mackenzie, et al. 1998) is a 14-item scale assessing caregiving areas that cause the carer difficulty, including handling incontinence, personal hygiene etc. Because these areas do not cover satisfaction with formal professional care this scale was excluded.
- The VOICES (Views of Informal Carers – Evaluation of Services) scale (Addington-Hall, 1998) consists of 160 items assessing various aspects of palliative care in the last year of life, including items on death place, nursing care, hospital, hospice, social services

provision and symptom control. Because it is administered retrospectively at 6 months after death of the care recipient, there are memory recall issues that suggest the results are imperfect (Addington-Hall and McPherson, 2001).

- Jacoby, et al.'s (1999) postal questionnaire of 36 items assessing carer satisfaction after the death of the care recipient. It was not reviewed for two reasons. It was designed for a specific palliative care intervention, and 12 of the items explicitly include reference to that service. Additionally, many of the items require the respondent to assess factual caring rather than satisfaction with caring (indeed, just one item assessed satisfaction per se (How satisfied with information given). Twenty-two of the items were dichotomous (e.g. Hospital staff did enough with the response categories Yes/ No).
- The PREPARED scale (Grimmer and Moss, 2001) was written to assess continuous quality improvement in the acute hospital setting. This scale reports on preparation of a patient for discharge from the hospital and was designed to assess the processes and outcomes of discharge planning. It has just one item assessing carer satisfaction (satisfied with community service).

The remaining scales or items are reviewed against the study criteria outlined below. They are (in alphabetical order):

- The Carer Satisfaction Questionnaire;
- The Carer Satisfaction with Community Services Questionnaire;
- The Carer Satisfaction Survey;
- The Consumer Expectations Perceptions and Satisfaction Scale (CEPAS);
- The FAMCARE (Family Satisfaction with Advanced Cancer Care) scale; and
- The Satisfaction with Care at the End of Life in Dementia Scale (SWC-EOLD).

In addition, single item assessments were reviewed.

### 11.2.3 The Review Criteria

The review criteria are those outlined Section 2 and in Table 1 of this report. Each criterion was weighted for its applicability to the Australian setting.

Although these criteria are used to rate each instrument, for ease of understanding the instrument review material has been organised to reflect basic psychometric axioms. Psychometric theory postulates that the valid and reliable measurement of a latent construct requires the construction of a manifest instrument that delivers an observed model which is isomorphic with the construct. To achieve this, the following axioms are widely accepted:

1. There should be a latent model of the construct, including an adequate description of its dimensions. For each dimension, there should be measurement items, such that the item content covers the dimension adequately. All items combined form the descriptive system of an instrument from which the manifest model is derived;
2. The resulting instrument should possess a nomological net of evidence suggesting validity (Cronbach and Meehl, 1955);
3. It should also be reliable and responsive; and
4. Instruments to be used with respondents who may be under stress or suffering mild cognitive impairment should be short and simple to minimise response burden (McGrath, et al. 2005).

Where there is a nomological net of evidence relating to each of these criteria, it may be inferred that an instrument is valid and reliable. Since validity and reliability are functions of both the instrument itself and the respondents who complete it, these are never fixed properties but may vary from sample to sample. The important corollary is that although there may be validity and/or reliability evidence for an instrument developed in, say, the USA, that same instrument may be invalid and/or unreliable in Australia due to cultural differences. It is accepted among psychometricians that this implies basic tests of validity and reliability need to be applied each time an instrument is used with a different population.

## **11.2.4 Review of Items and the Instruments**

### **11.2.4.1 Carer Satisfaction Questionnaire**

The Carer Satisfaction Questionnaire was developed for administration as part of the Victorian Department of Human Services' Consumer and Carer Satisfaction Survey – Aged Persons Mental Health Services (Gill and Maas, 2000).

The sections of the questionnaire cover availability of services (5 items), getting information (6 items), the service staff (5 items), treatment and assistance (5 items), participation (6 items), the hospital (8 items) and a general service rating section (3 items). The response set for all items is a Likert scale (very dissatisfied/ dissatisfied/ neither/ satisfied/ very satisfied). A separate don't know/ not applicable option is also provided. At the end of each section there is an open-ended question for additional comments. Scores across all sections can be factor weighted (so each section contributes appropriately), and then summed to form the Service Satisfaction Index, which is scored on a percentage scale. No psychometric details of the Carer Satisfaction Questionnaire were reported (Gill and Maas, 2000).

#### **Evidence of a latent construct**

No theory of carer satisfaction was reported.

#### **Validity evidence**

Content validity: No evidence was presented.

Construct validity: No evidence was presented.

Criterion validity: No evidence was presented.

#### **Reliability**

Reliability: No evidence was presented.

#### **Responsiveness**

Although mean scores were reported by service provider, no statistical analysis of these was undertaken and so it is uncertain if the Service Satisfaction Index is sensitive or not.

No significant differences in carer satisfaction were reported between two cross-sectional surveys carried out in 1999 and 2000 (Gill and Maas, 2000).

#### **Assessment against the study criteria**

No comparative data was reported. At 38 items this is a long scale. The instrument does not appear to have been used other than in the Victorian satisfaction surveys.

No psychometric data on the instrument is available.

There was no indication in the seminal paper of any copyright restriction on the instrument, so it may be assumed that it can be freely used, although acknowledgments should always be made. Administration and scoring costs are likely to be high given the structure of the instrument.

### 11.2.4.2 Carer Satisfaction with Community Services Questionnaire

Simon, et al.'s (2003) Carer Satisfaction with Community Services Questionnaire was developed to assess carers' satisfaction with community services providing support for carers caring for stroke patients at home. Interviews were conducted with carers caring for a stroke patient at home, and the transcribed interviews analysed to elicit key themes. Twenty-eight themes emerged, which were sorted into four dimensions: information and education, provision of practical help, convenience, coordination and adaptability of services, and consultation with and consideration of the carer. From the interview transcripts, 23 items representative of the themes were administered to a convenience sample of 40 carers; participating carers were re-interviewed two weeks later. Following psychometric analyses involving test of construct validity, test-retest, inter-rater reliability, internal consistency and factor analysis, 20 items were retained in 7 factors, being a mixture of positive and negative items.

The factors and items were: Information (satisfaction with given information about community services, understanding by the health professionals of the carer role, being consulted about the care recipient, information about financial help, where to get further information), health services (satisfied with total amount of help provided, choice in carer tasks, coordination among services used, satisfaction with given information about community services,<sup>37</sup> knowing what to do if reaching a crisis point), stroke information (satisfaction with information on stroke effects, satisfied with information on stroke), change and caring (changes made or equipment in place, health professionals concerned about the carer), listening to carer (confident services would adapt if the situation changed, opinion taken notice of), problem management (like more help applying for benefits and services, satisfied with help received), confidence in information (received information is accurate, knowing who to contact if there was a problem). The response scales were Likert scales (agree-disagree, coded 0 – 4). Missing data were assigned the neutral value (2).

#### Evidence of a latent construct

No theory of carer satisfaction was reported (Simon, et al. 2003), and the development of the items, although clearly informed by the literature, appears to have been purely in response to the themes elicited during the interviews.

#### Validity evidence

Content validity: No evidence is presented.

Construct validity: This was assessed through factor analysis of the items. The resulting model (7 factors) did not support the hypothesized model (4 dimensions). An additional difficulty was that of the 7 factors, 5 consisted of just two items.

Criterion validity: The correlation between the total scale score and a single item assessing carer satisfaction was  $r_s = 0.80$ .

#### Reliability

The reliability of the seven factor scales was Cronbach  $\alpha = 0.80, 0.76, 0.79, 0.55, 0.48, 0.73$  and  $0.46$ , respectively, for information, health services, stroke information, change and caring, listening to carer, problem management, and confidence in information. For the full scale  $\alpha = 0.86$ .

Test-retest at 2-weeks ( $n=17$  carers) was  $r_s = 0.89$ . Inter-rater reliability was  $r_s = 0.87$ .

#### Responsiveness

No evidence is presented.

#### Assessment against the study criteria

No comparative data was reported. At 20 items this is a moderately long scale – and the item stems are also long. Although the instrument was developed for use with stroke caregivers, most

---

<sup>37</sup> Item appears under both Information and Services factors.

of the item stems are generic. To use with other health conditions 3 items would need to be modified. The instrument does not appear to have been used since publication.

The reliability evidence is rather mixed. On the one hand, internal consistency is excellent for the whole scale; against this half of the sub-scales are unreliable. The very high test-retest and inter-rater reliability estimates may suggest that the scale is likely to be non-responsive in longitudinal studies.

The validity evidence is unsatisfactory. There are two problems with the factor analysis. The sample size ( $n=40$ ) was simply too small for a stable factor analysis: calculated from the data in the paper,  $Y = 0.16$  which exceeded the requirement for stability ( $Y \leq 0.10$ ) (Guadagnoli and Velicer, 1988) leading to data overfitting and the identification of non-generalisable factors. The second problem is the retention of factors with  $<3$  items. With few dissensions, since Thurstone (1947) it has been accepted that for stable factors in an R-model of factor analysis, such as that employed in the paper, the minimum number of items loading on a factor is 3 (based on the 'construct ratio' rule); it is therefore likely that the factors with 2 items are inherently unstable.

There was no indication in the seminal paper of any copyright restriction on the instrument, so it may be assumed that it can be freely used, although acknowledgments should always be made. Administration and scoring costs are likely to be moderate.

#### **11.2.4.3 The Carer Satisfaction Survey**

The 10-item British Carer Satisfaction Survey (also referred to as the carer version of the Pound Satisfaction Scale) was developed to assess carers' satisfaction with stroke services in the hospital and after discharge. The assumption behind the scale is that the patient will recover and be returned home (Pound, et al. 1993). The basis of the scale was interviews with 6 carers and a literature review. Nine items were drafted and completed by 15 carers. The nine items are divided into two scales, one for carers of care recipients still in hospital (4 items) and the other for when the care recipient has been discharged and returned home (5 items). There is also a general carer satisfaction item with a 7-point response set presented as a series of happy faces.

The hospital scale covers the care recipient being treated with kindness by hospital staff, having his/her needs attended to, that the staff did everything possible to make the care recipient well again, and that the hospital staff recognise the difficulties of caring for a person with stroke. The discharge home scale assesses the carer being given information about allowances and services he/she might need access to upon discharge of his/her care recipient, that the home was physically well prepared, that the carer is satisfied with hospital outpatient services, that the ambulance service was good, and that the carer received all the support he/she needed. All the scale items have a 4-point forced choice response scale (strongly agree/ agree/ disagree/ strongly disagree). Scores are summed across the items.

#### **Evidence of a latent construct**

No theory of carer satisfaction was reported (Pound, et al. 1993), and the development of the items, although clearly informed by the literature, appears to have been purely in response to the themes elicited during the 6 interviews.

#### **Validity evidence**

Content validity: No evidence is presented.

Construct validity: No evidence was reported.

Criterion validity: The correlation between the hospital scale and the global item was  $r = 0.59$  and  $r = 0.68$  for the discharge home scale. It correlated with the sleep subscale of the Nottingham Health Profile ( $r = -0.24$ ), and there was also a significant correlation between discharge home scale and the care recipient's Barthel Index ( $r = 0.25$ ).

## Reliability

The reliability was Cronbach  $\alpha = 0.87$  for the hospital sub-scale and 0.79 for the discharge scale (Pound, et al. 1993). Test-retest reliability at 2-weeks ( $n=17$  carers) was  $r_s = 0.89$ . Inter-rater reliability was  $r_s = 0.87$  (Pound, et al. 1993).

## Responsiveness

Validation of the questionnaire was through distribution to 219 carer households and 103 returns received. Of the 99 carers who completed the hospital scale, 76 reported being satisfied or very satisfied with all aspects of hospital care; and for the discharge home scale of the 75 respondents 29 carers reported being satisfied with all aspects of home care support; thus there was a lower level of satisfaction with home care support when assessed against hospital care.

Re-analysis of the original Pound, et al. dataset by Gompertz, et al. (1995) failed to show any significant difference between two area health services, despite differences in health care costs between the two districts. Similarly, no significant differences were reported among stroke carers by Mant, et al. (2000), Rodgers, et al. (1999), or at 12-month follow-up by Tilling, et al. (2005).

## Assessment against the study criteria

No comparative data was reported. At 10 items this is a short scale. Although the instrument was developed for use in stroke caregivers, all of the item stems are generic, except one which would need modification if the Carer Satisfaction Survey was to be used in dementia studies.

The evidence for the reliability of the Carer Satisfaction Survey is rather mixed. Although the seminal paper reported that, overall, it was good; the only other paper identified which reported reliability provided an estimate that was just satisfactory. No validity evidence was identified for the scale. No costs were identified for using the Carer Satisfaction Survey and administration and scoring costs would be very low.

### 11.2.4.4 The Consumer Expectations Perceptions and Satisfaction Scale (CEPAS)

The 20-item Australian CEPAS (Spear, 2003) was based on the results of a focus group of patients and carers, and has six scales assessing access, respect, reliability, responsiveness, empathy and participation which are collapsed into three scales assessing expectations, perceptions and satisfaction. There are two additional global items assessing meeting expectations and how the patient felt with his/her experience of the health care service. A typical item is "Did you have to wait too long for help". The original response scales are not reported; the mean satisfaction score was reported to be 98%. Following modification, the response scales are from dissatisfied/ to extremely satisfied; the mean scores on this modified response set are not reported.

A feature of the CEPAS is that each of the dimensions probes expectations, support and perceptions. For example, the three items under access are: Do you expect the service to be convenient for you?/ Did you get the help you wanted?/ and How do you feel about how easy it was to get help?

The authors report that the CEPAS is suitable for use by carers, although no evidence on this point is presented (Spear, 2003).

## Evidence of a latent construct

No evidence is presented. The authors note that the results of the focus group were similar to published literature reviews.

## Validity evidence

Content validity: No evidence is presented, other than the comparison with the literature review.

Construct validity: No evidence is presented. The correlation between the perception and satisfaction scales is reported to be 0.85, suggesting they are measuring the same construct. Expectations and perceptions were correlated  $r = 0.56$ .  
Criterion validity: The CEPAS correlated with the Client Satisfaction Questionnaire  $r = 0.67$ .

### **Reliability**

The reliability of the three scales was Cronbach  $\alpha = 0.80$ ,  $0.82$  and  $0.81$  for expectations, perceptions and satisfaction, respectively. The weighted kappa for individual items was between  $0.63$  and  $0.77$ .

### **Responsiveness**

No evidence is presented.

### **Assessment against the study criteria**

For comparative data, the CEPAS was correlated against the Client Satisfaction Questionnaire. At 20 items this is a moderate-length scale. The reliability evidence suggests the three sub-scales are reliable, there is, however, virtually no validity evidence to support the instrument. No evidence is presented that the scale is really suitable for use with carers.

No copyright restriction on the use of the instrument was identified. Administration and scoring costs should be low.

#### **11.2.4.5 The FAMCARE (Family Satisfaction with Advanced Cancer Care) Scale**

The FAMCARE (Family Satisfaction with Advanced Cancer Care) scale (Kristjanson, 1993) is a 20-item instrument measuring the degree to which the patient's family are satisfied with the health care provided and the providers' behaviours directed towards the care recipient and the family.

The FAMCARE scale was developed through interleaving the literature on family satisfaction with the findings from interviews with families of hospice patients (Kristjanson, 1993). The resulting item bank was administered to a sample of family caregivers who were asked to Q-sort the items from most to least important. The 20 most important items formed the basis of the FAMCARE. Cluster analysis was used to group these items into the sub-dimensions of care (described by Kristjanson as the referential level of measurement) were care availability, physical care, psychosocial care and information giving. The response scales for all items are 5-point Likert scales (very satisfied/ satisfied/ undecided/ dissatisfied/ very dissatisfied). None of the items are reversed. The 20 items cover satisfaction with pain relief, prognosis information, health professional's answers to questions, side effect information, specialist referrals, hospital bed availability, family conferences, treatment speed, the doctor's attention, the way tests and treatments are performed, availability of doctors and nurses to talk with the family, care coordination, time taken for a diagnosis, inclusion of family in decision-making, information given on pain management, information given on patient's tests, thoroughness of assessing the patient's symptoms, clinician follow-up of tests and treatments, and the availability of the clinician to the patient.

In a study of palliative home care support Jarvis, et al. (1996) used a modified version of the FAMCARE, removing items they felt were extraneous; these items were those covering clinician availability to the patient, specialist referrals, time taken for a diagnosis, and assessing the patient's symptoms. Of the remaining 16 items, just 4 were sensitive to palliative care compared with long-term care (side effect information, thoroughness of assessing the patient's symptoms, availability of doctors and nurses to talk with the family, and information given on patient's tests).

### **Evidence of a latent construct**

The FAMCARE is based on Ajzen and Fishbein's conceptualization of attitudes towards objects (Ajzen and Fishbein, 1980), which was operationalised in this instance as satisfaction with health care, specifically with advanced cancer care. This was operationalised through a literature review

and carer interviews and administration of the resulting item bank to a construction sample and cluster analysis used to identify sub-dimensions.

### **Validity evidence**

**Content validity:** As outlined above the content appears to excellent. Given the construction procedures, the FAMCARE has excellent ecological validity.

**Construct validity:** Discriminant analysis was used to identify the sub-dimensions of the construct. Insufficient detail is given in the seminal paper to assess the strength of the model, but several problems were reported by Kristjanson (1993), including redundancy of several items and that in the pilot test sample the original 4-sub-dimension structure was not entirely replicated; although the first three factors were identified, the 4<sup>th</sup> factor was concerned with pain rather than information.

Ringdal, et al. (2003) examined the internal structure of the FAMCARE scale in a sample of 181 family members, where the questionnaire was mailed out to families 1 month after the death of the care recipient. One of the items (#14: time taken to make a diagnosis) was excluded from analysis because it failed to correspond with the other items (e.g. it was the lowest loading item on the principal component and had a poor Loevinger H). Factor analysis suggested all remaining 19 items loaded >0.30 on the principal component, and when analysed for unidimensionality using Mokken analysis the Loevinger H was 0.59, indicating a strong unidimensional scale. On the basis of these results, Ringdal, et al. suggested that the FAMCARE could be revised and redundant items removed in future research.

**Criterion validity:** The FAMCARE was correlated  $r = 0.77$  and  $r = 0.80$  with the McCusker satisfaction scales. It was also correlated in different samples with a global satisfaction item,  $r = 0.62-0.64$  (Kristjanson, 1993). In a study of carers for those with metastatic cancer, the FAMCARE correlated with unmet carer needs  $r = -0.46$  and  $-0.48$  for spouse and non-spouse caregivers (Hwang, et al. 2003).

### **Reliability**

The reliability of the FAMCARE, as assessed by internal consistency, was reported to be Cronbach  $\alpha = 0.93$ , and for the sub-scales it was 0.82 (information), 0.84 (physical care), 0.83 (psychosocial care) and 0.73 (care availability) (Kristjanson, 1993). In a later study of caregivers across 4 Canadian states Kristjanson, et al. (1997) reported Cronbach  $\alpha = 0.90$ . Elsewhere the internal consistency has been reported to be Cronbach  $\alpha = 0.95-0.96$  (Hwang, et al. 2003; Ringdal, et al. 2003).

Test-retest reliability at 24-hours assessed by correlation was  $r = 0.92$  (Kristjanson, 1993).

### **Responsiveness**

Scores systematically vary on the FAMCARE by educational status, care recipient age and time since diagnosis (Kristjanson, 1993; Kristjanson, et al. 1997). A Norwegian study of palliative care versus conventional care showed that there were significant differences on 10/20 FAMCARE items, generally favouring the palliative care treatment carers; there was an overall statistically significant difference in total FAMCARE scores (Ringdal, et al. 2002).

No significant differences were reported between spouse and non-spouse caregivers in a study of metastatic cancer (Hwang, et al. 2003), or between a palliative care team intervention and telephone support (Hanks, et al. 2002).

### **Assessment against the study criteria**

There is some comparative data available correlating the FAMCARE with the McCusker Satisfaction Scales and a global satisfaction item. The correlations were suggestive of a reasonable relationship. At 20 items the FAMCARE is a moderately long instrument, although it

should be noted that at least 2 research teams have suggested it contains redundant items which could be usefully removed to shorten it.

The reliability evidence reviewed above is excellent; although it is possible this may in part reflect the redundancy described above. The validity evidence is consistent with this interpretation. Although Kristjanson, et al. (1993) reported a 4-dimensional structure for the FAMCARE, this has not been supported by other researchers. Indeed, it would appear that the FAMCARE is a unidimensional scale rather than a multidimensional one (Ringdal, et al. 2003).

Although written for cancer caregivers, there is no reason the FAMCARE could not be used in dementia studies. No copyright restriction on the use of the instrument was identified. Administration and scoring costs should be moderate.

#### **11.2.4.6 Satisfaction with Care at the End of Life in Dementia Scale (SWC-EOLD)**

Designed in response to the US MediCaring National Demonstration and Evaluation Project for those dying of chronic disease (e.g. congestive heart failure) (Skolnick, 1998), the SWC-EOLD scale assesses the quality of care during the last 90 days of life (Volicer, et al. 2001). Fifteen items were administered to a convenience sample of caregivers (n=156, response rate 27%) whose care recipient had died of terminal dementia within the previous year. Inspection of the results showed that 1 item was poorly endorsed and it was deleted. Factor analysis showed a unidimensional scale, and 4 items were deleted because the corrected item-total correlations were >0.80, i.e. they did not influence the scale's reliability.

The questionnaire itself consists of 10 items assessing being fully involved in decision-making, the provision of information, keeping the care recipient comfortable, the health professionals being sensitive to the needs/feelings of the carer, understanding the care recipient's condition, knowing which doctor/nurse was in charge of providing care, that the care recipient received all necessary nursing assistance, that medication issues were clearly explained, that the care recipient received all treatments/interventions that he/she could have benefited from, and that the care recipient received the best treatment at the end of his/her life. Three of the items were negative, and needed to be reversed prior to scoring. The response set was a forced choice scale (strongly disagree/ disagree/ agree/ strongly agree); and a not applicable option. The correlations between items have been reported to be 0.34 – 0.60 (Kiely, et al. 2006). The score range is 10-40, with higher scores indicating greater carer satisfaction. A feature of the scale is that scores approximate a normal distribution (Engel, et al. 2006; Kiely, et al. 2006; Volicer, et al. 2001).

#### **Evidence of a latent construct**

No evidence was presented in the seminal paper (Volicer, et al. 2001).

#### **Validity evidence**

Content validity: No evidence is presented. The authors state that most of the items were taken from other scales or the views of experts. No further information is given (Volicer, et al. 2001).

Construct validity: No evidence is presented.

Criterion validity: The SWC-EOLD was correlated with two other scales, measuring symptom management at the end of life in dementia (SM-EOLD) and comfort assessment in dying with dementia (CAD-EOLD); the correlations were 0.28 and 0.30, respectively (Volicer, et al. 2001). Kiely, et al. in a study of dementia care dyads report it correlated 0.81 with the Decision Satisfaction Inventory (a measure of satisfaction with medical decision-making).

#### **Reliability**

The reliability of the SWC-EOLD in the construction sample was Cronbach  $\alpha = 0.90$ , elsewhere among dementia care dyads it was 0.83 (Kiely, et al. 2006).

## **Responsiveness**

In a study of end-of-life care for those with advanced dementia, scores on the SWC-EOLD systematically varied by care planning, symptom management, dementia ward status and tube feeding (Engel, et al. 2006).

## **Assessment against the study criteria**

No comparative data were identified comparing the SWC-EOLD with another carer satisfaction scale. At 10 items this is a short, unidimensional scale. There is limited evidence on its reliability, although this evidence appears to be satisfactory. There is almost no satisfactory validity evidence available for the scale. As acknowledged by the authors, much further work needs to be done on this scale.

No copyright restriction on the use of the instrument was identified. Administration and scoring costs should be low.

### **11.2.4.7 Single Item Assessments**

Single item assessments can be grouped into closed and open questions.

#### **Closed questions**

Hancock, et al. (2003) used a single closed question for carers to complete asking about their satisfaction with the quality of care given in respect of four different aspects of nursing (physical care, psychosocial care, doctor's orders and discharge planning). The response scale was 1 = poor, through to 5 = excellent. Luscombe, et al. (1998) used a single item of satisfaction with services, with a 3-point rating scale: poor, fair, good. Nicoll, et al. (2002) used a single item "How satisfied are you with the respite care that the person you care for receives?" and the response set was a 7-point Likert response scale from very dissatisfied to very satisfied.

Bekelman, et al. (2005) asked four questions of carers whose care recipient had recently died: had the care recipient seen a doctor in the month prior to death, could more have been done to keep the deceased comfortable, if more information had been available would treatment decisions have changed, and an assessment of the overall treatment quality. Lincoln, et al. (2004) devised four questions covering satisfaction with knowledge, practical help, emotional support and overall satisfaction. Wellwood, et al. (1995) used 10 different questions to ask about satisfaction with care; no attempt was made to combine these into a scale.

Melzer, et al. (1996) used a combination of open-ended qualitative questions and a closed single item assessing the carer's opinion of the overall quality of services received. Although the response scale is not given the researchers state that it contained very good, very bad, moderate and good.

Chesterman, et al. (2001), in a longitudinal study, used three different versions of the same single item at each of three time points: satisfaction with the way social services assessed and tried to help with recent problems, satisfaction at time 2 with the level of services received, and experience of services at time 3 during the past 6 months. Each version had a different response set. At time 1 the response set was very satisfied/ satisfied/ mixed feelings/ dissatisfied/ very dissatisfied; and time 2 it was very satisfied/ satisfied/ neutral/ dissatisfied/ very dissatisfied, and at time 3 it was favourable/ mixed/ unfavourable.

#### **Open questions**

Schneider, et al. (1999) used an open-ended question to ask about the carer's perception of the support received. Simpson, et al. (1995) used open-ended questions to probe the best and worst aspects of hospital care. Montalto (1996) used four open items in a study of hospital-in-the-home

to probe what carers disliked about the program, liked about the program, and to offer the opportunity for carers to make comments in relation to nursing care and medication, and Koffman and Higginson (2001) in a study of district nursing used six questions to probe carer satisfaction covering reassurance and support, the district nurse providing enough time, the understanding of the general practitioner, the extent to which the deceased person had treatment choices, the doctors providing enough time, and if the carer was able to find out all the needed information regarding the care recipient's condition.

### **Use of a carer diary**

A variation on the above was the use of a carer diary (Simpson, 1997).

### **Assessment against the study criteria**

Examining the above papers, none report any psychometric or measurement properties for the items used in these studies. Plus none of these papers meet any of the study criteria. As this suggests, most researchers did not report the actual item used or the response or the frequency distributions. For these reasons single items are not considered further in this report.

## **11.2.5 Discussion and Recommendations**

This review has examined self-report instruments designed to assess carer satisfaction with health services. Importantly, it has not reviewed instruments designed for quality assurance of health care services. Although at first this distinction may seem somewhat artificial, as shown in the introduction it has a profound impact on the type of instrument used to assess satisfaction with health services.

The original spurs for the development of the instruments reviewed in this study were the perceived need for the provision of appropriate care, as assessed by family caregivers, for those with chronic incurable or terminal illnesses living in hospices or for community support or respite care for those carers providing informal family care at home.

Although researchers and clinicians have been aware of carer satisfaction from this perspective since the early 1960s, it wasn't until the 1980s that this was systematically explored. As several of the papers reviewed in this study have made clear, even today the concept is poorly developed, it is inadequately operationalised and there are very few instruments designed to measure it. The purpose of this review, within this general framework, was to assess and recommend carer satisfaction instruments for use in Australian studies of geriatric care for those with cognitive impairment, usually from dementia.

A search of the leading databases, MEDLINE, CINAHL and PsycINFO, led to the identification of six scales for review: the Carer Satisfaction Questionnaire (Gill and Maas, 2000), the Carer satisfaction with community services questionnaire (Simon, et al. 2003), the Carer Satisfaction Survey (Pound, et al. 1993), the Consumer Expectations Perceptions and Satisfaction Scale (CEPAS) (Spear, 2003); The FAMCARE (Family Satisfaction with Advanced Cancer Care) scale (Kristjanson, 1993) and the Satisfaction with Care at the End of Life in Dementia Scale (SWC-EOLD) (Volicer, et al. 2001). In addition, single item assessments were reviewed.

The literature for each scale was obtained and reviewed against the study criteria. The results are summarized in Table 62, where rankings against the criteria are weighted by their assessed importance in Australian settings.

**Table 62 Summary Assessing Carer Satisfaction Instruments Against the Study Criteria**

Criteria	Weight	Instrument					
		SWC	CSS	FAMC	CSCS	CEPAS	CSQ
Availability of comparison data	3	1	1	1	1	1	1
Length/feasibility of instrument for inclusion in battery	2	3	3	2	2	2	1
Complexity of administration/ cognitive burden	2	3	2	2	2	1	1
Cultural Appropriateness	1	1	1	1	1	1	1
Ease of obtaining score	2	3	3	2	2	2	1
Sensitivity to dementia	3	3	1	1	1	1	1
Reliability evidence	3	2	2	2	2	2	1
Validity evidence	3	1	1	2	1	1	1
Cost of the instrument	2	3	3	3	3	3	3
Cost of instrument administration	2	3	3	3	3	3	1
<b>Weighted Total</b>		<b>52</b>	<b>44</b>	<b>43</b>	<b>40</b>	<b>38</b>	<b>27</b>

### Availability of comparison data

No head-to-head comparative studies were identified; i.e. there is no information concerning whether the different instruments are measuring the same underlying construct or whether any one instrument outperforms other instruments.

### Instrument length

Parsimony is important for reasons of enabling inclusion in instrument batteries and for psychometric reasons related to clarity of what is being measured. The length of instruments reviewed ranged from 10-items to 38 items. The shortest instruments were the Carer Satisfaction Survey (Pound, et al. 1993) and the SWC-EOLD; each has just 10 items.

### Complexity and cognitive burden

Complex instruments should be avoided, especially where the respondents may be under stress or suffering mild cognitive impairment. This suggests that instruments should have simple and straightforward items and scoring systems for both ease of administration and to minimise cognitive burden.

Some of the reviewed questionnaires were unduly long and complex, for example the Carer Satisfaction Questionnaire (Gill and Maas, 2000) necessitated interviewer administration and contained items that were unnecessarily complex, such as: How satisfied were you with the information you were given about practical matters on the ward (for example, visiting times)? Similarly, the Carer Satisfaction with Community Services Questionnaire (Simon, et al. 2003) also had several long and complex items, like this: You feel you'd know where to get more information on any topic related to stroke or caring for a person who has had a stroke if you needed it.

The simplest instruments were the Carer Satisfaction Survey (Pound, et al. 1993), the CEPAS (Spear, 2003), the FAMCARE (Kristjanson, 1993), and the SWC-EOLD (Volicer, et al. 2001).

### **Cultural appropriateness**

There is no reported research involving any of these instruments examining whether the construct of carer satisfaction is culturally bound in any way; there does not appear to have been any cross-cultural validation work done on any of the measures reviewed. In short, there was no evidence for any instrument on this criterion referring to appropriate use by CALD or illiterate clients or with an interpreter. All instruments were therefore ranked similarly.

### **Ease of scoring**

Scoring ease will assist with instrument acceptance in the field by clinicians. Accordingly, those instruments with simple scoring algorithms which, if necessary, can be applied during interview are preferred. All the instruments appeared to be easy to score through simple summation of responses. The shortest instruments would be the easiest to score; these were the Carer Satisfaction Survey (Pound, et al. 1993) and the SWC-EOLD (Volicer, et al. 2001).

### **Sensitivity to dementia**

No studies, except one, were identified which reported the use of any of the instruments in samples of carers with mild cognitive impairment, yet there is evidence in the literature that a high proportion of older adult caregivers suffer mild dementia or cognitive impairment themselves. The exception to this was the SWC-EOLD (Volicer, 2001) which was developed in a population of carers for those with terminal dementia.

### **Reliability evidence**

In general the evidence on reliability was mixed. No evidence was available for some instruments (e.g. the Carer Satisfaction Questionnaire (Gill and Maas, 2000)), for others there was some rather mixed evidence available (e.g. the Carer satisfaction with community services questionnaire (Simon, et al. 2003) or the Carer Satisfaction Survey (Pound, et al. 1993)). The instruments with more consistent reliability evidence were the CEPAS (Spear, 2003) and the SWC-EOLD (Volicer, et al. 2001). Although the FAMCARE (Kristjanson, 1993) had the most reliability evidence of any of the instruments, there was evidence that this was obtained through redundancy.

### **Validity evidence**

As shown in the detailed instrument reviews above, there was considerable variation in the available validity evidence. What is striking about the literature is the, generally, limited approach adopted by instrument designers to the importance of this aspect of their work.

No validity evidence, or very little, was available for the Carer Satisfaction Questionnaire (Gill and Maas, 2000), Carer satisfaction with community services questionnaire (Simon, et al. 2003), the Carer Satisfaction Survey (Pound, et al. 1993), the CEPAS (Spear, 2003) or the SWC-EOLD (Volicer, et al 2001).

The only instrument for which there was a reasonable amount of validity evidence was the FAMCARE (Kristjanson, 1993) – and this evidence suggested that the internal structure of the instruments was not fully supportive of the original hypothesized model.

### **Instrument costs**

None of the instruments reviewed appears to have been commercialized; no commercial websites were identified for any of the instruments and no copyright costs were identified. All the instruments appear to be available free to users, subject to journal copyright permissions.

## Instrument administration costs

All the instruments were designed for self-completion, except for the Carer Satisfaction Questionnaire (Gill and Maas, 2000). Importantly, none of the instruments reviewed here appear to have been tested for the effects of administration mode or proxy-report.

### 11.2.6 Recommendations

Given the review findings, none of the reviewed instruments can be given an unqualified recommendation for use in Australian studies with carers of older adults who have cognitive impairment or dementia.

1. The most promising instrument appears to be the SWC-EOLD (Volicer, et al. 2001), and it is recommended that this instrument is used in an Australian study specifically designed to test its measurement properties. As part of this testing, the 4-point response set currently used with the instrument should be replaced with a conventional 5-point response set (strongly disagree/ disagree/ neither agree nor disagree/ agree/ strongly agree).<sup>38</sup>

2. The alternative would be to mount a specific carer satisfaction study, where all items from all reviewed instruments were pooled and tested. The explicit purpose would be identifying well performing items and/or the best performing instrument.

### 11.3 Other Informal Care Outcome Measures

The scope of this project has been confined to an examination of carer satisfaction with health services and thus a detailed review of a number of important informal carer outcome measures was not included in the scope of this report. It is recommended that such a review could form a follow-up project. This section provides an overview of the types of carer outcome measures that have been used in dementia research and briefly discusses some of the interrelationships between these measures.

People with dementia living in the community rely on a substantial informal care contribution (Langa, Chernew, Kabeto, Herzog, et al. 2001). It can therefore be expected that health and social care programs designed to prevent or defer institutionalisation of people with dementia are

---

<sup>38</sup> A brief note on the number of response options in an attitude measurement response set.

Volicer et al (2001) opted to use a 4-point response set for the SWC-EOLD to keep it consistent with previous work. Four-point response sets like this are known as 'forced choice' response sets because they force a respondent to either agree/disagree with the item stem. The limitation of this approach is that it equates (in scoring) those with a neutral or ambivalent position with those with a positive or negative position (i.e. those who do not hold a position very strongly).

Likert (1932, 21), who is credited with developing the conventional 5-point response set (strongly disagree/ disagree/ neutral/ agree/ strongly agree), argued that "...five-point statements... yielded a distribution resembling a normal distribution... it seems justifiable for experimental purposes to assume that attitudes are distributed fairly normally and to use this assumption for combining different statements". Likert et al (Likert et al 1934) described the non-neutral positions as the "extreme alternative" and "intermediate alternative" responses and referred to the neutral position as the "interrogation point". The semantic distance between the interrogation point and one of the alternatives marked how far away from a neutral position the respondent was.

Guttman (1954) demonstrated that responses to an attitude item comprised three distinct components: the response to the item content, the intensity with which this response is held, and the feelings (which Guttman called 'closure') which the respondent has about the item. He showed that there was a relationship between these that indicated a neutral response position around which the respondent had almost no feelings of intensity (the 'zero point'), and that intensity and closure were both symmetrical around this neutral point. In forced choice scales respondents are denied the zero point and the closures are systematically distorted thus leading to endorsement of a position that the respondent may not actually agree with (i.e. many people do not hold strong attitudinal positions). In his review of response sets, Foddy (1993, 111) come to the conclusion that "There is no justification for the practice of collecting and pooling answers that are uninterpretable in the sense that... they mean quite different things to different respondents. For this reason.... 'No opinion' options should always be included, because they generate additional information...".

In the interests of disentangling those with a neutral position from those with a positive attitude towards health care it would seem advisable to use a Likert scale rather than a forced choice scale.

likely to have consequences for their supporting family and friends. Thus, in addition to patient outcomes, program evaluation must take account of the impact on these informal carers. The Dementia Initiative which aims to support people with dementia to remain at home and to help people with dementia and their carers by enhancing their health outcomes and quality of life is such a program. Consequently, measurement of a comprehensive range of informal carer outcomes is essential for program evaluation. In addition, the on-going monitoring of carer outcomes would facilitate early intervention to protect carer welfare or prevent care recipient institutionalisation. Moreover, in the dementia context, informal care-giving and its consequences do not necessarily end with the institutionalisation of the care recipient, as carers continue to visit, support and have legal responsibility for the care recipient and may suffer on-going emotional and mental health consequences related (at least in part) to the institutionalisation decision (Schulz, Belle, Czaja, McGinnis, et al. 2004; Chene, 2006).

There are three main areas of research to measure informal carer outcomes encompassing measures of the carer's experience and perceptions of care-giving, measures of the carer's health and well-being and measures of the carer's satisfaction with services targeting the care recipient or the carer. These three types of measures are usually not unrelated and in practice are sometimes overlapping.

### 11.3.1 Carers' Experience

There are many measures which have been developed to quantify the experience of informal carers providing care and support for those with a chronic illness or disability. These have been based on different concepts which range from a focus on negative aspects of care-giving such as burden, hassles, stress or strain to the neutral appraisal or the positive aspects such as satisfaction, esteem or uplift (Hunt, 2003).

Carer burden is one of the earliest (for example Zarit, Reever and Bach-Peterson, 1980) and more widely applied concepts, defined as "the extent to which caregivers perceived their emotional or physical health, social life and financial status as suffering as a result of caring for their relative" (Zarit, Todd and Zarit, 1986). The burden concept has subsequently been developed to distinguish between objective and subjective burden. Objective burden comprises the observable demands on the carer while subjective burden encompasses the carer's feelings in response to those demands. However, empirically the two are usually confounded and, while most burden questionnaires include objective and subjective burden aspects, many cannot be scored separately (for example the Burden Interview, Zarit, Reever and Bach-Peterson, 1980). Further developments have led to the understanding of carer burden in terms of the stress paradigm (for example Pearlin, Mullan, Semple and Skaff, 1990), where the demands of care-giving, such as patient dependency or problem behaviours, are seen as primary stressors; family and role conflict as secondary stressors; and the carer's psychological well-being a stress outcome which is mediated by access to psychic, social and material resources.

Although the application of burden has commonly been as a negative concept, some authors have included positive subjective aspects. Carer satisfaction with the care-giving role is one of the more commonly applied positive concepts and is usually used to describe the benefits or positive consequences of care-giving (for example the Carer's Assessment of Satisfaction Index, Nolan, Grant and Keady, 1996), while caregiver appraisal refers to the carer's assessment of both the demands of her care-giving situation and her capacity to cope with those demands. The assessment may be positive, negative or neutral as is exemplified in the Caregiver Appraisal Measure (Lawton, Kleban, Moss, Rovine, et al. 1989).

While much of the development of measures of carer burden has taken place in the dementia context, their sensitivity to the effectiveness of interventions aimed at reducing carer burden in dementia has been limited (Acton and Kang, 2001). This is possibly related to the broad and varied way in which the concept has been defined and operationalised. See Vitaliano, et al.

(1991), for a review of burden measures used in dementia and their conceptualisation. Also, when statistically significant effects have been demonstrated, the interpretation and clinical importance of the effect has often been doubtful (Schulz, O'Brien, Czaja, Ory, et al. 2002; Sorensen, Pinquart and Duberstein, 2002).

### **11.3.2 Carer Health and Well-being**

There are a range of generic measures of health and well-being which have been applied as informal carer outcome measures. These include measures of: overall well-being, life satisfaction, quality of life, health related quality of life, mental health, physical health and health state utility. Studies of informal care in dementia have commonly applied at least one measure of psychological morbidity such as a depression scale.

While it is important to prevent or treat carer morbidity associated with care-giving, it might also be argued that the responsibility of health services to minimise the negative impacts of care-giving goes beyond impacts on health. Braithwaite (1992) argued that the impacts of care-giving for the elderly and disabled, viewed as frustration of carers' basic needs (for example sleep and rest, financial security, relationships) lead to justification of state interventions and support for carers on the basis of their welfare rights. Further, a population of carers which is to a large extent without disease may pose difficulties for demonstrating the effectiveness of beneficial interventions, if only health related outcome measures are used.

A study of the carers of patients with Alzheimer's disease (Bell, Araki and Neumann, 2001) suggests that generic health state measures might be insufficiently sensitive to differences in carer well-being in this context; the Health Utilities Index Mark-2 did not differ across patient disease stage or care setting and the SF-36 Mental Component Summary was slightly lower (less than 0.5 of a standard deviation) for carers of community residing patients with moderate and severe disease relative to carers of patients with mild disease and carers of institutionalised patients with all levels of severity. Recently, there has been a move to the development of instruments to measure care related quality of life (Brouwer, van Exel, van Gorp and Redekop, 2006) and care related quality of life in specific contexts such as palliative care (Cohen, Leis, Kuhl, Charbonneau, et al. 2006). This might be a potential area for future research in dementia.

### **11.3.3 Carer Satisfaction with Services**

Studies of carer interventions have reported high levels of satisfaction as has been indicated earlier in this section. However, it is noted that reported satisfaction with services may be confounded by the carer's relationship with the provider and, in the case of educational interventions, only satisfied carers tend to complete the intervention (Farran, 2001).

### **11.3.4 Relationships Among Measures**

A number of studies have investigated the relationships between the different informal carer outcome measures in the carers of people with dementia or carers of the elderly in populations which included a substantial proportion of dementia patients. These have principally focused on the relationship between carer burden and carer health, well-being or quality of life, with little available evidence to link carer satisfaction with services to other carer outcome measures. One study of carers of people with dementia found satisfaction with respite services was generally high and was associated with social support but not carer strain or depression, however the study response rate was low (20%; Nicoll, Ashworth, McNally and Newman, 2002).

Stull, et al. (1994) conceived carer burden and well-being as interim outcome measures to predict service use and found care-giving tasks to be correlated with carer burden but not with generic

well-being. Burden was a better predictor of the use of day care facilities and recently considering nursing home placement, while a generic measure of social participation was a better predictor of the use of homemaker services than burden.

Chappell and Reid (2002) conceived carer burden as a potential predictor of carer well-being and found that, while increased burden was associated with poorer well-being, self-esteem and informal care hours predicted both. Self esteem was positively associated with well-being and negatively associated with burden while the reverse applied to informal care hours. In addition, higher levels of care recipient behaviour problems predicted worse burden and higher levels of perceived social support predicted better well-being. Coen, et al. (2002) found that higher levels of carer burden were associated with more care recipient behaviour disturbance and worse carer quality of life. The quality of life dimensions associated with carer burden included satisfaction with time for self, finances and marriage but not satisfaction with health.

Fritz, et al. (1997) found that carer burden, life satisfaction and depression appeared to measure separate but related concepts. Care recipient memory impairment and disruptive behaviours predicted carer burden as did carer satisfaction with family support. The number of daily care tasks was associated with life satisfaction, and participation in group activities was associated with depression. All three measures were associated with carer self-rated health, and life satisfaction and depression were also associated with the frequency of telephone calls. Three of these four studies measured burden with the Burden Interview (Zarit, Reever and Bach-Peterson, 1980), a measure developed among carers of people with dementia, but all used different generic welfare measures such as well-being, satisfaction or quality of life. While carer burden was found to be associated with the generic measures of carer welfare, these studies suggest that the burden measures were also measuring other factors not captured by the generic welfare measures.

### 11.3.5 Conclusion

Carer burden, satisfaction with care-giving, quality of life, health and well-being are all important outcomes for the evaluation of dementia services and there are a large number of published measures to consider. Interpreting the meaning and importance of changes in carer burden and satisfaction with care-giving, and the development of dementia specific care related quality of life measures are all important areas for further research. The selection of informal carer outcome measures should be guided by the informal care context and the purpose and nature of the services to be evaluated. Generic measures of health and well-being allow comparisons with non-carer populations but may not be sufficiently sensitive to the impact of support services and may need to be used in conjunction with more specific measures. A systematic review of measures used in studies of the informal care provided to people with dementia is needed to identify the most appropriate measures.

### References

Abdellah FG and Levine E (1957) *Developing a measure of patient and personnel satisfaction with nursing care*. Nurs Res. Vol. 5, No.3, pp.100-8.

Abbott S, Meyer J, Copperman J, et al. (2005) *Quality criteria for patient advice and liaison services: what do patients and the public want?* Health Expect. Vol. 8, No.2, pp.126-37.

Addington-Hall J (1998) *VOICES (Views of Information Carers - Evaluation of Services)*. King's College School of Medicine and Dentistry.

Addington-Hall J and McPherson C (2001) *After-death interviews with surrogates/bereaved family members: some issues of validity*. Journal of Pain and Symptom Management. Vol. 22, No.3, pp.784-90.

- Aggarwal N, Vass AA, Minardi HA, et al. (2003) *People with dementia and their relatives: personal experiences of Alzheimer's and of the provision of care*. Journal of Psychiatric and Mental Health Nursing. Vol. 10, No.2, pp.187-97.
- AIHW (2004) *Australia's Health 2004: The Ninth Biennial Health Report of the Australian Institute of Health and Welfare*. Australian Institute of Health and Welfare. Report nr AIHW Cat AUS44. 544 p.
- Ajzen I and Fishbein M (1980) *Understanding Attitudes and Predicting Social Behaviour*. Prentice-Hall, Englewood Cliffs.
- Aragon SJ and Gesell SB (2003) *A patient satisfaction theory and its robustness across gender in emergency departments: a multigroup structural equation modeling investigation*. American Journal of Medical Quality. Vol. 18, No.6, pp.229-41.
- Aspinal F, Addington-Hall J, Hughes R, et al. (2003) *Using satisfaction to measure the quality of palliative care: a review of the literature*. Journal of Advanced Nursing. Vol. 42, No.4, pp.324-39.
- Attkisson CC and Zwick R (1982) *The client satisfaction questionnaire. Psychometric properties and correlations with service utilization and psychotherapy outcome*. Evaluation and Program Planning. Vol. 5, No.3, pp.233-7.
- Awad G, Hogan T, Voruganti L, et al. (1995) *Patients' subjective experiences on antipsychotic medications: implications for outcome and quality of life*. International Clinical Psychopharmacology. Vol. 10S, pp.123-132.
- Baker R (1990) *Development of a questionnaire to assess patients' satisfaction with consultations in general practice*. British Journal of General Practice. Vol. 40, No.December, pp.487-490.
- Baker R (1991) *The reliability and criterion validity of a measure of patients' satisfaction with their general practice*. Family Practice. Vol. 8, No.2, pp.171-177.
- Baker R, Mainous AG, 3rd, Gray DP, et al. (2003) *Exploration of the relationship between continuity, trust in regular doctors and patient satisfaction with consultations with family doctors*. Scandinavian Journal of Primary Health Care. Vol. 21, No.1, pp.27-32.
- Baker R and Whitfield M (1992) *Measuring patient satisfaction: a test of construct validity*. Quality in Health Care. Vol. 1, pp.104-109.
- Bashshur RL, Metzner CA and Worden C (1967) *Consumer satisfaction with group practice, the CHA case*. Am J Public Health Nations Health. Vol. 57, No.11, pp.1991-9.
- Bates EW and Brown BS (1988) *Geriatric care needs and HMO technology. A theoretical analysis and initial findings from the National Medicare Competition Evaluation*. Medical Care. Vol. 26, No.5, pp.488-98.
- Bauld L, Chesterman J and Judge K (2000) *Measuring satisfaction with social care amongst older service users: issues from the literature*. Health and Social Care in the Community. Vol. 8, No.5, pp.316-324.
- Beck EJ, Griffith R, Fitzpatrick R, et al. (1999) *Patient satisfaction with HIV service provision in NPMS hospitals: the development of a standard satisfaction questionnaire*. NPMS Steering Group. AIDS Care. Vol. 11, No.3, pp.331-43.
- Bekelman DB, Black BS, Shore AD, et al. (2005) *Hospice care in a cohort of elders with dementia and mild cognitive impairment*. Journal of Pain and Symptom Management. Vol. 30, No.3, pp.208-14.
- Bell CM, Araki SS and Neumann PJ (2001) *The association between caregiver burden and caregiver health-related quality of life in Alzheimer disease*. Alzheimer Disease and Associated Disorders. Vol. 15, No.3, pp.129-36.
- Ben-Sira Z (1976) *The function of the professional's affective behaviour in client satisfaction: a revised approach to social interaction theory*. Journal of Health and Social Behavior. Vol. 17, pp.3-11.

- Bindman J, Davies S, Taylor R, et al. (1996) *Developments in mental health policy in the United Kingdom*. Epidemiologia e Psichiatria Sociale. Vol. 5, No.2, pp.87-91.
- Birrell R and Hawthorne L (2001) *Doctor dilemmas: how GP medicine is practised in Australia*. People and Place. Vol. 9, No.3, pp.41-53.
- Blake H and Lincoln NB (2000) *Factors associated with strain in co-resident spouses of patients following stroke*. Clinical Rehabilitation. Vol. 14, No.3, pp.307-14.
- Bowers BJ (1988) *Family perceptions of care in a nursing home*. The Gerontologist. Vol. 28, No.3, pp.361-8.
- Braithwaite V (1992) *Caregiving Burden: Making the Concept Scientifically Useful and Policy Relevant*. Research on Ageing. Vol. 14, No.1, pp.3-27.
- Brouwer WBF, van Exel NJA, van Gorp B and Redekop WK (2006) *The CarerQol instrument: a new instrument to measure care-related quality of life of informal caregivers for use in economic evaluations*. Quality of Life Research. Vol. 15, No.6, pp.1005-21.
- Butters E, Higginson I, George R, et al. (1993) *Palliative care for people with HIV/AIDS: views of patients, carers and providers*. AIDS Care. Vol. 5, No.1, pp.105-16.
- Buttle F (1996) *SERVQUAL - review, critique, research agenda*. European Journal of Marketing. Vol. 30, No.1, pp.8-32.
- Camara EG (1991) *A psychiatry outpatient consultation-liaison clinic. Experience at the Cleveland Clinic Foundation*. Psychosomatics. Vol. 32, No.3, pp.304-8.
- Caplan EK and Sussman MB (1966) *Rank order of important variables for patient and staff satisfaction with outpatient service*. J Health Hum Behav. Vol. 7, No.2, pp.133-7.
- Chappell NL and Reid RC (2002) *Burden and well-being among caregivers: examining the distinction*. Gerontologist. Vol. 42, No.6, pp.772-80.
- Chene B (2006) *Dementia and Residential Placement: A View from the Carers' Perspective*. Qualitative Social Work. Vol. 5, No.2, pp.187-215.
- Chesterman J, Bauld L and Judge K (2001) *Satisfaction with the care-managed support of older people: an empirical analysis*. Health and Social Care in the Community. Vol. 9, No.1, pp.31-42.
- Chung CC, Ha JP, Tai YP, et al. (2002) *Double-blind, randomized trial comparing Harmonic Scalpel hemorrhoidectomy, bipolar scissors hemorrhoidectomy, and scissors excision: ligation technique*. Dis Colon Rectum. Vol. 45, No.6, pp.789-94.
- Clare J and Hofmeyer A (1998) *Discharge planning and continuity of care for aged people: indicators of satisfaction and implications for practice*. Australian Journal of Advanced Nursing. Vol. 16, No.1, pp.7-13.
- Coen RF, O'Boyle CA, Coakley D and Lawlor BA (2002) *Individual quality of life factors distinguishing low-burden and high-burden caregivers of dementia patients*. Dementia and Geriatric Cognitive Disorders. Vol. 13, No.3, pp.164-70.
- Cohen R, Leis AM, Kuhl D, Charbonneau C, et al. (2006) *QOLLTI-F: Measuring family carer quality of life*. Palliative Medicine. Vol. 20, No.8, pp.755-767.
- Cortina J (1993) *What is coefficient alpha? Examination of theory and applications*. Journal of Applied Psychology. Vol. 78, No.1, pp.98-104.
- Coucill W, Bryan S, Bentham P, et al. (2001) *EQ-5D in patients with dementia: an investigation of inter-rater agreement*. Medical Care. Vol. 39, No.8, pp.760-71.
- Cronbach J and Meehl P (1955) *Construct validity in psychological tests*. Psychological Bulletin. Vol. 52, No.4, pp.281-302.

- Cryns AG, Nichols RC, Katz LA, et al. (1989) *The hierarchical structure of geriatric patient satisfaction. An Older Patient Satisfaction Scale designed for HMOs.* Med Care. Vol. 27, No.8, pp.802-16.
- Cummins R (1995) *On the trail of the gold standard for subjective wellbeing.* Social Indicators Research. Vol. 35, pp.179-200.
- Dennis M, O'Rourke S, Slattery J, et al. (1997) *Evaluation of a stroke family care worker: results of a randomised controlled trial.* BMJ. Vol. 314, No.7087, pp.1071-6; discussion 1076-7.
- Diener E, Suh E, Lucas R, et al. (1999) *Subjective well-being: three decades of progress.* Psychological Bulletin. Vol. 125, pp.276-302.
- Donabedian A (1980) *Explorations in Quality Assessment and Monitoring: Vol 1 The Definition of Quality and the Approaches to its Assessment.* Arbor A, editor. Health Administration Press, MI.
- Donabedian A (1988) *The quality of care: how can it be assessed?* Journal of the American Medical Association. Vol. 260, No.12, pp.1743-48.
- Dozier AM, Kitzman HJ, Ingersoll GL, et al. (2001) *Development of an instrument to measure patient perception of the quality of nursing care.* Research in Nursing and Health. Vol. 24, No.6, pp.506-17.
- Elzinga RH and Barlow J (1991) *Patient satisfaction among the residential population of a psychiatric hospital.* International Journal of Social Psychiatry. Vol. 37, No.1, pp.24-34.
- Engel SE, Kiely DK and Mitchell SL (2006) *Satisfaction with end-of-life care for nursing home residents with advanced dementia.* Journal of the American Geriatrics Society. Vol. 54, No.10, pp.1567-72.
- Everaert K, De Ridder D, Baert L, et al. (2000) *Patient satisfaction and complications following sacral nerve stimulation for urinary retention, urge incontinence and perineal pain: a multicenter evaluation.* International Urogynecology Journal and Pelvic Floor Dysfunction. Vol. 11, No.4, pp.231-5; discussion 236.
- Farran CJ (2001) *Family caregiver intervention research: where have we been? Where are we going?* Journal of Gerontological Nursing. Vol. 27, No.7, pp.38-45.
- Firestone DN, Jimenez-Briceno L, Reimann JO, et al. (2004) *Predictors of diabetes-specific knowledge and treatment satisfaction among Costa Ricans.* Diabetes Education. Vol. 30, No.2, pp.281-292.
- Fitzpatrick R (1990) *Measurement of patient satisfaction.* In Hopkins D, Costain D, (eds.) Measuring the Outcomes of Medical Care. Royal College of Physicians and King's Fund Centre, London.
- Fitzpatrick R (1991) *Surveys of patient satisfaction: I - important general considerations.* British Medical Journal. Vol. 302, pp.887-889.
- Fitzpatrick R and Hopkins A (1983) *Problems in the conceptual framework of patient satisfaction research: an empirical exploration.* Social Health Illn. Vol. 5, No.3, pp.297-311.
- Foddy W (1993) *Constructing Questions for Interviews and Questionnaires: Theory and Practice in Social Research.* Cambridge University Press, Cambridge.
- Fox JG and Storms DM (1981) *A different approach to sociodemographic predictors of satisfaction with health care.* Social Science and Medicine. Vol. 15, No.5, pp.557-64.
- Fritz CL, Farver TB, Kass PH and Hart LA (1997) *Correlation among three psychological scales used in research of caregivers for patients with Alzheimer's disease.* Psychological Reports. Vol. 80, No.1, pp.67-80.
- Fulford M and Farhall J (2001) *Hospital versus home care for the acutely mentally ill? Preferences of caregivers who have experienced both forms of service.* Australian and New Zealand Journal of Psychiatry. Vol. 35, No.5, pp.619-25.

- Gaddini A, Ascoli M and Biscaglia L (2005) *Mental health care in Rome*. European Psychiatry. Vol. 20, Suppl 2, pp.S294-7.
- Gessert CE, Forbes S and Bern-Klug M (2000) *Planning end-of-life care for patients with dementia: roles of families and health professionals*. Omega: Journal of Death and Dying. Vol. 42, No.4, pp.273-91.
- Gill L and Maas F (2000) *Consumer and Carer Satisfaction Survey Aged Persons Mental Health Services - Summary Report*. ACNielsen. 123 p.
- Gompertz P, Pound P, Briffa J, et al. (1995) *How useful are non-random comparisons of outcomes and quality of care in purchasing hospital stroke services?* Age and Ageing. Vol. 24, No.2, pp.137-41.
- Gonzalez N, Quintana JM, Bilbao A, et al. (2005) *Development and validation of an in-patient satisfaction questionnaire*. International Journal for Quality in Health Care. Vol. 17, No.6, pp.465-72.
- Greehalgh J and Meadows K (1999) *The effectiveness of the use of patient-based measures of health in routine practice in improving the process and outcomes of patient care: a literature review*. Journal of Evaluation in Clinical Practice. Vol. 5, No.4, pp.401-16.
- Grol R and Wensing M (2000) *Patients Evaluate General/Family Practice: The EUROPEP Instrument*. The Task Force on Patient Evaluations of General Practice Care. 70 p.
- Grol R, Wensing M, Mainz J, et al. (1999) *Patients' priorities with respect to general practice care: an international comparison*. European Task Force on Patient Evaluations of General Practice (EUROPEP). Family Practice. Vol. 16, No.1, pp.4-11.
- Grimmer K and Moss J (2001) *The development, validity and application of a new instrument to assess the quality of discharge planning activities from the community perspective*. International Journal for Quality in Health Care. Vol. 13, No.2, pp.109-16.
- Grumbach K, Selby JV, Damberg C, et al. (1999) *Resolving the gatekeeper conundrum: what patients value in primary care and referrals to specialists*. JAMA. Vol. 282, No.3, pp.261-6.
- Grunfeld E, Coyle D, Whelan T, et al (2004) *Family caregiver burden: results of a longitudinal study of breast cancer patients and their principal caregivers*. Canadian Medical Association Journal. Vol. 170, No.12, pp.1795-801.
- Guadagnoli E and Velicer W (1988) *Relation of sample size to the stability of component patterns*. Psychological Bulletin. Vol. 103, No.2, pp.265-275.
- Guttman L (1954) *The principal components of scalable attitudes*. In Lazarfeld P, (ed.) Mathematical Thinking in the Social Sciences. The Free Press, Glencoe, Illinois, pp.216-257.
- Guyatt GH, Mitchell A, Molloy DW, et al. (1995) *Measuring patient and relative satisfaction with level or aggressiveness of care and involvement in care decisions in the context of life threatening illness*. Journal of Clinical Epidemiology. Vol. 48, No.10, pp.1215-24.
- Hancock K, Chang E, Chenoweth L, et al. (2003) *Nursing needs of acutely ill older people*. Journal of Advanced Nursing. Vol. 44, No.5, pp.507-16.
- Hanks GW, Robbins M, Sharp D, et al. (2002) *The imPaCT study: a randomised controlled trial to evaluate a hospital palliative care team*. British Journal of Cancer. Vol. 87, No.7, pp.733-9.
- Hara I, Kawabata G, Miyake H, et al. (2003) *Comparison of quality of life following laparoscopic and open prostatectomy for prostate cancer*. J Urol. Vol. 169, No.6, pp.2045-8.
- Hardy G, West M and Hill F (1996) *Components and predictors of patient satisfaction*. British Journal of Health Psychology. Vol. 1, pp.65-85.
- Hare R, Rogers H, Lester H, et al. (2006) *What do stroke patients and their carers want from community services?* Family Practice. Vol. 23, No.1, pp.131-6.

Harris R, Ashton T, Broad J, et al. (2005) *The effectiveness, acceptability and costs of a hospital-at-home service compared with acute hospital care: a randomized controlled trial.* Journal of Health Services Research and Policy. Vol. 10, No.3, pp.158-66.

Hasselkus BR (1988) *Meaning in family caregiving: perspectives on caregiver/professional relationships.* The Gerontologist. Vol. 28, No.5, pp.686-91.

Hawthorne G (2006) *Review of Patient Satisfaction Measures.* Australian Government Department of Health and Ageing. pp.225-289.

Hawthorne G and Harmer C (2000) *GUTSS: The Genito-Urinary Treatment Satisfaction Scale Study.* Centre for Health Program Evaluation, Melbourne. p 50.

Hawthorne G, Sansoni J, Hayes L, et al. (2006) *Measuring Patient Satisfaction with Incontinence Treatment.* Department of Psychiatry, The University of Melbourne, and the Centre for Health Service Development, University of Wollongong for the Commonwealth Department of Health and Ageing. 114 p.

Hawthorne L, Birrell B and Young D (2003) *The Retention of Overseas Trained Doctors in General Practice in Regional Victoria.* Faculty of Medicine, Dentistry and Health Sciences, University of Melbourne.

Hays RD, Davies AR and Ware JE (1987) *Scoring the Medical Outcomes Study Patient Satisfaction Questionnaire PSQ-III.* Medical Outcomes Trust.

Helmstadter GC (1964) *Principles of Psychological Measurement.* Appleton-Century-Crofts, New York.

Henderson C, Hales H and Ruggeri M (2003) *Cross-cultural differences in the conceptualisation of patients' satisfaction with psychiatric services--content validity of the English version of the Verona Service Satisfaction Scale.* Social Psychiatry and Psychiatric Epidemiology. Vol. 38, No.3, pp.142-8.

Hendriks AA, Vrieling MR, Smets EM, et al. (2001) *Improving the assessment of (in)patients' satisfaction with hospital care.* Medical Care. Vol. 39, No.3, pp.270-83.

Hilton BA, Budgen C, Molzahn AE, et al. (2001) *Developing and testing instruments to measure client outcomes at the Comox Valley Nursing Center.* Public Health Nursing. Vol. 18, No.5, pp.327-39.

Holcomb WR, Parker JC, Leong GB, et al. (1998) *Customer satisfaction and self-reported treatment outcomes among psychiatric inpatients.* Psychiatric Services. Vol. 49, No.7, pp.929-34.

Hudak PL, Hogg-Johnson S, Bombardier C, et al. (2004) *Testing a new theory of patient satisfaction with treatment outcome.* Medical Care. Vol. 42, No.8, pp.726-39.

Hunt CK (2003) *Concepts in Caregiver Research.* Journal of Nursing Scholarship. Vol. 35, No.1, pp.27-32.

Hwang SS, Chang VT, Alejandro Y, et al. (2003) *Caregiver unmet needs, burden, and satisfaction in symptomatic advanced cancer patients at a Veterans Affairs (VA) medical center.* Palliative Support Care. Vol. 1, No.4, pp.319-29.

Ingleton C, Morgan J, Hughes P, et al. (2004) *Carer satisfaction with end-of-life care in Powys, Wales: a cross-sectional survey.* Health and Social Care in the Community. Vol. 12, No.1, pp.43-52.

Jackson JL and Kroenke K (2001) *The effect of unmet expectations among adults presenting with physical symptoms.* Annals of Internal Medicine. Vol. 134, No.9 Pt 2, pp.889-97.

Jacoby A, Lecouturier J, Bradshaw C, et al. (1999) *Feasibility of using postal questionnaires to examine carer satisfaction with palliative care: a methodological assessment.* Palliative Medicine. Vol. 13, No.4, pp.285-98.

Jaglal S, Cockerill R, Lemieux-Charles L, et al. (2007) *Perceptions of the process of care among caregivers and care recipients in dementia care networks.* American Journal of Alzheimer's Disease and Other Dementias. Vol. 22, No.2, pp.103-11.

Jarvis H, Burge FI and Scott CA (1996) *Evaluating a palliative care program: methodology and limitations*. Journal of Palliative Care. Vol. 12, No.2, pp.23-33.

Johnson JE (1987) *Selecting nursing activities for hospitalized clients*. Journal of Gerontological Nursing. Vol. 13, No.10, pp.29-33.

James G (1967) *The medical school of the future and its role in the community*. JAMA. Vol. 202, No.5, pp.415-8.

Jenkins C (1992) *Assessment of outcomes of health intervention*. Social Science and Medicine. Vol. 35, p.367-375.

Johannsson HO, Graf W and Pahlman L (2002) *Long-term results of haemorrhoidectomy*. European Journal of Surgery. Vol. 168, No.8-9, pp.485-9.

Kane RL, Maciejewski M and Finch M (1997) *The relationship of patient satisfaction with care and clinical outcomes*. Medical Care. Vol. 35, No.7, pp.714-30.

Kealey P and McIntyre I (2005) *An evaluation of the domiciliary occupational therapy service in palliative cancer care in a community trust: a patient and carers perspective*. European Journal of Cancer Care (English Language Edition). Vol. 14, No.3, pp.232-43.

Kelleher M and Mannix M (2001) *Carers' views on services for the elderly mentally ill*. Irish Journal of Medical Science. Vol. 170, No.4, pp.239-43.

Kessler D, Peters TJ, Lee L, et al. (2005) *Social class and access to specialist palliative care services*. Palliative Medicine. Vol. 19, No.2, pp.105-10.

Kiely DK, Volicer L, Teno J, et al. (2006) *The validity and reliability of scales for the evaluation of end-of-life care in advanced dementia*. Alzheimer Disease and Associated Disorders. Vol. 20, No.3, pp.176-181.

Kikano GE, Goodwin MA and Stange KC (1998) *Physician employment status and practice patterns*. Journal of Family Practice. Vol. 46, No.6, pp.499-505.

Kinnersley P, Stott N, Peters T, et al. (1996) *A comparison of methods for measuring patient satisfaction with consultations in primary care*. Family Practice. Vol. 13, No.1, pp.41-51.

Koffman J and Higginson IJ (2001) *Accounts of carers' satisfaction with health care at the end of life: a comparison of first generation black Caribbeans and white patients with advanced disease*. Palliative Medicine. Vol. 15, No.4, pp.337-45.

Kring A, Kerr S, Smith D, et al. (1993) *Flat affect in schizophrenia does not reflect diminished subjective experience of emotion*. Journal of Abnormal Psychology. Vol. 102, pp.507-517.

Kristjanson LJ (1989) *Quality of terminal care: salient indicators identified by families*. Journal of Palliative Care. Vol. 5, No.1, pp.21-30.

Kristjanson LJ (1993) *Validity and reliability testing of the FAMCARE Scale: measuring family satisfaction with advanced cancer care*. Social Science and Medicine. Vol. 36, No.5, pp.693-701.

Kristjanson LJ, Leis A, Koop PM, et al. (1997) *Family members' care expectations, care perceptions, and satisfaction with advanced cancer care: results of a multi-site pilot study*. Journal of Palliative Care. Vol. 13, No.4, pp.5-13.

La Monica EL, Oberst MT, Madea AR, et al. (1986) *Development of a patient satisfaction scale*. Research in Nursing and Health. Vol. 9, No.1, pp.43-50.

Labarere J, Fourny M, Jean-Phillippe V, et al. (2004) *Refinement and validation of a French in-patient experience questionnaire*. International Journal of Health Care Quality Assurance. Vol. 17, No.1, pp.17-25.

- Langa KM, Chernew ME, Kabeto MU, Herzog AR, et al. (2001) *National estimates of the quantity and cost of informal caregiving for the elderly with dementia*. Journal of General Internal Medicine. Vol. 16, No.11, pp.770-8.
- Larsen D and Rootman I (1976) *Physician role performance and patient satisfaction*. Social Science and Medicine. Vol. 10, pp.29-32.
- Larsen DL, Attkisson CC, Hargreaves WA, et al. (1979) *Assessment of client/patient satisfaction: development of a general scale*. Evaluation and Program Planning. Vol. 2, No.3, pp.197-207.
- Lawton MP, Kleban MH, Moss M, Rovine M, et al. (1989) *Measuring caregiving appraisal*. Journal of Gerontology. Vol. 44, No.3, pp.61-71.
- Likert R (1932) *A technique for the measurement of attitudes*. Archives of Psychology. Vol. 140, pp.3-55.
- Likert R, Roslow S and Murphy G (1934) *A simple and reliable method of scoring the Thurstone attitude scales*. Journal of Social Psychology. Vol. 5, pp.228-238.
- Lincoln NB, Francis VM, Lilley SA, et al. (2003) *Evaluation of a stroke family support organiser: a randomized controlled trial*. Stroke. Vol. 34, No.1, pp.116-21.
- Lincoln NB, Walker MF, Dixon A, et al. (2004) *Evaluation of a multiprofessional community stroke team: a randomized controlled trial*. Clinical Rehabilitation. Vol. 18, No.1, pp.40-7.
- Linder-Pelz S (1982) *Toward a theory of patient satisfaction*. Social Science and Medicine. Vol. 16, pp.577-582.
- Linder-Pelz S and Struening EL (1985) *The multidimensionality of patient satisfaction with a clinic visit*. Journal of Community Health. Vol. 10, No.1, pp.42-54.
- Lindsey I, Smilgin-Humphreys MM, Cunningham C, et al. (2002) *A randomized, controlled trial of fibrin glue vs. conventional treatment for anal fistula*. Dis Colon Rectum. Vol. 45, No.12, pp.1608-15.
- Locker D and Dunt D (1978) *Theoretical and methodological issues in sociological studies of consumer satisfaction with medical care*. Social Science and Medicine. Vol. 12, pp.283-292.
- Lubart E, Leibovitz A, Shapira A, et al. (2004) *Satisfaction with the care of institutionalized psychogeriatric patients, as reflected by a survey of their relatives*. Archives of Gerontology and Geriatrics. Vol. 38, No.1, pp.45-50.
- Luscombe G, Brodaty H and Freeth S (1998) *Younger people with dementia: diagnostic issues, effects on carers and use of services*. International Journal of Geriatric Psychiatry. Vol. 13, No.5, pp.323-30.
- MacDonald L, Sibbald B and Hoare C (1988) *Measuring patient satisfaction with life in a long-stay psychiatric hospital*. International Journal of Social Psychiatry. Vol. 34, No.4, pp.292-304.
- Mackenzie A, Holroyd E and Lui MH (1998) *Community nurses assessment of the needs of Hong Kong family carers who are looking after stroke patients*. International Journal of Nursing Studies. Vol. 35, No.1, pp.132-140.
- Magaziner J (1997) *Use of proxies to measure health and functional outcomes in effectiveness research in persons with Alzheimer disease and related disorders*. Alzheimer Disease and Associated Disorders. Vol. 11, Suppl 6, pp.168-74.
- Manning WG, Leibowitz A, Goldberg GA, et al. (1984) *A controlled trial of the effect of a prepaid group practice on use of services*. New England Journal of Medicine. Vol. 310, No.23, pp.1505-10.
- Mant J, Carter J, Wade DT, et al. (2000) *Family support for stroke: a randomised controlled trial*. Lancet. Vol. 356, No.9232, pp.808-13.

- Margolis SA, Al-Marzouq S, Revel T, et al. (2003) *Patient satisfaction with primary health care services in the United Arab Emirates*. International Journal for Quality in Health Care. Vol. 15, No.3, pp.241-9.
- Mason A, Weatherly H, Spilsbury K, et al. (2007) *A systematic review of the effectiveness and cost-effectiveness of different models of community-based respite care for frail older people and their carers*. Health Technology Assessment. Vol. 11, No.15, pp.1-176.
- McCusker J (1984) *Development of scales to measure satisfaction and preferences regarding long-term and terminal care*. Medical Care. Vol. 22, No.5, pp.476-93.
- McGrath P, Moore A, McNaught M, et al. (2005) *Another form to fill in! Clients' reflections on the hospice use of questionnaires*. Support Care Cancer. Vol. 13, No.9, pp.691-701.
- McKinley RK, Fraser RC, Baker RH, et al. (2004) *The relationship between measures of patient satisfaction and enablement and professional assessments of consultation competence*. Medical Teacher. Vol. 26, No.3, pp.223-8.
- Meakin R and Weinman J (2002) *The 'Medical Interview Satisfaction Scale' (MISS-21) adapted for British general practice*. Family Practice. Vol. 19, No.3, pp.257-63.
- Meehan T, Bergen H and Stedman T (2002) *Monitoring consumer satisfaction with inpatient service delivery: the Inpatient Evaluation of Service Questionnaire*. Australian and New Zealand Journal of Psychiatry. Vol. 36, No.6, pp.807-11.
- Melzer D, Bedford S, Denning T, et al. (1996) *Carers and the monitoring of psychogeriatric community teams*. International Journal of Geriatric Psychiatry. Vol. 11, No.12, pp.1057-1-61.
- Montalto M (1996) *Patients' and carers' satisfaction with hospital-in-the-home care*. International Journal for Quality in Health Care. Vol. 8, No.3, pp.243-51.
- Munro BH, Jacobsen BS and Brooten DA (1994) *Re-examination of the psychometric characteristics of the La Monica-Oberst Patient Satisfaction Scale*. Research in Nursing and Health. Vol. 17, No.2, pp.119-25.
- Newsome PR and Wright GH (1999) *A review of patient satisfaction: 1. Concepts of satisfaction*. British Dental Journal. Vol. 186, No.4 Spec No, pp.161-5.
- Nguyen TD, Attkisson CC and Stegner BL (1983) *Assessment of patient satisfaction: development and refinement of a service evaluation questionnaire*. Evaluation and Program Planning. Vol. 6, No.3-4, pp.299-313.
- Nicoll M, Ashworth M, McNally L and Newman S (2002) *Satisfaction with respite care: a pilot study*. Health and Social Care in the Community. Vol. 10, No.6, pp.479-84.
- Nicoll M, Ashworth M, McNally L, et al. (2002) *Satisfaction with respite care: a pilot study*. Health and Social Care in the Community. Vol. 10, No.6, pp.479-84.
- Nolan M and Grant G (1992) *Regular Respite: An Evaluation of Hospital Rota Bed Scheme for Elderly People*. Age Concern Institute of Gerontology. Report nr Research Paper N 6.
- Nolan MR, Grant G and Ellis NC (1990) *Stress is in the eye of the beholder: reconceptualizing the measurement of carer burden*. Journal of Advanced Nursing. Vol. 15, No.5, pp.544-55.
- Nolan M, Grant G and Keady J (1996) *Understanding family care : a multi-dimensional model of caring and coping*. Open University Press, Philadelphia, PA.
- Nunnally JC and Bernstein IH (1994) *Psychometric Theory*. McGraw-Hill, New York 752 p.
- O'Connell B, Young J and Twigg D (1999) *Patient satisfaction with nursing care: a measurement conundrum*. International Journal of Nursing Practice. Vol. 5, No.2, pp.72-7.

Pang MW, Chan LW and Yip SK (2003) *One-year urodynamic outcome and quality of life in patients with concomitant tension-free vaginal tape during pelvic floor reconstruction surgery for genitourinary prolapse and urodynamic stress incontinence.* International Urogynecology Journal and Pelvic Floor Dysfunction. Vol. 14, No.4, pp.256-60; discussion 259-60.

Parasuraman A, Zeithaml VA and Berry L (1985) *A conceptual model of service quality and its implications for future research.* Journal of Marketing. Vol. 49, No.4, pp.41-50.

Parasuraman A, Zeithaml VA and Berry L (1988) *SERVQUAL: a multiple-item scale for measuring consumer perceptions of service quality.* Journal of Retailing. Vol. 64, No.1, pp.12-37.

Pearlin LI, Mullan JT, Semple SJ and Skaff MM (1990) *Caregiving and the stress process: an overview of concepts and their measures [see comments].* Gerontologist. Vol. 30, No.5, pp.583-94.

Pound P, Gompertz P and Ebrahim S (1993) *Development and results of a questionnaire to measure carer satisfaction after stroke.* Journal of Epidemiology and Community Health. Vol. 47, No.6, pp.500-5.

Pound P, Gompertz P and Ebrahim S (1994) *Patients' satisfaction with stroke services.* Clinical Rehabilitation. Vol. 8, pp.7-17.

Pritchard EJ and Dewing J (2001) *A multi-method evaluation of an independent dementia care service and its approach.* Ageing and Mental Health. Vol. 5, No.1, pp.63-72.

Quintana JM, Gonzalez N, Bilbao A, et al. (2006) *Predictors of patient satisfaction with hospital health care.* BMC Health Services Research. Vol. 6, pp.102.

Ringdal GI, Jordhoy MS and Kaasa S (2002) *Family satisfaction with end-of-life care for cancer patients in a cluster randomized trial.* Journal of Pain and Symptom Management. Vol. 24, No.1, pp.53-63.

Ringdal GI, Jordhoy MS and Kaasa S (2003) *Measuring quality of palliative care: psychometric properties of the FAMCARE Scale.* Quality of Life Research. Vol. 12, No.2, pp.167-76.

Rodgers H, Atkinson C, Bond S, et al. (1999) *Randomized controlled trial of a comprehensive stroke education program for patients and caregivers.* Stroke. Vol. 30, No.12, pp.2585-91.

Roberts RE, Pascoe GC and Attkisson CC (1983) *Relationship of service satisfaction to life satisfaction and perceived well-being.* Evaluation and Program Planning. Vol. 6, No.3-4, pp.373-83.

Ross C, Steward C and Sinacore J (1993) *The importance of patient preferences in the measurement of health care satisfaction.* Medical Care. Vol. 31, No.12, pp.1138-1149.

Ross CK, Frommelt G, Hazelwood L, et al. (1987) *The role of expectations in patient satisfaction with medical care.* J Health Care Mark. Vol. 7, No.4, pp.16-26.

Rouse MO (1967) *Organizing and delivering health care.* 3. JAMA. Vol. 202, No.4, pp.305.

Rowbotham MC (1953) *Centralized or decentralized service; Faulkner Hospital combines them for greater efficiency and patient satisfaction.* Mod Hosp. Vol. 80, No.6, pp.114-6.

Rubin HR, Gandek B, Rogers WH, et al. (1993) *Patients' ratings of outpatient visits in different practice settings. Results from the Medical Outcomes Study.* Journal of the American Medical Association. Vol. 270, No.7, pp.835-40.

Ruggeri M and Dall'Agnola R (1993) *The development and use of the Verona Expectations for Care Scale (VECS) and the Verona Service Satisfaction Scale (VSSS) for measuring expectations and satisfaction with community-based psychiatric services in patients, relatives and professionals.* Psychological Medicine. Vol. 23, pp.511-523.

Schneider J, Murray J, Banerjee S, et al. (1999) *EUROCARE: a cross-national study of co-resident spouse carers for people with Alzheimer's disease: I--Factors associated with carer burden.* International Journal of Geriatric Psychiatry. Vol. 14, No.8, pp.651-61.

- Schulz R, Belle SH, Czaja SJ, McGinnis KA, et al. (2004) *Long-term care placement of dementia patients and caregiver health and well-being*. Journal of the American Medical Association. Vol. 292, No.8, pp.961-967.
- Schulz R, O'Brien A, Czaja S, Ory M, et al. (2002) *Dementia caregiver intervention research: In search of clinical significance*. The Gerontologist. Vol. 42, No.5, pp.589-602.
- Schulz R and Williamson GM (1997) *The measurement of caregiver outcomes in Alzheimer disease research*. Alzheimer Disease and Associated Disorders. Vol. 11, Suppl 6, pp.117-24.
- Seale CF (1989) *What happens in hospices: a review of research evidence*. Social Science and Medicine. Vol. 28, No.6, pp.551-9.
- Shepperd S and Iliffe S (1998) *The effectiveness of hospital at home compared with in-patient hospital care: a systematic review*. Journal of Public Health Medicine. Vol. 20, No.3, pp.344-350.
- Shepperd S and Iliffe S (2005) *Hospital-at-home versus in-patient hospital care*. Cochrane Database of Systematic Reviews. No.3, pp.CD000356.
- Simon C, Little P, Birtwistle J, et al. (2003) *A questionnaire to measure satisfaction with community services for informal carers of stroke patients: construction and initial piloting*. Health and Social Care in the Community. Vol. 11, No.2, pp.129-37.
- Simpson RG (1997) *Carers as equal partners in care planning*. Journal of Psychiatric and Mental Health Nursing. Vol. 4, No.5, pp.345-54.
- Simpson RG, Scothern G and Vincent M (1995) *Survey of carer satisfaction with the quality of care delivered to in-patients suffering from dementia*. Journal of Advanced Nursing. Vol. 22, No.3, pp.517-27.
- Sitzia J (1999) *How valid and reliable are patient satisfaction data? An analysis of 195 studies*. International Journal for Quality in Health Care. Vol. 11, No.4, pp.319-28.
- Sitzia J and Wood N (1997) *Patient satisfaction: a review of issues and concepts*. Social Science and Medicine. Vol. 45, No.12, pp.1829-1843.
- Skolnick AA (1998) *MediCaring Project to demonstrate, evaluate innovative end-of-life program for chronically ill*. JAMA. Vol. 279, No.19, pp.1511-2.
- Sorensen S, Pinquart M and Duberstein P (2002) *How effective are interventions with caregivers? An updated meta-analysis*. Gerontologist. Vol. 42, No.3, pp.356-72.
- Soliman AA (1992) *Assessing the quality of health care: a consumerist approach*. Health Marketing Quarterly. Vol. 10, No.1-2, pp.121-41.
- Spear J (2003) *A new measure of consumer expectations, perceptions and satisfaction for patients and carers of older people with mental health problems*. Australasian Psychiatry. Vol. 11, No.3, pp.330.
- Stephenson C, Wilson S and Gladman JR (1995) *Patient and carer satisfaction in geriatric day hospitals*. Disability and Rehabilitation. Vol. 17, No.5, pp.252-5.
- Stull DE, Kosloski K and Kercher K (1994) *Caregiver burden and generic well-being: opposite sides of the same coin?* Gerontologist. Vol. 34, No.1, pp.88-94.
- Suchman EA (1965) *Social patterns of illness and medical care*. Journal of Health and Human Behavior. Vol. 106, pp.2-16.
- Teno JM, Landrum K and Lynn J (1997) *Defining and measuring outcomes in end-stage dementia*. Alzheimer Disease and Associated Disorders. Vol. 11, Suppl 6, pp.25-9.

- Thompson AG and Sunol R (1995) *Expectations as determinants of patient satisfaction: concepts, theory and evidence*. International Journal for Quality in Health Care. Vol. 7, No.2, pp.127-41.
- Thurstone L (1947) *Multiple Factor Analysis*. University of Chicago Press, Chicago.
- Tilling K, Coshall C, McKeivitt C, et al (2005) *A family support organiser for stroke patients and their carers: a randomised controlled trial*. Cerebrovascular Diseases. Vol. 20, No.2, pp.85-91.
- Vahey DC, Aiken LH, Sloane DM, et al. (2004) *Nurse burnout and patient satisfaction*. Medical Care. Vol. 42, No.2 Suppl, pp.II57-66.
- Van De Water MS, Kutner M, Parmelee PA, et al. (2003) *Which long-term care residents should be asked to complete a customer satisfaction survey?* Journal of the American Medical Directors Association. Vol. 4, No.5, pp.257-63.
- Vitaliano PP, Young HM and Russo J (1991) *Burden: a review of measures used among caregivers of individuals with dementia*. Gerontologist. Vol. 31, No.1, pp.67-75.
- Volicer L, Hurley AC and Blasi ZV (2001) *Scales for evaluation of end-of-life care in dementia*. Alzheimer Disease and Associated Disorders. Vol. 15, No.4, pp.194-200.
- Walker SA, Naji SA, Howie FL, et al. (2001) *The effectiveness of psychiatric discharge coordinators*. Primary Care Psychiatry. Vol. 6, No.4, pp.149-153.
- Ware JE, Jr. and Hays RD (1988) *Methods for measuring patient satisfaction with specific medical encounters*. Medical Care. Vol. 26, No.4, pp.393-402.
- Ware JE, Snyder MK, Wright WR, et al. (1983) *Defining and measuring patient satisfaction with medical care*. Evaluation and Program Planning. Vol. 6, pp.247-263.
- Wellwood I, Dennis M and Warlow C (1995) *Patients' and carers' satisfaction with acute stroke management*. Age and Ageing. Vol. 24, No.6, pp.519-24.
- Westaway MS, Rheeder P, Van Zyl DG, et al. (2003) *Interpersonal and organizational dimensions of patient satisfaction: the moderating effects of health status*. International Journal for Quality in Health Care. Vol. 15, No.4, pp.337-44.
- White MB (1972) *Importance of selected nursing activities*. Nursing Research. Vol. 21, No.1, pp.4-14.
- Wilkinson EK, Salisbury C, Bosanquet N, et al. (1999) *Patient and carer preference for, and satisfaction with, specialist models of palliative care: a systematic literature review*. Palliative Medicine. Vol. 13, No.3, pp.197-216.
- Williams B (1994) *Patient satisfaction: a valid concept?* Social Science and Medicine. Vol. 38, No.4, pp.509-516.
- Wilson A, Wynn A and Parker H (2002) *Patient and carer satisfaction with 'hospital at home': quantitative and qualitative results from a randomised controlled trial*. British Journal of General Practice. Vol. 52, No.474, pp.9-13.
- Wilson W and Goldschmidt P (1995) *Quality Management in Healthcare*. McGraw-Hill Book Company Australia, Sydney.
- Wolf MH, Putnam SM, James SA, et al. (1978) *The Medical Interview Satisfaction Scale: development of a scale to measure patient perceptions of physician behavior*. Journal of Behavioral Medicine. Vol. 1, No.4, pp.391-401.
- Wolf MH and Stiles WB (1981) *Further development of the Medical Interview Satisfaction Scale*. American Psychological Association Convention, Los Angeles.
- Wood J, Taylor S and Lichtman R (1985) *Social comparison in adjustment to breast cancer*. Journal of Personality and Social Psychology. Vol. 49, pp.1169-1183.

Wyrwich KW, Metz SM, Babu AN, et al. (2002) *The reliability of retrospective change assessments*. Quality of Life Research. Vol. 11, No.7, pp.636.

Zarit SH, Reever KE and Bach-Peterson J (1980) *Relatives of the impaired elderly: correlates of feelings of burden*. Gerontologist. Vol. 20, No.6, pp.649-55.

Zarit SH, Todd PA and Zarit JM (1986) *Subjective burden of husbands and wives as caregivers: a longitudinal study*. Gerontologist. Vol. 26, No.3, pp.260-6.

Zebiene E, Razgauskas E, Basys V, et al. (2004) *Meeting patient's expectations in primary care consultations in Lithuania*. International Journal of Quality in Health Care. Vol. 16, No.1, pp.83-9.

## 12 Measurement and Implementation Issues

### 12.1 Introduction

This section discusses some key measurement issues relevant to the use of these measures with dementia patients and their carers. The first of these is the issue of the use of proxy reports followed by a discussion of the level of cognitive impairment at which dementia patients may retain the capacity to self rate. These issues are particularly important to consider when assessing more subjective phenomena such as health related quality of life and social isolation but these measurement considerations are relevant to all areas of this report. For such phenomena self – rating is to be preferred where it is possible as is outlined below.

The applicability of these measures for particular population groups is also discussed. The issue of the applicability of the measures for those from Culturally and Linguistically Diverse (CALD) populations is considered as is the applicability of these measures for use with Aboriginal and Torres Strait Islander Groups.

This section concludes with a discussion of some implementation issues. Some instruments are more suitable in some settings rather than others, and there are training issues to be considered with the use of many of these measures.

### 12.2 Cognitive Impairment and Self-Report

The above review of proxy reporting issues would suggest that where possible subjective phenomenon such as health related quality of life, overall well-being and social isolation should be assessed through self report. The capacity to provide self-ratings on self-report instruments will be limited by the severity of cognitive impairment experienced by people with dementia. With respect to this discussion Wlodarczyk, et al. (2003) suggest the most commonly accepted score ranges for classifying the severity of cognitive impairment are: scale cut-points of <10 to indicate severe cognitive impairment, 10–14 moderate cognitive impairment, 15–19 mild to moderate cognitive impairment, and 20–24 mild impairment. A similar classification of severity and guidance for interpretation are outlined by the National Institute for Health and Clinical Excellence (NICE) (2007) in the United Kingdom.

One approach has been to adapt methods of instrument administration to facilitate completion of self report measures by those with mild to moderate cognitive impairments. However, a major issue is the determination of the level of cognitive impairment at which dementia patients retain the capacity to self rate.

#### 12.2.1 Cognitive Impairment and the Capacity to Self Rate

Given the above considerations it does seem important for there to be some guidelines regarding the severity of dementia or cognitive impairment below which self-report is probably undesirable.

In an early study, Berger (1980) classified increasing senility into 6-classes of progressive deterioration. He then dichotomized these between Classes III and IV for the performance of basic activities. Those in Classes I-III were patients who could complete tasks if asked to; those in Classes IV-VI needed someone to complete the task for them. Based on this classification, self-report would be meaningful so long as a patient can respond appropriately to instructions; proxy report would be preferred for those in Classes IV-VI. More recently, Mozley, et al. (1999), using the MMSE with a cutpoint of >9 for study inclusion, investigated the reliability of self-response to interview. They reported by MMSE scores the proportions who were interviewable, being 78% (MMSE score of 10+), 86% (12+), 95% (18+) and 100% (26+). The conclusion was that many

older people with cognitive impairment could provide meaningful answers to self-report interviewer-administered questionnaires.

Others have examined this problem using estimates of reliability, as assessed by test-retest, internal consistency or standard deviation increase or decrease. Naglie, et al. (2006) reported on the test-retest reliability of three quality of life measures at 13 days, reporting that for those with MMSE scores 19-26 the test-retests were within an acceptable range (0.70 – 0.81). For those with MMSE scores 10-18 the test-retest coefficients varied by instrument length and complexity. The conclusion was that patients with mild Alzheimer's disease could rate their own QoL, and that those with moderate impairment could do this with a facilitated interview. This finding is consistent with a study of the AQoL measure in those with Alzheimer's disease where the data was collected in interview. For those with MMSE scores 10+ the AQoL means and standard deviations were consistent and monotonic with MMSE scores, whereas for those with MMSE scores in the range 0-10 the standard deviation was extremely broad (Wlodarczyk, 2004). In another study, using the Nottingham Health Profile (NHP) Baro, et al. (2006) reported that under interviewer-administered conditions, the NHP could be successfully administered to those with moderate cognitive impairment (defined as  $\geq 10$ -points on the MMSE), whereas for those with severe impairment ( $< 10$ -points) it could not be meaningfully administered and similar findings were reported by Novella, et al. (2001) with regard to the interview administration of the Duke Health Profile. With regard to self rating without interview administration Novella, et al. (2001) suggest an MMSE score of  $>$  than 15 may be required.

The interpretation across these studies would be that for those with MMSE  $< 10$  that self-completion was problematic. These findings are consistent with Folstein, et al's original work (Folstein, 1975). The mean MMSE score of dementia patients was 10 for one sample and 12 for another; at 28-day test-retest the correlation was 0.98. The implication is that, in interview, patients with moderate dementia, as defined by an MMSE score of  $\geq 10$  can provide insight and complete short self-report measures, especially where these are interviewer-facilitated.

However, the capacity for cognitively impaired patients to self rate will depend on the structure, length, design and complexity of each questionnaire (Naglie, et al. 2006; Riemsma, et al. 2001). Shorter and less complex questionnaires would appear to be more suitable. It is suggested that a follow up study be undertaken to assess the required MMSE-3MS scores required for the recommended self report questionnaires under different modes of administration.

### 12.2.2 Methods to Facilitate Self-Completion

An alternative to proxy-completion is administration in an interviewer-facilitated setting, i.e. where the interviewer reads the questionnaire out to the participant (Ankri, 2003). Recommendations for interview-administration are a function of impairment level (Naglie, et al. 2006; Ankri, 2003). In a study of the EQ-5D quality of life measure in those with mild, moderate or severe dementia, Coucill, et al. (2001) concluded that the EQ-5D could be patient-completed when interviewer-administered, but that there was little evidence to support patient self-rated completion.

There is, however, an important caveat to facilitated interview completion. It is often assumed that a cognitively disabled person who has difficulty reading and responding to a complex questionnaire on his/her own can be verbally administered a questionnaire and the results accepted as valid. This assumption, however, is challengeable for three important reasons. First, the setting (interviewer reading and respondent selecting a verbal option) may lead to acquiescent response bias, which is where a respondent provides an answer that he/she deems acceptable (usually on the grounds that he/she is trying to please the interviewer in some way) (Sigelman, 1981a, 1981b; Foddy, 1993). Second, where material is poorly understood and is rephrased by the interviewer the rephrasing may represent the interviewer's beliefs about the question and what the response should be (Antaki, 1996, 1999; Rapley and Antaki, 1996). Third, there is some evidence that during facilitation interviewers lead the respondent into particular responses (Antaki, 1999). It

follows that interviewer-facilitated data collected using a non-standardized interview schedule may result in data that neither represents the views of the respondent nor that is comparable with other data from the same study (e.g. that collected through self-report).

Given the above issues it is important to develop standardized interview schedules for any instruments that may need to be administered to patients with more severe cognitive impairments. A related approach has also been to use an 'interview assisted' administration mode – for example to provide verbal cues/ standard prompts and using cards to assist in response choices etc. It is recommended that a study be undertaken to assess the recommended self report tools by self report administration, interview administration and assisted interview administration to identify the best approach for assessing the HRQOL and other subjective phenomena of dementia patients with more severe cognitive impairments. As will be detailed in the following section there is also a need to identify clearly at what level of cognitive impairment (e.g. MMSE score) the capacity to self-rate is impaired in relation to these modes of administration. It may also be that consideration may need to be given to the development of shorter and simplified forms of such measures for the more severely impaired patients.

### **12.2.3 Recommendations**

Where it is possible and feasible HRQOL (and other subjective phenomena) should be assessed by patient self report rather than by proxy report. However, an interim recommendation (awaiting the results of further recommended research) is that self rating report (by non interview administration) should not be considered for patients with MMSE scores below 15.

For patients with MMSE scores ranging from 10-15 an interview administration or an interview assisted administration of these self-report measures could be considered.

For patients with an MMSE score less than 10 it is suggested that data be collected via proxy reporting. Where a specific proxy form has been developed this should be utilised.

It is recommended that a study be undertaken to assess the recommended self report tools by self report administration, interview administration and assisted interview administration to identify the best approach for assessing the HRQOL and other subjective phenomena of dementia patients with more severe cognitive impairments.

As the capacity for cognitively impaired patients to self rate will depend on the structure, length, design and complexity of each questionnaire it is suggested that a follow up study be undertaken to assess the required MMSE-3MS scores required for the recommended self report questionnaires under different modes of administration.

## **12.3 Proxy Measurement**

This section discusses the issues surrounding the proxy or informant (or surrogate) measurement of a person who has dementia. It covers the following areas: definition of proxy measurement; the importance of direct measurement; highlights of recent research; the advantages and disadvantages of proxy measurement; characteristics affecting patient scores and proxy ratings; suitable domains of proxy measurement; and a useful list of proxy / informant instruments. It includes some recommendations when using proxy measures and areas of further research.

### **12.3.1 Definition of Proxy Measurement**

Snow, et al. (2005a) make the important distinction between proxy data and other-rater data "Proxy data refer to those collected from someone who speaks for a patient who cannot, will not, or is unavailable to speak for him or herself, whereas we use the term other-rater data to refer to

situations in which the researcher collects ratings from a person other than the patient to gain multiple perspectives on the assessed construct.” This is then related back to the measurement model underlying the data collection: Proxy measures need to be in accord with patient self reports, while other-rater data measures need to be in accord with an overarching construct where the patient and the other-rater data are component parts.

### 12.3.2 The Importance of Direct Measurement

At the outset this section supports the viewpoint that proxy measurement should be seen as complementary to the direct assessment of patients. Where possible the direct assessment of people with dementia should be attempted. The recent study by Byrne-Davis, et al. (2006) where a small group of dementia patients (n=25) could talk about their quality of life in focus groups, “challenges the heavy reliance of proxy completion QOL measures for this population” (page 863). Mozley, et al. (1999) also undertook direct interviews with dementia patients and recommended the following criteria for interviewability: “minimum level of orientation to place, attention and language skill” (page 782). In 2005, Snow, et al. (2005b) reported when examining depression measurement that the presence of dementia does not predict inaccurate depression self-reports but it seems to be that deficit unawareness or lack of insight may be a more important factor.

In their major review of general health status measurement / HRQoL for people with learning disability and acquired brain injury, Riemsma, et al. (2001), found few measurement studies of this issue. They recommended: “Studies should include a large number of respondents with different levels of cognitive impairment, so that differences in the instrument’s validity for different groups of people with cognitive impairment can be assessed (page 30).

As Harper (2000) said “More research is needed on the natural history of confusion for those people in early stages of dementia to determine when direct interviews can most appropriately be conducted . . . In general more research is needed on patient-focus measures (direct / interview assessment). This is particularly important for the development of standards to determine who is appropriate to interview for a particular domain” (page 509).

This section, by supporting further research into the direct assessment of people with dementia, also supports to the call to action made by Lezzoni in *Medical Care* in 2002 that test developers should be using universal design principles so that everybody can be accommodated when using the test regardless of age or disability (see also Wingfield, 1999). Wingfield (1999) sets out the key issues to watch out for when assessing the elderly: auditory acuity, the capacity of working memory and the rate at which speech input can be processed; though individual differences can be wide. Or as Park (1999) said:

“As we age, there are deficits in speed of information processing, working memory function, and sensory function. These deficits can have a substantial impact on the responses made by survey respondents such that age difference reflect difference in the ability to process and respond to the question rather than reflecting opinions and beliefs regarding a given survey question. It is important, as well, to recognize that a respondent’s behaviour occurs in a context and that the meaning of resource declines for everyday life will be somewhat blunted for highly familiar situations where older adults can rely on familiarity and automatic process to guide their behaviours” (page 67).

Schechter, et al. (1999) also points out that the presence of chronic disease or disability (e.g. arthritis) complicates health status measurement. O’Rourke, et al. 1999, Herzog, et al. 1999 and Schechter, et al. (1999) give further guidance in the area of designing surveys for older populations.

For those interested in alternative methods, Harper (2000) also includes other types of measurement including: observational assessment (laboratory or naturalistic); medical chart review, physiological (e.g. blood pressure or saliva testing) and technological (e.g. video cameras).

The following two sections list the advantages and disadvantages of proxy measurement and are based on the following works: Harper (2000) in Kane and Kane (2000) and Neumann, et al. (2000) which examine aging and the assessment of older adults; and dementia specific papers by Novella, et al. (2001), Novella, et al. (2006); as well as the review of the IQCODE by Jorm (2004).

### 12.3.3 Highlights of Recent Research

The following highlights recent research in the scientific literature. This includes papers by Novella, et al. (2006), Snow et al. (2005b), and Edelman, et al. (2005) which applied outcome measures to people with dementia. Novella, et al. (2006) used the SF-36 (interview format) with people with Alzheimer's disease; while Edelman, et al. (2005) examined the perspectives of staff, observers and residents in disease specific quality of life measurement (using the QOL-AD, DQoL, ADRQL, DCM and MMSE). Snow, et al. (2005b) examined the issue of cognitive status (ADAS-Cog) and self-report accuracy for depression (GDS Yesavage) in people with dementia, suggesting it is issue of deficit awareness, rather than the presence of dementia, which affects the accuracy of reports.

Yasuda, et al. (2004), examined the abilities of proxies to detect changes in functional status; and Eslinger, et al. (2005) examined self-awareness deficits in patients with Frontotemporal Dementia. Both of these papers used innovative methods to analyse proxy data.

In other research, Novella, et al. (2001) found that nurses aides had the worst agreement amongst professionals with patient scores for the measurement of health related quality of life (using the 17 item Duke Health profile), supporting the notion that there are professional differences in rating due to training. In this regard, the literature review by Neumann, et al. (2000) notes that while clinicians can act as proxies "few content areas appear to have been studied extensively" (page 1652).

Recently, Watson, et al. (2004) found that a simple question about "memory loss" with family caregivers had a poor relationship to cognitive impairment (as measured by screening + CERAD neuropsychological battery) in a community survey. This supports the view that subjective matters are harder to measure with proxy raters.

In terms of measurement issues, Neumann, et al. (2000) found in their literature review that co-residence seemed to improve the agreement between proxy rating and care recipient's self-report. Whilst Magaziner (1997) found better agreement with symptom present or absent judgements than those concerning symptom intensity ratings (from Snow, et al. 2005b).

Some Australian based work has also been undertaken in this area. Waite, et al. (1999) used an informant based measurement using the Clinical Dementia Rating (CDR) and Kemp, et al. (2002) used the CAMDEX – CAMCOG and informant judgements on recent memory, delayed memory, language and concentration. Kemp, et al. (2002) found that 40% of informant gave discrepant responses in one of these areas. They found that under-reporting was associated with milder dementia or MCI, lower levels of patient education, and poorer delayed remote memory; while over-reporting tended to occur when the criteria for dementia were met. This paper also supports the view that internal or subjective matters are harder to measure with proxies.

### 12.3.4 Advantages of Proxy Measurement

The advantages of proxy measures include:

- They are a way to examine premorbid ability (McDowell, 2006 commenting on the IQCODE).
- With regard to longitudinal studies – they are useful when a patient dies and there is a need to ascertain whether the patient developed dementia before their death (Waite, et al. 1999). As

Neuman, et al. (2000) said, proxies also permit longer follow-up periods as data collection is not dependent on client's capacity to respond.

- The use of proxy measures may improve the response rate for surveys and studies, as it allows individuals to be included who otherwise may be left out of the research (Harper, 2000) - either because of acute illness, lack of co-operation, death or low education and literacy (Jorm, 2004).
- In many cases there are well-established validities for proxy responses (Harper, 2000).
- They are not as time-consuming and expensive when compared to laboratory tests and naturalistic observation (Harper, 2000).
- Data can be constructed in the same format for proxy and direct responders, thus making analysis easier (Harper, 2000). Novella, et al. (2006) also points out that when using different instruments between informants and patients this will have an effect on scores.
- Proxy results are unaffected by the patient's education or premorbid ability and proficiency in the culture's dominant language (Jorm, 2004). (However, in response to the comments in Jorm (2004) - are the results affected by the proxy's premorbid ability and proficiency in the culture's dominant language? Usually spouse or child proxies come from the same background as the patient).

Another key advantage of proxy measurement is that when used as complementary piece of data to brief cognitive tests they can improve screening accuracy (Jorm, 2004; and a specific example is provided by Tierney, et al. 2003). Li, et al. (2006) also successfully combined informant measurement with the assessment of IADLs). This is also the logic behind the development of the GPCOG (Brodaty, et al. 2000) and the CSI-D (Hall, et al. 2000).

### 12.3.5 Disadvantages of Proxy Measurement

The disadvantages of proxy measurement include:

- The requirement for the informant to have significant exposure to the patient and the "best" sources may not be available (Harper, 2000).
- Some sampling bias as some groups more likely to have proxies than others (Harper, 2000).
- There is a potential for conflict of interest between the patient and the proxy. These depend on the environmental and social context (e.g. fear of nursing home admission) (Harper, 2000).
- Proxy measures rely on good recall of events (Harper, 2000).
- Formal caregivers as proxies may be subject to a number of biases, including time effects due to the time of their shift (e.g. afternoon shift vs. night shift), labelling of the patient as difficult, and the numbing to the behaviours shown on the ward.
- Phone interviews with proxies are more likely to result in missing information (Harper, 2000).
- They are based on impressions, not observations (Harper, 2000). Or as Harper (2000) says "Proxies tend to make inferences based on dispositional (in terms of the subject's likes and dislikes) rather than situational (based on situations that may influence behaviour) information." (page 487).
- Observers give more weight to negative rather than positive information (Novella, et al. 2006).
- There is a high degree of inference involved in proxy assessments (Harper, 2000).
- Proxy ratings subject to caregiver burden and stress (Novella, et al. 2006; Sands, et al. 2003).
- There is greater disagreement in scores between people with dementia and their proxies when the patient has worse health and worse cognitive status (Novella, et al. 2006).

- Proxy assessment are also subject to bias and motivation depending on the context (environmental and social) where informal caregivers adapt to the care giving load, or respond in socially desirable ways to questioning. As Jorm (2004) says scores are affected by the mental health of the informant and the quality of the relationship.

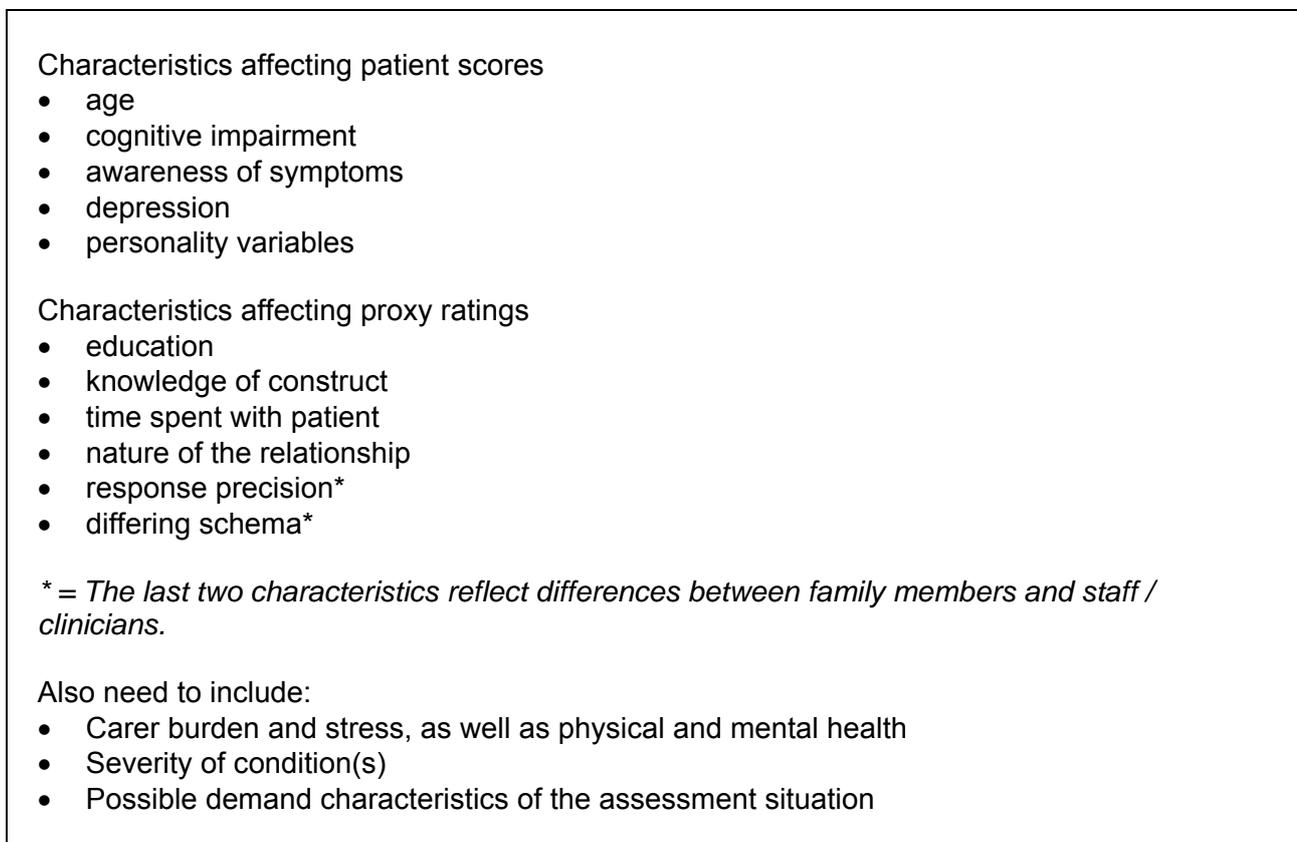
Jorm (2004) summed up the issue of potential biases well: “The major weakness of the IQCODE is that some informants provide less valid data than others. However, little is known about which informants provide the best data. More information is needed on how validity is affected by variables like age, education, frequency of contact, and not living with the subject. Furthermore, little is known about how the purpose of the screening might affect informant ratings. For example, in a clinical situation where a carer wants support services, they might overrate cognitive decline, whereas in a community screening situation they may be reluctant to support a diagnosis of dementia in a loved one. There is also a need for the development of approaches to handling any lowered validity, whether by exclusion of certain informants or by adjustment of IQCODE ratings” (page 15).

See also the section below on the suitability of proxy measurement for the various assessment domains. For some domains proxy measurement is not supported in the literature at present.

### 12.3.6 Characteristics Affecting Scores for Patients and Proxies

Snow, et al. (2005a) neatly summarizes the characteristics affecting scores for patients and proxies (from pages 1685-1687). These are outlined in the figure below.

**Figure 4 Characteristics Affecting Patient Scores and Proxy Ratings as Outlined by Snow, et al. (2005a)**



The additional issues of carer burden and stress, as well as physical and mental health have also been added to this Figure. Also included is the severity of the condition(s) and possible demand characteristics of the assessment situation.

Novella, et al. (2006) concerning the rating the SF-36 studied the agreement between different types of proxies (family members: care staff) with the dementia patient's own ratings. In this study although all proxies rated the patient's health status as poorer on almost all dimensions, the care staff ratings actually had greater agreement with the patient ratings than did those provided by family members. In the selection of the most appropriate proxy to use perhaps the key issue to consider is the closeness to the patient and the frequency of their interaction. Although Cummins (2002) suggests that proxies where used should be partners or peers in a nursing home setting it is quite possible that family members may not relate to the dementia patient as often as a formal carer and there may be substantial differences between family carers concerning how often they visit their family member. Naglie, et al. (2006) only used family/informal carer proxies who visited their relative with dementia 3 or more times per week and this might be used as a rule of thumb. However, in community settings the partner or close family member may be the best proxy, and in some cases, they may be the only viable source of information. There appears to be a need for further research to provide advice concerning the selection of appropriate proxies across settings and with regard to the severity of illness of the person with dementia.

### 12.3.7 Suitable Domains of Proxy Measurement

From the scientific evidence it is clear that proxy measurement is more suitable to certain domains than to others. Neumann, et al. (2000) make the following findings of their view of the literature using proxy data. These are presented in the table below.

**Table 63 Main Findings from the Literature Review of Neumann, et al. (2000)**

<ul style="list-style-type: none"> <li>• <i>Proxy and subject reports are often comparable in describing levels of functioning, although proxies tend to identify more impairment</i></li> </ul>
<ul style="list-style-type: none"> <li>• <i>Researchers generally report good agreement in comparisons of proxy and subject assessments in describing overall health, chronic physical conditions, and physical symptoms</i></li> </ul>
<ul style="list-style-type: none"> <li>• <i>Relatively little is known about the comparability of proxy and subject reports on healthcare utilization or in the concordance between responses and data from medical records or claims</i></li> </ul>
<ul style="list-style-type: none"> <li>• <i>Limited evidence shows high agreement between proxies and patient preferences for type or setting of care and lower agreement for preferences for health states</i></li> </ul>
<ul style="list-style-type: none"> <li>• <i>There is low to moderate agreement between proxies' and subjects' reports of depressive symptoms and psychological well-being, with proxies describing more problems</i></li> </ul>
<ul style="list-style-type: none"> <li>• <i>Proxies are often in agreement with subjects on reports of cognitive status, although proxies may overestimate cognitive abilities</i></li> </ul>
<ul style="list-style-type: none"> <li>• <i>Proxies tend to describe more functional impairment among persons with dementia compared with self-reports, especially with respect to instrumental functions</i></li> </ul>
<ul style="list-style-type: none"> <li>• <i>Spouses, children, and other close family members tend to be capable proxies, although proxy reports may be influenced by caregiving burden</i></li> </ul>

In summary, Neumann, et al. (2000) found that there is evidence to support the use of proxies for the measurement of function, physical health and cognition; while there are some problems with the measurement of emotional / behavioural symptoms and depression. Proxy ratings, on the whole, were more negative for functioning and mental health, than those provided by the care recipient. This pattern of reporting was most apparent for care recipients with cognitive impairment and for proxies who reported more caregiver burden and stress.

This view of the literature is supported by Snow, et al. (2005a) who stress that the more objective the construct the more amenable it is for the proxy measurement approach (see page 1682). They argue this is why there is less discrepancy between proxy and patient reports for physical symptoms and functional activity, than for depression symptoms and quality of life. This view makes a certain amount of sense as the proxy rater can see the outward signs of depression (e.g. irritability, tiredness) but does not have access to the patient's inner life, therefore making it harder for them to report on the condition properly. This is not the same for measures of functional activity where the proxy rater can see / experience the patient's performance directly on everyday tasks.

Snow, et al. (2005a) also outlines analysis techniques and interpretation guides for the level of agreement. As well they examine the theories that attempt to explain differences between proxy and direct assessment (including response shift, favourable ratings of self, cognitive dissonance theory, self-awareness theory and self-schema theory).

### 12.3.8 Proxy / Informant Instruments

Table 64 is a list of proxy instruments which can be used as a resource for informant questions. It should be noted that this list focuses on published measures rather than measures used for one off studies or papers.

**Table 64 List of Proxy / Informant Instruments**

Instrument	Original Article Cite Author(s) + Publication Date
<b>Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) *</b>	Jorm, et al. (1989)
<b>AD8</b>	Galvin, et al. (2006)
<b>BPSD Checklist</b>	Snellgrove and Hecker (2005)
<b>Community Screening Instrument for Dementia (CSI-D) Interview</b>	Hall, et al. (2000)
<b>Activities of daily living questionnaire (ADLQ)</b>	Johnson, et al. (2004)
<b>Revised Memory and Behaviour Problem Checklist</b>	Teri, et al. (1992)
<b>Psychogeriatric Assessment Scales (PAS) Informant interview – Cognitive Decline and Behaviour Change *</b>	Jorm, et al. (1995)
<b>GPCOG Informant Interview</b>	Brodaty, et al. (2000)

\* = related instruments

The IQCODE, the AD8 and GPCOG Informant Interview have been described in Section 6: Instruments for the Assessment of Cognitive Status.

In addition to this list there are those measures with informant sections or interviews used in diagnostic application, most notably the Cambridge Mental Disorders of the Elderly Examination (CAMDEX) - Informant Section (Roth, et al. 1986) and the Canberra Interview for the Elderly (CIE) Informant Interview – Cognitive Decline and Behaviour Change (Henderson, et al. 1992) (The CIE is also related to the IQCODE). Harper (2000) also notes that there are also informant versions of numerous BPSD instruments like the Cornell Scale for Depression in Dementia (CSDD) and the Geriatric Depression Scale (GDS Yesavage) (see Snow, et al. 2005b); as well as for measures of function like the Blessed, the Cleveland ADL scale, Functional Assessment Staging and Lawton and Brody's IADL instrument (see Kemp, et al. 2002).

### 12.3.9 Recommendations when using Proxy Measures

Where it is possible and feasible HRQOL and other subjective phenomena should be assessed by patient self report rather than by proxy report. The previous section has provided a discussion of cognitive impairment in relation to the capacity to self rate. However, it is understood that where patient's experience severe dementia it may not be possible to assess such phenomena by patient self report and the use of proxy measures may be unavoidable.

Below are a number of recommendations when using proxy measures:

- Proxy reports should be examined for three potential biases: (1) the cognitive status of the proxy (as many elderly people are cared for by an elderly spouse carer, who may themselves be impaired or unwell, but to a lesser degree); (2) the health status of the proxy; and (3) the level of carer burden and stress (Harper, 2000).
- There is usually a trade-off between those “with the greatest amount of contact and those with more training” (Harper, 2000, page 488). However, generally, where a proxy report is used information should be collected from the family member/carer or care staff member that is closest to the patient and has the greatest degree of interaction with the patient.
- Proxy reports should be based on usual behaviour rather than extreme or rare behaviours (Harper, 2000).
- Proxy reports should be based on observable phenomena like physical symptoms and functioning, rather than subjective phenomena like depression, social isolation and quality of life (Snow, 2005a).

### 12.3.10 Areas for Further Research

From this review of the literature a number of areas for further research stand out:

- It is necessary to examine whether the training of proxies to make structured observations, improves the quality of their ratings (Harper, 2000).
- It is necessary to examine how the framing of questions, the use of terminology and the administration of instruments influences the results of proxy reporting. For instance, asking a proxy about how the person with dementia performs an everyday task, like using the telephone, could be either broad and general, or broken down into a number of specific and observable component activities (i.e. getting the telephone number, dialling the telephone number, etc). These two different approaches to the question might generate different proxy answers based on how the proxy rater interprets the term “using the telephone”. The same definitional problems apply when professional terms are used to make severity ratings as these may not be fully understood by proxy carers (Neuman, et al. 2000).
- There is a need to compare proxy reports with performance based measures and information from medical records and health care utilization (Harper, 2000; Neuman, et al. 2000).

- Many of the recent papers use single or dual item informant measures (e.g. Tierney, et al. 2003; Watson, et al. 2004; Li, et al. 2006). Further research is required to ascertain whether these items have the requisite accuracy as compared to the longer proxy measures.

## **12.4 Dementia Measurement Issues with Culturally and Linguistically Diverse (CALD) Populations**

### **12.4.1 Introduction**

As in many other developed countries, a growing cultural diversity in Australia highlights the importance of developing and providing health care services that are culturally appropriate and sensitive to the needs of those who do not have the same cultural backgrounds as the predominant culture in Australia. According to Access Economics in 2005, 33 per cent of older Australians (over 60 years) were born overseas, while 16 per cent of older people spoke a language other than English at home. It is estimated that approximately one in eight people with dementia in Australia (12.4%), i.e. about 25,400 of 204,800 people with dementia, do not speak English at home, of which the indigenous population comprises approximately 0.1 per cent. The report also points out a significant increase in a number of people with dementia whose language spoken at home is either an Asian or Middle Eastern language in the next four decades (Access Economics, 2006, pp. 2-5).

Cultural diversity entails a complex fabric of interactions between people and environment, beyond racial, ethnic and linguistic differences. Andary, et al. (2003, p. 27) define culture as “a process arising out of shared ethnicity, religion, beliefs, language, knowledge, values, meanings and rules, which enable members of a given society to communicate, live, work, anticipate and interpret each other’s behaviour and motives”. Culture is heterogeneous and there is a great deal of variations within CALD communities (Howe, 2006). Hence treating people in a particular way based on their racial or ethnic background or a country of origin, or perceiving their behaviours as a means of explaining a particular culture poses a risk of stereotyping without valuing their individuality. It is argued “race, ethnicity, and culture are interrelated, complex, and sometimes ‘loaded’ concepts. Defining groups in these terms ignores enormous within-group heterogeneity and overshadows group differences on factors such as education, vocabulary, reading level, and acculturation that directly affect test performance” (Manly and Teng, 2005, p. 269).

There are commonalities shared by all populations, regardless of their cultural backgrounds, in terms of their need for support and care in maintaining quality of life and wellbeing. How the need for support and care of CALD communities is met, however, begs a special consideration. Whilst acknowledging differences between and within CALD communities there are unique challenges they experience because of the differences they have from those of the mainstream populations. The most commonly addressed challenge relates to accessing appropriate health care services. People from CALD communities have lower rates of access to various aged care and community services, for example Aged Care Assessment Teams, Home and Community Care programs and Community Aged Care Packages, than those born in Australia (Karmel, et al. 2003, cited in Bartlett, Rao and Warburton, 2006; Lister and Benson, 2006). Grounds for this are multifaceted, but common determinants are likely to include language difficulties, often resulting in lack of knowledge of the existing health care services, religious belief and observance, ritual and ethnic practices, and family involvement (Iliffe and Manthorpe, 2004; Lister and Benson, 2006). Language in particular appears to have a major influence in service access, shown in numerous literatures reporting the notion of differential rates of access to service between people from English speaking and non English speaking backgrounds (Davis, et al. 1996; Hassett, et al. 1999; LoGuidice, et al. 2001).

How health care services, models of care and relevant policies are, and should be, developed to meet the unique needs of CALD communities is an important issue to be addressed, however, is

beyond the scope of this project. For the purpose of the DOMS project, this section will focus on issues of CALD populations within the arena of dementia measurement and provide recommendations accordingly. This section summarises current assessment issues, latest research developments, guidelines for the assessment of non-English speaking people with dementia, analysis of DOMS instruments for suitability with CALD populations, and recommendations for the assessment of CALD populations.

### 12.4.2 Assessment Issues

Cultural competency of practitioners and interpreters such as sufficient language skills and cultural awareness is critical in conducting valid and reliable dementia assessment. One of the outstanding problems with dementia assessment in CALD populations relates to its reliance on instruments originated in Western culture that are developed and validated predominantly in developed countries such as USA, UK and Canada. Daker-White, et al. (2002) in their review of the diagnosis and prevalence studies conducted in the 1990s noted that 65% of the articles they found on dementia and ethnic minority groups originated from USA, reflecting issues affecting that culture. In the recent report at the International Psychogeriatric Association Consensus meeting Chiu and Lam (2007) argue the notion of 'one size fits all' in relation to dementia outcome measures originating from Western culture is not only inappropriate for those from Eastern culture, but also is potentially misleading. They point out barriers such as linguistic and cultural differences, high illiteracy rates, and lack of resources and time further complicate the use of the existing outcome measures in developing countries. As discussed by Shah, Dalvi and Thompson (2005) the issue extends to those ethnic minority groups in developed countries, which is the case in Australia.

These issues are not insignificant and require test users, adaptors and developers in multicultural Australia to undertake detailed data collections for CALD groups. As the accepted, international guidelines state:

"9.1 Testing practice should be designed to reduce threats to the reliability and validity of test score inferences that may arise from language differences.

9.2 When credible research evidence reports that test scores differ in meaning across subgroups of linguistically diverse test takers, then to the extent feasible, test developers should collect for each linguistic subgroup studied the same form of validity evidence collected for the examinee population as a whole." (page 97)

American Educational Research Association, American Psychological Association, and the National Council on Measurement in Education (1999) *Standards for educational and psychological testing*.

This section attempts to cover the key aspects of assessment by providing: some important issues when using and interpreting assessment tools for CALD populations; some examples of cultural and language differences in dementia assessment; and general principles and recommendations of dementia assessment for CALD groups.

#### 12.4.2.1 Using and Interpreting Assessment Tools for CALD populations

The following figure is a list of issues to consider when using and interpreting assessment tools for CALD populations, especially when using cognitive screening instruments that require direct questioning. It is based on the work *Culturally appropriate dementia assessment* by the Centre for Applied Gerontology – Bundoora Extended Care Centre (1996) with some additions.

**Figure 5 List of Issues when Using and Interpreting Assessment Tools for CALD Populations**

Assessment instruments

- Most instruments have been designed for western, white and English speaking populations - which Teng and Manly (2005) refer to as the “majority culture”.
- Lack of norms for cultural and educational sub-groups.
- More than a literal translation is required for different language groups (see the examples below).
- Some items related to acculturation – e.g. In which month is Australia Day?

Assessment of language

- Are you looking at real impairment or a fragile second language?
- The present native language may differ from the language clients have been taught to read and write with (see Howe, 2006).

Language testing issues / use of interpreters

- Different dialects.
- Creolisation of the original language with English.

Cultural differences

- Each ethnic / cultural group is not homogeneous but culturally and socio economically diverse.
- Individuals differ with the amount of time in Australia and the extent they have retained the culture of their country of origin.

*(These issues are in addition to the standard assessment issues like test-taker anxiety, practice effects, fatigue effects and motivation)*

Manly and Teng (2005), in their review of neuropsychological testing for people with dementia, highlight many of the above issues, adding concerns about the fact that many older people may have little practical experience with the testing situation; the poor ecological validity of tests e.g. remembering lists of words and abstract thinking tests; and the effect of the test administrator on test performance. Hargrave (2006) also comments on the US based research on the concept of ‘stereotype threat’ which is the “effect of attention diverting from a task at hand to the concern that one’s performance will confirm a negative stereotype about one’s group” (page 39).

**12.4.2.2 Some Examples of Cultural and Language Differences in Dementia Assessment**

During this review some examples of cultural and language differences in assessment were found, including:

- In some Asian cultures it is a sign of politeness to say ‘Yes’. To say ‘No’ is rude (Dragans, 1984, cited in Centre of Applied Gerontology, 1996).

- The time-conscious and clinical approach of doctors is not always understood by the Vietnamese elderly. They may interpret the ‘professional’ behaviour of doctors as representing disinterest, aloofness, apathy and most importantly, disrespect (Thomas, 1991, cited in Centre of Applied Gerontology, 1996).
- Chinese elders may have been taught to write with brushes instead of pens which require a different form of motor control (Dick, et al. 2006).
- ‘Memorise’ is often translated in Italian to a word which means ‘learn’. Perhaps a better translation would be ‘commit to memory’. However, this is problematic when single words are required to be read out (Centre of Applied Gerontology, 1996).
- In the MMSE the phrase ‘No ifs, ands or buts’ would not be familiar to someone unfamiliar with the expression. Also a purely literal translation would not capture the phrase’s ‘articulatory complexity’ (Centre of Applied Gerontology, 1996).
- In Italy for example the terms Dementia or Alzheimer’s disease denote craziness or insanity (Centre of Applied Gerontology, 1996).
- The English word ‘blue’ is sometimes used as a lay term for depression. In Vietnamese it means “hope” or a state of ‘calmness’. In Russian slang, ‘blue’ means drunk, while in German ‘blue’ can be used to refer to someone who is ‘gay’ (Andary, et al. 2003, cited in Bartlett, Rao and Warburton, 2006).
- In Chinese remembering the months of the year backwards is relatively easy as the names of the months are month 1, month 2, month 3 etc. Thus you only have to recite from 12 to 1 backwards. (Manly and Teng, 2005).
- In digit span tests where subjects repeat back numbers to the examiner, in Spanish the majority of digits (1 to 9) have two syllables, while in Vietnamese the numbers are monosyllabic (Dick, et al. 2002, cited in Manly and Teng, 2005). This complexity is also mirrored in the syllables used in animal names, which affect a commonly used test to list all the animal names you can think of in one minute.

The above examples illustrate potential and real problems with adopting dementia assessment instruments without considering cultural and linguistic diversity.

### 12.4.3 Current Research Developments

Two important research developments are worth noting, namely the role of differential item of functioning to examine the responses of CALD groups, and the development in Australia of the RUDAS.

#### 12.4.3.1 Differential Item Functioning

Differential Item Functioning (DIF) is a common method for examining the measurement equivalence or non equivalence of the performance of items and instruments with different groups of people. Some examples, provided above, highlight the importance of this issue. Recent papers in Medical Care have also examined DIF or measurement equivalence of the MMSE with English and Spanish samples (Ramirez Diaz, et al. 2005; Morales, et al. 2006; and Orlando, Edelen, et al. 2006) and similar work needs to be undertaken for major language versions of the MMSE-3MS in Australia.

“Attention to measurement equivalence is not an esoteric, psychometric issue that has little or no consequences for science, policy, or medicine. Understanding and assessing measurement equivalence is fundamental to science—to developing outcomes instruments, to theory building, to testing hypotheses, to screening and diagnosing individuals, and to evaluating health service delivery programs. Culturally fair patient outcome assessment (with ‘culture’ defined broadly,

including gender, age, racial, ethnic, socioeconomic, geographic, and language variations) is crucial when individual decisions are in the balance such as with mental or physical screening, diagnostic, and referral decisions. If items in outcome measures are biased, detection rates can be biased (overestimated or underestimated), leading to over- and under detection and over- and under treatment. At the group level, measurement bias can have consequences in terms of errors in hypothesis testing and, thus, internal and external validity” (Mc Horney and Fleishman, 2006, p. S205).

McHorney and Fleishman (2006) suggest the importance of: looking behind the data to examine why DIF occurs; better understanding of a test taken from a personality and motivation perspective using qualitative techniques; and developing recommendations on what to do when DIF is detected in established instruments.

It is suggested that further studies analysing the measurement equivalence of the core recommended measures be undertaken for major language groups within Australia.

#### **12.4.3.2 Development of the RUDAS**

In terms of instrument development with CALD populations, the Rowland Universal Dementia Assessment Scale (RUDAS) (Storey, Rowland, Basic, Conforti, and Dickson, 2004) has been reviewed here (see Section 6) as a promising new instrument for multi-cultural cognitive screening in Australia. However, it should be noted that this new scale has not been studied in comparison with other more established cross cultural instruments like the Fuld Object Memory Evaluation (FOME) (Wall, et al. 1998, Mast, et al. 2001), the Cognitive Abilities Screening Instrument (CASI) (Teng, et al. 1994), the Cross-Cultural Cognitive Examination (CCCE) (Glosser, et al. 1993) and the Community Screening Instrument for Dementia (CSI-D) (Hall, et al. 1996). Here the work of Klimidis, et al. (2004), in developing a brief 4 item functional English proficiency measure for health surveys, may also be useful. This measure has been used in Turkish, Macedonian, Spanish, Italian and Greek speaking communities.

It is suggested that validity studies of this kind be undertaken with the RUDAS.

#### **12.4.4 Guidelines for Assessment of Non-English Speaking People with Dementia**

A recent report from Alzheimer’s Australia – National Cross Cultural Dementia Network (NCCDN) (Grypma, Mahajani and Tam, 2007, pp.3-9) provides guidelines for practitioners, service managers and policy makers in assessing non-English speaking people with dementia. It recommends:

- History taking when screening or assessing patients from non-English speaking backgrounds should seek the same information as from a patient from an English speaking background, and should do so in a culturally appropriate and sensitive manner. To do this effectively the practitioner needs to establish an understanding of the culturally relevant issues for each patient.
- Practitioners need to utilise a range of communication strategies to maximise two way information flow during screening, assessment or management of people with cognitive impairment. When working with patients from non-English speaking backgrounds and a third party such as an interpreter, the practitioner also needs to monitor information flow to and from the interpreter, and non-verbal cues.
- When screening, assessing or treating a patient with cognitive impairment from a non-English speaking background, it should be undertaken using the approach that will maximise communication and cultural awareness. The best way for this to occur is through a competent practitioner who is fluent in the patient’s language. Efforts should be made to establish and promote a referral network of competent bi- and multi-lingual practitioners that can be readily accessed to undertake appropriate screening and assessment.

- Environments used for screening, assessing and treating people with cognitive impairment should be communication-friendly. Additional time should be allocated for sessions with patients from non-English speaking backgrounds, particularly when a third party such as an interpreter is used.
- Practitioners need to be aware of the limitations of some existing screening and assessment tools for cognitive impairment in their use with patients from non-English speaking backgrounds, and of emerging research validating tools that may be more appropriate for use with patients from non-English speaking backgrounds.

The report also offers recommendations for promoting early assessment and prevention of dementia as well as suggestions for system relevant strategies and policy development.

#### **12.4.5 Valid and Reliable Dementia Outcome Measures for CALD Populations**

Table 65 provides a summary of the recommended DOMS instrument with regards to cultural applicability and cultural adaptations. This analysis reflects a face validity assessment of the instruments and is separate to the issues related to the accurate and feasible translation of certain items.

**Table 65 Recommended DOMS Instruments - Analysing Items for Acculturation and other issues**

Domain	Instrument Name + Abbrev.	Cultural Applicability and Cultural Adaptations	Description	Issues to be considered
Dementia Specific Assessment Instrument	Global Deterioration Scale (GDS)	A German study did not report any difficulties with translating the GDS for use in the clinical setting and showed no differences in scoring or interpretation by translation. Diaz (2005) states that the GDS is used widely in memory clinics throughout the European Union (38% of all measures used to screen for and determine dementia severity). It has been reported in the international literature, however, it is unknown if the measure is used in culturally diverse health populations.	Clinical rating scale	-
Dementia Specific Assessment Instrument	Clinical Dementia Rating (CDR)	Instrument is available in Chinese for Taiwan, Czech, Dutch, Dutch for Belgium, English for Australia, English for the UK, Finnish, French, French for Belgium, French for Canada, German, German for Austria, Hebrew, Polish, Spanish, Spanish for Argentina, Swedish.	Clinical rating scale	-
Dementia Specific Assessment Instrument	Dementia Severity Rating Scale (DSRS)	Mostly American studies identified, but its simplicity makes it suitable for translation.	Proxy rating scale	-

Domain	Instrument Name + Abbrev.	Cultural Applicability and Cultural Adaptations	Description	Issues to be considered
Dementia Specific Health Related Quality of Life Instrument	Quality of Life in Alzheimer's Disease (QOL-AD)	Validated translations are available in French, Japanese, Mandarin, Portuguese, Danish, German, Italian, Spanish, Swedish and Greek.	Self-report interview or paper version	Issue of whether it is culturally appropriate to disclose such information to others.
Dementia Specific Health Related Quality of Life Instrument	DEMQOL	This is a new instrument, so as yet there are no other translations.	Self-report interview or paper version	Issue of whether it is culturally appropriate to disclose such information to others.
Dementia Specific Health Related Quality of Life Instrument	Quality of Life in Late Stage Dementia (QUALID)	The instrument is available in Swedish, Finnish, German, and Lithuanian. Details can be obtained from the author.	Clinical rating scale with informant / proxy	Issue of whether it is culturally appropriate to disclose such information to others.

Domain	Instrument Name + Abbrev.	Cultural Applicability and Cultural Adaptations	Description	Issues to be considered
Instruments for the Assessment of Cognitive Status	Modified Mini-Mental State Exam (3MS)	Similar to the MMSE, the 3MS in its original format may not be culturally sensitive. However, adaptations of/adjustment to the 3MS have been made over the years in various translated versions appropriate to the specific culture, with moderate to high successful outcomes reported, including,: French (Bravo & Hebert, 1997a, 1997b; Cappeliez et al., 1996; Patenaude & Baillargeon, 1996; Viscogliosi, Desrosiers, Gauthier, & Beauchemin, 2000), Korean (Jeong, Cho, & Kim, 2004), German (Alexopoulos, Perneckzy, Cramer, Grimmer, & Kurz, 2006; Sandholzer, Breull, & Fischer, 1999); Nigerian population (Baker, Ogunniyi, & Osuntokun, 1995; Ogunniyi, Osuntokun, Lekwauwa, & Falope, 1992); Hungarian (Merkli, Pal, & Horvathne, 2001; Tariska & Paksy, 2000); and Mexican American population (Miller et al., 2003; Wu et al., 2003; Wu et al., 2003).	Cognitive test	Note the language translation issue eg. "No ifs, ands or buts"
Instruments for the Assessment of Cognitive Status	General Practitioner Cognition Scale (GP-COG)	No information published regarding culture or language bias. The instrument has also not as yet been translated into other languages.	Cognitive test	Some problem noted. Cultural factors may apply to the memory questions 1, 5 and 6: requiring the learning of an address and discussing "something that happened in the news".
Instruments for the Assessment of Cognitive Status	Minimum Dataset - Cognition (MDS-Cog)	There is no information available about cultural applicability. Translations are as yet not available.	Clinical Rating Scale	-

Domain	Instrument Name + Abbrev.	Cultural Applicability and Cultural Adaptations	Description	Issues to be considered
Multi-attribute Utility Measure	EQ-5D (formerly the EuroQoL)	<p>The EQ5D is available in the following languages: Afrikaans, Armenian, Basque, Bulgarian, Catalan, Chinese, Croatian, Czech, Danish, Dutch, English, Estonian, Filipino, Finnish, French, German, Greek, Hebrew, Hungarian, Icelandic, Indonesian, Italian, Japanese, Korean, Latvian, Lithuanian, Malay, Norwegian, Polish, Portuguese, Romanian, Russian, Sesotho, Shona, Slovakian, Slovenian, Spanish, Swedish, Thai, Tongan, Turkish, Xhosa, Zulu.</p> <p>The EQ5D was developed cross-culturally across the UK, the Netherlands, Norway, Finland and Sweden (Anderson et al., 1993). It was developed for use in a battery of other measures and for use in postal surveys hence its shortness (Brazier et al., 1993).</p>	Self-report instrument or interview version or proxy	Issue of whether it is culturally appropriate to disclose such information to others.
Multi-attribute Utility Measure	Assessment of Quality of Life (AQoL)	The AQoL has been translated into Canadian French and Danish. No particular difficulties were reported. However, there have been no tests of this in the reported literature.	Self-report instrument or interview version	<p>Minimal problem noted.</p> <p>Cultural factors may apply to questions on independent living and household tasks (questions 4 and 5) - the "can do" and "do do" problem - may require a change of wording.</p> <p>Issue of whether it is culturally appropriate to disclose such information to others.</p>
Measures of Social Function and Participation	De Jong Gierveld Loneliness Scale	<p>The de Jong Gierveld Loneliness Scale is available in English, Dutch, French and Italian.</p> <p>No evidence appears to have been published examining cross-cultural issues.</p>	Self-report instrument	<p>Minimal problem noted.</p> <p>Cultural factors may apply to Question 11 - "I can call on my friends whenever I need them".</p> <p>Issue of whether it is culturally appropriate to disclose such information to others.</p>

Domain	Instrument Name + Abbrev.	Cultural Applicability and Cultural Adaptations	Description	Issues to be considered
Patient Satisfaction Measures	Short Assessment of Patient Satisfaction (SAPS)	There is evidence throughout the literature that patient satisfaction is culturally specific. It cannot be assumed that an instrument that is relevant, valid and reliable in one culture retains those properties in another culture. Thus instruments developed overseas may not be appropriate in Australian settings.	Self-report instrument	Issue of whether it is culturally appropriate to disclose such information to others.
Carer Satisfaction Measures	Satisfaction with Care at the End of Life in Dementia Scale (SWC-EOLD)	There is no reported research involving any of these instruments examining whether the construct of carer satisfaction is culturally bound in any way; there does not appear to have been any cross-cultural validation work done on any of the measures reviewed. In short, there was no evidence for any instrument on this criterion referring to appropriate use by CALD or illiterate clients or with an interpreter. All instruments were therefore ranked similarly.	Self-report instrument for carer	Issue of whether it is culturally appropriate to disclose such information to others.
Measures of Functional Status - Generic	Functional Independence Measure (FIM)	NA	Clinical rating scale	Items include stairs and wheelchairs.

Domain	Instrument Name + Abbrev.	Cultural Applicability and Cultural Adaptations	Description	Issues to be considered
Measures of Functional Status - Generic	Barthel Index	NA	Clinical rating scale	Items include stairs and wheelchairs.
Measures of Functional Status - Generic	Katz	This scale has been used in the Scandinavian countries and German research, hence there are some translations available but not from a central source.	Clinical rating scale	-
Measures of Functional Status - Generic	Lawton & Brody IADL	NA	Clinical rating scale	Reflects activities suitable for Western and urbanised cultures (e.g. telephone, handle finances). Items applied according to gender role stereotypes.

Domain	Instrument Name + Abbrev.	Cultural Applicability and Cultural Adaptations	Description	Issues to be considered
Measures of Functional Status - Generic	OARS-IADL	NA	Clinical rating scale	Reflects activities suitable for Western and urbanised cultures (e.g. telephone, handle your own money).
Measures of Functional Status - Dementia Specific	Alzheimer's Disease Co-operative Study - ADL (ADCS-ADL)	Some clinical trials using the ADCS-ADL were noted to have occurred in Spain (Arrieta 2006), Sweden (Winblad et al. 2006), Latvia (Doody et al. 2004) and Bosnia and Hercegovina (Rustembegovi et al. 2003).	Informant / Proxy report	Reflects activities suitable for Western and urbanised cultures (e.g. watched television, microwaved food).
Measures of Functional Status - Dementia Specific	Disability Assessment for Dementia Scale (DAD)	Developed for the English and French languages (Gelinas et al. 1999). Korean (Suh et al. 2004b) and Chinese (Mok et al. 2005) versions are available; and papers have been noted to have been undertaken in Finland, Germany, Norway (Jones et al. 2004), Denmark (Stokholm et al. 2005) and Hong Kong (Lam et al. 2006).	Informant / Proxy report	Reflects activities suitable for Western and urbanised cultures (e.g. telephoning, finances and correspondence).

Domain	Instrument Name + Abbrev.	Cultural Applicability and Cultural Adaptations	Description	Issues to be considered
Measures of Functional Status - Dementia Specific	Cleveland Scale for Activities of Daily Living (CSADL)	No information found on this aspect of the scale's development. However, Fritsch et al. 2002 did find an effect for ethnicity on functional decline when comparing caregivers of European American and African American patients.	Clinical rating scale	Reflects activities suitable for Western and urbanised cultures (e.g. telephoning, money management).  Use of the descriptive terms appropriate and acceptable, may allow for value laden judgments by raters.
Global BPSD	Behavioural Pathology in Alzheimer's Disease Rating Scale (BEHAVE-AD)	The BEHAVE-AD has been translated into French (see Sclan, 1996), Swedish (Midlov, Bondesson et al. 2002), German (Auer, Hampel et al. 2000), Dutch (Engelborghs, Maertens et al. 2005), Spanish (Boada, Tarraga et al. 2006), Chinese (Chan, Lam et al. 2001), and Korean (Suh, Son et al. 2004).	Clinical rating scale	-
Global BPSD	Consortium to Establish a Registry for Alzheimer's Disease – Behavioral Rating Scale for Dementia (CERAD-BRSD)	The 48 item version of the instrument has been translated into French and Spanish.	Clinical rating scale	-

Domain	Instrument Name + Abbrev.	Cultural Applicability and Cultural Adaptations	Description	Issues to be considered
Delirium	Confusion Assessment Methodology (CAM)	The CAM is translated in several languages and used in various countries, demonstrating its satisfactory adaptations. Some examples of the studies include: in Chile, Chilean (Carrasco et al., 2005) and Spanish (Gonzalez & Barros, 2000); in Germany, German (Bickel et al., 2004; Galanakis et al., 2001); in Mexico (Villalpando-Berumen et al., 2003); in Italy (Caraceni et al., 2000; Grassi et al., 2001); in Brazil, Portuguese (Fabbri et al., 2001; Furlaneto & Garcez-Leme, 2006); in Spain (Diaz et al., 2001); in Argentina, Spanish (Regazzoni, Aduriz, & Recondo, 2000; Vazquez et al., 2000); in Belgium, Dutch (Lemiengre et al., 2006; Milisen et al., 2005); in Japan (Kawaguchi et al., 2006; Kudoh et al., 2004; Shigeta et al., 2001); in France (Bourdel-Marchasson et al., 2004); in the Netherlands (Kalisvaart et al., 2005; Kalisvaart et al., 2006).	Clinical rating scale	-
Delirium	Delirium Rating Scale (DRS-R-98)	The DRS-R-98 has been used successfully in various non-English background countries, including Japan (Takeuchi et al., 2007); Korea (Pae et al., 2004); Spain (Fonseca et al., 2005); and the Netherlands (de Jonghe et al., 2005; de Jonghe et al., 2007; de Rooij et al., 2006; Kalisvaart et al., 2005). The Spanish (Fonseca et al., 2005) and Dutch (de Rooij et al., 2006) versions of the DRS-R-98 have been the subject of validation demonstrating various psychometric properties of the translated versions. The original DRS is available in 11 languages, French, Italian, Spanish, Dutch, Mandarin Chinese, Korean, Swedish, Japanese, German, and Indian-language translations, which have been successfully applied in a variety of ethnicities and countries.	Clinical rating scale	Minimal problem noted.  The long term memory item (Question 12) should use either a formal memory question about 3 verbal or visual items or personal history information, rather than "general information that is culturally relevant".
Aggression	Rating Scale for Aggressive Behaviour in the Elderly (RAGE)	The RAGE is employed largely in English speaking countries, such as the USA, Canada, and UK. Cross-cultural reliability and validity has been established for the Chinese population (Lam, Chui & Ng, 1997) and in Scandinavia (Patel & Hope, 1992b). Translations of the English version of the RAGE are reliable and valid (Lam, Chiu & Ng, 1996; Patel & Hope, 1992b).	Clinical rating scale	Contains some items or terms which could be seen as western value judgements (eg. argumentative, shouted, used abusive language, critical).

Domain	Instrument Name + Abbrev.	Cultural Applicability and Cultural Adaptations	Description	Issues to be considered
Agitation	Cohen Mansfield Agitation Inventory (CMAI)	The CMAI has been translated into numerous languages. It is available in the following European languages: Dutch, Danish, French, German, Greek, Norwegian and two Spanish versions. In Asia and the Middle East it is available in Chinese, Korean and Japanese and Hebrew. Information about how to obtain these translations is available in the Instruction manual available from the authors.	Clinical rating scale able to be used by informant / proxy	Contains some items or terms which could be seen as western value judgements (eg. cursing, complaining, hoarding, use of the word inappropriate).
Anxiety	Rating Anxiety in Dementia (RAID)	Information not found	Clinical rating scale	Contains some items or terms which could be seen as western value judgements (eg. worry over finances, health).
Apathy	Apathy Evaluation Scale (AES)	Papers found using the AES in European languages of Dutch, Italian and German (van der Wurff 2003, Ravizza et al 1995, Isella et al 1998, Lueken et al 2006).	Clinical rating scale with informant / proxy and self-report versions	Contains some items or terms which could be seen as western value judgements (eg. Approaches life with intensity; Getting things done during the day is important to her / him; Seeing a job through to the end is important to her / him).

Domain	Instrument Name + Abbrev.	Cultural Applicability and Cultural Adaptations	Description	Issues to be considered
Depression	Cornell Scale for Depression in Dementia (CSDD)	Used with Japanese (Schreiner et al 2003), Korean (Shah et al 2005), Spanish (Ownby et al 2001, Harwood et al 2000), French (Camus et al 1995), Chinese (Lam et al 2004) and Turkish (Amuk et al 2003) speaking patients or community populations.	Clinical rating scale	-
Depression	Geriatric Depression Scale (GDS)	Numerous language versions including Chinese, Italian, Turkish, Vietnamese and Spanish. For a full list see the PROQOLID database. (Though users are advised to check for accuracy - Bowling 2005; GDS web-site)	Self-report instrument or interview version	<p>Contains some items or terms which could be seen as western value judgements (eg. Do you prefer to stay at home, rather than going out and doing new things ? Do you think it is wonderful to be alive now ? Do you worry a lot about the past ? Do you think that most people are better off than you are?).</p> <p>Issue of whether it is culturally appropriate to disclose such information to others</p>

This analysis shows the following:

- The DOMS reviews for the cultural applicability section focus mainly on the language translations of instruments.
- There were few problems identified with DOMS recommended instruments in terms of acculturation. Those with minimal problems require a minor wording change e.g. NPI long version – Irritability / Lability section. Those with some problems, like the GPCOG, require further investigation to test this face validity assessment. (NB: Traditionally, face validity is regarded as the lowest form of validity evidence and needs to be supplemented with empirical data).
- To some extent issues of cultural bias in this area, are a function of the type of test being examined. For example, acculturation is most likely to apply in direct assessment of cognitive tasks. While self-report, quality of life instruments will be affected by the cultural issues surrounding of the disclosure of personal information and feelings. Clinical rating scales, because they try to examine observed or reported signs or symptoms by trained clinical raters, are less susceptible to acculturation effects (i.e. they are not about patients trying to interpret biased questions, rather they are about trained coders trying describe their patients and their behaviours).
- The newer instruments examining the associated symptoms of dementia, like agitation, aggression, anxiety, depression and apathy seem to have some items containing western value judgements. These issues apply to lower level or milder items of these instruments and may be because they are looking at less psychiatry / neurology / medically grounded constructs.
- Two additional points need further elaboration. They are the cultural biases involved in using clinical rating scales and the issue of western value judgements or dominant culture stereotypes are contained in some items. These are outlined below.

#### 12.4.5.1 Cultural Biases Involved in Using Clinical Rating Scales

The majority of instruments recommended by the DOMS project are clinical rating scales. Clinical rating scales are based on a semi-structured interviews or reports, or by spending time with client and carer. They are not usually subject to acculturation effects as these scales mainly try to rate observable events or signs from a trained clinician's perspective. However, clinical raters need to be aware of the possible effects of cultural differences in behaviour when rating items or behaviours using these scales. For example grief reactions in certain cultural groups after the death of a love one; different degrees of support with personal care or home activities (e.g., dressing, cooking) from a spouse; a patient's reactions to the likelihood of being a burden on the family. In essence, clinical raters should be aware of possible culture differences when make ratings (Purnell and Paulanka, 2005).

From this review the better clinical rating scales to use:

- Are shorter using broad descriptions (e.g., BEHAV-AD, CSDD) or provide clear and detailed behavioural descriptors (e.g., CERAD-BRSD, CAM).
- Are grounded in psychiatric or neurological symptom frameworks (e.g., BEHAV-AD, NPI).
- Clearly relate the questions to changes in pre-morbid behaviours or personality (i.e. the normal self) (e.g., NPI most sections, CDR, DSRS).

Finally, these comments about clinical rating scales have ignored a potential form of cultural bias, based on the power dynamic that goes on between a rater and a "ratee", a clinician and a patient. This assumes that clinicians will not abuse their power relationship when treating or assessing someone. For this ethical reason, this paper dislikes the use of the ill-defined terms like 'inappropriate' (e.g. CMAI), as they may beg the question of 'inappropriate' for whom.

#### 12.4.5.2 Western Value Judgements Contained in Some Items

Some of the associated symptom instruments and HRQOL instruments could be interpreted as containing western value judgements. These implied values include:

- Positivism / optimism / happiness and action – one must move forward and not be stuck in the past, one must not complain or worry, one must be cheerful (e.g., GDS Yesavage, NPI, DEMQOL).
- The protestant work ethic – one must achieve things and not waste time (e.g., AES).
- Rules of polite society – one must not be critical or argumentative or swear or be loud (e.g., CMAI, RAGE).

While some functional assessment instruments and cognitive instruments may contain some dominant culture stereotypes or assumptions in item construction. For example:

- Gender role stereotypes may be present in instrumental activities of daily living (IADL) instruments, like the original version of Lawton and Brody's IADL instrument (Lawton and Brody, 1969) which excluded items on food preparation, housekeeping and laundering for men.
- Fillenbaum (1985) reports that IADL instruments may have different cultural elements or assumptions, for example: items from Britain include making a cup of tea or carrying a tray, while instruments from the Netherlands include items on bed making, while in New Zealand they have items on gardening.
- The RUDAS (Question 5 – Judgement) which is about traffic lights and a busy street, may reflect a degree of acculturation to the dominant Western and urban based culture.

By resting on these implied values or cultural stereotypes some items may be biased against those individuals not from the dominant culture. In other words, these items may be susceptible to cross cultural differences. *NB: This theoretical analysis requires supporting empirical evidence and acknowledges that implied values or cultural bias can not be completely avoided.*

#### 12.4.6 Recommendations for the Assessment of CALD Populations

Following the analysis of the cultural appropriateness of the recommended outcome measures within the scope of DOMS project it is recommended that:

- Use of the DOMS selected tools can be interpreted with less confidence if used by practitioners and interpreters who are not culturally competent. For an outline of the application of culturally competent assessment see the guidelines proposed by Alzheimer's Australia – National Cross Cultural Dementia Network (Grypma, Mahajani and Tam, 2007).
- A further project is necessary to ensure a more comprehensive database intended for dementia outcome measures solely for CALD communities - where translated versions of the DOMS selected measures are further reviewed and made available if possible.
- Further studies analysing the measurement equivalence of the core recommended measures be undertaken for major language groups within Australia.
- Research to further examine new instruments developed in Australia such as the RUDAS, the GPCOG and the KICA-Cog is supported to ensure their validity and reliability in different groups of CALD populations.

---

## **12.5 Dementia Assessment Issues for Aboriginal and Torres Strait Islander People**

### **12.5.1 Introduction**

There is a general paucity of information relating to the mental health of indigenous Australians. In his review of mental health Hunter (2003) highlights two areas which are particularly under researched; these are alcohol and pregnancy and cognitive decline in the elderly. Hunter argues that mental health issues in the elderly are often invisible, due to the fact that it is often a less urgent problem, than many of the other health issues that clinics have to deal with (2003: 148). He also acknowledges the difficulties in the identification of cognitive impairment, particularly in remote Aboriginal communities.

It is also important to note that Indigenous people in Australia live in a wide variety of settings, from urban through to very remote, traditional communities and have a wide variety of beliefs and experiences relating to mental health. For example recent research is just beginning to explore the intergenerational trauma of the effects of colonization, dispossession, racism and the stresses of poverty (Paradies, 2007:75).

There are undoubtedly special considerations for people who live in very remote communities, where local beliefs about the causation of illness may influence people's understanding of mental health problems and their willingness to report them. In her study of health beliefs in a remote Aboriginal community in Arnhem Land, Senior (2003), found a general reluctance to talk about mental health issues, and a high level of community anxiety about people who exhibited symptoms. Individuals with mental health problems were viewed by some, especially older people, as people to be avoided, or hidden away. She provides some examples of community members' reactions to people with mental health problems:

“As soon as they (other community members) see people with mental health problems, they run to the house and shut the doors” (carer of a person with a mental health problems”.

“They should find a home, a long way away for the mad ones, long way from everyone”.

The stigma associated with mental health problems appeared to stem from both the outward display of symptoms that the individual displayed, including aggressive and unpredictable behaviour and also because of concerns due to the causation of the problem. In this particular community, sorcery was considered to be a major cause of mental health problems, and a person who has been ensorcelled was considered to be fundamentally dangerous because they are situated outside the normal boundaries of acceptable behaviour.

### **12.5.2 Assessment Issues**

Indigenous clients exhibiting symptoms of dementia may be younger than non-Indigenous clients and there may also be a strong association between alcohol-related cognitive impairment and dementia (Hunter, 2003: 148). Hunter, citing the research of Zann (1994) comments that “Alcohol problems probably contribute to indigenous dementia in old age...alcohol related dementia (Korsakoff's syndrome) was the major diagnosis encountered in a north Queensland survey of dementia among those over 65” (2003: 148). This particular study found that about 20% of the age class had dementia compared with a national average of 5.4% (Zann, 1994: 5).

There are also some important considerations about the context in which assessments are made and the familiarity of the person making the assessments with the individual, their family and their circumstances. Many indigenous people become extremely anxious when they are taken away from their homes and families to a clinical environment. This anxiety has been linked to the very

high levels of non-compliance with treatment and patients absconding from hospital (Reid and Dhamarrandji, 1978). A highly anxious patient, whose concerns are further heightened by dealing with an unfamiliar clinician, may exhibit very different symptoms to those that they may display in a more familiar environment. Assessment is also influenced by language barriers and the degree of cultural knowledge of the person undertaking the assessment.

### 12.5.3 General Difficulties with Dementia Measures in Indigenous Settings

The difficulties of using standardized outcomes measures that have been developed for western populations with indigenous populations have been well described (Sansoni and Senior, 1998). Some of the most obvious problems with many instruments include:

- Concepts of functioning being related to career and employment.
- Concepts of independence as being a positive value (rather than valuing the level dependence an individual may have on their family).
- Measures that include concepts of time (last week, last year) and also volume (a lot; a little).
- Examples that may have little meaning, especially in a remote context (solving financial affairs, remembering the name of the high school from which they graduated).

### 12.5.4 Issues with the Tools for Assessment

A table is provided in Appendix 14 which provides comments on the suitability of the recommended measures for the assessment of the associated symptoms of Dementia with Indigenous people. Many of the issues raised in this table (inappropriate language, concepts, timeframes etc) also apply to the other instruments that are discussed below.

The cognitive tests that are part of many dementia rating scales pose a significant barrier to individuals who have had limited education, may have limited literacy and numeracy, and significant problems understanding English. Many people in remote Indigenous communities may have problems with questions asking about the day of the week and the month of the year (Mini Mental State Examination, ADAS, Dementia Severity Rating Scale) Other 'simple questions' such as the date of World War 1, World War 2 and the name of the monarch (Blessed Dementia Scale) may have very little meaning. The RUDAS which is a multicultural cognitive assessment scale may be more useful in Indigenous contexts, with the possible exception of the question about judgment: many older people in remote indigenous communities would not have much experience with crossing busy streets, pedestrian crossings and traffic lights. The Kimberley Indigenous Cognitive Assessment tool (KICA-Cog) was developed with Indigenous health and aged care organizations to provide a tool to assess cognition in older Indigenous Australians (LoGiudice, et al. 2005; refer Appendix 15) and thus avoids some of the problems raised above. This is a promising new tool, with reasonable psychometric properties reported by the authors (LoGiudice, et al. 2005). However, as yet there is limited published information concerning its validation and the replication of its psychometric properties in other samples. Despite these limitations it may be a more appropriate tool to use for the cognitive assessment of Indigenous people, particularly those located in rural and remote settings.

Recently Cairney (2007; personal communication) has also used some of the tests from CogState to assess the cognitive status of Indigenous people – mainly to assess cognitive impairment arising from petrol sniffing in remote communities. The CogState tests involve computerized card game tests which are designed to provide culture free assessments of attention, memory and learning. CogState tests are a proprietary product and some elements of this battery are also currently being assessed as screening measures for Mild Cognitive Impairment (MCI) in Australia. There are few publications as yet concerning the psychometric properties of this battery, or tests within it, although the website ([www.cogstate.com](http://www.cogstate.com)) refers to two studies where the CogState Health associate learning task has been shown to be more sensitive to change than other similar

measures for the detection of subtle memory decline over 1-2 years. More detailed assessment of this innovative approach will be required as more published literature becomes available concerning the psychometric properties of cognitive tests within this battery.

Many of the clinician rating scales contain material that is based on what is considered to be normal functioning in a non-Aboriginal context. For example the Global Deterioration Scale contains an item “Co workers become aware of patient’s relatively poor performance” and “patient may read a passage or book and retain relatively little material”. The Clinical Dementia Rating asks about whether the patient “handles their financial affairs well”. The Sandoz Clinical Assessment asks about ‘motivations’ (lack of spontaneous interest in completing task) which may well be a question which is more pertinent in cultures which place a strong emphasis on work ethic. This underlying concept also underpins the questions in the Apathy Evaluation Scale, which has questions of ‘seeing a job through’, initiative and motivation. These are very difficult concepts to assess in communities which are based on collective rather than individual decision making.

The clinician ratings also ask about behaviours such as anxiety and irritability. As previously stated, patients may exhibit high anxiety and irritability because of the context of their treatment and their fear of medical procedures and thus their mental state may be difficult to interpret. Finally many of the tools that address associated symptoms such as the NPI, Behave AD, CERAD-BRS, Confusion Assessment Method, Delirium Rating Scale, Rating Scale for Aggressive Behavior in the Elderly, ask questions about hallucinations. In traditional cultures, where there is a strong belief in spirits and the supernatural, “seeing things which are not there” may be difficult to explain and may not be regarded as aberrant behaviour. Readers are referred to Appendix 14 for further comments on these scales.

Of the self report tools, the simplest is possibly the EQ-5D, this should not pose too many difficulties even in a remote context, except the VAS question regarding a persons’ health state today (from best imaginable to worst imaginable) if this is used. Given the responses that have been found in other surveys, for example the National Aboriginal and Torres Strait Islander Survey (1994) where indigenous people (particularly from remote communities) were asked to rate their own health, the ability of this question to reflect people’s actual health state is uncertain.

The questions in the DEMQOL appear to be fairly unproblematic, especially in a non-remote context. However the differences between “quite a bit” and “a little” may be very difficult to explain. This is an example of a tool that maybe enhanced by pictorial demonstrations of quantity (for example circles in descending sizes to illustrate volume) as was done in the Kimberly Aboriginal Health Survey (Spark, Donovan and Binns, 1992).

Finally, the concept of loneliness in the De Jong Gierveld loneliness scales may also be very difficult to assess, as people from remote communities may not have a concept of loneliness and ideas about being surrounded by a circle of friends may not be relevant in a place where everyone is related and has a cultural obligation to care for relatives.

### **12.5.5 Discussion**

Many of the above scales may have applications among urban Indigenous people. It will be necessary however to develop some focus groups in these settings to discuss how appropriate they are to members of these communities. There is very limited application for these tools for people from remote Aboriginal communities. A notable exception to this is the Kimberley Indigenous Cognitive Assessment tool (LoGiudice, et al. 2005) which had been designed for use with Indigenous people in remote locations and it is suggested this is used to assess the cognitive status of remote indigenous peoples rather than the MMSE-3MS (refer Section 6.4.6).

Clinician ratings may have more application than the self report or the proxy administered forms, as some of the ratings can be made through observation, rather than attempting to elicit answers

from the patient. Cognitive assessment will be extremely difficult in many remote settings, and especially if the patient speaks and understands limited English. Clinical assessments may be improved if other confounding factors are removed, such as unfamiliarity of the clinician and environment. A clinician, who is familiar to the individual and has a good knowledge of their life, may be in a position to make a more informed judgment.

While it may be possible to use some of the simpler tools in a remote setting, especially with modifications to pictorially demonstrate concepts such as volume, questions will still remain about what the answers that individuals supply actually mean. There needs to be further detailed research on the meaning of dementia in Indigenous communities, and how to ask questions which capture the experience of living with dementia in an Indigenous community.

## **12.6 Implementation Issues**

### **12.6.1 Introduction**

Although some issues pertaining to implementation have been discussed throughout this report there are a number of key areas to address. These are:

- The issue of mandating the recommended measures;
- The application of the instruments in different settings and for different stages of dementia;
- Training issues;
- A dissemination strategy; and
- Identified research gaps;

These are discussed below.

### **12.6.2 Should Measures be Mandated or Recommended?**

The National Expert Panel thought clarification should be sought from the Department of Health and Ageing concerning the proposed implementation strategy for the recommended measures. The DOMS-NEP inquired about the future employment of the recommended measures in practice, and whether the measures and routine data collection will be mandated across dementia services as occurs in Mental Health. If so, the Department would need to consider the required education and training for clinicians/staff to employ the measures.

It was suggested that Jan Sansoni and Marc Budge meet with relevant officers of the Department of Health and Ageing to seek clarification and discuss this issue. At a meeting on 18 August 2006 they were advised mandating was not a consideration at this time as the Dementia Outcomes Measurement Suite was a first-stage project to assess key gaps and tools. It was noted that the Dementia CRCs and Study Centres may promote the use of particular tools agreed as a result of the DOMS-NEP project; however, this would be as best practice, rather than to mandate.

It would be difficult to mandate the use of the recommended measures without full consideration of the training requirements and the burden on staff time for all service settings to implement these measures. If routine data collection and analysis is desired, with a view to benchmarking the outcomes of similar services, then careful thought must be given to the design of such systems and the phased implementation of such an approach. To adopt such an approach will also require a considerable financial investment by the Department of Health and Ageing as has occurred with mental health services.

A thorough examination of the advantages and disadvantages of mandating the use of the recommended measures and establishing routine data collection systems is beyond the scope of this project. If the Department of Health and Ageing desired to mandate these measures in the future it is recommended that a thorough scoping exercise be undertaken to assess the implications of such an approach. In this respect the experience gained in mental health settings through such bodies as the Australian Mental Health Outcomes and Casemix Collection Network would be invaluable. Until such a study was undertaken it is suggested that these measures be recommended rather than mandated.

As indicated above there are ways to facilitate the use of the measures such as the Dementia CRCs and Study Centres promoting the use of these tools. Similarly, the use of appropriate measures can be a consideration in the selection of research projects under the National Dementia Initiative. As with the Continence Outcomes Measurement Suite Project (Thomas, et al. 2006) consideration could be given to the funding of a Dissemination Strategy Project to provide manuals, booklets, brochures, web materials and presentations to facilitate the adoption of these measures in the field.

### **12.6.3 The Application of the Instruments in Different Settings and for Different Stages of Dementia**

In applying these instruments in practice the following factors must be considered for each of the instruments:

- The purpose of the assessment – whether this be for initial assessment and screening, for more comprehensive diagnosis and the assessment of prognosis, or for outcomes monitoring and outcomes evaluation.
- The service setting for the assessment – whether this be for primary and community care settings, specialist dementia settings or for research settings.
- The stage/severity of dementia. Some instruments, such as self –report measures for example, may be inappropriate for those experiencing severe dementia.

Given consideration to these factors, the following recommendations are made concerning each category of measures. Figure 6 also provides a guide for the application of the measures in different settings.

#### **Dementia Staging and Descriptive Measures**

Burns et al (2004) indicates these measures are widely used as staging measures in descriptive and intervention studies. It is noted that specialist clinicians are less likely to use these global staging instruments than other clinical or research personal. Such instruments may not be particularly useful for fine differentiation at an early stage of dementia. However, global functional scales like the GDS and the CDR have their place in broadly describing people with dementia; particularly for research purposes and in residential care and community care settings.

The Global Deterioration Scale (GDS) and the Clinical Dementia Rating (CDR) scale are the preferred instruments with the best psychometric properties. These instruments are applicable for initial assessment, prognosis and outcomes evaluation and for all stages of dementia. They can be used in all care settings. However, the GDS may be preferred in routine care settings given its ease of use. The CDR may be more applicable in research and specialist settings.

The DSRS may also be used in community settings.

#### **Health Status and Health Related Quality of Life Measures**

The preferred instruments were the Quality of Life- Alzheimer's Disease (QOL-AD) and the DEMQOL for mild to moderate dementia and the QUALID for late stage dementia. All these

measures have both proxy and patient forms. It is likely that most of these instruments will be used in research applications to evaluate and monitor the outcomes of intervention programs rather than in routine care. However, residential care services could be encouraged to undertake health related quality of life assessments of their residents.

No generic health status measures were recommended. In research settings consideration could be given to using instruments such as the SF-36V2 or SF-36V1 for people with mild dementia. The latter instruments may also be quite useful in research applications where there is a desire to assess health related quality of life of carers.

### **Cognitive Status Measures**

The instruments with the highest scores were the MMSE-3MS and the ADAS-Cog. The 3MS was considered the best of the MMSE family for routine settings and also has less proprietary issues. The ADAS-Cog may be preferred if a more in depth assessment is required (e.g. in clinical research or specialist settings). It is noted that the scope of this project excluded a consideration of more detailed and specialist neuropsychological measures that may be used for in depth cognitive assessment.

In terms of applying the MMSE-3MS in Australia, the authors of this work caution against the setting of mandated, a priori targets or change scores on instruments used for people with dementia, in order that they can obtain access to treatment (or further treatment). This approach places a lot of pressure on the accuracy and contextual validity of instrument cut-scores, as well as allowing demand characteristics and possible biases to influence clinical ratings or self-reports. At this stage of their psychometric development assessment instruments should be allowed to describe the person's stage of illness and care experience. The present state of knowledge in the dementia field requires a great deal of further clinical and normative data for the Australian health care context.

The GPCOG was considered most appropriate for primary care as it has been designed for that setting and it may be useful as an initial screening instrument. The MDS-Cog can also be considered for use in residential care settings where time constraints may preclude the use of more detailed assessment instruments. Despite the total score concerning its psychometric properties being slightly lower than the other instruments, the individual attributes are more than adequate.

The RUDAS is a new instrument that was designed to enable the easy translation of the items into other languages and to be culture fair. There are relatively few papers published as yet concerning its psychometric properties (especially construct validity) but in the interim it is recommended for use with those from Culturally and Linguistically Diverse backgrounds. The RUDAS, however, contains an item on judgement that may be inappropriate for remote Indigenous people (refer below).

An interim recommendation is to use the Kimberley Indigenous Cognitive Assessment (KICA-Cog) tool for the cognitive assessment of rural and remote Indigenous people. The KICA is a new instrument and although there is little published evidence concerning this tool available as yet, and further research is required, this instrument has been designed for use with Indigenous people.

### **Multi-attribute Utility Measures**

These instruments will largely be used in research settings for the purpose of evaluating the cost effectiveness or cost-utility of particular treatments. They are unlikely to be used in routine care settings. Although no instrument is recommended, rather a program of research (refer Section 7), the EQ-5D and the AQoL would currently be the preferred instruments to consider for such research.

However, in their current forms the EQ-5D and the AQoL could only be considered for use with people with mild dementia who retain the capacity to self rate or for use with those with moderate dementia if interview administration was utilised. Clinician proxy forms would need to be utilised with those with severe dementia. As the AQoL is a longer instrument it is thought it may present more difficulties for people with moderate dementia unless it is shortened as has been recommended.

### **Measures of Social Isolation**

There were issues concerning the theoretical basis of the construct being measured for many of the measures. Some measures focus on social isolation – others on actual and perceived social support – and others were a mix of both dimensions.

Those instruments that attempted to measure actual social support had numerous cultural issues associated with them and many were too complex to consider with people with dementia. Most of the instrument with the better psychometric properties focused on social isolation. Only the De Jong–Gierveld Loneliness Scale was selected for more comprehensive review and this will need minor adaptation for use in Australia. It should be noted that as this is a short self-report measure it would not be suitable for people with severe dementia. Following adaptation it may be used to assess social isolation in primary and community care settings to assist with care planning and the identification of support services that may be required by people with dementia living in the community. Thus, while these instruments appear to be more relevant in community and research contexts, and social isolation is rarely assessed in residential care, it is an issue of significance for these settings. If social isolation appears to be a presenting issue for a resident, and the resident retains the cognitive capacity to respond to such self report questions, the assessment of social isolation should be considered. Such instruments may also provide some useful contextual information for treating clinicians.

### **Associated Symptoms: Global Measures of Behavioural and Psychological Disorder**

The NPI and the Behave-AD were the instruments with the highest ratings. NPI produces both a global assessment and has good coverage of most symptoms. These instruments are applicable across a range of settings although are they less likely to be utilized in primary and community care settings. However, where patients in residential care or community care are exhibiting some symptoms of BPSD it is recommended that an assessment using the NPI be undertaken as this has a good coverage of the range of associated symptoms.

### **Associated Symptoms: Differential Diagnosis**

As dementia may need to be differentiated from delirium for some individuals, these delirium instruments may be considered when undertaking initial diagnosis and assessment.

The two instruments selected for comprehensive review were the Confusion Assessment Method (CAM) and the Delirium Rating Scale – Revised 98 (DRS-R98). CAM is a useful screening measure and DSR-R98 provides a more in depth coverage should this be required.

### **Associated Symptoms: Specific Symptoms**

Measures that assessed the specific symptoms of aggression, agitation, anxiety, apathy and depression were also examined to cover situations, in all settings, where a more detailed coverage of this aspect may be required. For example it may be necessary to determine whether the primary diagnosis for a person is dementia or depression and to ascertain whether depression is a major co-morbidity for the patient. The identification of depression may also have ramifications for patient medication. It may also be important to identify whether a patient is suffering from depression, as against apathy, as this may well influence the care plan.

In many instances the use of a global BPSD measure such as the Neuropsychiatric Inventory may suffice in identifying major behavioural and psychological symptoms. However, if a patient presents with one of the symptoms below a third stage of assessment may be undertaken using

one of the recommended instruments below. These instruments, with the exception of the Geriatric Depression Scale are all clinical rating scales and thus can be used with people with dementia at all levels of severity and in all settings.

The recommended instruments are as follows:

Aggression: Rating Scale for Aggression in the Elderly (RAGE)

Agitation: Cohen Mansfield Agitation Inventory (CMAI) and Pittsburgh Agitation Scale (PAS)

Anxiety: Rating Anxiety in Dementia (RAID)

Apathy: Apathy Evaluation Scale (AES)

Depression: Cornell Scale for Depression in Dementia (CSDD) and Geriatric Depression Scale (GDS Yesavage) – the latter only for community settings or those with mild dementia.

It is likely that the above measures will also be relevant in research applications.

### **Other Omnibus Measures: BPSD**

The HoNOS 65+ clinical rating scale was not reviewed under global measures of BPSD as there were issues with its total score (refer Section 9.5). It is thus considered more as an omnibus set of measures. It does, however, have a good coverage of ICF domains and includes environmental aspects such as living conditions which are rarely assessed by other measures.

However, it should be noted that NPI can be considered both as a global measure and an omnibus measure and appears to have better psychometric properties and to be more appropriate to dementia vs. mental health settings.

There is fairly limited published evidence concerning the psychometric properties of HoNOS 65+ scales and there appear to be problems with some scales (e.g. depression; relationships). There is a need for more published evidence on its psychometric properties and its applicability to dementia as well as mental health settings. For these reasons this instrument is not recommended at this stage for other than mental health settings. Its use is currently mandated in mental health services for older persons within Australia.

### **Measures of Function**

The assessment of function is important when decisions concerning residential placement and the use of support services need to be made. Many of these instruments may be used in primary and community care settings but it may also be useful to monitor function/ functional decline in residential care settings. In these settings a functional assessment might identify the need for care programs to assist with the maintenance of function of the resident and/or identify necessary changes to the care plan for the resident. In the latter cases a more systematic form of assessment, using tools outlined below, may be useful to consider.

The Functional Independence Measure (FIM), the Barthel Index and the Lawton and Brody IADL and the Older Americans Resources and Services (OARS-IADL) instruments were chosen as generic measures of ADL and IADL respectively. These instruments have been reviewed elsewhere recently (Eagar, et al. 2001, 2006; Thomas, et al. 2006) have good psychometric properties and have been used in geriatric settings.

With regard to the activities of daily living the FIM is probably more appropriate for acute care and high level residential care settings and with people with more severe dementia but it is noted that accredited training is required for its use. However, it is already widely used in acute care rehabilitation settings within Australia. The Barthel Index is an easier to use measure and may be more appropriate for use in primary and community care settings with people with mild to moderate forms of dementia. Although the Katz ADL instrument has been quite widely used in dementia settings the review of this instrument by Thomas, et al. (2006) indicated it had weak psychometric properties and thus it is not recommended for use.

The Lawton and Brody IADL and the Older Americans Resources and Services (OARS-IADL) are recommended as generic instruments for the assessment of instrumental activities of daily living (IADL). The OARS-ADL is preferred as it is an advance on the Lawton and Brody IADL scale with improved psychometric properties and less reliance on gender role stereotypes; and it has been adapted for use in primary and community care settings in Australia (see Green, et al. 2006).

The recommended dementia specific instruments for the assessment of function (ADL and IADL) for people with dementia include both proxy measures and clinical rating scales. While it is acknowledged that proxy reports have their limitations (refer Section 12), they will generally be used where assessment by interview or self rating is no longer possible due to the degree of cognitive impairment of the person with dementia. Proxy measures are also extremely useful in primary and community care settings in order to monitor the maintenance of functional status or its decline, in connection with drug therapy or in terms of care management as the disease progresses. The direct observation rating scale may be more appropriate for acute care and residential care settings. By recommending both proxy and direct observation rating scales different practice settings and clinical situations (e.g. a person with dementia may not have a carer) can be addressed.

The Alzheimer's Disease Co-operative Study – ADL (ADCS-ADL) and Disability Assessment for Dementia Scale (DAD) are the two proxy report instruments that are recommended.

For the direct observation of functioning the Cleveland Scale for Activities of Daily Living (CSADL) is recommended.

### **Measures of Patient and Carer Satisfaction**

The Short Assessment of Patient Satisfaction (SAPS) scale, a generic measure of patient satisfaction, was developed from a study by Hawthorne, et al. (2006) which selected the best items from a range of commonly used generic patient satisfaction measures to produce a shorter patient satisfaction scale. The SAPS contains only seven items (one for each theoretical dimension of patient satisfaction) and had better psychometric properties than the instruments from which it was derived. It also had excellent internal consistency with a Cronbach's alpha of 0.86. Although SAPS needs to be further tested in other samples and populations (e.g. including people with dementia and carers) it is the interim recommendation concerning a generic measure for the assessment of patient satisfaction. The SAPS may be most appropriate for people with mild to moderate dementia that are receiving a particular treatment intervention and in research settings. For people with severe dementia who no longer retain the capacity to self rate a proxy version of this scale needs to be developed.

A single item for the assessment of patient satisfaction would be attractive for use in routine care settings. It is recommended that community and residential services routinely assess satisfaction with the care they provide. Six generic items measuring global satisfaction were identified from the Hawthorne, et al. (2006) study and were analysed concerning their appropriateness as a single measure for immediate assessment of patient satisfaction. Two of these items had better psychometric properties. One of these items (how satisfied are you with the outcome of your treatment?) could be used where a particular treatment/ change to care plan is being assessed. A second item (how satisfied are you with the amount of help received?) needs modification to its response categories and further analysis of its psychometric properties. An item such as this may be more useful for the assessment of everyday care practices such as in residential settings.

## **Impediments to assessment**

It is noted that some people with dementia will have other disabilities (e.g. vision and hearing disabilities) that may form an impediment to assessment. Where the disability is mild modifications to the administration method may overcome some of these limitations. For example, a person with mild dementia and mild vision impairment may be able to complete a self report form if it is converted to a large print size. These design and implementation issues are further outlined in section 12.3.2

Where direct assessment is used, such as in a clinical rating scale, assessors need to be mindful of other disabilities or illnesses which the person with dementia may have and the effect these may have on the domains of assessment (e.g. depression, arthritis or COPD).

Some research studies with the elderly have used telephone interview techniques rather than person to person interviews. This research appears to have been largely based on studies of older people without dementia or those with mild cognitive impairment or mild dementia. Given the limited research on the applicability of these methods for people with dementia at this time it is suggested that person-to-person assessment methods are to be preferred.

## **12.6.4 Further Issues Concerning the Application of the Instruments**

This section of the report further examines the application of the DOMS recommended instruments according to a staged approach to assessment and the assessment of dementia in different practice settings.

### **12.6.4.1 Stages of Assessment**

For the first stage of assessment it is suggested that in many settings the Global Deterioration Scale (GDS) could be used to assess the stage and severity of dementia. This should be used in conjunction with a measure to assess cognitive impairment. This would normally be the MMSE-3MS although in General Practice settings the GPCOG may be preferred. In residential care if time/staffing constraints are a major issue the Minimum Data Set – Cog might be used. For the assessment of CALD populations RUDAS should be considered and for rural and remote Indigenous people the Kimberley Indigenous Cognitive Assessment (KICA) may be the most appropriate tool.

At this stage of assessment it may also be useful to clarify issues concerning diagnosis and to ascertain whether residential services are required. If there is the possibility that delirium may be involved it is suggested that the Confusion Assessment Method (CAM) be used to screen for delirium. If a more detailed assessment of delirium and repeated assessments are required, as may occur in the acute setting, then the Dementia Rating Scale (R98) could be used. If it is unclear whether the primary diagnosis may be depression or dementia then the assessment of depression may be required. Assessment of function may be very relevant if residential placement is an issue.

In the second stage of assessment the focus is on identifying particular issues for the patient which may inform the care plan. For the second stage of assessment it is noted the majority of people diagnosed with dementia also experience symptoms of behavioural and psychological disorder. If it is suspected that the patient has symptoms of BPSD a more detailed examination of these symptoms using the NPI would seem appropriate. It may be relevant to assess function if this has not already been assessed. At this stage of assessment issues such as social isolation and health related quality of life may be assessed.

In specialist and research settings there may be the desire for further information concerning particular cognitive deficits of the patient in which case an assessment may be undertaken using the more detailed ADAS-Cog tool or other more specialist neuropsychological batteries (it is noted

that the latter have not been reviewed in this report but could be the subject of a follow up consultancy).

In the third stage of assessment there is more detailed examination of particular presenting problems for the patient. The assessments recommended for the associated symptoms of apathy, anxiety, agitation, aggression and depression would be relevant measures for this stage of assessment. For example if a patient is presenting with apparent depression it may be useful to use measures to assess for both depression and apathy to determine whether pharmaceutical treatment for depression may be required. Similarly a more in depth assessment of symptoms such as agitation and anxiety might be useful for those patients who appear to be presenting with these symptoms and again the information derived from these scales can be used to inform the care plan and potential treatment interventions. Similarly if issues of physical function have been identified at the second stage of assessment there are a range of specialist function tests which may be utilized by the appropriate professionals (e.g. physiotherapists, speech therapists etc).

Two other general symptoms (not specific to dementia) are worthy of mention but are beyond the scope of this report. Many patients with dementia may also experience incontinence and it is suggested that the readers refer to reports by Sansoni, et al. (2006) and O'Connell, et al. (2007) for tools appropriate for the assessment of incontinence across the range of care settings.

The systematic assessment of pain may also be necessary on occasion. For example if a patient is presenting with apparent agitation it may be necessary to differentiate between agitation deriving from a pain state as against generalized agitation. It is generally accepted that pain is under diagnosed and under treated for people with dementia (Fries, et al. 2001; Williams, et al. 2005) and that psychological and behavioural symptoms of dementia (BPSD) may be exacerbated due to the treatment of pain (Weiner, et al. 2002).

The Australian Pain Society (2004) tested a variety of scales suitable for use for people with dementia and their 2004 publication outlines effective management strategies. McDowell (2006) also provides a discussion of a number of generic pain scales that may be relevant in this context.

One can see from this discussion how the different instruments can be combined according to the purpose of assessment – for example, moving from general screening to a more detailed assessment of specific issues like depression, apathy or anxiety. This is known as tiered assessment approach, where scores on a particular instrument or item trigger the use of further instruments according to the particular purpose of the assessment. With further research the staged assessment approach outlined above has the potential to be developed into a tiered assessment approach. For an example of how tiered assessment works in the field see the report by Owen et al. (2004) describing the Queensland Ongoing Needs Identification (ONI) tool for community care clients.

#### **12.6.4.2 Assessment in Dementia Settings**

As can be seen above the assessments that may be undertaken will also vary depending on the practice setting. A brief consideration of settings is outlined below and this should be used in conjunction with the information provided in Figure 6 below.

##### **Acute Care**

In acute care settings people with dementia may be receiving treatment for other co-morbid conditions or acute injuries, or are in hospital care for the stabilisation of medical and behavioural symptoms related to dementia. The key issues in relation to measurement are: (1) the determination of the presence of delirium and/or pre-existing dementia or depression – with the consideration of delirium first, prior to a diagnosis of dementia or depression; (2) the identification, reporting and management of any associated behavioural symptoms of dementia; and (3) the documentation of health status for discharge planning to community care or residential care. In this

setting, the use of the MMSE-3MS, Global Deterioration Scale, the Neuropsychiatric Inventory (NPI) and the two delirium instruments (CAM and DRS-R-98) would seem to be the most important application of the DOMS recommended instruments.

### **Primary and Community Care**

In primary care settings, General Practitioners are interested in identifying whether dementia is present in their elderly clients within the community. A brief assessment of cognition (e.g. GPCOG or other culturally appropriate tools recommended for specific groups e.g. RUDAS) may be useful in screening for dementia and may form the basis for a referral to specialist practitioners or to a memory clinic for a more comprehensive assessment. Other short screening instruments like the Mini-Cog and the Memory Impairment Screen could also be considered, however their memory items would require more training of General Practitioners to administer, score and interpret these instruments correctly.

General Practitioners also manage elderly patients within residential care settings. For patients demonstrating behavioural and psychological symptoms associated with dementia, or where delirium may be suspected, an evaluation of these symptoms and the assessment of underlying or co-morbid medical conditions is advised. Prior to prescribing medication for symptoms such as depression, agitation and other behavioural and psychological symptoms it is suggested that a standardised assessment of these symptoms be undertaken either by appropriately trained residential care staff and GPs or through referral to specialist practitioners / multi-disciplinary teams.

In community care settings, community nurses and Aged Care Assessment Teams (ACATs) play an important role in connecting older adults with relevant aged care services and supports, as well as making recommendations about eligibility and the appropriate level of care required. In the majority of cases access to home and community services is derived through ACAT or if not referral is usually made back to ACAT e.g. services through Aged Care Packages will only be obtained following ACAT assessment. ACATs are multidisciplinary teams often comprised of social workers, occupational and physiotherapists, as well as nurses and doctors. They are usually based in hospitals or community health centres, as are community nurses. With this team based approach and given the multiple service provider environment that is community care, it is vital that clinicians and services use a common language when they are talking about their clients and developing pathways of care. The recommended instruments described here can be used and adapted at the local level by ACATs, community nurses and service providers for this purpose. For instance, the functional instruments could be streamlined and adapted to look at functional independence and safety around the home of the person with dementia. If introduced properly across the local system it should also reduce the number of times a person and/or carer has to undergo assessment and “repeat their story”.

Increasingly, in the community area, specialist mental health assessment services for older people (SMHOPS, BASIS and DBMAS) are taking a greater role in the assessment and management of people with dementia with associated behavioural symptoms. Greater alignment is required with the DOMS recommended tools (e.g. NPI, CMAI, Global Deterioration Scale) and the aged care mental health measurement tools developed for mental health settings (as in the MH-OAT program in NSW).

### **Residential Care**

In residential care the focus is on the monitoring of the cognitive status and functional abilities of patients with dementia in relation to their care plan and medical decision making. In this setting a premium is placed on clinical rating instruments, which require the assessor's in-depth knowledge of the resident's status in an on-going manner. Unlike other care settings the majority of care in residential care facilities is provided by unlicensed care workers, hence easily adoptable tools such as the Global Deterioration Scale for assessing the staging of dementia can provide an opportunity for staff to recognise key changes in the resident's status associated with dementia

that may require a different approach of care and management. The provision of training in order to obtain standardized observations of symptoms and inform health professional decision making is required.

Also in this setting, the assessment of depression, apathy, health related quality of life and social isolation, in the context of person-centred and / or relationship centred are emerging issues. Such an approach is necessary to provide high quality care, however it is an often ignored area of clinical practice.

In policy terms residential care has been the target of improved and consistent assessment processes for the person with dementia, most particularly with regard to resource allocation that best meets the person's patient centred health needs. There has also been a strong emphasis on standardized assessment in the acute care and community sectors for the older person. Assessments recommended include reliable and valid screening measures of cognitive, physical and psychosocial functioning. These screening measures are included in the Australian Care Funding Instrument (ACFI) for the residential care sector and to a lesser extent in the Australian Community Care Needs Assessment (ACCNA) in the community care sector. However, key issues remain in terms of the acceptability and implementation of these funding instruments across the sectors. In this context the DOMS recommended tools may require some alignment with these more general instruments especially the ACFI. There is some overlap with the DOMS instruments, as the ACFI uses the Cornell Scale for Depression in Dementia, however it does also rely on the Psychogeriatric Assessment Scale (PAS) to assess cognitive impairment. This measure was considered but not recommended by this project, though some items from the PAS are used in the GPCOG.

### **Specialist and More Detailed Assessment**

More detailed assessment or specialist setting assessment is often required in memory clinics and by geriatric multi-disciplinary teams. These clinics or teams may be associated with teaching hospitals or are privately funded. The staff may include: geriatricians, psycho-geriatricians, neuropsychologists, clinical psychologists, nurses, occupational and physiotherapists and social workers. Memory clinics recognise the need for a systematic, multidisciplinary and specialized approach to the diagnosis and management of cognitive disorders. Usually a comprehensive assessment of both medical and psychological factors is undertaken (including neuropsychological testing, brain scans and blood tests).

Many of the instruments recommended in this report may be used by memory clinics and multi-disciplinary teams to assess both the core and the associated symptoms of dementia. It would be desirable if memory clinics and specialist geriatric assessment teams examine the findings of this report and adopt the use of the recommended instruments to assist with the standardisation and comparability of data across this field.

A range of more detailed neuropsychological tests may also be used by these clinics to provide an in depth analysis of the cognitive function of patients. In Australia, it is noted a specialist must confirm the diagnosis of probable Alzheimer's disease in order for a person to be eligible for subsidised Alzheimer's medications. Given the DOMS recommendations on cognitive measures it would be prudent to use the 3MS, rather than the MMSE, as an initial indicator for the Pharmaceutical Benefits Scheme (PBS) eligibility criteria.

### **Research Settings**

In research settings the recommended instruments can be used as a guide to "current industry practice" or standard in terms of dementia instruments. In this way, they can be useful to researchers to help consider: (1) additional secondary outcome measures to the primary outcome studied in clinical trials (for instance, including a patient reported outcome, health related quality of life instrument in a drug treatment trial designed to slow functional decline on a functional

instrument); and (2) as a set of standard comparison or benchmark measures when new instruments are being developed and validated.

From this initial work examining different practice settings, one can see how the DOMS dementia tool-kit of instruments moves toward the development of screening and assessment modules or assessment tiers (from initial screening to comprehensive assessment) – across settings, patient groups and dementia severity.

Finally Figure 6 below outlines the appropriate application of each of the recommended measures for different service settings and for the different stages of dementia.

**Figure 6 A Matrix Model for the Recommended Instruments**

Recommended Instruments	Settings							Severity of Dementia		
	Acute	P&CC	Resi. Care	More Detailed Assessment	Research	Indig.	CALD	Mild	Mod	Severe
<b>Dementia Assessment</b>										
Global Deterioration Scale	✓	✓	✓		✓	?	?	✓	✓	✓
Clinical Dementia Rating	✓	✓	✓		✓	?	?	✓	✓	✓
Dementia Severity Rating Scale		✓	?		✓	?	?	✓	✓	✓
<b>Cognitive Assessment *</b>										
MMSE-3MS	✓	✓	✓		✓	?	?	✓	✓	✓
ADAS-Cog		✓		✓	✓	?	?	✓	✓	
MDS-Cog	?	✓	✓		✓	?	?		✓	✓
GP-Cog	?	✓	✓		✓	?	?	✓	✓	
RUDAS	?	✓	✓		✓	?	✓	✓	✓	
KICA-Cog	?	✓	✓		✓	✓	?	✓	✓	
<b>Global BPSD</b>										
NPI	✓	✓	✓	✓	✓	?	?	✓	✓	✓
BEHAVE-AD	✓	✓	✓	✓	✓	?	?	✓	✓	✓
CERAD	✓	✓	✓	✓	✓	?	?	✓	✓	✓
<b>Diff. Diagnosis Delirium</b>										
CAM	✓		✓		✓	?	?	✓	✓	✓
DRS-R-98	✓			✓	✓	?	?	✓	✓	✓
<b>Individual Symptoms</b>										
RAGE - aggression		✓	✓	✓	✓	?	?	✓	✓	✓
CMAI - agitation	✓		✓	✓	✓	?	?	✓	✓	✓
PAI - agitation		✓	✓	✓	✓	?	?	✓	✓	✓
RAID - anxiety		✓	✓	✓	✓	?	?	✓	✓	✓
AES - apathy		✓	✓	✓	✓	?	?	✓	✓	✓
Cornell - Depression	?	✓	✓	✓	✓	?	?	✓	✓	✓
Geriatric Depression Scale	✓	✓	✓		✓	?	?	✓	✓	
<b>Function - Generic</b>										
FIM (ADL)	✓	✓	✓		✓	?	?	✓	✓	✓
Barthel Index (ADL)	✓	✓	✓		✓	?	?	✓	✓	✓
Lawton & Brody (IADL)		✓	✓		✓	?	?	✓	✓	✓
OARS - IADL		✓	✓		✓	?	?	✓	✓	✓
<b>Function - Dementia Spec.</b>										
ADCS-ADL		✓	✓		✓	?	?	✓	✓	✓
DAD		✓	✓		✓	?	?	✓	✓	✓
CSADL		✓	✓		✓	?	?	✓	✓	✓
<b>HRQOL- Dementia Spec.</b>										
DEMQL		✓	✓		✓	?	?	✓	✓	
QOL-AD		✓	✓		✓	?	?	✓	✓	
QUALID			✓		✓	?	?			✓
<b>Multi-attribute Utility</b>										
AQoL					✓	?	?	✓	?	
EQ-5D					✓	?	?	✓	✓	
<b>Social Isolation</b>										
De Jong Gieveland Loneliness		✓	✓		✓	?	?	✓	✓	
<b>Patient Satisfaction</b>										
SAPS		✓	✓		✓	?	?	✓	✓	
<b>Carer Satisfaction</b>										
SWC-EOLD		✓	✓		✓	?	?	✓	✓	✓

- ✓ = Yes, some evidence for the instruments potential application in Australia  
 ? = Unsure, minimal evidence requiring further research in the Australian context

\* These are cognitive tests for people with dementia, if an informant / proxy measure is required use the IQCODE

### 12.6.5 Training Issues

Generally measures that were shorter, easier to use and require less training have been recommended. However any systematic assessment requires some training to ensure that ratings/reports derived from any instrument are consistent across settings and across raters. There is a need for training on the principles of undertaking systematic assessment as well as the specific characteristics of a recommended instrument. It is thought that a project to develop modules addressing these issues, which could be incorporated in the curricula for relevant professional and para-professional groups, should be considered. In-service workshops could also be considered as is detailed in the discussion of a dissemination strategy below.

A major issue for some services may be the receipt of adequate funding so that staff can be released to undertake the necessary training. The time required for more comprehensive assessment may need to be factored into the workload and this will also have financial implications for services. An issue the Department of Health and Ageing might consider is how to best provide funding incentives for facilities that incorporate a more systematic approach to the assessment of people with dementia.

If mandating of the measures were to be considered at a later point then funding for training, and associated activities such as information technology, data collection, analysis and reporting would require a substantial investment.

### 12.6.6 A Dissemination Strategy

As with the Continence Outcomes Measurement Suite (Thomas, et al. 2006) it is recommended that a dissemination strategy project be funded to facilitate the uptake of the recommended tools. This strategy could incorporate the following elements:

- Development of an instrument kit/user manual for the recommended tools. This would include a copy of each tool (with approval from the authors), instructions for administration and scoring, appropriate applications for each tool, and the review of each tool.
- Development of web site materials about the project and the recommended instruments to go on the National Dementia website and related websites.
- The Department of Health and Ageing to facilitate access to the Final Report within Commonwealth and State jurisdictions.
- Development of brochures concerning the project and the recommended instruments for dissemination amongst professional groups.
- Conference presentations and training workshops at key conferences in Australia and overseas. Consideration could be given to the sponsoring of dedicated sessions on dementia research undertaken as part of the National Dementia Initiative at relevant conferences.
- Development of training videos to demonstrate how the instruments are administered.
- Development of in-service training workshops for clinical and care staff, managers and service providers.
- Development of newsletter articles for dementia, ageing and health publications.
- Development of peer reviewed journal articles concerning aspects of this project.

As mentioned above it is thought that a separate project to develop training modules addressing assessment issues, which could be incorporated in the curricula for relevant professional and para-professional groups, should be considered.

## 12.6.7 Identified Research Gaps

During the course of the project a number of research gaps have been identified. These are outlined below and are presented for consideration to the Department. These identified research gaps have also been judged in terms of their order of priority. Priority 1 research recommendations are important for the practical application and implementation of the DOMS recommended instruments in Australia. Priority 2 research recommendations either broaden the scope of the DOMS project or are more technical in nature and can therefore be conducted at a later date or in parallel to the basic research and development activities outlined as Priority 1. The final priority ratings were determined by averaging the individual priority ratings of 5 members of the EMG. (It is also noted that for items 15 and 17, concerning Indigenous communities, that there are NHMRC projects currently addressing these issues.)

1. Some measures need pilot testing in Australia to obtain reference data (e.g. Dementia Specific HRQOL measures, associated symptom measures). (Priority 1)
2. Further research work is required to assess the point at which people can no longer self-rate (e.g. the relevant MMSE-3MS score) that applies under different modes of administration (e.g. self report, interview, interview assisted). This will be required for all recommended self-report instruments. (Priority 1)
3. Future research might also address how training, the framing of questions, the terminology used and the administration of instruments influences the results of proxy assessment and more research is needed to compare proxy reports with performance-based measures as well as information from medical records and health utilization data. (Priority 2)
4. Many of the recent papers on proxy assessment use single or dual item informant measures (e.g. Tierney, et al. 2003; Watson, et al. 2004; Li, et al. 2006). Further research is required to ascertain whether these items have the requisite accuracy compared to longer proxy measures. (Priority 2)
5. Further research activities are required to address identified problems with Multi-attribute Utility measures: AQoL (shorten) and/or EQ-5D (scoring and distribution issues). (Priority 2)
6. A linguistic validation study be undertaken for the De Jong Gierveld Loneliness Scale to develop response categories more appropriate to the Australian context. (Priority 1)
7. It is recommended that the three social isolation instruments which performed relatively well (the De Jong Gierveld Loneliness Scale, the Friendship Scale and the Medical Outcomes Study Social Support Survey) could be trialled in at least one large dementia study for the explicit purpose of identifying the instrument to be recommended for future use. This trial should also assess proposed modifications to these instruments. From this study a statistically-derived single item measure could also be identified for use in everyday clinical consultations. (Priority 2)
8. Social function / social support areas may need follow up research if there is a wish to focus on social participation as well as social isolation but this could be combined with research outlined in recommendation 7 above. It might also include an examination of social support items from relevant ABS Surveys. (Priority 2)
9. In the absence of a research consensus for the measurement of function in dementia, and given a high degree of overlap in items, there is a clear need for a streamlining the various functional instruments and items across each of the practice settings (Spector, 1997). A study including a large group of dementia patients could examine and calibrate functional

items from the short-listed functional status instruments (both generic and dementia specific) to create a comprehensive item bank. This dementia item bank could then be used to examine item redundancy and coverage across the range of severity levels and could be used to develop new tools or provide cross-calibration between the existing instruments. This project would also need to examine the relationship of these items with the recommended cognitive and functional assessment staging instruments. (Priority 1)

10. It is recommended that a study be undertaken to test the Short Assessment of Patient Satisfaction scale and the two global patient satisfaction items identified with samples of people with dementia and their carers. It is noted that all these items would require minor rewording to make them suitable for use with an informant/carer. (Priority 1)
11. Carer satisfaction with services has been addressed in this project but an examination of the instruments used to assess carer burden, carer appraisal and carer wellbeing was outside the scope of this project. It is recommended that a more detailed follow up project be undertaken to examine issues relating to the assessment of instruments used to assess carer burden, carer appraisal and carer wellbeing. (Priority 2)
12. A further project is necessary to ensure the development of a more comprehensive database of dementia outcome measures solely for use with CALD communities - where translated versions of the DOMS selected measures are further reviewed and made available if possible. (Priority 1)
13. Further studies analysing the measurement equivalence of the core recommended measures be undertaken for major language groups within Australia. (Priority 2)
14. Research be undertaken to further examine instruments developed in Australia such as the RUDAS, the GPCOG, the KICA-Cog to ensure their validity and reliability in different groups of CALD populations. (Priority 1)
15. There needs to be further detailed research on the meaning of dementia in Indigenous communities, and how to ask questions which capture the experience of living with dementia in an Indigenous community. (Priority 1)
16. Further research be undertaken to adapt the recommended tools, as necessary, for Aboriginal and Torres Strait Islander Groups – particularly for rural and remote populations. (Priority 1)
17. It is recommended that further research be undertaken to assess the psychometric properties of the KICA-Cog and its' appropriateness for the assessment of cognitive impairment with both urban and remote Indigenous people. (Priority 1)
18. Individual assessment methods such as Goal Attainment Scaling, recently advocated by Rockwood (2007), to individualise outcome measurement for people with dementia, have not been examined in this project. It is recommended that a systematic review be undertaken to assess these methods. (Priority 2)
19. A decision was made by the DOMS-EMG that the project should focus on the instruments/tools that are available for use in routine care and this would exclude many of the more detailed neuropsychological instruments or instruments that require specialist training for their administration and interpretation. It is recommended that a further study could examine neuropsychological and specialist tests for people with dementia, in association with the relevant professional groups. (Priority 2)

20. That further investigation in the field be undertaken to fine tune the recommendations for the use of these instruments with people with Late Stage Dementia. This could also include an examination of instruments used in palliative care settings. (Priority 2)
21. There is a need for training on the principles of undertaking systematic assessment as well as the specific characteristics of a recommended instrument. It is thought that a project to develop modules addressing these issues, which can be incorporated in the curricula for relevant professional and para-professional groups, should be undertaken. (Priority 1)
22. Issues concerning safety / risk assessment are outside the scope of this project. It is recommended that a further project be undertaken to examine risk assessment issues (e.g. elder abuse, aggression, self harm etc) for people with dementia. (Priority 2)

## References

- Access Economics (2006) *Dementia prevalence and incidence among people who do not speak English at home*. Access Economics.
- American Educational Research Association, American Psychological Association, and the National Council on Measurement in Education (1999) *Standards for educational and psychological testing*. American Educational Research Association, American Psychological Association, National Council on Measurement in Education, Washington DC.
- Ankri J, Beaufile B, Novella JL, et al. (2003) *Use of the EQ-5D among patients suffering from dementia*. Journal of Clinical Epidemiology. Vol. 56, No.11, pp.1055-63.
- Antaki C (1999) *Interviewing persons with a learning disability: How setting lower standards may inflate well-being scores*. Qualitative Health Research. Vol. 9, No.4, pp.437-454.
- Antaki C and Rapley M (1996) 'Quality of Life' talk: *The liberal paradox of psychological testing*. Discourse and Society. Vol. 7, No.3, pp.293-316.
- Australian Pain Society (2004) *Residential Aged Care Pain Management Guidelines*.
- Baro E, Ferrer M, Vazquez O, et al. (2006) *Using the Nottingham Health Profile (NHP) among older adult inpatients with varying cognitive function*. Quality of Life Research. Vol. 15, No. 4, pp.575-585.
- Bartlett H, Rao DV and Warburton J (2006) *Ageing and cultural diversity in Queensland: Working together to make a difference*. Australasian Centre on Ageing, The University of Queensland.
- Berger EY (1980) *A system for rating the severity of senility*. Journal of the American Geriatrics Society. Vol. 28, No.5, pp.234-236.
- Borson S, Brush M, Gil E, et al. (1999) *The Clock Drawing Test: utility for dementia detection in multiethnic elders*. Journals of Gerontology Series A-Biological Sciences and Medical Sciences. Vol. 54, No.11, pp.M534-40.
- Borson S, Scanlan JM, Watanabe J, et al. (2006) *Improving identification of cognitive impairment in primary care*. International Journal of Geriatric Psychiatry. Vol. Vol. 21, No.4, pp.349-355.
- Brodsky H, Pond D, Kemp NM, et al. (2000) *The GPCOG: A new screening test for dementia designed for general practice*. Journal of the American Geriatrics Society. Vol. 50, pp.530-534.
- Byrne-Davis LMT, Bennett PD and Wilcock GK (2006) *How are quality of life ratings made ? Toward a model of quality of life in people with dementia*. Quality of Life Research. Vol. 15, pp.855-865.
- Centre for Applied Gerontology - Bundoora Extended Care Centre (1996) *Culturally appropriate dementia assessment*. Centre for Applied Gerontology - Bundoora Extended Care Centre, Victoria.

- Coucill W, Bryan S, Bentham P, et al. (2001) *EQ-5D in patients with dementia: an investigation of inter-rater agreement*. Medical Care. Vol. 39, No.8, pp.760-71.
- Cummins RA (2002) *Proxy responding for subjective well-being: A review*. International Review of Research in Mental Retardation. Vol. 25, pp.183-207.
- Daker-White G, Beattie AM, Gilliard J, et al. (2002) *Minority ethnic groups in dementia care: a review of service needs, service provision and models of good practice*. Ageing and Mental Health. Vol. 6, No.2, pp.101-8.
- Davis C, Wilson A and McCarthy S (1996) *Ethnicity and aged care assessment teams in Queensland*. Australian and New Zealand Journal of Public Health. Vol. 20, No.1, pp.33-40.
- Dick MB, Dick-Muehlke C and Teng EL (2006) *Assessment of cognitive status in asians*. In Yeo G, Gallagher-Thompson D, (eds). Ethnicity and the Dementias. 2nd ed. Routledge, Taylor and Francis Group, New York, pp.71-86.
- Edelman P, Fulton BR, Kuhn D, et al. (2005) *A comparison of three methods of measuring dementia-specific quality of life: perspectives of residents, staff, and observers*. Gerontologist. Vol. 45, Spec No.1, pp.27-36.
- Eslinger PS, Dennis K, Moore P, et al. (2005) *Metacognitive deficits in frontotemporal dementia*. J Neuro Neurosurg Psychiatry. Vol. 76, pp.1630-1635.
- Fillenbaum GG (1985) *Screening the elderly: A brief instrumental activities of daily living measure*. JAGS. Vol. 33, No.10, pp.698-706.
- Foddy W (1993) *Constructing Questions for Interviews and Questionnaires: Theory and Practice in Social Research*. Cambridge University Press, Cambridge.
- Folstein M, Folstein S and McHugh P (1975) *Mini-Mental State: a practical method for grading the cognitive state of patients for the clinician*. Journal of Psychiatric Research. Vol. 12, pp.189-198.
- Fountoulakis KN, Kaprinis SG and Kaprinis GS (2002) *Special characteristics affecting the neuropsychologic assessment of the elderly in Greece*. American Journal of Alzheimer's Disease and Other Dementias. Vol. 17, No.5, pp.273-276.
- Fries BE, Simon SE, Morris JN, et al. (2001) *Pain in US nursing homes; Validating a pain scale for the Minimum Data Set*. The Gerontologist. Vol. 41, pp.173-179
- Gallagher-Thompson D (2006) *The family as the unit of assessment and treatment in work with ethnically diverse older adults with dementia*. In Yeo G, Gallagher-Thompson D, (eds). Ethnicity and the Dementias. 2nd ed. Routledge, Taylor and Francis Group, New York, pp.71-86.
- Glosser G, Wolfe N, Albert ML, et al. (1993) *Cross-cultural cognitive examination: Validation of a dementia screening instrument for neuroepidemiological research*. JAGS. Vol. 41, pp.931-939.
- Graham C, Howard R and Ha Y (1998) *Dementia and ethnicity*. International Psychogeriatrics. Vol. 10, No.2, pp.183-91.
- Grypma R, Mahajani S and Tam E (2007) *Screening and diagnostic assessment of non-english speaking people with dementia: Guidelines and system recommendations for practitioners, service managers and policy makers*. Alzheimer's Australia.
- Hall KS, Ogunniyi AO, Hendrie HC, et al. (1996) *A cross-cultural community based study of dementias: methods and performance of the survey instrument in Indianapolis, U.S.A. and Ibadan, Nigeria*. International Psychogeriatrics. Vol. 6, pp.129-142.
- Hargrave R (2006) *Neuro-cognitive assessment of dementia in African American Elders*. In Yeo G, Gallagher-Thompson D, (eds). Ethnicity and the Dementias. 2nd ed. Routledge, Taylor and Francis Group, New York, pp.33-53.

- Harper GJ (2000) *Assessing older adults who cannot communicate*. In Kane RL, Kane RA, (eds). Assesing older persons: Measures, meaning, and practical applications. Oxford University Press, Oxford.
- Harwood DG and Ownby RL (2000) *Ethnicity and dementia*. Current Psychiatry Reports. Vol. 2, No.1, pp.40-5.
- Hassett A, George K and Harrigan S (1999) *Admissions of elderly patients from English-speaking and non-English-speaking backgrounds to an inpatient psychogeriatric unit*. Australian and New Zealand Journal of Psychiatry. Vol. 33, No.4, pp.576-82.
- Herzog AR and Rodgers WL (1999) *Cognitive performance measures in survey research on older adults*. In Schwarz N, Park DC, Knauper B, Sudman S, (eds). Cognition, ageing, and self reports. Psychology Press, Taylor and Francis Group, Philadelphia.
- Howe AL (2006) *Cultural diversity, ageing and HACC: Trends in Victoria in the next 15 years*. Department of Human Services Victoria.
- Hunter, E (2003) *Mental Health* in N Thomson (ed), The Health of Indigenous Australians, Oxford University Press, Melbourne, pp127-157.
- Iezzoni LI (2002) *Toward universal design in assessing health care experiences*. Medical Care. Vol. 40, No.9, pp.725-728.
- Iliffe S and Manthorpe J (2004) *The debate on ethnicity and dementia: from category fallacy to person-centred care?* Ageing and Mental Health. Vol. 8, No.4, pp.283-292.
- Johnson N, Barion A, Rademaker A, et al. (2004) *The Activities of Daily Living Questionnaire: A validation study in patients with dementia*. Alzheimer Disease and Associated Disorders. Vol. 18, No.4, pp.223-230.
- Jorm AF (2004) *The Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE): A review*. International Psychogeriatrics. Vol. 16, No.3, pp.275-293.
- Kemp NM, Brodaty H, Pond D, et al. (2002) *Diagnosing dementia in primary care: The accuracy of informant reports*. Alzheimer Disease and Associated Disorders. Vol. 16, No.3, pp.171-176.
- Klimidis S, Reddy P, Minas IH, et al. (2004) *Brief functional English proficiency measure for health survey research*. Australian Psychologist. Vol. 39, No.2, pp.154-165.
- Lam LC, Chiu HF, Ng KO, et al. (1998) *Clock-face drawing, reading and setting tests in the screening of dementia in Chinese elderly adults*. Journals of Gerontology Series B-Psychological Sciences and Social Sciences. Vol. 53, No.6, pp.P353-7.
- Lampléy-Dallas VT (2002) *Research issues for minority dementia patients and their caregivers: what are the gaps in our knowledge base?* Alzheimer Disease and Associated Disorders. Vol. 16 Suppl 2, pp.S46-9.
- Lawton M and Brody E (1969) *Assessment of older people: Self-maintaining and instrumental activities in daily living*. Gerontologist. Vol. 9, pp.179-186.
- Li M, Ng TP, Kua EH, et al. (2006) *Brief Informant Screening Test for Mild Cognitive Impairment and Early Alzheimer's Disease*. Dementia and Geriatric Cognitive Disorders. Vol. 21, No.5-6, pp.392-402.
- Lister S and Benson C (2006) *Comparative analysis of dementia and ethnicity in the New South Wales Aged Care Assessment Program: 1996 and 2001*. Australasian Journal on Ageing. Vol. 25, No.1, pp.24-30.
- LoGiudice D, Hassett A, Cook R, et al. (2001) *Equity of access to a memory clinic in Melbourne? Non-English speaking background attenders are more severely demented and have increased rates of psychiatric disorders*. International Journal of Geriatric Psychiatry. Vol. 16, No.3, pp.327-34.
- LoGiudice D, Smith K, Thomas J, et al. (2005) *Kimberley Indigenous Cognitive Assessment tool (KICA): Development of a cognitive assessment tool for older indigenous Australians*. International Psychogeriatrics. Vol. 18, No.2, pp.269-280.

Mackin RS, Depp C, Arean P, et al. (2006) *Overview of psychiatric assessment with dementia patients*. In Yeo G, Gallagher-Thompson D, (eds). Ethnicity and the Dementias. 2nd ed. Routledge, Taylor and Francis Group, New York, p 71-86.

Magaziner J (1997) *Use of proxies to measure health and functional outcomes in effectiveness research in persons with Alzheimer disease and related disorders*. Alzheimer Disease and Associated Disorders. Vol. 11, Suppl 6, pp.168-74.

Mast BT, Fitzgerald J, Steinberg J, et al. (2001) *Effective screening for Alzheimer's disease among older African Americans*. Clinical Neuropsychologist. Vol. 15, No.2, pp.196-202.

McCracken CF, Boneham MA, Copeland JR, et al. (1997) *Prevalence of dementia and depression among elderly people in black and ethnic minorities*. British Journal of Psychiatry. Vol. 171, pp.269-73.

McDowell I (2006) *Measuring health: A guide to rating scales and questionnaires*. Oxford University Press, Oxford.

McHorney CA and Fleishman JA (2006) *Assessing and understanding measurement equivalence in health outcome measures*. Medical Care. Vol. 44, No.11 Suppl 3, pp.S205-S210.

Morales LS, Flowers C, Gutierrez P, et al. (2006) *Item and scale differential functioning of the Mini-Mental State Exam assessed using the Differential Item and Test Functioning (DFIT) Framework*. Medical Care. Vol. 44, No.11 Suppl 3, pp.S143-51.

Mozley CG, Huxley P, Sutcliffe C, et al. (1999) *'Not knowing where I am doesn't mean I don't know what I like': Cognitive impairment and quality of life responses in elderly people*. International Journal of Geriatric Psychiatry. Vol. 14, No.9, pp.776-83.

Mungas D (2006) *Neuropsychological assessment of hispanics elders*. In Yeo G, Gallagher-Thompson D, (eds). Ethnicity and the Dementias. 2nd ed. Routledge, Taylor and Francis Group, New York, p 71-86.

Munsey C (2006) *Testing for impairment: Researchers are seeking culturally valid methods to assess cognitive functioning in the minority elderly*. Monitor on Psychology (APA). Vol. 37, No.10, pp.54.

Naglie G, Tomlinson G, Tansey C, Irvine J, Ritvo P, Black SE, Freedman M, Silberfeld M and Krahn M (2006) *Utility-based quality of life measures in Alzheimer's disease*. Quality of Life Research. Vol. 15, No.4, pp.631-43.

National Institute for Health and Clinical Excellence (NICE) (2007) *TA111 Alzheimer's disease - donepezil, galantamine, rivastigmine (review) and memantine: guidance (amended)*. National Institute for Health and Clinical Excellence, National Health Service, London. Available at <http://guidance.nice.org.uk/TA111/guidance/pdf/English/download.dspix>, accessed 27 September 2007.

Neumann PJ, Araki SS and Gutterman EM (2000) *The use of proxy respondents in studies of older adults: Lessons, challenges and opportunities*. JAGS. Vol. 48, pp.1646-1654.

Novella JL, Boyer F, Jochum C, et al. (2006) *Health status in patient's with Alzheimer's disease: An investigation of inter-rater agreement*. Quality of Life Research. Vol. 15, pp.811-819.

Novella JL, Jochum C, Ankri J, et al. (2001) *Measuring general health status in dementia: practical and methodological issues in using the SF-36*. Journal Ageing-Clinical and Experimental Research. Vol. 13, pp.362-9.

O'Connell B, et al. (2007) *Continence Assessment Tool for Residential Aged Care (Draft)*. School of Nursing, Deakin University.

O'Connor D, Phinney A, Smith A, et al. (2007) *Personhood in dementia care: Developing a research agenda for broadening the vision*. Dementia. Vol. 6, No.1, pp.121-142.

- Orlando Edelen MO, Thissen D, Teresi JA, et al. (2006) *Identification of differential item functioning using item response theory and the likelihood-based model comparison approach. Application to the Mini-Mental State Examination*. Medical Care. Vol. 44, No.11 Suppl 3, pp.S134-42.
- O'Rourke D, Sudman S, Johnson T, et al. (1999) *Cognitive testing of cognitive functioning questions*. In Schwarz N, Park DC, Knauper B, Sudman S, (eds). Cognition, ageing, and self reports. Psychology Press, Taylor and Francis Group, Philadelphia.
- Owen A, Eagar K and Ramsay L (2004) *Ongoing Needs Identification in Community Care: How to use the Tier 1 Assessment and Referral Tools*. Centre for Health Service Development, University of Wollongong.
- Pandav R, Fillenbaum G, Ratcliff G, et al. (2002) *Sensitivity and specificity of cognitive and functional screening instruments for dementia: The Indo-U.S. dementia epidemiology study*. Journal of the American Geriatrics Society. Vol. 50, No.3, pp.554-561.
- Paradies Y, (2007) *Racsim*. In Carson B, Dunbar T, Chenhall RD and Bailie R (eds.) The Social Determinants of Indigenous Health. Allen and Unwin, Sydney, pp.65-86.
- Park DC (1999) *Cognitive aging, processing resources, and self-report*. In Schwarz N, Park DC, Knauper B, Sudman S, (eds.) Cognition, ageing, and self reports. Psychology Press, Taylor and Francis Group, Philadelphia.
- Powell AL (2002) *On issues pertinent to Alzheimer disease and cultural diversity*. Alzheimer Disease and Associated Disorders. Vol. 16, No.Suppl 2, pp.S43-S45.
- Prince M, Acosta D, Chiu H, et al. (2004) *Effects of education and culture on the validity of the Geriatric Mental State and its AGE-CAT algorithm*. British Journal of Psychiatry. Vol. 185, No.5, pp.429-436.
- Purnell LD and Paulanka BJ (2005) *Guide to culturally competent health care*. F.A. Davis and Company, Philadelphia.
- Ramirez M, Teresi JA, Holmes D, et al. (2006) *Differential item functioning (DIF) and the Mini-Mental State Examination (MMSE). Overview, sample, and issues of translation*. Medical Care. Vol. 44, No.11 Suppl 3, pp.S95-S106.
- Rapley M and Antaki C (1996) *A conversation analysis of the 'acquiescence' of people with learning disabilities*. Journal of Community and Applied Social Psychology. Vol. 6, No.3, pp.207-227.
- Reid J and Dhamarrandji M (1978) *Curing not caring: why Aboriginal patients 'abscond'*. New Doctor: The Journal of the Doctors Reform Society. Vol. 8, pp.27-32.
- Riemsma RP, Forbes CA, Glanville JM, et al. (2001) *General health status measures for people with cognitive impairment : Learning disability and acquired brain injury*. Health Technology Assessment 2001. Vol. 5, No.6.
- Rockwood K (2007) *The measuring, meaning and importance of activities of daily living (ADLs) as an outcome*. International Psychogeriatrics. Vol. 19, No.3, pp.467-482.
- Sands LP, Ferreira P, Stewart AL, et al. (2004) *What explains differences between dementia patients' and their caregivers' ratings of patients' quality of life*. Am J Geriatric Psychiatry. Vol. 12, No.3, pp.272-280.
- Sansoni J, Marosszeky N, Sansoni E, et al. (2006) *Refining Continence Measurement Tools*. Centre for Health Service Development, University of Wollongong and the Department of Psychiatry, The University of Melbourne.
- Sansoni, J and Senior K (1998) *Health outcomes measurement and the coordinated care trials for Aboriginal and Torres Straits islander peoples*. Health Outcomes. Bulletin 9.

---

Schechter S, Beatty P and Willis GB (1999) *Asking survey respondents about health status: Judgment and response issues*. In Schwarz N, Park DC, Knauper B, Sudman S, (eds.) Cognition, ageing, and self reports. Psychology Press, Taylor and Francis Group, Philadelphia.

Senior K (2003) *A gudbala Laif? Health and well being in a remote Aboriginal community-what are the problems and where lies responsibility?* PhD thesis, Australian National University, Canberra.

Shah A, Dalvi M and Thompson T (2005) *Behavioural and psychological signs and symptoms of dementia across cultures: current status and the future*. International Journal of Geriatric Psychiatry. Vol. 20, No.12, pp.1187-95.

Sigelman CK, Budd EC, Spanhel CL and Schoenrock CJ (1981a) *Asking questions of retarded persons: a comparison of yes-no and either-or formats*. Applied Research in Mental Retardation. Vol. 2, No.4, pp.347-57.

Sigelman CK, Budd EC, Spanhel CL and Schoenrock CJ (1981b) *When in doubt, say yes: acquiescence in interviews with mentally retarded persons*. Mental Retardation. Vol. 19, No.2, pp.53-8.

Snellgrove CA and Hecker JR (2005) *Brief checklist for non-cognitive symptoms of dementia*. Australasian Journal on Ageing. Vol. 24, No.2, pp.88-93.

Snow AL, Cook KF, Lin P-S, et al. (2005a) *Proxies and other external raters: Methodological considerations*. Health Services Research. Vol. 40, No.5, pp.1676-1693.

Snow AL, Kunik ME, Molinari VA, et al. (2005b) *Accuracy of self-reported depression in persons with dementia*. JAGS. Vol. 53, pp.389-396.

Spark R, Donovan RJ and Binns C (1992) *Aboriginal peoples' perceptions of their own and their community's health*. Health Promotion Journal of Australia. pp.60-61.

Storey JE, Rowland JTJ, Basic D, et al. (2002) *Accuracy of the clock drawing test for detecting dementia in a multicultural sample of elderly Australian patients*. International Psychogeriatrics. Vol. 14, No.3, pp.259-71.

Storey JE, Rowland JTJ, Conforti DA, et al. (2004) *The Rowland Universal Dementia Assessment Scale (RUDAS): a multicultural cognitive assessment scale*. International Psychogeriatrics. Vol. 16, No.1, pp.13-31.

Storey JE, Rowland JTJ, Conforti DA, et al. (2004) *The Rowland Universal Dementia Assessment Scale (RUDAS): a multicultural cognitive assessment scale*. International Psychogeriatrics. Vol. 16, No.1, pp.13-31.

Teng EL, Hasegawa K, Homma A, et al. (1994) *The Cognitive Abilities Screening Instrument (CASI): A practical test for cross-cultural epidemiological studies of dementia*. International Psychogeriatrics. Vol. 6, pp.45-58.

Teng EL and Manly JJ (2005) *Neuropsychological testing: helpful or harmful?* Alzheimer Disease and Associated Disorders. Vol. 19, No.4, pp.267-271.

Teri L, Truax P, Logsdon R, Uomoto J, Zarit S and Vitaliano PP (1992) *Assessment of behavioral problems in dementia: the revised memory and behavior problems checklist*. Psychology and Ageing. Vol. 7, No.4, pp.622-631.

Tierney MC, Herrmann N, Geslani DM, et al. (2003) *Contribution of Informant and Patient Ratings to the Accuracy of the Mini-Mental State Examination in Predicting Probable Alzheimer's Disease*. Journal of the American Geriatrics Society. Vol. 51, No.6, pp.813-818.

Torti FM Jr., Gwyther LP, Reed SD, et al. (2004) *A multinational review of recent trends and reports in dementia caregiver burden*. Alzheimer Disease and Associated Disorders. Vol. 18, No.2, pp.99-109.

Victorian Transcultural Psychiatry Unit (2006) *Guidelines for working effectively with interpreters in mental health settings*. Victorian Transcultural Psychiatry Unit.

---

Waite L, Grayson D, Jorm AF, et al. (1999) *Informant-based staging of dementia using the Clinical Dementia Rating*. Alzheimer Disease and Associated Disorders. Vol. 13, No.1, pp.34-37.

Wall JR, Deshpande SA, MacNeill SE, et al. (1998) *The Fuld Object Memory Evaluation, a useful tool in the assessment of urban geriatric patients*. Clinical Gerontologist. Vol. 19, No.1, pp.39-49.

Watson LC, Lewis CL and Fillenbaum GG (2004) *Asking family about memory loss: Is it helpful?* J General Internal Medicine. Vol. 20, pp.28-32.

Weiner DK, Herr K and Rudy TE (eds.) (2002) *Persistent pain in older adults*. Springer Publishing Company, New York.

Williams C, Zimmerman S, Sloane P, Reed P (2005) *Characteristics associated with pain in long-term residents with dementia*. The Gerontologist. Vol. 45, pp.68-73.

Wingfield A (1999) *Comprehending spoken questions: Effects of cognitive and sensory change in adult aging*. In Schwarz N, Park DC, Knauper B, Sudman S, (eds). Cognition, ageing, and self reports. Psychology Press, Taylor and Francis Group, Philadelphia.

Wlodarczyk, JH, Brodaty H and Hawthorne G (2004) *The relationship between quality of life, Mini-Mental State Examination, and the Instrumental Activities of Daily Living in patients with Alzheimer's Disease*. Archives of Gerontology and Geriatrics. Vol. 39, No.1, pp.25-33.

Wolfe N (2002) *Cross-cultural neuropsychology of aging and dementia: An update*. In Ferraro, Richard (ed.) Minority and Cross-Cultural Aspects of Neuropsychological Assessment, Taylor and Francis.

Yasuda N, Zimmerman S, Hawkes WG, et al. (2004) *Concordance of proxy-perceived change and measured change in multiple domains of function in older persons*. JAGS. Vol. 52, pp.1157-1162.

Yeo G and Gallagher-Thompson D (2006) *Ethnicity and the Dementias*. Routledge, Taylor and Francis Group, New York.

## 13 Conclusions and Recommendations

This report, the Final Project Report for the Dementia Outcomes Measurement Suite, provides a summary of work undertaken since the DOMS Project commenced in April 2006. The sections below provide details of the recommendations from each section of this report.

### 13.1 *Clinical Terminology and Classifications Systems*

Section 3 of this report provides a detailed discussion of these issues. The recommendations below have been based on the review of literature, clinical feedback and these recommendations have also been ratified by the National Expert Panel.

It is recommended that:

- The ICD-10-AM is used to inform the diagnostic classifications for dementia and its subtype given this system is already in place in collecting national data in Australia.
- The ICD-10-AM and ICD-10 are used for diagnostic criteria for dementia and AD. Following consultation it seemed appropriate to recommend the ICD-10 instead of DSM-IV. Clinicians do not necessarily follow either of the classifications as they often rely on their clinical judgement. Given that the majority of the health related information is collected based on the ICD-10 and the ICD-10-AM it is more efficient for clinicians to use one system rather than two (i.e. DSM-IV diagnostic criteria and ICD-10 for coding exercise).
- For research, the DSM-IV is preferred as it is more inclusive of mild to moderate dementia and most epidemiological studies use the DSM-IV because of ease of comparison with prior studies. However this is not mandatory, providing the study states the type of the classification used, as there is no evidence available to say the DSM-IV is superior to the ICD-10.
- In terms of differential diagnosis (DD) and diagnoses of frontotemporal dementia (FTD) and dementia with Lewy bodies (LBD), additional criteria are used: the National Institute of Neurologic Disorders and Stroke and the Association Internationale pour la Recherche et l'Enseignement en Neurosciences (NINDS-AIREN) (Roman, et al. 1993) for DD of Vascular dementia from Alzheimer's type; the Lund-Manchester criteria for FTD (1994) and the consensus criteria for LBD (McKeith, et al. 2005).
- Mild cognitive impairment (MCI) is not to be included in this project as a diagnostic entity, however screening measures for those who are suspected of cognitive impairment need to be considered.
- For assessing the severity of dementia, the CDR scale has been used for two main reasons: the AIHW recommends this and, in addition to three stages of dementia, the CDR allows room to record abnormal cognitive function without necessarily labelling it as MCI. It is well validated and widely recognised. Similarly the GDS has also been widely used to assess the severity of dementia. A detailed review of these instruments is provided in the following section and in Appendix 5.
- The ICF may be used as a conceptual framework for classification of measurement scales. However, given its early developmental status as a classification system in Australia, hence its unfamiliarity among clinicians and researchers, and lack of evidence relating to validity and reliability of the classification, it is deemed beyond the scope of the DOMS project to provide a definite recommendation on this subject.
- Behavioural and psychological symptoms of dementia (BPSD) are an integral part of dementia outcome measures. The guidelines provided by the International Psychogeriatric Association (IPA) are to be used for the definitions. Whilst the AIHW recommends Caldwell and Bird's

guideline for the severity of BPSD, it has been suggested that a more widely recognised measure is selected for this project.

### **13.2 Recommended Assessment Instruments for Dementia**

The review of selected measures for the following categories of instruments has been completed and the recommendations are below:

#### **13.2.1 Dementia Staging and Descriptive Instruments**

Burns et al (2004) indicates these measures are widely used as staging measures in descriptive and intervention studies. It is noted that specialist clinicians are less likely to use these global staging instruments than other clinical or research personal. Such instruments may not be particularly useful for fine differentiation at an early stage of dementia. However, global functional scales like the GDS and the CDR have their place in broadly describing people with dementia; particularly for research purposes and in residential care and community care settings.

- The most highly recommended instrument across the range of settings and measurement purposes was the Global Deterioration Scale (GDS).
- The Clinical Dementia Rating (CDR) scale is also recommended for a more in depth assessment and it is also often used in clinical research settings.
- The Dementia Severity Rating Scale (DSRS) may be applicable for use in community settings and for obtaining information from caregivers.

#### **13.2.2 Health Related Quality of Life and Health Status Instruments**

Both generic health status measures and dementia specific health related quality of life measures were examined. Many of the generic instruments are lengthy, complex and contain items that may not be relevant for this group and require self-completion. For these reasons no generic health status measures were recommended.

- The recommended dementia specific health related quality of life instruments were the Quality of Life in Alzheimer's Disease (QOL-AD) and the DEMQOL for mild and moderate dementia and the Quality of Life in Late Stage Dementia (QUALID) for late stage dementia only.
- It is noted that none of these instruments have published Australian reference data and thus it is recommended that such data be collected in an Australian field test of these instruments.

#### **13.2.3 Instruments for the Assessment of Cognitive Status**

- The most highly recommended instrument across the range of settings and measurement purposes was the Modified Mini-Mental State Exam (3MS).
- The General Practitioner Cognition Scale (GPCOG) was also recommended for use in General Practice settings.

- The Minimum Data Set - Cognition (MDS-Cog) rating scale was recommended for use with patients with severe dementia in the residential care (nursing home) context.
- The RUDAS is a new instrument that was designed to enable the easy translation of the items into other languages and to be culture fair. There are relatively few papers published as yet concerning its psychometric properties (especially construct validity) but in the interim it is recommended for use with those from Culturally and Linguistically Diverse backgrounds. The RUDAS, however, contains an item on judgement that may be inappropriate for remote Indigenous people (refer below).
- An interim recommendation is to use the Kimberley Indigenous Cognitive Assessment (KICA-Cog) tool for the cognitive assessment of rural and remote Indigenous people. The KICA-Cog is a new instrument and although there is little published evidence concerning this tool available as yet, and further research is required, this instrument has been designed for use with Indigenous people.

### 13.2.4 Multi-attribute Utility Measures

With regard to the economic evaluation of dementia interventions the three instruments that score most highly on the study criteria are the EQ-5D, the AQoL and the HUI-3. However both the HUI-3 and the AQoL are lengthier instruments which may place considerable cognitive burden on patients with dementia. It is noted that the HUI-3 does not score as highly on these criteria as the AQoL and the EQ-5D instruments for dementia settings and there are also considerable costs associated with the use of the HUI-3 which may also preclude its adoption.

- It is recommended that the EQ-5D and the AQoL are to be the preferred instruments when undertaking economic evaluation of dementia interventions.

The obvious instrument of choice for use in dementia studies might be the EQ-5D because of the simplicity of the descriptive system. There are, however, very good technical reasons which provide caveats to its widespread use, including competing scoring algorithms, ceiling effects, inconsistent utility scores and poor score distribution.

- It is recommended that an Australian study into these aspects of the EQ-5D with a view to validating and/or revising existing EQ-5D scoring algorithms.

Based on the scoring criteria, the next best-performing MAU-instrument was the AQoL. There are, however, two important caveats to recommending it as the instrument of choice. Although the AQoL's descriptive system is simple, the wording of items is stilted. The second caveat is in relation to the number of items needed to score the AQoL (12-items) which may explain higher rates of missing data when compared with the EQ-5D, and inconsistent scores for those with severe cognitive impairment. Theoretically, given the factorial structure of the AQoL it could be shortened through removal of 4 items (1 from each dimension) leaving it as an 8-item instrument.

- It is recommended that a study be undertaken to examine the effect of simplifying the AQoL items and removing four items to make it more appropriate for use in dementia research.

A single MAU-instrument could be recommended as the preferred instrument of choice for routine use at the clinician- and specialist-levels. This instrument should be short, easy to administer and score and population norms could be made available for easy reference. If such a policy was

adopted, it would be in light of the limitations outlined in this report and there would be no guarantee that results obtained would be comparable with results obtained elsewhere using another instrument. Indeed, where QALYs were computed as the result of a treatment, it is likely these would reflect instrument choice as much as treatment effect.

- It is recommended that two MAU-instruments could be included in any particular research or evaluation study, and that researchers be encouraged to provide both sets of results. One of the recommended instruments should be that recommended for clinician use. This strategy would have the benefit of reducing the bias inherent in a one-instrument strategy, and it would produce a range of estimated benefits from interventions, thus acknowledging the limitations of relying upon any particular existing MAU-instrument. Given that, inevitably, comparisons will be made with dementia studies overseas, this strategy would have the further benefit of enabling cross-cultural comparisons. An important limitation of this strategy is that it would increase the cognitive burden for those with moderate to severe cognitive impairment. It may also lead to interviewer-facilitated or proxy completions, with all the implications of mixed-methods data collection.

### 13.2.5 Measures of Perceived Social Isolation and Social Support

None of the reviewed instruments can be given an unqualified recommendation for use in Australian studies with older adults who have cognitive impairment or dementia. Subject to this finding, the standout instrument was the De Jong Gierveld Loneliness Scale. The reasons were that it was carefully conceived over a very substantial period of time, that it was developed in population samples (including older adults), and that there is a very substantial body of evidence supporting its reliability and validity. The reason the De Jong Gierveld Loneliness Scale, especially the short 6-item version, cannot be recommended outright is that the response categories may be inappropriate for use in Australian samples of people with cognitive impairment. However, a study could easily be completed to undertake a linguistic validation of this instrument for Australian use and this is recommended

The two other instruments that performed relatively well against the criteria were the Friendship Scale and the Medical Outcomes Study Social Support Survey. Given this situation, it is recommended that the three instruments which performed well (the De Jong Gierveld Loneliness Scale, the Friendship Scale and the Medical Outcomes Study Social Support Survey) be trialled in at least one large dementia study for the explicit purpose of identifying the instrument to be recommended for future use. This would enable many of the questions raised in this report regarding the validity of these instruments to be thoroughly investigated in an Australian context.

It was noted that the above instruments largely focus on perceived social isolation. There were issues concerning the theoretical basis of the construct being measured for many of the measures identified. Some measures focus on social isolation – others on actual and perceived social support – and others were a mix of both dimensions. Those instruments that attempted to measure actual social support had numerous cultural issues associated with them and many were too complex to consider with people with dementia. If a measure of perceived social support is required, it is suggested that a number of additional items be trialed in the field study suggested above to derive a short measure that could tap this domain.

It is further recommended that:

- The three instruments which performed well (the De Jong Gierveld Loneliness Scale, the Friendship Scale and the Medical Outcomes Study Social Support Survey) be trialled in at least one large dementia study for the explicit purpose of identifying the instrument to be recommended for future use. This would enable many of the questions raised in this report regarding the validity of these instruments to be thoroughly investigated in an Australian

context. It may be possible to derive a better short measure by selecting the items with the best psychometric properties from these scales.

- Explicit modifications to the De Jong Gierveld Loneliness Scale and the Medical Outcomes Study Social Support Survey be tested. These modifications are a revision of the De Jong Gierveld Loneliness Scale response categories, and a reduction in the number of items in the Medical Outcomes Study Social Support Survey (which would need to be tested in the study recommended above).
- The three instruments which performed well be tested in a trial for the effect of administration mode on scores given that there are good reasons for limiting self-completion among those with moderate or severe cognitive impairment. Three methods of administration should be directly compared (self-completion without assistance, interviewer-assisted completion, and proxy-completion) both cross-sectionally and longitudinally in order to develop algorithms for weighting enabling score equivalence across administration mode. This would overcome issues related to the cognitive impairment of respondents and meet the need to collect outcome efficacy data relating to program evaluation.
- From any study carried out under the recommendations above, a statistically-derived single item measure be identified for use in everyday clinical consultations.

### **13.2.6 Measures of the Associated Symptoms of Dementia**

'Associated symptoms of dementia' relate to characteristics of dementia that are not historically considered as major features such as cognitive impairment and related functional consequences, yet have a significant impact on the well-being of the persons with dementia and their family and caregivers. Measuring outcomes of care, service, treatment and interventions related to the associated symptoms of dementia is an important aspect. For the purpose of the DOMS project, the assessment of associated symptoms of dementia comprises:

- 1) Measures of global behavioural and psychological symptoms of dementia (BPSD Global, henceforward);
- 2) Measures of delirium, which is one of the two most frequently mistaken features requiring differential diagnosis from dementia (the other commonly mistaken feature is depression); and
- 3) Measures of particular symptoms of BPSD including aggression, agitation, anxiety, apathy, and depression.

#### **13.2.6.1 Recommendations Concerning BPSD Global Instruments**

A number of global measures of behavioural and psychological disturbance (Global BPSD) have been reviewed. The examination of key attributes and psychometric properties of the five final instruments of BPSD Global, measured against the weighting criteria indicates the NPI and the BEHAVE-AD as the best measures for assessment of BPSD, followed by the CERAD-BRSD, the DBDS and the NRS. Based on these reviews it is recommended that:

- The NPI and the BEHAVE-AD be used in both clinical and research settings for assessment of Global BPSD. These instruments both have well established psychometric properties.

- The CERAD-BRSD is recommended for research rather than routine practice given its cost and the time required for its administration. A 17 item abbreviated version may be considered better for clinical utility, but limited evidence on this version is currently available.

### 13.2.6.2 Recommendations Concerning Measures of Delirium

A number of delirium measures were also assessed in order to aid in the differential diagnosis of dementia and delirium. The Confusion Assessment Method (CAM) is the most widely utilised screening/diagnostic tool for detecting delirium internationally among older people with or without dementia. Less well known, however, the Delirium Rating Scale (DRS-R-98) is also a widely recognised and well validated measure. Whilst the CAM is superior in its utility to the DRS-R-98, it does not capture severity of delirium symptoms hence is not appropriate for repeated measures of delirium severity. The DRS-R-98 is designed for assessment of both the presence and the severity of delirium symptoms. Limitations of the DRS-R-98, and the DRS, include that they are time taxing and require sufficient training, especially for those who do not have psychiatric background. The DRS-R-98 is not appropriate for use in the community setting given its requirement for observation over a 24 hour period. However, it allows for comprehensive assessment of individuals who are at risk or suspected of developing delirium in institutional care settings.

For the purpose of the DOMS project it is recommended both measures be included as they have two distinct, yet equally important functions.

- It is recommended that the Confusion Assessment Method is used to assess the presence of delirium across most service settings.
- It is recommended that the Delirium Rating Scale (DRS-R-98) is used where a more comprehensive assessment of both the presence and severity of delirium is required. It is noted this instrument is not appropriate for use in community settings.

### 13.2.6.3 Recommendations Concerning Measures of Particular Symptoms of BPSD

In many cases the use of Global BPSD measures such as the NPI may suffice for the assessment of the associated symptoms of dementia. However, if a more detailed assessment of a particular symptom is required the following recommendations are made:

Aggression: Rating Scale for Aggressive Behaviour in the Elderly (RAGE)  
 Agitation: Cohen Mansfield Agitation Inventory (CMAI); Pittsburgh Agitation Scale (PAS)  
 Anxiety: Rating Anxiety in Dementia (RAID)  
 Apathy: Apathy Evaluation Scale (AES)  
 Depression: Cornell Scale for Depression in Dementia (CSDD)  
 Geriatric Depression Scale (GDS Yesavage) - less severe cases and in community settings)

A full discussion of these measures and their assessment can be found in Section 9 and Appendix 10 of this report.

### 13.2.7 Measures of Function

The Functional Independence Measure (FIM) and the Barthel Index, and the Lawton and Brody IADL and the Older Americans Resources and Services (OARS-IADL) instruments, were chosen as generic measures of ADL and IADL respectively. These instruments have been reviewed

elsewhere recently (Eagar, et al. 2001, 2006; Thomas, et al. 2006), have good psychometric properties and have been used in geriatric settings.

With regard to the activities of daily living, the FIM is probably more appropriate for acute care and high level residential care settings but it is noted that accredited training is required for its use. However, it is already widely used in acute care rehabilitation settings within Australia. The Barthel Index is an easier to use measure and may be more appropriate for use in primary and community care settings with people with mild to moderate forms of dementia. Although the Katz ADL instrument has been quite widely used in dementia settings the review of this instrument by Thomas, et al. (2006) indicated it had weak psychometric properties and thus it is not recommended for use.

- The Functional Independence Measure (FIM) and the Barthel Index are recommended as the generic measures of ADL.
- The Lawton and Brody IADL and the Older Americans Resources and Services (OARS-IADL) are recommended as generic instruments for the assessment of instrumental activities of daily living (IADL). The OARS-ADL is preferred as it is an advance on the Lawton and Brody IADL scale with improved psychometric properties and less reliance on gender role stereotypes; and it has been adapted for use in primary and community care settings in Australia (see Green, et al. 2006).

The recommended dementia specific instruments for the assessment of function (ADL and IADL) for people with dementia include both proxy measures and clinical rating scales. While it is acknowledged that proxy reports have their limitations (refer Section 12), they will generally be used where assessment by interview or self rating is no longer possible due to the degree of cognitive impairment of the person with dementia. Proxy measures are also extremely useful in primary and community care settings in order to monitor the maintenance of functional status or its decline, in connection with drug therapy or in terms of care management as the disease progresses. The direct observation rating scale may be more appropriate for acute care and residential care settings. By recommending both proxy and direct observation rating scales different practice settings and clinical situations (e.g. a person with dementia may not have a carer) can be addressed (refer Table 10).

- The Alzheimer's Disease Co-operative Study – ADL (ADCS-ADL) and Disability Assessment for Dementia Scale (DAD) are the two proxy report instruments that are recommended.
- For the direct observation of functioning the Cleveland Scale for Activities of Daily Living (CSADL) is recommended.

The discussion of measures of functional status in Section 10 highlights a number of measurement problems with regard to the assessment of function with people with dementia. It is clear there is an urgent need for a program of research and development in this area. It is recommended that:

- In the absence of a research consensus for the measurement of function in dementia, and given a high degree of overlap in items, there is a clear need for a streamlining of the various functional instruments and items across each of the practice settings (Spector, 1997). The work of Lindeboom, et al. (2003) in the Amsterdam Liner Disability Score Project using IRT to calibrate ADL instruments in neurology could be used as a guide. A similar study with a large group of people with dementia could examine and calibrate functional items from the short-listed instruments (both generic and dementia specific) to create a comprehensive item bank. This dementia item bank could then be used to examine item redundancy and coverage across the range of severity levels and could be

used to develop new tools or provide cross-calibration between the existing instruments. This project would also need to examine the relationship of these items with recommended cognitive and functional assessment staging instruments.

### 13.2.8 Measures of Patient and Carer Satisfaction

The Short Assessment of Patient Satisfaction (SAPS) scale, a generic measure of patient satisfaction, was developed from a study by Hawthorne, et al. (2006) which selected the best items from a range of commonly used generic patient satisfaction measures to produce a shorter patient satisfaction scale. The SAPS contains only seven items (one for each theoretical dimension of patient satisfaction) and had better psychometric properties than the instruments from which it was derived. It also had excellent internal consistency with a Cronbach's alpha of 0.86. Although SAPS needs to be further tested in other samples and populations (e.g. including persons with dementia and carers) it is the recommended generic measure for the assessment of patient satisfaction. The SAPS may be most appropriate for people with dementia that are receiving a particular treatment intervention for dementia and in research settings.

Six generic items measuring global satisfaction were identified from the Hawthorne, et al. (2006) study and were analysed concerning their appropriateness as a single measure for immediate assessment of patient satisfaction. Two of these items had better psychometric properties and were less prone to differential item functioning. These items were a) How satisfied are you with the outcome of your treatment? and b) How satisfied are you with the amount of help received? Item a) was chosen as the single item for satisfaction with incontinence treatment (Hawthorne, et al. 2006) given its better psychometric properties. However, this item may be less appropriate for dementia settings where often general care services are provided rather than specific treatment interventions per se. Further consideration should be given to this issue.

It is recommended that a study be undertaken to test the SAPS and the two single patient satisfaction items identified above with samples of people with dementia and their carers. It is noted that all these items would require minor rewording to make them suitable for use with an informant/carer.

With regard to measures of carer satisfaction given the findings of the review in Section 11.2 and Table 12 (Executive Summary), none of the reviewed instruments can be given an unqualified recommendation for use in Australian studies with carers of older adults who have cognitive impairment or dementia. The following recommendations are made:

- The most promising instrument appears to be the SWC-EOLD (Volicer, et al. 2001), and it is recommended that this instrument is used in an Australian study specifically designed to test its measurement properties.
- The alternative would be to mount a specific carer satisfaction study, where all items from all reviewed instruments were pooled and tested. The explicit purpose would be identifying well performing items and/or the best performing instrument.

A brief discussion is provided in Section 11.3 concerning carer satisfaction with services and its relationship to the related domains of carer burden and carer wellbeing. Carer satisfaction with services has been addressed in this project but an examination of carer burden, carer appraisal and carer wellbeing was outside the scope of this project. Although a number of recent studies (Brodaty, et al. 2002; Ramsay, et al. 2006) have examined issues relating to carer burden, a comparison of the leading instruments used to assess carer burden is yet to be undertaken.

### **13.3 Measurement Issues**

Some key measurement issues relevant to the use of these measures with dementia patients and their carers are outlined. The first of these is the issue of the use of proxies (formal and informal carers) for the assessment of the person with dementia. People with severe dementia may not be able to be assessed directly and may be unable to provide a self report where this may be required. This is followed by a discussion of the level of cognitive impairment at which people with dementia may lose the capacity to self rate. These issues are most important to consider when assessing more subjective phenomena such as health related quality of life and social isolation.

The applicability of these measures for particular population groups is also discussed. The issue of the applicability of the measures for those from Culturally and Linguistically Diverse (CALD) populations is considered as is the applicability of these measures for use with Aboriginal and Torres Strait Islander Groups. The recommendations pertaining to these issues are outlined below.

#### **13.3.1 Recommendations Concerning Cognitive Impairment and the Capacity to Self Rate**

Section 12 provides a more detailed discussion of this issue. Where it is possible and feasible HRQOL (and other subjective phenomena) should be assessed by patient self report rather than by proxy report. Sometimes this is not possible with persons with severe dementia and thus the following recommendations are made:

- An interim recommendation (awaiting the results of further recommended research) is that self rating report (by non interview administration) should not be considered for patients with MMSE scores below 15.
- For patients with MMSE scores ranging from 10-15 an interview administration or an interview assisted administration of these self-report measures could be considered.
- For patients with an MMSE score less than 10 it is suggested that data be collected via proxy reporting. Where a specific proxy form has been developed this should be utilised.
- It is recommended that a study be undertaken to assess the recommended self report tools by self report administration, interview administration and assisted interview administration to identify the best approach for assessing the HRQOL and other subjective phenomena of people with dementia with more severe cognitive impairments.
- As the capacity for cognitively impaired individuals to self rate will depend on the structure, length, design and complexity of each questionnaire it is suggested that a follow up study be undertaken to assess the required MMSE-3MS scores required for the recommended self report questionnaires under different modes of administration.

#### **13.3.2 Recommendations Concerning Proxy Assessment**

Section 12 provides a discussion of the issues concerning using proxy assessment where direct assessment of the dementia patient is not possible. Below are a number of recommendations when using proxy measures:

- Proxy reports should be examined for three potential biases: (1) the cognitive status of the proxy (as many elderly people are cared for by an elderly spouse carer, who may themselves be impaired or unwell, but to a lesser degree); (2) the health status of the proxy; and (3) the level of carer burden and stress (Harper, 2000).
- There is usually a trade-off between those “with the greatest amount of contact and those with more training” (Harper, 2000, page 488). However, generally, where a proxy report is used

information should be collected from the family member/carer or care staff member that is closest to the patient and has the greatest degree of interaction with the patient.

- Proxy reports should be based on usual behaviour rather than extreme or rare behaviours (Harper, 2000).
- Proxy reports should be based on observable phenomena like physical symptoms and functioning, rather than subjective phenomena like depression, social isolation and quality of life (Snow, 2005a).

### **13.3.3 Recommendations Concerning Assessment with Culturally and Linguistically Diverse (CALD) Populations**

- The use of the DOMS selected tools can be interpreted with less confidence if used by practitioners and interpreters who are not culturally competent. For an outline of the application of culturally competent assessment see the guidelines proposed by Alzheimer's Australia – National Cross Cultural Dementia Network (Grypma, Mahajani and Tam, 2007).
- A further project is necessary to ensure a more comprehensive database intended for dementia outcome measures solely for CALD communities - where translated versions of the DOMS selected measures are further reviewed and made available if possible.
- Further studies analysing the measurement equivalence of the core recommended measures be undertaken for major language groups within Australia.
- Research to further examine new instruments developed in Australia such as the RUDAS, KICA-Cog and the GPCOG is supported to ensure their validity and reliability in different groups of CALD populations.

### **13.3.4 Recommendations Concerning Assessment with Aboriginal and Torres Strait Islander Populations**

Many of the recommended scales may have applications among urban Indigenous people but this needs to be ascertained.

- It is recommended that some focus groups in urban settings are developed to discuss how appropriate the recommended instruments are to members of these communities.

There is very limited application for these tools for people from remote Aboriginal communities. A notable exception to this is the Kimberley Indigenous Cognitive Assessment tool (LoGiudice, et al. 2005) which had been designed for use with Indigenous people in remote locations.

- There is an interim recommendation, pending further research, that the KICA is used to assess the cognitive status of rural and remote indigenous peoples rather than the MMSE-3MS.

Clinician ratings may have more application than the self report or the proxy administered forms, as some of the ratings can be made through observation, rather than attempting to elicit answers from the patient. Cognitive assessment will be extremely difficult in many remote settings, and especially if the patient speaks and understands limited English. Clinical assessments may be improved if other confounding factors are removed, such as unfamiliarity of the clinician and environment. A clinician, who is familiar to the individual and has a good knowledge of their life, may be in a position to make a more informed judgment. While it may be possible to use some of the simpler tools in a remote setting, especially with modifications to pictorially demonstrate

concepts such as volume, questions will still remain about what the answers that individuals supply actually mean.

- It is recommended that there needs to be further detailed research on the meaning of dementia in Indigenous communities, and how to ask questions which capture the experience of living with dementia in an Indigenous community.
- It is recommended that a project be undertaken to examine the adaptations required to the recommended tools to make them more appropriate to Indigenous peoples.
- It is recommended that further research be undertaken to assess the psychometric properties of the KICA-Cog and its' appropriateness for the assessment of cognitive impairment with both urban and remote Indigenous people.

### **13.4 Implementation Issues**

Although issues pertaining to implementation have been discussed throughout this report and particularly in Section 12.6 a number of key areas to address are identified. These are:

- The issue of mandating the recommended measures;
- The application of the instruments in different settings and for different stages of dementia;
- Training issues;
- A dissemination strategy; and
- Identified research gaps.

With regard to a discussion of the issue of mandating the recommended measures the reader is referred to Section 12.6.2 of this report. Advice received from the Department of Health and Ageing in August 2006 indicated there was no desire to mandate the recommended instruments at this stage. The project team was advised mandating was not a consideration at this time as the Dementia Outcomes Measurement Suite was a first-stage project to assess key gaps and tools. It was noted the CRCs and Study Centres may promote the use of particular tools agreed as a result of the DOMS-NEP project; however, this would be as best practice, rather than to mandate.

Given the use of the measurement tools is to be recommended rather than mandated, and more comprehensive assessment produces an increased burden on staff, there may need to be some consideration of financial incentives for services that adopt the use of the recommended tools.

It would be difficult to mandate the use of the recommended measures without full consideration of the training requirements and the burden on staff time for all service settings to implement these measures. If routine data collection and analysis is desired, with a view to benchmarking the outcomes of similar services, then careful thought must be given to the design of such systems and the phased implementation of such an approach. To adopt such an approach will also require a considerable financial investment by the Department of Health and Ageing as has occurred with mental health services. It is suggested a scoping exercise would need to be undertaken should the mandating of the measures be considered in the future.

Section 12.6.3 and Figure 6 provide a discussion of the appropriate application of each of the recommended measures for different service settings and for different stages of dementia. This is supplemented by a discussion of a staged approach to assessment in Section 12.6.4. and the potential for the adoption of a tiered assessment approach is also discussed.

A dissemination strategy, to facilitate the adoption of the recommended tools has been outlined in Section 12.6.6. While this could include the development of an instrument toolkit, presentations

and training workshops at conferences and the development of web materials, brochures, journal articles:

- It is recommended that a dissemination strategy project be undertaken to facilitate the dissemination and uptake of findings from this report.

Notwithstanding the above, the provision of more formal education and training will also be of paramount importance:

- It is recommended that a project be sponsored to a) ascertain coverage of assessment and the use of recommended tools in current curricula and b) to develop appropriate education modules for insertion in the training curricula of relevant professional and paraprofessional groups.

### 13.4.1 Identified Research Gaps

During the course of the project a number of research gaps have been identified. These are outlined below in a summary form and more detailed information is provided in the relevant Sections of the report. Priority rankings as outlined in section 12.6.7 are also provided. These are presented for consideration by the Department:

- Some measures need pilot testing in Australia to obtain reference data (e.g. Dementia Specific HRQOL measures, associated symptom measures) (refer Sections 5 and 9). (Priority 1)
- Further research work is required to assess the point at which people can no longer self-rate (e.g. MMSE-3MS score) under different modes of administration (e.g. self report, interview, interview assisted). This will be required for all recommended self-report instruments (refer Section 12.2). (Priority 1)
- Future research might also address how training, the framing of questions, the terminology used and the administration of instruments influences the results of proxy assessment and more research is needed to compare proxy reports with performance-based measures as well as information from medical records and health utilization data (refer Section 12.3). (Priority 2)
- Many of the recent papers on proxy assessment use single or dual item informant measures (e.g. Tierney, et al. 2003; Watson, et al. 2004; Li, et al. 2006). Further research is required to ascertain whether these items have the requisite accuracy compared to longer proxy measures (refer Section 12.3). (Priority 2)
- Further research activities are required to address identified problems with Multi-attribute Utility measures: AQoL (shorten) and/or EQ-5D (scoring and distribution issues) (refer Section 7). (Priority 2)
- A linguistic validation study be undertaken for the De Jong Gierveld Loneliness Scale to develop response categories more appropriate to the Australian context (refer Section 8). (Priority 1)
- It is recommended that the three social isolation instruments which performed relatively well (the De Jong Gierveld Loneliness Scale, the Friendship Scale and the Medical Outcomes Study Social Support Survey) be trialled in at least one large dementia study for the explicit purpose of identifying the instrument to be recommended for future use. This trial should also assess proposed modifications to these instruments. From this study a statistically-derived single item measure could also be identified for use in everyday clinical consultations (refer Section 8). (Priority 2)
- Social function / social support areas may need follow up research if there is a wish to focus on social participation as well as social isolation but this could be combined with the

recommendation above. It might also include an examination of social support items from relevant ABS Surveys (refer Section 8). (Priority 2)

- In the absence of a research consensus for the measurement of function in dementia, and given a high degree of overlap in items, there is a clear need for a streamlining the various functional instruments and items across each of the practice settings (Spector, 1997). A study including a large group of dementia patients could examine and calibrate functional items from the short-listed functional status instruments (both generic and dementia specific) to create a comprehensive item bank. This dementia item bank could then be used to examine item redundancy and coverage across the range of severity levels and could be used to develop new tools or provide cross-calibration between the existing instruments. This project would also need to examine the relationship of these items with the recommended cognitive and functional assessment staging instruments (refer Section 10). (Priority 1)
- It is recommended that a study be undertaken to test the Short Assessment of Patient Satisfaction scale and the two global patient satisfaction items identified with samples of people with dementia and their carers. It is noted that all these items would require minor rewording to make them suitable for use with an informant/carers (refer Section 11.1) (Priority 1)
- Carer satisfaction with services has been addressed in this project but an examination of the instruments used to assess carer burden, carer appraisal and carer wellbeing was outside the scope of this project. It is recommended that a more detailed follow up project be undertaken to examine issues relating to the assessment of instruments used to assess carer burden, carer appraisal and carer well-being (refer Section 11.2). (Priority 2)
- A further project is necessary to ensure the development of a more comprehensive database of dementia outcome measures solely for use with CALD communities - where translated versions of the DOMS selected measures are further reviewed and made available if possible (refer Section 12.4). (Priority 1)
- Further studies analysing the measurement equivalence of the core recommended measures be undertaken for major language groups within Australia (refer Section 12.4). (Priority 2)
- Research be undertaken to further examine new instruments developed in Australia such as the RUDAS, KICA-Cog and the GPCOG to ensure their validity and reliability in different groups of CALD populations (refer Section 12.4). (Priority 1)
- There needs to be further detailed research on the meaning of dementia in Indigenous communities, and how to ask questions which capture the experience and limitations of living with dementia in an Indigenous community (refer Section 12.5). (Priority 1)
- Further research be undertaken to adapt the recommended tools, as necessary, for Aboriginal and Torres Strait Islander Groups – particularly for rural and remote populations (refer Section 12.5). (Priority 1)
- It is recommended that further research be undertaken to assess the psychometric properties of the KICA-Cog and its' appropriateness for the assessment of cognitive impairment with both urban and remote Indigenous people (refer Section 12.5). (Priority 1)
- Individual assessment methods such as Goal Attainment Scaling, recently advocated by Rockwood (2007), to individualise outcome measurement for people with dementia, have not been examined in this project. It is recommended that a systematic review be undertaken to assess these methods. (Priority 2)
- A decision was made by the DOMS-EMG that the project should focus on the instruments/tools that are available for use in routine care and this would exclude many of the more detailed neuropsychological instruments or instruments that require specialist training for their administration and interpretation. It is recommended that a further study could examine and review neuropsychological and specialist tests for people with dementia, in association with the relevant professional groups. (Priority 2)

- That further investigation in the field be undertaken to fine tune the recommendations for use of these instruments with people with Late Stage Dementia. This could also include an examination of instruments used in palliative care settings. (Priority 2)
- There is a need for training on the principles of undertaking systematic assessment as well as the specific characteristics of a recommended instrument. It is thought that a project to develop modules addressing these issues, which can be incorporated in the curricula for relevant professional and para-professional groups, should be undertaken (refer Section 12.6). (Priority 1)
- Issues concerning safety/ risk assessment are outside the scope of this project. It is recommended that a further project be undertaken to examine risk assessment issues (e.g. elder abuse, aggression, self harm, etc) for people with dementia. (Priority 2)

### 13.5 Conclusion

While further research may need to be undertaken to clarify some assessment issues (as indicated above) this report provides a useful review of the best measures to assess the status and symptoms of people with dementia. The project has identified a set of recommended measures/tools for routine use in the assessment, diagnosis, screening and outcomes monitoring of dementia conditions and the evaluation of treatments that are applicable for the Australian health care context. By developing this set of recommended measures it is hoped to standardise the assessment and evaluation procedures used in this field to enhance comparability of findings across research and practice settings.

The DOMS-NEP and the DOMS-EMG have provided considerable advice throughout the project concerning these areas of assessment. These recommendations were ratified by DOMS-NEP and DOMS-EMG at their final meetings on the 17<sup>th</sup> August 2007.

### References

- Brodsky H, Green A, Banerjee S, Mittelman M, Schulz R, Whitehouse P, Harvey R, Powell M, Prince M, Rios D, Zarit S (2002) *Towards harmonisation of Caregiver Outcome Measures*. Brain Aging. Vol. 2, No. 4, pp.3-12.
- Grypma R, Mahajani S and Tam E (2007) *Screening and diagnostic assessment of non-english speaking people with dementia: Guidelines and system recommendations for practitioners, service managers and policy makers*. Alzheimer's Australia.
- Harper GJ (2000) *Assessing older adults who cannot communicate*. In Kane RL, Kane RA, (eds). Assessing older persons: Measures, meaning, and practical applications. Oxford University Press, Oxford.
- Hawthorne G, Sansoni J, Hayes L, et al. (2006) *Measuring Patient Satisfaction with Incontinence Treatment*. Department of Psychiatry, The University of Melbourne, and the Centre for Health Service Development, University of Wollongong for the Commonwealth Department of Health and Ageing.
- Li M, Ng TP, Kua EH, et al. (2006) *Brief Informant Screening Test for Mild Cognitive Impairment and Early Alzheimer's Disease*. Dementia and Geriatric Cognitive Disorders. Vol. 21, No.5-6, pp.392-402.
- McKeith IG, Dickson DW, Lowe J, Emre M, O'Brien JT, Feldman H, et al. (2005) *Diagnosis and management of dementia with Lewy bodies: third report of the DLB Consortium*. [erratum appears in *Neurology*. 2005 Dec 27; 65(12):1992] [summary for patients in *Neurology*. 2005 Dec 27; 65(12):E26-7; PMID: 16380603]. Neurology. Vol. 65, No.12, pp.1863-1872.
- Ramsay L, Owen A, Eagar K, Marosszeky N, Samsa P, Rix M, Fildes D, Willmott L (2006) *An Australian Carer Needs Assessment Instrument: The Recommended Approach*. Centre for Health Service Development, University of Wollongong.

Rockwood K (2007) *The measuring, meaning and importance of activities of daily living (ADLs) as an outcome.* International Psychogeriatrics. Vol. 19, No.3, pp.467-482.

Roman GC, Tatemichi TK, Erkinjuntti T, Cummings JL, Masdeu JC, Garcia JH, et al. (1993) *Vascular dementia: diagnostic criteria for research studies. Report of the NINDS-AIREN International Workshop. [see comment].* Neurology. Vol. 43, No.2, pp.250-260.

Snow AL, Cook KF, Lin P-S, et al. (2005a) *Proxies and other external raters: Methodological considerations.* Health Services Research. Vol. 40, No.5, pp.1676-1693.

Spector WD (1997) *Measuring functioning in daily activities for persons with dementia.* Alzheimer Disease and Associated Disorders. Vol. 11, No. Suppl 6, pp.81-90.

Tierney MC, Herrmann N, Geslani DM, et al. (2003) *Contribution of Informant and Patient Ratings to the Accuracy of the Mini-Mental State Examination in Predicting Probable Alzheimer's Disease.* Journal of the American Geriatrics Society. Vol. 51, No.6, pp.813-818.

Watson LC, Lewis CL and Fillenbaum GG (2004) *Asking family about memory loss: Is it helpful ?* J General Internal Medicine. Vol. 20, pp.28-32.

## 14 Report Reference List

Aalten P, van Valen E, de Vugt ME, et al. (2006) *Awareness and behavioral problems in dementia patients: a prospective study*. International Psychogeriatrics. Vol. 18, No.1, pp.3-17.

Abdellah FG and Levine E (1957) *Developing a measure of patient and personnel satisfaction with nursing care*. Nurs Res. Vol. 5, No.3, pp.100-8.

Abbott S, Meyer J, Copperman J, et al. (2005) *Quality criteria for patient advice and liaison services: what do patients and the public want?* Health Expect. Vol. 8, No.2, pp.126-37.

Access Economics (2006) *Dementia prevalence and incidence among people who do not speak English at home*. Access Economics.

Adams KB, Sanders S and Auth EA (2004) *Loneliness and depression in independent living retirement communities: risk and resilience factors*. Ageing and Mental Health. Vol. 8, No.6, pp.475-85.

Addington-Hall J and Kalra L (2001) *Who should measure quality of life?* BMJ. Vol. 322, No.7299, pp.1417-20.

Addington-Hall J (1998) *VOICES (Views of Information Carers - Evaluation of Services)*. King's College School of Medicine and Dentistry.

Addington-Hall J and McPherson C (2001) *After-death interviews with surrogates/bereaved family members: some issues of validity*. Journal of Pain and Symptom Management. Vol. 22, No.3, pp.784-90.

Adshead F, Cody DD and Pitt B (1992) *BASDEC: a novel screening instrument for depression in elderly medical inpatients*. [see comment]. BMJ. Vol. 305, No.6850, pp.397.

Aggarwal N, Vass AA, Minardi HA, et al. (2003) *People with dementia and their relatives: personal experiences of Alzheimer's and of the provision of care*. Journal of Psychiatric and Mental Health Nursing. Vol. 10, No.2, pp.187-97.

AIHW (2004) *Australia's Health 2004: The Ninth Biennial Health Report of the Australian Institute of Health and Welfare*. Australian Institute of Health and Welfare. Report nr AIHW Cat AUS44. 544 p.

Aisen PS, Schafer KA, Grundman M, et al. (2003) *Effects of rofecoxib or naproxen vs. placebo on Alzheimer disease progression: a randomized controlled trial*. [see comment]. JAMA. Vol. 289, No.21, pp.2819-26.

Ajzen I and Fishbein M (1980) *Understanding Attitudes and Predicting Social Behaviour*. Prentice-Hall, Englewood Cliffs.

Albert M and Cohen C (1992) *The Test for Severe Impairment: An instrument for the assessment of patients with severe cognitive dysfunction*. Journal of the American Geriatrics Society. Vol. 40, pp.449-453.

Albert MS, Levkoff SE, Reilly C, Liptzin B, Pilgrim D, Cleary PD, et al. (1992) *The delirium symptom interview: an interview for the detection of delirium symptoms in hospitalized patients*. Journal of Geriatric Psychiatry and Neurology. Vol. 5, No.1, pp.14-21.

Albert SM, Del Castillo-Castaneda C, Sano M, Jacobs DM, Marder K, Bell K, Bylsma F, Lafleche G, Brandt J, Albert M and Stern Y (1996) *Quality of life in patients with Alzheimer's disease as reported by patient proxies*. Journal of the American Geriatrics Society. Vol. 44, No.11, pp.1342-7.

Alemi F, Stephens R, Llorens S, et al. (2003) *The Orientation of Social Support measure*. Addictive Behaviors. Vol. 28, No.7, pp.1285-98.

Alexopoulos GS, Abrams RC, Young RC, et al. (1988) *Cornell scale for depression in dementia*. Biological Psychiatry. Vol. 23, No.3, pp.271-284.

- Allen NH, Gordon S, Hope T and Burns A (1996) *Manchester and Oxford Universities Scale for the Psychopathological Assessment of Dementia (MOUSEPAD)*. British Journal of Psychiatry. Vol. 169, No.3, pp.293-307.
- Alpass FM and Neville S (2003) *Loneliness, health and depression in older males*. Aging and Mental Health. Vol. 7, No.3, pp.212-216.
- Alwahhabi F (2003) *Anxiety symptoms and generalized anxiety disorder in the elderly: a review*. Harvard Review of Psychiatry. Vol. 11, No.4, pp.180-193.
- Alzheimer's Australia (2004) *Dementia Terminology Framework*. Position Paper 4, available at [http://www.alzheimers.org.au/upload/Terminology\\_Position\\_Paper\\_041.pdf](http://www.alzheimers.org.au/upload/Terminology_Position_Paper_041.pdf), accessed 1 June 2006.
- Alzheimer's Disease International (2003) *Behavioural and psychological symptoms of dementia*. 10/66 Pilot Studies. Findings. p.1, available at [www.alz.co.uk/1066/research/pilot/bpsd1.html](http://www.alz.co.uk/1066/research/pilot/bpsd1.html), accessed 14 March 2007.
- American Educational Research Association, American Psychological Association, and the National Council on Measurement in Education (1999) *Standards for educational and psychological testing*. American Educational Research Association, American Psychological Association, National Council on Measurement in Education, Washington DC.
- American Psychiatric Association (1987) *Diagnostic and statistical manual of mental disorders, third edition, revised: DSM-III-R*. American Psychiatric Association, Washington DC.
- American Psychiatric Association (1994) *Diagnostic and Statistical Manual of Mental Disorders DSM-IV (4th ed.)* American Psychiatric Association, Washington DC.
- American Psychiatric Association (1998) *Guidelines for the Evaluation of Dementia and Age-Related Cognitive Decline*, available at <http://www.apa.org/practice/dementia.html>, accessed 10 May 2006.
- American Psychiatric Association (1999) *Practice Guidelines for the treatment of patients with Alzheimer's disease and other dementias of later life*. American Journal of Psychiatry. Vol. 154, No.5, pp.1-39.
- American Psychiatric Association (2000) *Diagnostic and statistical manual of mental disorders, fourth edition, text revision*. American Psychiatric Association, Washington DC.
- Andersen CK, Wittrup-Jensen KU, Lolk A, Andersen K and Kragh-Sorensen P (2004) *Ability to perform activities of daily living is the main factor affecting quality of life in patients with dementia*. Health and Quality of Life Outcomes. Vol. 2, No.52.
- Andersen L, Kristiansen I, Falch J and Aursnes I (1995) *Cost-effectiveness of Alendronate for the prevention of osteoporotic fractures in Norwegian women*. Statens Institute for Folkehelse, Folkehelse, Oslo.
- Anderson J, Kaplan R, Berry C, Bush J and Ruben R (1989) *Interday reliability of function assessment for a health status measure: The Quality of Well-Being scale*. Medical Care. Vol. 27, No.11, pp.1076-1084.
- Andresen EM, Rothenberg BM and Kaplan RM (1998) *Performance of a self-administered mailed version of the Quality of Well-Being (QWB-SA) questionnaire among older adults*. Medical Care. Vol. 36, No.9, pp.1349-60.
- Andrews G, Peters L and Teeson M (1994) *The Measurement of Consumer Outcome in Mental Health: A Report to the National Mental Health Information Strategy Committee*. Clinical Research Unit for Anxiety Disorders, Sydney.
- Ankri J, Beaufils B, Novella JL, Morrone I, Guillemin F, Jolly D, Ploton L and Blanchard F (2003) *Use of the EQ-5D among patients suffering from dementia*. Journal of Clinical Epidemiology. Vol. 56, No.11, pp.1055-63.
- Antaki C (1999) *Interviewing persons with a learning disability: How setting lower standards may inflate well-being scores*. Qualitative Health Research. Vol. 9, No.4, pp.437-454.

- Antaki C and Rapley M (1996) *"Quality of Life" talk: The liberal paradox of psychological testing.* Discourse and Society. Vol. 7, No.3, pp.293-316.
- Applebaum F (1978) *Loneliness: a taxonomy and psychodynamic view.* Clinical Social Work Journal. Vol. 6, No.1, pp.13-20.
- Aragon SJ and Gesell SB (2003) *A patient satisfaction theory and its robustness across gender in emergency departments: a multigroup structural equation modeling investigation.* American Journal of Medical Quality. Vol. 18, No.6, pp.229-41.
- Aspinal F, Addington-Hall J, Hughes R, et al. (2003) *Using satisfaction to measure the quality of palliative care: a review of the literature.* Journal of Advanced Nursing. Vol. 42, No.4, pp.324-39.
- Asthill LF (2004) *Staff training and challenging behaviour in a day hospital.* Dementia: The International Journal of Social Research and Practice. Vol. 3, No.3, pp.384-392.
- Attkisson CC and Zwick R (1982) *The client satisfaction questionnaire. Psychometric properties and correlations with service utilization and psychotherapy outcome.* Evaluation and Program Planning. Vol. 5, No.3, pp.233-7.
- Auer S and Reisberg B (1996) *Reliability of the Modified Ordinal Scales of Psychological Development: a cognitive assessment batter for severe dementia.* International Psychogeriatrics. Vol. 8, No.2, pp. 225-31.
- Auer SR, Montreiro IM et al. (1996) *The Empirical Behavioral Pathology in Alzheimer's disease (E-BEHAVE-AD) Rating Scale.* International Psychogeriatrics. Vol. 8, No.2, pp. 247-66.
- Auer S, Hampel H, et al. (2000) *Translations of measurements and scales: Opportunities and diversities.* International Psychogeriatrics. Vol. 12, Suppl.1, pp. 391-394.
- Austin BA (1983) *Factorial structure of the UCLA Loneliness Scale.* Psychological Reports. Vol. 53, No.3, Pt 1, pp.883-9.
- Australian Bureau of Statistics (1997) *1995 National Health Survey: SF-36 Population Norms.* Australia. Cat No 4399.0. ABS, Canberra.
- Australian Bureau of Statistics (1998) *Mental Health and Wellbeing: Profile of Adults Australia 1997.* Australia, Cat No 4326.0 ABS, Canberra.
- Australian Institute of Health and Welfare (2003) *ICF Australian User Guide Version 1.0.* Disability Series (No. AIHW Cat. No. DIS 33.) AIHW, Canberra.
- Australian Institute of Health and Welfare (AIHW) (2007) *Final report – Australian Dementia Data Analysis and Development* (In press) AIHW, Canberra.
- Australian Pain Society (2004) *Residential Aged Care Pain Management Guidelines.*
- Awad G, Hogan T, Voruganti L and Heslegrave R (1995) *Patients' subjective experiences on antipsychotic medications: implications for outcome and quality of life.* International Clinical Psychopharmacology. Vol. 10S, pp.123-132.
- Awad G and Voruganti P (1999) *Cost-utility analysis in schizophrenia.* Journal of Clinical Psychiatry. Vol. 60, Suppl.3, pp.22-28.
- Ayis SA, Bowling A, Goberman-Hill R, et al. (2007) *The effect of definitions of activities of daily living on estimates of changing ability among older people.* International Journal of Rehabilitation Research. Vol. 30, No.1, pp.39-46.
- Baker R (1990) *Development of a questionnaire to assess patients' satisfaction with consultations in general practice.* British Journal of General Practice. Vol. 40, No.December, pp.487-490.
- Baker R (1991) *The reliability and criterion validity of a measure of patients' satisfaction with their general practice.* Family Practice. Vol. 8, No.2, pp.171-177.

Baker R, Mainous AG, 3rd, Gray DP, et al. (2003) *Exploration of the relationship between continuity, trust in regular doctors and patient satisfaction with consultations with family doctors*. Scandinavian Journal of Primary Health Care. Vol. 21, No.1, pp.27-32.

Baker R and Whitfield M (1992) *Measuring patient satisfaction: a test of construct validity*. Quality in Health Care. Vol. 1, pp.104-109.

Bakker T, Duivenvoorden H, van der Lee J, Schudel W (2004) *"Life expectancy following Psychogeriatric Reactivation" Identification of Prognostic Characteristics of Survival Assessed on Admission*. Dementia and Geriatric Cognitive Disorders. Vol. 18, pp.24-31.

Bal S, Crombez G, Van Oost P, et al. (2003) *The role of social support in well-being and coping with self-reported stressful events in adolescents*. Child Abuse and Neglect. Vol. 27, No.12, pp.1377-95.

Baladi J, Bailey P, et al. (2000) *Effectiveness of Rivastigmine for Alzheimer's disease: Canadian Interpretation of intermediate outcome measures and cost implications*. Clinical Therapeutics. Vol. 22, No.12, pp.1549-61.

Balandin S, Berg N and Waller A (2006) *Assessing the loneliness of older people with cerebral palsy*. Disability and Rehabilitation. Vol. 28, No.8, pp.469-79.

Ballard C, O'Brien J, James I, et al. (2001) *Quality of life for people with dementia living in residential and nursing home care: the impact of performance on activities of daily living, behavioral and psychological symptoms, language skills, and psychotropic drugs* International Psychogeriatrics. Vol. 13, No.1, pp.93-106.

Bamford CE and Bruce E (2000) *Defining the outcomes of community care: the perspectives of older people with dementia and their carers*. Ageing and Society. Vol. 20, pp.543-570.

Banerjee S, Smith SC, Lamping DL, et al. (2006) *Quality of life in dementia: more than just cognition. An analysis of associations with quality of life in dementia*. Journal of Neurology, Neurosurgery and Psychiatry. Vol. 77, No.2, pp.146-8.

Banks MR and Banks WA (2002) *The effects of animal-assisted therapy on loneliness in an elderly population in long-term care facilities*. J Gerontol A Biol Sci Med Sci. Vol. 57, No.7, pp.M428-32.

Baro E, Ferrer M, Vazquez O, et al. (2006) *Using the Nottingham Health Profile (NHP) among older adult inpatients with varying cognitive function*. Quality of Life Research. Vol. 15, No.4, pp.575-85.

Barron CR, Foxall MJ, Von Dollen K, et al. (1994) *Marital status, social support, and loneliness in visually impaired elderly people*. Journal of Advanced Nursing. Vol. 19, No.2, pp.272-80.

Bartlett H, Rao DV and Warburton J (2006) *Ageing and cultural diversity in Queensland: Working together to make a difference*. Australasian Centre on Ageing, The University of Queensland.

Barton GR, Bankart J and Davis AC (2005) *A comparison of the quality of life of hearing-impaired people as estimated by three different utility measures*. International Journal of Audiology. Vol. 44, No.3, pp.157-63.

Barton GR, Bankart J, Davis AC and Summerfield QA (2004) *Comparing utility scores before and after hearing-aid provision: results according to the EQ-5D, HUI3 and SF-6D*. Applied Health Economics and Health Policy. Vol. 3, No.2, pp.103-5.

Bashshur RL, Metzner CA and Worden C (1967) *Consumer satisfaction with group practice, the CHA case*. Am J Public Health Nations Health. Vol. 57, No.11, pp.1991-9.

Bates EW and Brown BS (1988) *Geriatric care needs and HMO technology. A theoretical analysis and initial findings from the National Medicare Competition Evaluation*. Medical Care. Vol. 26, No.5, pp.488-98.

Bathareethan M and Shah A (2000) *Aggressive behaviour, staff attitude and staff perception of patients on two continuing care psychogeriatric wards*. Ageing and Mental Health. Vol. 4, No.1, pp. 66-71.

- Bauld L, Chesterman J and Judge K (2000) *Measuring satisfaction with social care amongst older service users: issues from the literature*. Health and Social Care in the Community. Vol. 8, No.5, pp.316-324.
- Baumeister RF and Leary MR (1995) *The need to belong: desire for interpersonal attachments as a fundamental human motivation*. Psychological Bulletin. Vol. 117, No.3, pp.497-529.
- Baumgarten M, Becker R, et al. (1990) *Validity and reliability of the dementia behavior disturbance scale*. Journal of the American Geriatrics Society. Vol. 38, No.3, pp.221-6.
- Baxter LC, Sparks DL, Johnson SC, et al. (2006) *Relationship of cognitive measures and gray and white matter in Alzheimer's disease*. Journal of Alzheimer's Disease. Vol. 9, No.3, pp.253-60.
- Becchi A, Rucci P, Placentino A, Neri G and de Girolamo G (2004) *Quality of life in patients with schizophrenia--comparison of self-report and proxy assessments*. Social Psychiatry and Psychiatric Epidemiology. Vol. 39, No.5, pp.397-401.
- Beck AT, Epstein N, Brown G, et al. (1988) *An inventory for measuring clinical anxiety: psychometric properties*. J Consult Clin Psychol. Vol. 56, No.6, pp.893-7.
- Beck AT, Ward CH, Mendelson M, et al. (1961) *An inventory for measuring depression*. Archives of General Psychiatry. Vol. 4, pp.561-71.
- Beck EJ, Griffith R, Fitzpatrick R, et al. (1999) *Patient satisfaction with HIV service provision in NPMS hospitals: the development of a standard satisfaction questionnaire*. NPMS Steering Group. AIDS Care. Vol. 11, No.3, pp.331-43.
- Behl P, Lanctot KL, Streiner DL, et al. (2006) *Cholinesterase inhibitors slow decline in executive functions, rather than memory, in Alzheimer's disease: a 1 year observational study in the Sunnybrook dementia cohort*. Current Alzheimer Research. Vol. 3, No.2, pp.147-156.
- Bekelman DB, Black BS, Shore AD, et al. (2005) *Hospice care in a cohort of elders with dementia and mild cognitive impairment*. Journal of Pain and Symptom Management. Vol. 30, No.3, pp.208-14.
- Bell CM, Araki SS and Neumann PJ (2001) *The association between caregiver burden and caregiver health-related quality of life in Alzheimer disease*. Alzheimer Disease and Associated Disorders. Vol. 15, No.3, pp.129-36.
- Bennett K and Torrance G (1996) *Measuring health state preferences and utilities: Rating scale, time trade-off, and standard gamble techniques*. In Spilker, B (ed.) (2004) Quality of Life and Pharmacoeconomics in Clinical Trials. Lippincott-Raven Publishers, Philadelphia, pp. 253-265. *title*
- Ben-Sira Z (1976) *The function of the professional's affective behaviour in client satisfaction: a revised approach to social interaction theory*. Journal of Health and Social Behavior. Vol. 17, pp.3-11.
- Berg L, McKeel DW, Jr., Miller JP, et al. (1993) *Neuropathological indexes of Alzheimer's disease in demented and nondemented persons aged 80 years and older*. Archives of Neurology. Vol. 50, No.4, pp.349-58.
- Berg L, McKeel DWJ, Miller JP, et al. (1998) *Clinicopathologic studies in cognitively healthy aging and Alzheimer's disease: relation of histologic markers to dementia severity, age, sex, and apolipoprotein E genotype*. Archives of Neurology. Vol. 55, No.3, pp.326-35.
- Berg L, Miller JP, Storandt M, et al. (1988) *Mild senile dementia of the Alzheimer type: 2. Longitudinal assessment*. Annals of Neurology. Vol. 23, No.5, pp.477-84.
- Berger EY (1980) *A system for rating the severity of senility*. Journal of the American Geriatrics Society. Vol. 28, No.5, pp.234-236.
- Berkman L and Syme L (1979) *Social networks, host resistance and mortality: a nine-year follow-up study of Alameda County residents*. American Journal of Epidemiology. Vol. 109, No.2, pp.186-204.

- 
- Bergman-Evans B (2004) *Beyond the basics. Effects of the Eden Alternative model on quality of life issues.* Journal of Gerontological Nursing. Vol. 30, No.6, pp.27-34.
- Bergner M, Bobbitt RA, Kressel S (1976) *The Sickness Impact Profile: conceptual formulation and methodology for the development of a health status measure.* Int J Health Serv. Vol. 6, pp. 393-415.
- Bergner M, Bobbitt RA, Carter WB and Gilson BS (1981) *The Sickness Impact Profile: development and final revision of a health status measure.* Med Care. Vol. 19, No.8, pp.787-805.
- Bindman J, Davies S, Taylor R, et al. (1996) *Developments in mental health policy in the United Kingdom.* Epidemiologia e Psichiatria Sociale. Vol. 5, No.2, pp.87-91.
- Birrell R and Hawthorne L (2001) *Doctor dilemmas: how GP medicine is practised in Australia.* People and Place. Vol. 9, No.3, pp.41-53.
- Black BS, Rabins PV and Kasper JD (2000) *Alzheimer Disease Related Quality of Life User's manual*, : DEMeasure, Baltimore, Maryland.
- Blake H and Lincoln NB (2000) *Factors associated with strain in co-resident spouses of patients following stroke.* Clinical Rehabilitation. Vol. 14, No.3, pp.307-14.
- Bleichrodt H and Johannesson M (1997) *An experimental test of a theoretical foundation for rating-scale valuations.* Medical Decision Making. Vol. 17, No.2, pp.208-16.
- Blessed G, Tomlinson BE and Roth M (1968) *The association between quantitative measures of dementia and of senile change in cerebral grey matter of elderly subjects.* British Journal of Psychiatry. Vol. 114, pp.797-811.
- Blomfeldt R, Tornkvist H, Ponzer S, Soderqvist A and Tidermark J (2005) *Internal fixation versus hemiarthroplasty for displaced fractures of the femoral neck in elderly patients with severe cognitive impairment.* J Bone Joint Surg Br. Vol. 87, No.4, pp.523-9.
- Blumel JE, Castelo-Branco C, Cancelo MJ, et al. (2004) *Relationship between psychological complaints and vasomotor symptoms during climacteric.* Maturitas. Vol. 49, No.3, pp.205-10.
- Boada M, Cejudo JC, Tarraga L, et al. (2002) *Neuropsychiatric Inventory Questionnaire (NPI-Q): Spanish validation of an abridged form of the Neuropsychiatric Inventory (NPI).* Neurologica. Vol. 17, No.6, pp.317-323.
- Boada M, Tarraga L, et al. (2006) *Behavioral pathology in Alzheimer's disease Rating Scale (BEHAVE-AD): Spanish validation.* Neurologia. Vol. 21, No.1, pp.19-25.
- Bohner H, Friedrichs R, Habel U, Muller EE, Sandmann W and Schneider F (2003) *Delirium increases morbidity and length of stay after vascular surgery operations. Results of a prospective study.* Chirurg. Vol. 74, No.10, pp. 931-936.
- Bombardier C and Raboud J (1991) *A comparison of health-related quality-of-life measures for rheumatoid arthritis research. The Auranofin Cooperating Group.* Controlled Clinical Trials. Vol. 12, No.4 Suppl, pp.243S-256S.
- Bootsma-van der Wiel A, Gussekloo J, de Craen AJM, et al. (2001) *Disability in the oldest old: "Can do" or "Do do".* JAGS. Vol. 49, No.7, pp.909-914.
- Borson S, Brush M, Gil E, et al. (1999) *The Clock Drawing Test: utility for dementia detection in multiethnic elders.* Journals of Gerontology Series A-Biological Sciences and Medical Sciences. Vol. 54, No.11, pp.M534-40.
- Borson S, Scanlan JM, Brush M, et al. (2000) *The Min-Cog: A cognitive 'vital signs' measure for dementia screening in multi-lingual elderly.* International Journal of Geriatric Psychiatry. Vol. 15, pp.1021-1027.
-

- 
- Borson S, Scanlan JM, Watanabe J, et al. (2006) *Improving identification of cognitive impairment in primary care*. International Journal of Geriatric Psychiatry. Vol. 21, No.4, pp.349-355.
- Botwinick J, Storandt M and Berg L (1986) *A longitudinal, behavioral study of senile dementia of the Alzheimer type*. Archives of Neurology. Vol. 43, No.11, pp.1124-7.
- Botwinick J, Storandt M, Berg L, et al. (1988) *Senile dementia of the Alzheimer type. Subject attrition and testability in research*. Archives of Neurology. Vol. 45, No.5, pp.493-6.
- Boustani M, Peterson B, Hanson L, Harris R and Lohr K (2003) *Summary of the Evidence: Screening for dementia in primary care*, available at <http://www.ahcpr.gov/clinic/3rduspstf/dementia/demsum2.htm>, accessed 21 June 2006.
- Bova C (2001) *Adjustment to chronic illness among HIV-infected women*. Journal of Nursing Scholarship. Vol. 33, No.3, pp.217-23.
- Bovier PA, Chamot E and Perneger TV (2004) *Perceived stress, internal resources, and social support as determinants of mental health among young adults*. Quality of Life Research. Vol. 13, No.1, pp.161-70.
- Bowers BJ (1988) *Family perceptions of care in a nursing home*. The Gerontologist. Vol. 28, No.3, pp.361-8.
- Bowlby J (1971) *Attachment*. Pelican, London.
- Bowling A (2001) *Measuring disease (2nd ed.)* Open University Press, Buckingham England.
- Bowling A (2005) *Measuring health: a review of quality of life measurement scales (3rd ed.)* Open University Press, Berkshire, England.
- Boyle MH, Furlong W, Feeny D, Torrance GW and Hatcher J (1995) *Reliability of the Health Utilities Index - Mark III used in the 1991 cycle 6 Canadian General Social Survey Health Questionnaire*. Quality of Life Research. Vol. 4, pp.249-257.
- Bradlyn AS, Harris CV, Warner JE, Ritchey AK and Zaboy K (1993) *An investigation of the validity of the quality of Well-Being Scale with paediatric oncology patients*. Health Psychology. Vol. 12, No.3, pp.246-50.
- Brady MJ, Peterman AH, Fitchett G, et al. (1999) *A case for including spirituality in quality of life measurement in oncology*. Psycho-Oncology. Vol. 8, No.5, pp.417-28.
- Bravo G and Hebert R (1997) *Age and education-specific reference values for the Mini-Mental and modified Mini-Mental State Examinations derived from a non-demented elderly population*. International Journal of Geriatric Psychiatry. Vol. 12, No.10, pp.1008-1018.
- Braithwaite V (1992) *Caregiving Burden: Making the Concept Scientifically Useful and Policy Relevant*. Research on Aging. Vol. 14, No.1, pp.3-27.
- Brandt J, Spencer M, Folstein M (1988) *The telephone interview for cognitive status*. Neuropsychiatry Neuropsychol Behav Neurol. Vol. 1, pp.111-117.
- Brazier J and Deverill M (1999) *A checklist for judging preference-based measures of health related quality of life: learning from psychometrics*. Health Econ. Vol. 8, No.1, pp.41-51.
- Brazier J and Deverill M, Green C, Harper R and Booth A (1999) *A review of the use of health status measures in economic evaluation*. Health Technology Assessment. Vol. 3, No.9, pp.1-165.
- Brazier J, Jones N and Kind P (1993) *Testing the validity of the EuroQoL and comparing it with the SF-36 health survey questionnaire*. Quality of Life Research. Vol. 2, pp.169-180.
- Brazier J, Roberts J and Deverill M (2002) *The estimation of a preference-based measure of health from the SF-36*. Journal of Health Economics. Vol. 21, pp.271-292.
-

- 
- Brazier J, Roberts J, Tsuchiya A and Busschbach J (2004) *A comparison of the EQ-5D and SF-6D across seven patient groups*. Health Economics. Vol. 13, No.9, pp.873-84.
- Brazier J, Usherwood T, Harper R and Thomas K (1998) *Deriving a preference-based single index from the UK SF-36 Health Survey*. Journal of Clinical Epidemiology. Vol. 51, No.11, pp.1115-1128.
- Brazier JE and Roberts J (2004) *The estimation of a preference-based measure of health from the SF-12*. Medical Care. Vol. 42, No.9, pp.851-9.
- Breitbart W, Rosenfeld B, Roth A, Smith MJ, Cohen K and Passik S (1997) *The Memorial Delirium Assessment Scale*. Journal of Pain and Symptom Management. Vol. 13, No.3, pp.128-137.
- Broadhead W (1988) *Duke-UNC functional social support questionnaire (DUFSS): instructions for use*. Duke University Medical Centre. Report nr 919-286-9896.
- Broadhead WE, Gehlbach SH, de Gruy FV, et al.(1988) *The Duke-UNC Functional Social Support Questionnaire. Measurement of social support in family medicine patients*. Medical Care. Vol. 26, No.7, pp.709-23.
- Broadhead WE, Gehlbach SH, de Gruy FV, et al. (1989) *Functional versus structural social support and health care utilization in a family medicine outpatient practice*. Medical Care. Vol. 27, No.3, pp.221-33.
- Broadhead WE, Kaplan BH, James SA, et al (1983) *The epidemiologic evidence for a relationship between social support and health*. American Journal of Epidemiology. Vol. 117, No.5, pp.521-37.
- Brod M, Stewart AL, Sands L and Walton P (1999) *Conceptualization and measurement of quality of life in dementia: the dementia quality of life instrument (DQoL)*. Gerontologist. Vol. 39, No.1, pp.25-35.
- Brodaty H, Pond D, Kemp NM, et al. (2000) *The GPCOG: A new screening test for dementia designed for general practice*. Journal of the American Geriatrics Society. Vol. 50, pp.530-534.
- Brodaty H, Green A, Banerjee S, Mittelman M, Schulz R, Whitehouse P, Harvey R, Powell M, Prince M, Rios D, Zarit S (2002) *Towards harmonisation of Caregiver Outcome Measures*. Brain Aging. Vol. 2, No. 4, pp.3-12.
- Brodaty H, Draper BM and Low LF (2003) *Behavioural and psychological symptoms of dementia: a seven-tiered model of service delivery*. MJA. Vol. 178, pp.231-234.
- Brodaty H, Kemp NM and Low LF (2004) *Characteristics of the GPCOG, a screening tool for cognitive impairment*. International Journal of Geriatric Psychiatry. Vol. 19, No.9, pp.870-4.
- Brodaty H, Low LF, Gibson L and Burns K (2006) *What is the best dementia screening instrument for general practitioners to use?* American Journal of Geriatric Psychiatry. Vol. 14 No.5, in press.
- Brodaty H, Ames D, et al. (2003) *A randomized placebo-controlled trial of risperidone for the treatment of aggression, agitation, and psychosis of dementia*. Journal of Clinical Psychiatry. Vol. 64, No.2, pp.134-43.
- Brodaty H, Ames D, et al. (2005) *Risperidone for psychosis of Alzheimer's disease and mixed dementia: results of a double-blind, placebo-controlled trial*. International Journal of Geriatric Psychiatry. Vol. 20, No.12, pp.1153-7.
- Brodaty H, Ames D, Snowdon J, Woodward M, et al. (2003) *A randomized placebo-controlled trial of risperidone for the treatment of aggression, agitation, and psychosis of dementia*. Journal of Clinical Psychiatry. Vol. 64, No.2, pp.134-43.
- Brodaty H, Corey-Bloom J, Potocnik FCV, et al. (2005) *Galantamine prolonged-release formulation in the treatment of mild to moderate Alzheimer's disease*. Dementia and Geriatric Cognitive Disorders. Vol. 20, No.2-3, pp.120-132.
-

- Brodaty H, Draper B, et al. (2001) *Psychosis, depression and behavioural disturbances in Sydney nursing home residents: prevalence and predictors*. International Journal of Geriatric Psychiatry. Vol. 16, No.5, pp.504-12.
- Brodaty H, Draper BM, et al. (2003) *Randomized controlled trial of different models of care for nursing home residents with dementia complicated by depression or psychosis*. Journal of Clinical Psychiatry. Vol. 64, No.1, pp.63-72.
- Brouwer WBF, van Exel NJA, van Gorp B and Redekop WK (2006) *The CarerQol instrument: a new instrument to measure care-related quality of life of informal caregivers for use in economic evaluations*. Quality of Life Research. Vol. 15, No.6, pp.1005-21.
- Brown LM, Schinka JA, Mortimer JA and Graves AB (2003) *3MS normative data for elderly African Americans*. Journal of Clinical and Experimental Neuropsychology. Vol. 25, No.2, pp.234-241.
- Bryan S, Hardyman W, Bentham P, Buckley A and Laight A (2005) *Proxy completion of EQ-5D in patients with dementia*. Quality of Life Research. Vol. 14, No.1, pp.107-18.
- Buckingham K (1995) *The Research Strategy*. Methodology Workshop, EUROQoL Conference, London.
- Buettner LL, Lundegren H, Lago D, Farrell P, et al. (1996) *Therapeutic recreation as an intervention for persons with dementia and agitation: an efficacy study*. American Journal of Alzheimer's Disease and Other Dementias. Vol. 11, No.5, pp.4-12.
- Bullinger M, Azouvi P, Brooks N, Basso A, Christensen AL, Gobiet W, Greenwood R, Hutter B, Jennett B, Maas A, Truelle JL and von Wild KR (2002) *Quality of life in patients with traumatic brain injury-basic issues, assessment and recommendations*. Restorative Neurology and Neuroscience. Vol. 20, No.3-4, pp.111-24.
- Burch EA, Jr. and Andrews SR (1987) *Comparison of two cognitive rating scales in medically ill patients*. International Journal of Psychiatry in Medicine. Vol. 17, No.2, pp.193-200.
- Burgoyne RW and Saunders DS (2000) *Perceived support in newly registered HIV/AIDS clinic outpatients*. AIDS Care. Vol. 12, No.5, pp.643-50.
- Burke WJ, Miller JP, Rubin EH, et al. (1988) *Reliability of the Washington University Clinical Dementia Rating*. Archives of Neurology. Vol. 45, No.1, pp.31-2.
- Burke JR and Morgenlander JC (1999) *Managing common behavioural problems in dementia*. Post-graduate Medicine. Vol. 106, No.5, pp.131-140.
- Burns A, Rosso RM, Hecker J, et al. (1999) *The effects of donepezil in Alzheimer's disease - results from a multinational trial*. Dementia and Geriatric Cognitive Disorders. Vol. 10, No.3, pp.237-44.
- Burns A, Lawlor B and Craig S (2004) *Assessment scales in old age psychiatry (2nd ed.)* Taylor and Francis, London.
- Burns A, Beevor A, Lelliott P, Wing J, et al. (1999a) *Health of the Nation Outcome Scales for elderly people (HoNOS 65+)*. British Journal of Psychiatry. Vol. 174, pp.424-427.
- Burns A, Folstein S, Brandt J and Folstein M (1996) *Clinical assessment of irritability, aggression and apathy in Huntington and Alzheimer's disease*. Journal of Nervous and Mental Disease. Vol. 178, pp.20-26.
- Burns A, Spiegel R, et al. (2004) *Efficacy of rivastigmine in subjects with moderately severe Alzheimer's disease*. International Journal of Geriatric Psychiatry. Vol. 19, No.3, pp.243-9.
- Buschke H, Kuslansky G, Katz M, et al. (1999) *Screening for dementia with the Memory Impairment Screen*. American Academy of Neurology. Vol. 52, No.2, pp.231-238.
- Butters E, Higginson I, George R, et al. (1993) *Palliative care for people with HIV/AIDS: views of patients, carers and providers*. AIDS Care. Vol. 5, No.1, pp.105-16.

- 
- Buttle F (1996) *SERVQUAL - review, critique, research agenda*. European Journal of Marketing. Vol. 30, No.1, pp.8-32.
- Byrne-Davis LMT, Bennett PD and Wilcock GK (2006) *How are quality of life ratings made ? Toward a model of quality of life in people with dementia*. Quality of Life Research. Vol. 15, pp.855-865.
- Cacioppo JT, Hughes ME, Waite LJ, et al. (2006) *Loneliness as a specific risk factor for depressive symptoms: cross-sectional and longitudinal analyses*. Psychology and Aging. Vol. 21, No.1, pp.140-51.
- Cadet B (1994) *History of the construction of a health indicator integrating social preference: the Quality of Well-Being Scale*. 7th Meeting, International Network on Health Expectancy (REVES). Canberra, 1994.
- Caine E (1996) *Behavioural disturbances of dementia in our current nomenclature system: Diagnostic classification of neuropsychiatric signs and symptoms in patients with dementia*. International Psychogeriatrics. Vol. 8, Suppl.3, pp.273-279.
- Calkin PA, Kunik ME, Orengo CA, Molinari V, et al. (1997) *Tolerability of clonazepam in demented and non-demented geropsychiatric patients*. International Journal of Geriatric Psychiatry. Vol. 12, No.7, pp.745-9.
- Callahan CM, Unverzagt FW, Hui SL, Perkins AJ and Hendrie HC (2002) *Six-item screener to identify cognitive impairment among potential subjects for cognitive research*. Med Care. Vol. 40, No.9, pp.771-781.
- Calman KC (1984) *Quality of life in cancer patients--an hypothesis*. Journal of Medical Ethics. Vol. 10, No.3, pp.124-7.
- Calvert MM (1989) *Human-pet interaction and loneliness: a test of concepts from Roy's adaptation model*. Nursing Science Quarterly. Vol. 2, No.4, pp.194-202.
- Camara EG (1991) *A psychiatry outpatient consultation-liaison clinic. Experience at the Cleveland Clinic Foundation*. Psychosomatics. Vol. 32, No.3, pp.304-8.
- Cantor N and Sanderson C (1999) *Life task participation and well-being: the importance of taking part in daily life*. In Kahneman D, Diener E, Schwarz N, (eds). Well-Being: The Foundations of Hedonic Psychology. New York: Russell Sage Foundation. pp.230-243.
- Caplan EK and Sussman MB (1966) *Rank order of important variables for patient and staff satisfaction with outpatient service*. J Health Hum Behav. Vol. 7, No.2, pp.133-7.
- Caro J, Ward A, Ishak K, et al. (2002) *To what degree does cognitive impairment in Alzheimer's disease predict dependence of patients on caregivers?* BMC Neurology. Vol. 2, pp.6.
- Carstensen LL (1992) *Social and emotional patterns in adulthood: support for socioemotional selectivity theory*. Psychology and Aging. Vol. 7, No.3, pp.331-8.
- Cassel J (1976) *The contribution of the social environment to host resistance: the Fourth Wade Hampton Frost Lecture*. American Journal of Epidemiology. Vol. 104, No.2, pp.107-23.
- Cattell R (1952) *Factor Analysis: an Introduction and Manual for the Psychologist and Social Scientist*. Harper and Row, New York.
- Centre for Applied Gerontology - Bundoora Extended Care Centre (1996) *Culturally appropriate dementia assessment*. Centre for Applied Gerontology - Bundoora Extended Care Centre, Victoria.
- Chacko RC, Hurley RA, et al. (1995) *Clozapine for acute and maintenance treatment of psychosis in Parkinson's disease*. Journal of Neuropsychiatry and Clinical Neurosciences. Vol. 7, No.4, pp.471-5.
- Challis D, Mozley CG, Sutcliffe C, et al. (2000) *Dependency in older people recently admitted to care homes*. Age and Ageing. Vol. 29, No.3, pp.255-260.
- Chan CW, Hon HC, Chien WT, et al. (2004) *Social support and coping in Chinese patients undergoing cancer surgery*. Cancer Nursing. Vol. 27, No.3, pp.230-6.
-

- Chan WC, Lam LC, et al. (2001) *A double-blind randomised comparison of risperidone and haloperidol in the treatment of behavioural and psychological symptoms in Chinese dementia patients*. International Journal of Geriatric Psychiatry. Vol. 16, No.12, pp.1156-62.
- Chapman SB, Weiner MF, Rackley A, et al. (2004) *Effects of cognitive-communication stimulation for Alzheimer's disease patients treated with donepezil*. Journal of Speech Language and Hearing Research. Vol. 47, No.5, pp.1149-63.
- Chappell NL and Reid RC (2002) *Burden and well-being among caregivers: examining the distinction*. Gerontologist. Vol. 42, No.6, pp.772-80.
- Charles ST, Reynolds CA and Gatz M (2001) *Age-related differences and change in positive and negative affect over 23 years*. Journal of Personality and Social Psychology. Vol. 80, No.1, pp.136-51.
- Chene B (2006) *Dementia and Residential Placement: A View from the Carers' Perspective*. Qualitative Social Work. Vol. 5, No.2, pp.187-215.
- Cherbuin N, Anstey KJ and Lipnicki D (2006) *Dementia Screening Instruments: Recommendations for the National Dementia Website*. Report prepared for the Consortium for the National Dementia Website including PTG Global, Morphem, Alzheimer's Australia, Vision Australia and the Department of Health and Ageing.
- Chesterman J, Bauld L and Judge K (2001) *Satisfaction with the care-managed support of older people: an empirical analysis*. Health and Social Care in the Community. Vol. 9, No.1, pp.31-42.
- Chien WT, Chan SW and Morrissey J (2007) *The perceived burden among Chinese family caregivers of people with schizophrenia*. Journal of Clinical Nursing. Vol. 16, No.6, pp.1151-61.
- Chipuer HM, Bramston P and Pretty G (2003) *Determinants of subjective quality of life among rural adolescents: a developmental perspective*. Social Indicators Research. Vol. 61, No.1, pp.79-95.
- Chiu MJ, Chen TF, et al. (2006) *Behavioral and psychologic symptoms in different types of dementia*. Journal of the Formosan Medical Association. Vol. 105, No.7, pp.556-62.
- Choi SH, Na DL, Kwon HM, Yoon SJ, Jeong JH and Ha CK. (2000) *The Korean Version of the Neuropsychiatric Inventory: a scoring tool for neuropsychiatric disturbance in dementia patients*. Journal of Korean Medical Science. Vol. 15, No.6, pp.609-15.
- Chong D (1995) *Measurement of Instrumental Activities of Daily Living in Stroke*. Stroke. Vol. 26, pp.1119 - 1122.
- Chou KL, Jun LW and Chi I (2005) *Assessing Chinese older adults' suicidal ideation: Chinese version of the Geriatric Suicide Ideation Scale*. Aging and Mental Health. Vol. 9, No.2, pp.167-71.
- Choy CN, Lam LC, Chan WC, Li SW, et al. (2001) *Agitation in Chinese elderly: validation of the Chinese version of the Cohen-Mansfield Agitation Inventory*. International Psychogeriatrics. Vol. 13, No.3, pp.325-35.
- Chu LW, Chiu KC, Hui SL, et al. (2000) *The reliability and validity of the Alzheimer's disease Assessment Scale Cognitive Subscale (ADAS-Cog) among the elderly Chinese in Hong Kong*. Annals of the Academy of Medicine, Singapore. Vol. 29, No.4, pp.474-85.
- Chui HC, Mack W, Jackson J, Mungas D, Reed BR, Tinklenberg J, et al. (2000) *Clinical criteria for the diagnosis of vascular dementia: A multicenter study of comparability and interrater reliability*. Archives of Neurology. Vol. 57, No.2, pp.191-196.
- Chui HC, Victoroff JI, Margolin D, Jagust W, Shankle R and Katzman R (1992) *Criteria for the diagnosis of ischemic vascular dementia proposed by the State of California Alzheimer's disease Diagnostic and Treatment Centres*. Neurology. Vol. 42, pp.473-480.

Chung CC, Ha JP, Tai YP, et al. (2002) *Double-blind, randomized trial comparing Harmonic Scalpel hemorrhoidectomy, bipolar scissors hemorrhoidectomy, and scissors excision: ligation technique*. Dis Colon Rectum. Vol. 45, No.6, pp.789-94.

Cifu DX, Keyser-Marcus L, et al. (1997). *Acute predictors of successful return to work 1 year after traumatic brain injury: a multicenter analysis*. Archives of Physical Medicine and Rehabilitation. Vol. 78, No.2, pp.125-31.

Clark C and Ewbank D (1996) *Performance of the Dementia Severity Rating Scale: A Caregiver Questionnaire for Rating Severity in Alzheimer Disease*. Alz Dis and Ass Dis. Vol. 10, No.1, pp.31-39.

Clark CM, Ewbank D, Lerner A, Doody R, Henderson VW, Panisset M, Morris JC, Fillenbaum GG and Heyman A (1997) *The relationship between extra-pyramidal signs and cognitive performance in patients with Alzheimer's disease enrolled in the CERAD study*. Neurology. Vol. 49, No.1, pp.70-75.

Clarke DE, Reekum R, Simard M, et al. (2007) *Apathy in dementia: an examination of the psychometric properties of the apathy evaluation scale*. Journal of Neuropsychiatry and Clinical Neurosciences Vol.19, No.1, pp.57-64.

Clare J and Hofmeyer A (1998) *Discharge planning and continuity of care for aged people: indicators of satisfaction and implications for practice*. Australian Journal of Advanced Nursing. Vol. 16, No.1, pp.7-13.

Cobb S (1976) *Social support as a moderator of life stress*. Psychosomatic Medicine. Vol. 38, pp.300-314.

Coen RF, O'Boyle CA, Coakley D and Lawlor BA (2002) *Individual quality of life factors distinguishing low-burden and high-burden caregivers of dementia patients*. Dementia and Geriatric Cognitive Disorders. Vol. 13, No.3, pp.164-70.

Coen RF, Swanwick GR, et al. (1997) *Behaviour disturbance and other predictors of carer burden in Alzheimer's disease*. International Journal of Geriatric Psychiatry. Vol. 12, No.3, pp.331-336.

Cohen R, Leis AM, Kuhl D, Charbonneau C, et al. (2006) *QOLTI-F: Measuring family carer quality of life*. Palliative Medicine. Vol. 20, No.8, pp.755-767.

Cohen-Mansfield J, Marx MS and Rosenthal AS (1989) *A description of agitation in a nursing home*. Journal of Gerontology. Vol. 44, pp.M77-M84.

Cohen-Mansfield J, Taylor L, McConnell D, et al. (1999) *Estimating the cognitive ability of nursing home residents from the minimum data set*. Outcomes Management for Nursing Practice. Vol. 3, No.1, pp.43-6.

Cohen-Mansfield J (1986) *Agitated behaviors in the elderly: II Preliminary results in the cognitively deteriorated*. Journal of the American Geriatrics Society. Vol. 34, pp.722-27.

Cohen-Mansfield J (1991) *Instruction manual for the Cohen-Mansfield Agitation Inventory (CMAI)*. The Research Institute of the Hebrew Home of Greater Washington, Rockville.

Cohen-Mansfield J (1999) *Measurement of inappropriate behaviour associated with dementia*. Journal of Gerontological Nursing. Vol. 25, No.2, pp.42-51.

Cohen-Mansfield J and Billing N (1986) *Agitated behaviours in the elderly. A conceptual review*. Journal of the American Geriatrics Society. Vol. 34, pp.711-721.

Cohen-Mansfield J and Libin A (2004) *Assessment of agitation in elderly patients with dementia: correlations between informant rating and direct observation*. International Journal of Geriatric Psychiatry. Vol. 19, No.9, pp.881-91.

Cohen-Mansfield J and Werner P (1998) *Predictors of aggressive behaviour: a longitudinal study of senior day care centres*. Journal of Gerontology. Vol. 53B, pp.300-310.

Cohen-Mansfield J, Taylor L and Werner P (1998) *Delusions and hallucinations in an adult day care population. A longitudinal study*. American Journal of Geriatric Psychiatry. Vol. 6, No.2, pp.104-21.

- Cohen-Mansfield J, Werner P, Watson V and Pasis S (1995) *Agitation among elderly persons at adult day-care centers: the experiences of relatives and staff members*. International Psychogeriatrics. Vol. 7, No.3, pp.447-58.
- Collin C, Wade DT, Davis S, et al. (1988) *The Barthel ADL Index: A reliability study*. International Disability Studies. Vol. 10, pp.61-63.
- Connelly CD (1998) *Hopefulness, self-esteem, and perceived social support among pregnant and nonpregnant adolescents*. Western Journal of Nursing Research. Vol. 20, No.2, pp.195-209.
- Conner-Spady B and Suarez-Almazor ME (2003) *Variation in the estimation of quality-adjusted life-years by different preference-based instruments*. Medical Care. Vol. 41, No.7, pp.791-801.
- Cook K, Ashton C, Byrne M, Brody B, Geraci J, Giesler R, Hanita M, Soucek J and Wray N (2001) *A psychometric analysis of the measurement level of the rating scale, time trade-off, and standard gamble*. Social Science and Medicine. Vol. 53, pp.1275-1285.
- Coons S and Kaplan R (1993) *Quality of life assessment: understanding its use as an outcome measure*. Hospital Formulary. Vol. 28, pp.486-498.
- Cortes F, Gillette-Guyonnet S, Nourhashemi F, et al (2005) *Recent data on the natural history of Alzheimer's disease: results from the REAL.FR Study*. Journal of Nutrition, Health and Aging. Vol. 9, No.2, pp.86-93.
- Cortina J (1993) *What is coefficient alpha? Examination of theory and applications*. Journal of Applied Psychology. Vol. 78, No.1, pp.98-104.
- Costa PT, Williams TE, Somerfield M, et al. (1996) *Recognition and initial assessment of Alzheimer's disease and related dementias*. Clinical Practice Guidelines. No.19. Rockville, MD.: US Department of Health and Human Services, Public Health Service, Agency for Health Care Policy and Research.
- Coster WJ, Haley SM, Andres PL, et al. (2004) *Refining the conceptual basis for rehabilitation outcome measurement: Personal care and instrumental activities domain*. Medical Care. Vol. 42, No.1, pp.162-172.
- Cotter EM, Burgio LD, Stevens AB, et al. (2002) *Correspondence of the functional independence measure (FIM) self-care subscale with real-time observations of dementia patients' ADL performance in the home*. Clinical Rehabilitation. Vol. 16, No.1, pp.36-45.
- Coucoll W, Bryan S, Bentham P, Buckley A and Laight A (2001) *EQ-5D in patients with dementia: an investigation of inter-rater agreement*. Medical Care. Vol. 39, No.8, pp.760-71.
- Coyle N, Goldstein ML, Passik S, et al. (1996) *Development and validation of a patient needs assessment tool (PNAT) for oncology clinicians*. Cancer Nursing. Vol. 19, No.2, pp.81-92.
- Craig AH, Cummings JL, Rairbanks L, et al. (1996) *Cerebral blood flow correlates of apathy in Alzheimer's disease*. Arch Neurol. Vol. 53, pp.1116-1120.
- Cronbach J and Meehl P (1955) *Construct validity in psychological tests*. Psychological Bulletin. Vol. 52, No.4, pp.281-302.
- Cryns AG, Nichols RC, Katz LA, et al. (1989) *The hierarchical structure of geriatric patient satisfaction. An Older Patient Satisfaction Scale designed for HMOs*. Med Care. Vol. 27, No.8, pp.802-16.
- Cullum C, Saine K, Chan LD, et al. (2001) *Performance-based instrument to assess functional capacity in dementia: The Texas Functional Living Scale*. Neuropsychiatry, Neuropsychology and Behavioral Neurology. Vol. 14, No.2, pp.103-108.
- Cummings JL, Mega M, Gray K, Rosenberg-Thompson S, Carusi DA and Gornbein J (1994) *The Neuropsychiatric Inventory: comprehensive assessment of psychopathology in dementia*. Neurology. Vol. 44, No.12, pp.2308-2314.

---

Cummings JL (1996) *Theories behind scales and measurements: Theories behind existing scales for rating behaviour in dementia.* International Psychogeriatrics. Vol. 8, suppl.3, pp.293-300.

Cummings JL (1997) *The Neuropsychiatric Inventory: assessing psychopathology in dementia patients.* Neurology. Vol. 48, suppl.6, pp.S10-16.

Cummings JL, Diaz C, Levy ML, et al. (1996) *Behavioural syndromes in neurodegenerative diseases: frequency and significance.* Semin. Neuropsychiatry. Vol. 1, pp.241-247.

Cummings JL, McRae T, et al. (2006) *Effects of donepezil on neuropsychiatric symptoms in patients with dementia and severe behavioral disorders.* American Journal of Geriatric Psychiatry. Vol. 14, No.7, pp.605-12.

Cummings JL, Mega M, Gray K, et al. (1994) *The Neuropsychiatric Inventory: comprehensive assessment of psychopathology in dementia.* Neurology. Vol. 44, No.12, pp.2308-2314.

Cummings JL, Mega M, Gray K, Rosenberg-Thompson S, Carusi DA and Gornbein J (1994) *The Neuropsychiatric Inventory: comprehensive assessment of psychopathology in dementia.* Neurology. Vol. 44, pp.2308-2314.

Cummings JL, Raman R, Ernstrom K, et al. (2006) *ADCS Prevention Instrument Project: Behavioral Measures in Primary Prevention Trials.* Alzheimer Disease and Associated Disorders Vol. 20, Suppl.3, pp.S147-S151.

Cummings JL, Street J, Masterman D and Clarke WS (2002) *Efficacy of olanzapine in the treatment of psychosis in dementia with lewy bodies.* Dementia and Geriatric Cognitive Disorders. Vol. 13, No.2, pp.67-73.

Cummins R (1995) *On the trail of the gold standard for subjective wellbeing.* Social Indicators Research. Vol. 35, pp.179-200.

Cummins RA (2002) *Proxy responding for subjective well-being: a review.* International Review of Research in Mental Retardation. Vol. 25, pp.183-207.

Cutler NR, Shrottriya RC, Sramek JJ, et al. (1993) *The use of the Computerized Neuropsychological Test Battery (CNTB) in an efficacy and safety trial of BMY 21,502 in Alzheimer's disease.* Annals of New York Academy of Sciences. Vol. 695, pp.332-336.

Daker-White G, Beattie AM, Gilliard J, et al. (2002) *Minority ethnic groups in dementia care: a review of service needs, service provision and models of good practice.* Aging and Mental Health. Vol. 6, No.2, pp.101-8.

D'Attilio JP, Campbell BM, Lubold P, et al. (1992) *Social support and suicide potential: preliminary findings for adolescent populations.* Psychological Reports. Vol. 70, No.1, pp.76-8.

Davis P, Morris J and Grant E (1990) *Brief screening tests versus clinical staging in senile dementia of the Alzheimer type.* Journal of the American Geriatric Society. Vol. 38, No.2, pp.129-135.

Davis C, Wilson A and McCarthy S (1996) *Ethnicity and aged care assessment teams in Queensland.* Australian and New Zealand Journal of Public Health. Vol. 20, No.1, pp.33-40.

de Deyn PP, Katz IR, Brodaty H, Lyons B, et al (2005) *Management of agitation, aggression, and psychosis associated with dementia: a pooled analysis including three randomized, placebo-controlled double-blind trials in nursing home residents treated with risperidone.* Clinical Neurology and Neurosurgery. Vol. 107, No.6, pp.497-508.

de Deyn PP, Rabheru K, Rasmussen A, Bocksberger JP, et al. (1999) *A randomized trial of risperidone, placebo, and haloperidol for behavioral symptoms of dementia.[see comment].* Neurology. Vol. 53, No.5, pp.946-55.

- de Guise E, Feyz M, et al. (2005) *Overview of traumatic brain injury patients at a tertiary trauma centre*. Canadian Journal of Neurological Sciences. Vol. 32, No.2, pp.186-93.
- De Jong Gierveld J (1971) *Social isolation and the image of the unmarried*. Sociologia Neerlandica. Vol. 7, pp.1-14.
- De Jong Gierveld J (1978) *The construct of loneliness: components and measurement*. Essence. Vol. 2, No.4, pp.221-237.
- De Jong Gierveld J (1984) *The experience of loneliness: a test of a theoretical model of loneliness*. Nederlands Tijdschrift voor de Psychologie en haar Grensgebieden. Vol. 39, No.7, pp.376-384.
- De Jong Gierveld J (1987) *Developing and testing a model of loneliness*. Journal of Personality and Social Psychology. Vol. 53, No.1, pp.119-28.
- De Jong Gierveld J (1989) *Personal relationships, social support, and loneliness*. Journal of Social and Personal Relationships. Vol. 6, No.2, pp.197-221.
- De Jong Gierveld J and Havens B (2004) *Cross-national comparisons of social isolation and loneliness: introduction and overview*. Canadian Journal on Aging. Vol. 23, No.2, pp.109-13.
- De Jong Gierveld J and Kamphuis F (1985) *The development of a Rasch-type loneliness scale*. Applied Psychological Measurement. Vol. 9, No.3, pp.289-299.
- De Jong Gierveld J and Raadschelders J (1982) *Types of loneliness*. In Peplau LA, Perlman D, (eds.) Loneliness: A Sourcebook of Current Theory, Research and Therapy. New York: Wiley. pp.105-119.
- De Jong Gierveld J and Tilburg TV (1999) *Manual of the Loneliness Scale*. Vrije Universiteit Amsterdam, Faculty of Social Sciences, Department of Sociology. p.18.
- De Jong Gierveld J and Tilburg TV (2006) *A 6-item scale for overall, emotional and social loneliness: confirmatory tests on survey data*. Research on Aging. Vol. 28, No.5, pp.582-598.
- De Jong Gierveld J and van Tilburg T (1999) *Living arrangements of older adults in the Netherlands and Italy: coresidence values and behaviour and their consequences for loneliness*. Journal of Cross Cultural Gerontology. Vol. 14, pp.1-24.
- De Jong R, Osterlund OW and Roy GW (1989) *Measurement of quality-of-life changes in patients with Alzheimer's disease*. Clinical Therapeutics. Vol. 11, pp.545-554.
- De Jonghe JF and Kat MG (1996) *Factor structure and validity of the Dutch version of the Cohen-Mansfield Agitation Inventory (CMAI-D)*. Journal of the American Geriatrics Society. Vol. 44, No.7, pp.888-9.
- De Jonghe JF, Kalisvaart KJ, Dijkstra M, van Dis H, Vreeswijk R, Kat MG, et al. (2007) *Early symptoms in the prodromal phase of delirium: a prospective cohort study in elderly patients undergoing hip surgery*. Am J Geriatr Psychiatry. Vol. 15, No.2, pp.112-121.
- De Jonghe JFM, Kalisvaart KJ, Timmers JFM, Kat MG and Jackson JC (2005) *Delirium-O-Meter: a nurses' rating scale for monitoring delirium severity in geriatric patients*. International Journal of Geriatric Psychiatry. Vol. 20, No.12, pp.1158-1166.
- Dennis M, O'Rourke S, Slattery J, et al. (1997) *Evaluation of a stroke family care worker: results of a randomised controlled trial*. BMJ. Vol. 314, No.7087, pp.1071-6; discussion 1076-7.
- de Rooij SE, van Munster BC, Korevaar JC, Casteelen G, Schuurmans MJ, van der Mast RC, et al. (2006) *Delirium subtype identification and the validation of the Delirium Rating Scale--Revised-98 (Dutch version) in hospitalized elderly patients*. Int J Geriatr Psychiatry. Vol. 21, No.9, pp.876-882.
- Department of Veterans Affairs, University Health System Consortium, Federal Government Agency (1997) *Dementia identification and assessment. Guidelines for primary care practitioners*. National Guidelines Clearinghouse. (on line at [www.guideline.gov/index.asp](http://www.guideline.gov/index.asp))

- Devanand DP, Miller L, Richards M, Marder K, Bell K, Mayeux R and Stern Y (1992) *Columbia University Scale for Psychopathology in Alzheimer's disease*. Archives of Neurology. Vol. 49, pp.371-6.
- Dick MB, Dick-Muehlke C and Teng EL (2006) *Assessment of cognitive status in asians*. In Yeo G, Gallagher-Thompson D, (eds). Ethnicity and the Dementias. 2nd ed. Routledge, Taylor and Francis Group, New York, p 71-86.
- Diener E, Suh E, Lucas R and Smith H (1999) *Subjective well-being: three decades of progress*. Psychological Bulletin. Vol. 125, pp.276-302.
- DiTommaso E and Spinner B (1997) *Social and emotional loneliness: a re-examination of Weiss' typology of loneliness*. Personality and Individual Differences. Vol. 22, No.3, pp.417-427.
- Doble SE, Fisk JD, MacPherson KM, et al. (1997) *Measuring functional competence in older persons with Alzheimer's disease*. International Psychogeriatrics. Vol. 9, No.1, pp.25-38.
- Doble SE, Fisk JD and Rockwood K (1999) *Assessing the ADL functioning of persons with Alzheimer's disease: comparison of family informants' ratings and performance-based assessment findings*. International Psychogeriatrics. Vol. 11, No.4, pp.399-409.
- Dolan P (1997) *Modeling valuations for EuroQol health states*. Medical Care. Vol. 35, No.11, pp.1095-108.
- Dolan P, Gudex C, Kind P and Williams A (1995) *Social tariff for EUROQoL: results from a UK general population survey*. Centre for Health Economics, University of York, York.
- Dolan P, Gudex C, Kind P and Williams A (1996) *Valuing health states: a comparison of methods*. Journal of Health Economics. Vol. 15, pp.209-231.
- Dombovy M L and A C Olek (1997) *Recovery and rehabilitation following traumatic brain injury*. Brain Injury. Vol. 11, No.5, pp.305-18.
- Donabedian A (1980) *Explorations in Quality Assessment and Monitoring: Vol 1 The Definition of Quality and the Approaches to its Assessment*. Arbor A, editor. Health Administration Press, MI.
- Donabedian A (1988) *The quality of care: how can it be assessed?* Journal of the American Medical Association. Vol. 260, No.12, pp.1743-48.
- Dooneief G, Marder K, Tang MX, et al. (1996) *The Clinical Dementia Rating scale: community-based validation of "profound" and "terminal" stages*. Neurology. Vol. 46, No.6, pp.1746-9.
- Doraiswamy PM, Bieber F, Kaiser L, et al. (1997a) *Memory, language, and praxis in Alzheimer's disease: norms for outpatient clinical trial populations*. Psychopharmacology Bulletin. Vol. 33, No.1, pp.123-8.
- Doraiswamy PM, Bieber F, Kaiser L, et al. (1997b) *The Alzheimer's disease Assessment Scale: patterns and predictors of baseline cognitive performance in multicenter Alzheimer's disease trials.[see comment]*. Neurology. Vol. 48, No.6, pp.1511-7.
- Doraiswamy PM, Kaiser L, Bieber F, et al. (2001) *The Alzheimer's disease Assessment Scale: evaluation of psychometric properties and patterns of cognitive decline in multicenter clinical trials of mild to moderate Alzheimer's disease.[erratum appears in Alzheimer Dis Assoc Disord 2002 Jan-Mar;16(1):48]*. Alzheimer Disease and Associated Disorders. Vol. 15, No.4, pp.174-83.
- Dorman P, Slattery J, Farrell B, Dennis M and Sandercock P (1998) *Qualitative comparison of the reliability of health status assessments with the EuroQol and SF-36 questionnaires after stroke. United Kingdom Collaborators in the International Stroke Trial*. Stroke. Vol. 29, No.1, pp.63-68.
- Dorman P, Slattery J, Farrell B, Dennis MS and Sandercock PA (1997) *A randomised comparison of the EuroQol and Short Form-36 after stroke. United Kingdom collaborators in the International Stroke Trial*. British Medical Journal. Vol. 315, No.7106.

- Dozier AM, Kitzman HJ, Ingersoll GL, et al. (2001) *Development of an instrument to measure patient perception of the quality of nursing care*. Research in Nursing and Health. Vol. 24, No.6, pp.506-17.
- Drachman DA, Swearer JM, O'Donnell BF, Mitchell AL and Maloon A (1992) *The caretaker obstreperous-behaviour rating assessment (COBRA) scale*. Journal of the American Geriatric Society. Vol. 40, pp.463-470.
- Draper B and Turner J (2003) *Outcome of treatment of vocally disruptive behaviour in nursing home residents*. Vol. 2, pp.80-4.
- Dykstra P, van Tilburg TG and de Jong Gierveld J (2005) *Changes in older adult loneliness: results from a seven-year longitudinal study*. Research on Aging. Vol. 27, No.6, pp.725-747.
- Dykstra PA and de Jong Gierveld J (2004) *Gender and marital-history differences in emotional and social loneliness among Dutch older adults*. Can J Aging. Vol. 23, No.2, pp.141-55.
- Eagar K, Owen A, Cromwell D, et al. (2001) *Towards a national measure of functional dependency for home and community care services in Australia - Stage 1 report of the HACCC Dependency Data Items Project*. Centre for Health Service Development, University of Wollongong.
- Eagar K, Owen A, Marosszeky N, et al. (2006) *Towards a measure of function for home and community care services in Australia: Part 1 - Development of a Standard National Approach*. Australian Journal of Primary Health. Vol. 12, No.1, pp.73-81.
- Edelman P, Fulton BR and Kuhn D (2004) *Comparison of dementia-specific quality of life measures in adult day centers*. Home Health Care Services Quarterly. Vol. 23, No.1, pp.25-42.
- Edelman P, Fulton BR, Kuhn D, et al. (2005) *A comparison of three methods of measuring dementia-specific quality of life: perspectives of residents, staff, and observers*. Gerontologist. Vol. 45 Spec No.1, No.1, pp.27-36.
- Ellaway A, Wood S and MacIntyre S (1999) *Someone to talk to? The role of loneliness as a factor in the frequency of GP consultations*. British Journal of General Practice. Vol. 49, pp.363-367.
- Ellis P and Hickie I (2001) *What causes mental illness?* In Bloch S, Singh B, (eds.) Foundations of Clinical Psychiatry. Melbourne: Melbourne University Press. pp.43-62.
- Elvik R (1995) *The validity of using health state indexes in measuring the consequences of traffic injury for public health*. Social Science and Medicine. Vol. 40, No.10, pp.1385-98.
- Ely EW, Inouye SK, Bernard GR, Gordon S, Francis J, May L, et al. (2001) *Delirium in mechanically ventilated patients: validity and reliability of the confusion assessment method for the intensive care unit (CAM-ICU)*. JAMA. Vol. 286, No.21, pp.2703-2710.
- Ely EW, Margolin R, Francis J, May L, Truman B, Dittus R, et al. (2001) *Evaluation of delirium in critically ill patients: validation of the Confusion Assessment Method for the Intensive Care Unit (CAM-ICU)*. Critical Care Medicine. Vol. 29, No.7, pp.1370-1379.
- Elzinga RH and Barlow J (1991) *Patient satisfaction among the residential population of a psychiatric hospital*. International Journal of Social Psychiatry. Vol. 37, No.1, pp.24-34.
- Engel SE, Kiely DK and Mitchell SL (2006) *Satisfaction with end-of-life care for nursing home residents with advanced dementia*. Journal of the American Geriatrics Society. Vol. 54, No.10, pp.1567-72.
- Engelborghs S, Maertens K, et al. (2005) *Neuropsychiatric symptoms of dementia: cross-sectional analysis from a prospective, longitudinal Belgian study*. International Journal of Geriatric Psychiatry. Vol. 20, No.11, pp.1028-37.
- Erkinjuntti T, Ostbye T, Steenhuis R and Hachinski V (1997) *The effect of different diagnostic criteria on the prevalence of dementia*. New England Journal of Medicine. Vol. 337, No.23, pp.1667-1674.

Erzigkeit H, Lehfeld H, Pena-Casanova J, Bieber F, Yekrangi-Hartman C, Rupp M, Rappard F, Arnold K, and Hindmarch I (2001) *The Bayer-Activities of Daily Living Scale (B-ADL): Results from a validation study in three European countries.* Dementia and Geriatric Cognitive Disorders. Vol. 12, pp. 348-358.

Eslinger PS, Dennis K, Moore P, et al. (2005) *Metacognitive deficits in frontotemporal dementia.* J Neurol Neurosurg Psychiatry. Vol. 76, pp.1630-1635.

EuroQol Group (1990) *EuroQol: a new facility for measurement of health-related quality of life.* Health Policy. Vol. 16, pp.199-208.

Everaert K, De Ridder D, Baert L, et al. (2000) *Patient satisfaction and complications following sacral nerve stimulation for urinary retention, urge incontinence and perineal pain: a multicenter evaluation.* International Urogynecology Journal and Pelvic Floor Dysfunction. Vol. 11, No.4, pp.231-5; discussion 236.

Farias ST, Harrell E, Neumann C, et al. (2003) *The relationship between neuropsychological performance and daily functioning in individuals with Alzheimer's disease: ecological validity of neuropsychological tests.* Archives of Clinical Neuropsychology. Vol. 18, No.6, pp.655-672.

Farlow M, Gracon S, Hershey L, et al. (1992) *A controlled trial of tacrine in Alzheimer's disease. The Tacrine Study Group. [see comment].Clinical Trial. Multicenter Study. Randomized Controlled Trial.* JAMA. Vol. 268, No.18, pp.2523-9.

Farlow MR, He Y, Tekin S, et al. (2004) *impact of APOE in mild cognitive impairment.* Neurology. Vol. 63, No.10, pp.1898-1901.

Farmer I, Meyer P, Ramsey D, et al. (1996) *Higher levels of social support predict greater survival following acute myocardial infarction: the Corpus Christi Heart Project.* Behavioural Medicine. Vol. 22, pp.59-66.

Farran CJ (2001) *Family caregiver intervention research: where have we been? Where are we going?* Journal of Gerontological Nursing. Vol. 27, No.7, pp.38-45.

Fayers PM and Machin D (2000) *Quality of Life: Assessment, Analysis and Interpretation.* John Wiley and Sons Ltd, pp.404.

Feeny D, Furlong W and Torrance G (1996a) *Health Utilities Index Mark 2 and Mark 3 (HUI2/3) 15-item questionnaire for self-administered, self-assessed usual health status,* Centre for Health Economics and Policy Analysis. McMaster University, Hamilton.

Feeny D, Torrance G and Furlong W (1996) *Health Utilities Index* In Spilker, B (ed.). Quality of Life and Pharmacoeconomics in Clinical Trials. Lippincott-Raven Publishers, Philadelphia.

Feldman HH, Sauter A, Donald A, et al. (2001) *The disability assessment for dementia scale: a 12-month study of functional ability in mild to moderate severity Alzheimer disease.* Alzheimer Disease and Associated Disorders. Vol. 15, No.2, pp.89-95.

Feldman HH, Schmitt FA, Olin JT, et al. (2006) *Activities of daily living in moderate-to-severe Alzheimer disease: An analysis of the treatment effects of memantine in patients receiving stable donepezil treatment.* Alzheimer Disease and Associated Disorders. Vol. 20, No.4, pp.263-268.

Feldman HH, Van Baelen B, Kavanagh SM, et al. (2005) *Cognition, function, and care giving time patterns in patients with mild-to-moderate Alzheimer disease: a 12-month analysis.* Alzheimer Disease and Associated Disorders. Vol. 19, No.1, pp.29-36.

Feldman HH, Gauthier S, Hecker J, et al. (2003) *Efficacy of donepezil on maintenance of activities of daily living in patients with moderate to severe Alzheimer's disease and the effect on caregiver burden.* JAGS. Vol. 51, No.6, pp.737-744.

Fichter MM, Meller I, Schroppe H, and Steinkirchner R (1995) *Dementia and cognitive impairment in the oldest old in the community: prevalence and comorbidity.* British Journal of Psychiatry. Vol. 166, pp.621-629.

- Fillenbaum GG (1985) *Screening the elderly: A brief instrumental activities of daily living measure*. JAGS. Vol. 33, No.10, pp.698-706.
- Fillenbaum GG (1988) *Multidimensional functional assessment of older adults: the Duke Older Americans Resources and Services procedures*. Lawrence Erlbaum Associates, Hillsdale, NJ.
- Fillenbaum GG, Peterson B and Morris JC (1996) *Estimating the validity of the clinical Dementia Rating Scale: the CERAD experience. Consortium to Establish a Registry for Alzheimer's Disease*. Ageing-Clinical and Experimental Research. Vol. 8, No.6, pp.379-85.
- Fillenbaum GG and Smyer MA (1981) *The development, validity, and reliability of the OARS multidimensional functional assessment questionnaire*. Journal of Gerontology. Vol. 36, No.4, pp.428-34.
- Finkel SI (1998) *The signs of behavioural and psychological symptoms of dementia*. Clinician. Vol. 16, No.1, pp.33-42.
- Finkel S, Lyons J and Anderson R (1992) *Reliability and validity of the Cohen-Mansfield Agitation Inventory in institutionalized elderly*. International Journal of Geriatric Psychiatry. Vol. 7, pp.487-90.
- Finkel SI, Lyons JS, et al. (1993) *A brief agitation rating scale (BARS) for nursing home elderly*. Journal of the American Geriatrics Society. Vol. 41, No.1, pp.50-2.
- Firestone DN, Jimenez-Briceno L, Reimann JO, et al. (2004) *Predictors of diabetes-specific knowledge and treatment satisfaction among Costa Ricans*. Diabetes Education. Vol. 30, No.2, pp.281-292.
- First MB and Tasman A (eds.) (2004) *DSM-IV-TR Mental Disorders: Diagnosis, Etiology and Treatment*. John Wiley and Sons, LTD, West Sussex.
- Fitzpatrick R (1990) *Measurement of patient satisfaction*. In Hopkins D, Costain D, (eds.) Measuring the Outcomes of Medical Care. Royal College of Physicians and King's Fund Centre, London.
- Fitzpatrick R (1991) *Surveys of patient satisfaction: I - important general considerations*. British Medical Journal. Vol. 302, pp.887-889.
- Fitzpatrick R and Hopkins A (1983) *Problems in the conceptual framework of patient satisfaction research: an empirical exploration*. Social Health Illn. Vol. 5, No.3, pp.297-311.
- Fleming R and Kramer T (1995) *Dementia Care Planning: individual, integrated and effective. A national action plan for dementia project*. Hammond Care Group.
- Foddy W (1993) *Constructing Questions for Interviews and Questionnaires: Theory and Practice in Social Research*. Cambridge University Press, Cambridge.
- Foglia JP, Pollock BG, et al. (1997) *Plasma levels of citalopram enantiomers and metabolites in elderly patients*. Psychopharmacology Bulletin. Vol. 33, No.1, pp.109-12.
- Folstein M, Folstein S and McHugh P (1975) *'Mini-Mental State': a practical method for grading the cognitive state of patients for the clinician*. Journal of Psychiatric Research. Vol. 12, pp.189-198.
- Fonseca F, Bulbena A, Navarrete R, Aragay N, Capo M, Lobo A, et al. (2005) *Spanish version of the Delirium Rating Scale-Revised-98: reliability and validity*. Journal of Psychosomatic Research. Vol. 59, No.3, pp.147-151.
- Fountoulakis KN, Kaprinis SG and Kaprinis GS (2002) *Special characteristics affecting the neuropsychologic assessment of the elderly in Greece*. American Journal of Alzheimer's Disease and Other Dementias. Vol. 17, No.5, pp.273-276.
- Fox JG and Storms DM (1981) *A different approach to sociodemographic predictors of satisfaction with health care*. Social Science and Medicine. Vol. 15, No.5, pp.557-64.

- 
- Fratiglioni L, Wang HX, Ericsson K, et al. (2000) *Influence of social network on occurrence of dementia: a community-based longitudinal study.* Lancet. Vol. 355, No.9212, pp.1315-9.
- Frank LB (2006) *Commentary on "Meta-analysis of memantine: summary and commentary on the Cochrane Collaboration's systematic review." Meaning and measurement in cognitive impairment.* Alzheimer's and Dementia. Vol. 2, pp.93-95.
- Franulic A, Carbonell CG, et al. (2004) *Psychosocial adjustment and employment outcome 2, 5 and 10 years after TBI.* Brain Injury. Vol. 18, No.2, pp.119-29.
- Franulic A, Horta E, et al. (2000) *Organic personality disorder after traumatic brain injury: cognitive, anatomic and psychosocial factors. A 6 month follow-up.* Brain Injury. Vol. 14, No.5, pp.431-9.
- Fries BE, Simon SE, Morris JN, et al. (2001) *Pain in US nursing homes; Validating a pain scale for the Minimum Data Set.* The Gerontologist. Vol. 41, pp.173-179
- Frisoni GB, Rozzini L, Gozzetti A, Binetti G, Zanetti O, Bianchetti A and Trabucchi M (1999) *Behavioural syndromes in Alzheimer's disease: description and correlates.* Dementia Geriatr Cogn Discord. Vol. 10, pp.130-138.
- Fritz CL, Farver TB, Kass PH and Hart LA (1997) *Correlation among three psychological scales used in research of caregivers for patients with Alzheimer's disease.* Psychological Reports. Vol. 80, No.1, pp.67-80.
- Froberg DG and Kane RL (1989) *Methodology for measuring health-state preferences--I: Measurement strategies.* Journal of Clinical Epidemiology. Vol. 42, No.4, pp.345-54.
- Fromm-Reichmann F (1959) *Loneliness.* Psychiatry. Vol. 22, No.1, pp.1-15.
- Fryback DG (2005) *A US valuation of the EQ-5D.* Medical Care. Vol. 43, No.3, pp.199-200.
- Fujikawa T, Takahashi T, Kinoshita A, Kajiyama H, et al. (2004) *Quetiapine treatment for behavioral and psychological symptoms in patients with senile dementia of Alzheimer type.* Neuropsychobiology. Vol. 49, No.4, pp.201-4.
- Fulford M and Farhall J (2001) *Hospital versus home care for the acutely mentally ill? Preferences of caregivers who have experienced both forms of service.* Australian and New Zealand Journal of Psychiatry. Vol. 35, No.5, pp.619-25.
- Furlong W, Feeny D, Torrance G, Goldsmith C, DePauw S, Zhu Z, Denton M and Boyle M (1998) *Multiplicative Multi-attribute Utility Function for the Health Utilities Index Mark 3 (HUI3) System: A Technical Report.* Centre for Health Economics and Policy Analysis, McMaster University, Hamilton.
- Furlong WJ, Feeny DH, Torrance GW and Barr RD (2001) *The Health Utilities Index (HUI) system for assessing health-related quality of life in clinical studies.* Annals of Medicine. Vol. 33, pp.375-384.
- Fuh JL and Wang SJ (2006) *Assessing quality of life in Taiwanese patients with Alzheimer's disease.* International Journal of Geriatric Psychiatry. Vol. 21, No.2, pp.103-7.
- Furukawa T, Yokouchi T, Hirai T, et al. (1999) *Parental loss in childhood and social support in adulthood among psychiatric patients. Group for Longitudinal Affective Disorders Study (GLADS).* Journal of Psychiatric Research. Vol. 33, No.2, pp.165-9.
- Gaddini A, Ascoli M and Biscaglia L (2005) *Mental health care in Rome.* European Psychiatry. Vol. 20, Suppl 2, pp.S294-7.
- Galasko D, Bennett D, Sano M, et al. (1997) *An inventory to assess activities of daily living for clinical trials in Alzheimer's disease.* Alzheimer Disease and Associated Disorders. Vol. 11, No.Suppl 2, pp.S33-39.
- Galasko D, Kershaw PR, Schneider LS, et al. (2004) *Galantamine maintains ability to perform activities of daily living in patients with Alzheimer's disease.* JAGS. Vol. 52, No.7, pp.1070-1076.
-

---

Gallagher LA (ed.) (2004) *Thesaurus of Psychological Index Terms (1974 – 2004) Tenth Edition 2004*. American Psychological Association, Washington DC.

Gallagher-Thompson D (2006) *The family as the unit of assessment and treatment in work with ethnically diverse older adults with dementia*. In Yeo G, Gallagher-Thompson D, (eds). Ethnicity and the Dementias. 2nd ed. Routledge, Taylor and Francis Group, New York, p.71-86.

Galvin JE, Roe CM, Xiong C, Morris JC (2006) *Validity and reliability of the AD8 informant interview in dementia*. Neurology. Vol. 67, pp. 1942-1948.

Ganz PA, Rowland JH, Desmond K, et al. (1998) *Life after breast cancer: understanding women's health-related quality of life and sexual functioning*. Journal of Clinical Oncology. Vol. 16, No.2, pp.501-14.

Gaudreau JD, Gagnon P, Harel F, Tremblay A and Roy MA (2005) *Fast, systematic, and continuous delirium assessment in hospitalized patients: the nursing delirium screening scale*. Journal of Pain and Symptom Management. Vol. 29, No.4, pp.368-375.

Geldmacher DS (2002) *Cost-effective recognition and diagnosis of dementia*. Seminars in Neurology. Vol. 22, No.1, pp.63-70.

Gelinas I, Gauthier L, McIntyre M, et al. (1999) *Development of a functional measure for persons with Alzheimer's disease: The Disability Assessment for Dementia*. American Journal of Occupational Therapy. Vol. 53, pp.471-481.

Geller JS (2004) *Loneliness and pregnancy in an urban Latino community: associations with maternal age and unscheduled hospital utilization*. Journal of Psychosomatic Obstetrics and Gynaecology. Vol. 25, No.3-4, pp.203-9.

Gerdner LA and Swanson EA (1993) *Effects of individualized music on confused and agitated elderly patients*. Archives of Psychiatric Nursing. Vol. 7, No.5, pp.284-91.

Gerson AC and Perlman D (1979) *Loneliness and expressive communication*. Journal of Abnormal Psychology. Vol. 88, No.3, pp.258-61.

Gessert CE, Forbes S and Bern-Klug M (2000) *Planning end-of-life care for patients with dementia: roles of families and health professionals*. Omega: Journal of Death and Dying. Vol. 42, No.4, pp.273-91.

Gibbons LE, Teri L, Logsdon RG, et al. (2006) *Assessment of Anxiety in Dementia: An Investigation into the Association of Different Methods of Measurement*. Journal of Geriatric Psychiatry and Neurology. Vol.19, No.4, pp.202-208.

Gigliotti E (2002) *A confirmation of the factor structure of the Norbeck Social Support Questionnaire*. Nursing Research. Vol. 51, No.5, pp.276-84.

Gigliotti E (2006) *A confirmatory factor analysis of situation-specific Norbeck Social Support Questionnaire items*. Nursing Research. Vol. 55, No.3, pp.217-22.

Gill L and Maas F (2000) *Consumer and Carer Satisfaction Survey Aged Persons Mental Health Services - Summary Report*. ACNielsen. 123 p.

Gill T, Williams CS, Richardson E, et al. (1997) *A predictive model for ADL dependence in community-living older adults based on a reduced set of cognitive status items*. JAGS. Vol. 45, No.4, pp.441-445.

Gill TM and Feinstein AR (1994) *A critical appraisal of the quality of quality-of-life measurements*. Journal of the American Medical Association. Vol. 272, No. 8, pp. 619-26.

Glosser G, Wolfe N, Albert ML, et al. (1993) *Cross-cultural cognitive examination: Validation of a dementia screening instrument for neuroepidemiological research*. JAGS. Vol. 41, pp.931-939.

Goddaer J and Abraham IL (1994) *Effects of relaxing music on agitation during meals among nursing home residents with severe cognitive impairment*. Archives of Psychiatric Nursing. Vol. 8, No.3, pp.150-8.

---

Gold G, Bouras C, Canuto A, Bergallo MF, Herrmann FR, Hof PR, et al. (2002) *Clinicopathological validation study of four sets of clinical criteria for vascular dementia. [see comment]. American Journal of Psychiatry.* Vol. 159, No.1, pp.82-87.

Goldney RD, Fisher LJ, Wilson DH and Cheek F (2001) Suicidal ideation and health-related quality of life in the community. *Medical Journal of Australia.* Vol. 175, No.10, pp.546-549.

Goldstein FC, Strasser DC, Woodard JL, et al. (1997) *Functional outcome of cognitively impaired hip fracture patients on a geriatric rehabilitation unit. JAGS.* Vol. 45, No.1, pp.35-42.

Gomez-Isla T, Price JL, McKeel Jr DW, et al. (1996) Profound Loss of Layer II Entorhinal Cortex Neurons Occurs in Very Mild Alzheimer's disease. *Journal of Neuroscience.* Vol. 16, No.14, pp.4491-4500.

Gompertz P, Pound P, Briffa J, et al. (1995) *How useful are non-random comparisons of outcomes and quality of care in purchasing hospital stroke services? Age and Ageing.* Vol. 24, No.2, pp.137-41.

Gonzalez N, Quintana JM, Bilbao A, et al. (2005) *Development and validation of an in-patient satisfaction questionnaire. International Journal for Quality in Health Care.* Vol. 17, No.6, pp.465-72.

Gonzalez-Salvador T, Lyketsos CG, Baker A, Hovanec L, et al. (2000) *Quality of life in dementia patients in long-term care. International Journal of Geriatric Psychiatry.* Vol. 15, No.2, pp.181-9.

Gormley N, Lyons D and Howard R (2001) *Behavioural management of aggression in dementia: a randomised controlled trial. Age and Ageing.* Vol. 30, pp.141-145.

Gormley N, Rizwan M and Lovestone S (1998) *Clinical predictors of aggressive behaviours in Alzheimer's disease. Int J Geriatric Psychiatry.* Vol. 15, pp.109-115.

Grabowski TJ and Damasio AR (2004) *Definition, clinical features and neuroanatomical basis of dementia.* In Esiri MM, Lee VM and Trojanowski JQ (eds.) (2004) *The Neuropathology of Dementia (2nd ed.)* Cambridge University Press, Cambridge.

Grace J and Malloy PF (2001) *Frontal Systems Behavior Scale (FrSBe): Professional Manual.* Psychological Assessment Resources.

Grace SL, Evindar A, Kung TN, et al. (2004) *Automatic referral to cardiac rehabilitation. Medical Care.* Vol. 42, No.7, pp.661-9.

Graham C, Howard R and Ha Y (1998) *Dementia and ethnicity. International Psychogeriatrics.* Vol. 10, No.2, pp.183-91.

Grant C, Goodenough T, Harvey I, et al. (2000) *A randomised controlled trial and economic evaluation of a referrals facilitator between primary care and the voluntary sector. British Medical Journal.* Vol. 320, No.7232, pp.419-23.

Grasel E (2002) *When Home Care Ends-Changes in the Physical Health of Informal Caregivers Caring for Dementia Patients: A Longitudinal Study. JAGS.* Vol. 50, pp.843-849.

Grassi L, Caraceni A, Beltrami E, Borreani C, Zamorani M, Maltoni M, et al. (2001) *Assessing delirium in cancer patients: the Italian versions of the Delirium Rating Scale and the Memorial Delirium Assessment Scale. Journal of Pain and Symptom Management.* Vol. 21, No.1, pp.59-68.

Greehalgh J and Meadows K (1999) *The effectiveness of the use of patient-based measures of health in routine practice in improving the process and outcomes of patient care: a literature review. Journal of Evaluation in Clinical Practice.* Vol. 5, No.4, pp.401-16.

Green CR, Mohs RC, Schmeidler J, et al. (1993) *Functional decline in Alzheimer's disease: a longitudinal study. JAGS.* Vol. 41, No.6, pp.654-661.

- Green J, Eagar K, Owen A, et al. (2006) *Towards a measure of function for home and community care services in Australia: Part 2 - Evaluation of the Screening Tool and Assessment Instruments*. Australian Journal of Primary Health. Vol. 12, No.1, pp.82-90.
- Green MS, JA Kaye, et al. (2000) *The Oregon brain aging study: neuropathology accompanying healthy aging in the oldest old*. Neurology. Vol. 54, No.1, pp.105-13.
- Grimmer K and Moss J (2001) *The development, validity and application of a new instrument to assess the quality of discharge planning activities from the community perspective*. International Journal for Quality in Health Care. Vol. 13, No.2, pp.109-16.
- Grol R and Wensing M (2000) *Patients Evaluate General/Family Practice: The EUROPEP Instrument*. The Task Force on Patient Evaluations of General Practice Care. 70 p.
- Grol R, Wensing M, Mainz J, et al. (1999) *Patients' priorities with respect to general practice care: an international comparison*. *European Task Force on Patient Evaluations of General Practice (EUROPEP)*. Family Practice. Vol. 16, No.1, pp.4-11.
- Grube M (2004) *Which types of aggressive behaviour are associated with suicidal and self-injurious behaviour at the time of admission?* Psychopathology. Vol. 37, No.1, pp.41-49.
- Gruber-Baldini AL, Zimmerman SI, Mortimore E, et al. (2000) *The validity of the minimum data set in measuring the cognitive impairment of persons admitted to nursing homes*. Journal of the American Geriatrics Society. Vol. 48, No.12, pp.1601-6.
- Grumbach K, Selby JV, Damberg C, et al. (1999) *Resolving the gatekeeper conundrum: what patients value in primary care and referrals to specialists*. JAMA. Vol. 282, No.3, pp.261-6.
- Grundman M, Petersen RC, Ferris SH, Thomas RG, Aisen PS, Bennett DA, et al. (2004) *Mild cognitive impairment can be distinguished from Alzheimer disease and normal aging for clinical trials*. Archives of Neurology. Vol. 61, No.1, pp.59-66.
- Grunfeld E, Coyle D, Whelan T, et al (2004) *Family caregiver burden: results of a longitudinal study of breast cancer patients and their principal caregivers*. Canadian Medical Association Journal. Vol. 170, No.12, pp.1795-801.
- Grypma R, Mahajani S and Tam E (2007) *Screening and diagnostic assessment of non-english speaking people with dementia: Guidelines and system recommendations for practitioners, service managers and policy makers*. Alzheimer's Australia.
- Guadagnoli E and Velicer W (1988) *Relation of sample size to the stability of component patterns*. Psychological Bulletin. Vol. 103, No.2, pp.265-275.
- Gudex C, Kind P, van Dalen H, Durand M, Morris J and Williams A (1993) *Comparing Scaling Methods for Health State Valuations - Rosser Revisited*, Centre for Health Economics, University of York, York.
- Guttman L (1954) *The principal components of scalable attitudes*. In Lazarfeld P, (ed.) Mathematical Thinking in the Social Sciences. The Free Press, Glencoe, Illinois, pp.216-257.
- Guy W (ed) (1976) *Clinical Global Impressions (CGI)*. ECDEU Assessment Manual for Psychopharmacology, revised (DHEW Publ No ADM 76-338.) National Institute for Mental Health, Rockville, MD, pp. 218-222.
- Guyatt GH, Mitchell A, Molloy DW, et al. (1995) *Measuring patient and relative satisfaction with level or aggressiveness of care and involvement in care decisions in the context of life threatening illness*. Journal of Clinical Epidemiology. Vol. 48, No.10, pp.1215-24.
- Hall JN (1977) *The content of ward rating scales for long-stay patients*. Br. J Psychiatr. Vol. 130, pp.287-293.

- 
- Hall JN (1980) *Ward rating scales for long-stay patients*. Psychol.Med. Vol. 10, pp.277-288.
- Hall KM, Mann N, High WM, et al. (1996) *Functional measures after traumatic brain injury: ceiling effects of FIM, FIM+FAM, DRS, and CIQ*. J Head Trauma Rehabil. Vol. 11, No.5, pp.27-39.
- Hall KS, Ogunniyi AO, Hendrie HC, et al. (1996) *A cross-cultural community based study of dementias: methods and performance of the survey instrument in Indianapolis, U.S.A. and Ibadan, Nigeria*. International Psychogeriatrics. Vol. 6, pp.129-142.
- Hamilton BB, Granger CV, Sherwin FS, et al. (1987) *A uniform national data system for medical rehabilitation*. In Fuhrer M (ed.) Rehabilitation outcomes: analysis and measurement. P.H. Brookes, Baltimore. pp.137-147.
- Hamilton M (1959) *The assessment of anxiety states by rating*. British Journal of Medical Psychology. Vol. 32, No.1, pp.50-55.
- Hamilton M (1960) *A rating scale for depression*. Journal of Neurology, Neurosurgery and Psychiatry. Vol. 23, pp.56-62.
- Hancock K, Chang E, Chenoweth L, et al. (2003) *Nursing needs of acutely ill older people*. Journal of Advanced Nursing. Vol. 44, No.5, pp.507-16.
- Hannesdottir K and Snaedal J (2002) *A study of the Alzheimer's disease Assessment Scale-Cognitive (ADAS-Cog) in an Icelandic elderly population*. Nordic Journal of Psychiatry. Vol. 56, No.3, pp.201-6.
- Hanks GW, Robbins M, Sharp D, et al. (2002) *The imPaCT study: a randomised controlled trial to evaluate a hospitial palliative care team*. British Journal of Cancer. Vol. 87, No.7, pp.733-9.
- Hara I, Kawabata G, Miyake H, et al. (2003) *Comparison of quality of life following laparoscopic and open prostatectomy for prostate cancer*. J Urol. Vol. 169, No.6, pp.2045-8.
- Hardy G, West M and Hill F (1996) *Components and predictors of patient satisfaction*. British Journal of Health Psychology. Vol. 1, pp.65-85.
- Hare R, Rogers H, Lester H, et al. (2006) *What do stroke patients and their carers want from community services?* Family Practice. Vol. 23, No.1, pp.131-6.
- Hargrave R (2006) *Neuro-cognitive assessment of dementia in African American Elders*. In Yeo G, Gallagher-Thompson D, (eds). Ethnicity and the Dementias. 2nd ed. Routledge, Taylor and Francis Group, New York, pp.33-53.
- Harrell LE, Marson D, Chatterjee A and Parrish JA (2000) *The Severe Mini-Mental State Examination: a new neuropsychologic instrument for the bedside assessment of severely impaired patients with Alzheimer disease*. Alzheimer Disease and Associated Disorders. Vol. 14, No.3, pp.168-175.
- Haroutunian V, Perl DP, Purohit DP, et al. (1998) *Regional distribution of neuritic plaques in the nondemented elderly and subjects with very mild Alzheimer disease*. Archives of Neurology. Vol. 55, No.9, pp.1185-91.
- Harper GJ (2000) *Assessing older adults who cannot communicate*. In Kane RL, Kane RA, (eds). Assesing older persons: Measures, meaning, and practical applications. Oxford University Press, Oxford.
- Harris R, Ashton T, Broad J, et al. (2005) *The effectiveness, acceptability and costs of a hospital-at-home service compared with acute hospital care: a randomized controlled trial*. Journal of Health Services Research and Policy. Vol. 10, No.3, pp.158-66.
- Hartmaier SL, Sloane PD, Guess HA, et al. (1994) *The MDS Cognition Scale: a valid instrument for identifying and staging nursing home residents with dementia using the minimum data set.[see comment]*. Journal of the American Geriatrics Society. Vol. 42, No.11, pp.1173-9.
-

- Hartshorne TS (1993) *Psychometric properties and confirmatory factor analysis of the UCLA loneliness scale*. Journal of Personality Assessment. Vol. 61, No.1, pp.182-95.
- Harwood DG and Ownby RL (2000) *Ethnicity and dementia*. Current Psychiatry Reports. Vol. 2, No.1, pp.40-5.
- Harwood DG, Ownby RL, et al. (1998) *The behavioral pathology in Alzheimer's disease Scale (BEHAVE-AD): factor structure among community-dwelling Alzheimer's disease patients*. International Journal of Geriatric Psychiatry. Vol. 13, No.11, pp.793-800.
- Harwood DG, Sultzer DL, et al. (2000) *Impaired insight in Alzheimer disease: association with cognitive deficits, psychiatric symptoms, and behavioral disturbances*. Neuropsychiatry, Neuropsychology, and Behavioral Neurology. Vol. 13, No.2, pp.83-8.
- Harwood DG, Sultzer DL, et al. (2005) *Frontal lobe hypometabolism and impaired insight in Alzheimer disease*. Am J Geriatr Psychiatry. Vol. 13, No.11, pp.934-41.
- Harwood RH, Gompertz P and Ebrahim S (1994) *Handicap one year after stroke: validity of a new scale*. Journal of Neurology, Neurosurgery and Psychiatry. Vol. 57, No.7, pp.825-829.
- Havens B, Hall M, Sylvestre G, et al. (2004) *Social isolation and loneliness: differences between older rural and urban Manitobans*. Canadian Journal on Aging. Vol. 23, No.2, pp.129-40.
- Harvey P, Moriarty P, et al. (2005) *The Validation of a Caregiver Assessment of Dementia*. Alz Dis Ass Disord. Vol. 9, No.4, pp.186-194.
- Harvey PD, Moriarty PJ, Kleinman L, et al. (2005) *The validation of a caregiver assessment of dementia: the Dementia Severity Scale*. Alzheimer Disease and Associated Disorders. Vol. 19, No.4, pp.186-94.
- Hasselkus BR (1988) *Meaning in family caregiving: perspectives on caregiver/professional relationships*. The Gerontologist. Vol. 28, No.5, pp.686-91.
- Hassett A, George K and Harrigan S (1999) *Admissions of elderly patients from English-speaking and non-English-speaking backgrounds to an inpatient psychogeriatric unit*. Australian and New Zealand Journal of Psychiatry. Vol. 33, No.4, pp.576-82.
- Hatoum HT, Lin S-J, Arcona S, Thomas SK, Koumaris B and Mirski D (2005) *The use of the occupational disruptiveness scale of the neuropsychiatric inventory-nursing home version to measure the impact of rivastigmine on the disruptive behaviour of nursing home residents with Alzheimer's disease*. Journal of the American Medical Directors Association. Vol. 6, No.4, pp.238-45.
- Havranek EP and Steiner JF (2005) *Valuation of health states in the US versus the UK: two measures divided by a common language?* Medical Care. Vol. 43, No.3, pp.201-2.
- Hawes C, Morris JN, Phillips CD, et al. (1997) *Development of the Nursing Home Resident Assessment Instrument in the USA*. Age and Ageing. Vol. 26, No. Suppl 2, pp.19-25.
- Hawkey LC, Browne MW and Cacioppo JT (2005) *How can I connect with thee? Let me count the ways*. Psychological Science. Vol. 16, No.10, pp.798-804.
- Hawthorne G (2006) *Measuring social isolation in older adults: development and initial validation of the Friendship Scale*. Social Indicators Research. Vol. 77, No.3, pp.521-548.
- Hawthorne G (2003) *The effect of different methods of collecting data: mail, telephone and filter data collection issues in utility measurement*. Quality of Life Research. Vol. 12, pp.1081-1088.
- Hawthorne G (2006) *Review of Patient Satisfaction Measures*. Australian Government Department of Health and Ageing. pp.225-289.

- 
- Hawthorne G, Cheek F, Goldney R and Fisher L (2003a) *The excess cost of depression in South Australia: a comparative study of two methods of calculating burden*. Australian and New Zealand Journal of Psychiatry. Vol. 37, No.3, pp.362-373.
- Hawthorne G and Harmer C (2000) *GUTSS: The Genito-Urinary Treatment Satisfaction Scale Study*. Centre for Health Program Evaluation, Melbourne.
- Hawthorne G, Osborne R, McNeil H and Richardson J (1996) Centre for Health Program Evaluation, Melbourne.
- Hawthorne G, Osborne RH and Elliott P (2003b) *Commentary on: a psychometric analysis of the measurement level of the rating scale, standard gamble and time-trade off, by Cook et al*. Social Science and Medicine. Vol. 56, pp.895-897.
- Hawthorne G and Richardson J (1995) *An Australian MAU/QALY Instrument: Rationale and Preliminary Results*. Proceedings Health Outcomes and Quality of Life Measurement Conference, Canberra.
- Hawthorne G and Richardson J (2001) *Measuring the value of program outcomes: a review of utility measures*. Expert Review of Pharmacoeconomics Outcomes Research. Vol. 1, No.2, pp.215-228.
- Hawthorne G, Richardson J and Day N (2000a) *A comparison of five multi-attribute utility instruments*. Australian Health Economics Society Conference: Economics and Health: 2000 Proceedings of the Twenty-First Australian Conference of Health Economists, Gold Coast, Australia, pp.151-179.
- Hawthorne G, Richardson J and Day N (2000b) *A comparison of the Assessment of Quality of Life (AQoL) with four other generic utility instruments*. XII Medical Symposium "Quality of Life Measurement in Clinical Studies", Helsinki, Finland, pp.358-3760.
- Hawthorne G, Richardson J and Day N (2001a) *A comparison of five multi-attribute utility instruments*. Australian Studies in Health Services Administration. Vol. 89, pp.151-179.
- Hawthorne G, Richardson J and Day N and McNeil H (2000c) *Life and Death: Theoretical and Practical Issues in Utility Measurement*. Centre for Health Program Evaluation, Melbourne.
- Hawthorne, Richardson J, Osborne R, Day N and McNeil H (2000d) *Construction and Utility Scaling of the Assessment of Quality of Life (AQoL) Instrument*. Centre for Health Program Evaluation, Melbourne.
- Hawthorne G, Richardson J and Day NA (2001b) *A comparison of the Assessment of Quality of Life (AQoL) with four other generic utility instruments*. Annals of Medicine. Vol. 33, No.5, pp.358–370.
- Hawthorne G, Richardson J and Osborne R (1999) *The Assessment of Quality of Life (AQoL) instrument: a psychometric measure of health related quality of life*. Quality of Life Research. Vol. 8, pp.209-224.
- Hawthorne G, Richardson J, Osborne R and McNeil H (1997) *The Assessment of Quality of Life (AQoL) instrument: construction, initial validation and utility scaling*. Centre for Health Program Evaluation, Melbourne.
- Hawthorne G, Sansoni J, Hayes L, et al. (2006) *Measuring Patient Satisfaction with Incontinence Treatment*. Department of Psychiatry, The University of Melbourne, and the Centre for Health Service Development, University of Wollongong for the Commonwealth Department of Health and Ageing.
- Hawthorne L, Birrell B and Young D (2003) *The Retention of Overseas Trained Doctors in General Practice in Regional Victoria*. Faculty of Medicine, Dentistry and Health Sciences, University of Melbourne.
- Hayden KM, Khachaturian AS, Tschanz JT, Corcoran C, Nortond M, Breitner JCS, et al. (2003) *Characteristics of a two-stage screen for incident dementia*. Journal of Clinical Epidemiology. Vol. 56, No.11, pp.1038-1045.
- Hayes V, Morris J, Wolfe C and Morgan M (1995) *The SF-36 health survey questionnaire: Is it suitable for use with older adults?* Age and Aging, Vol. 24, pp 120-125.
-

- Hays RD, Davies AR and Ware JE (1987) *Scoring the Medical Outcomes Study Patient Satisfaction Questionnaire PSQ-III*. Medical Outcomes Trust. 18 p.
- Hays RD and DiMatteo MR (1987) *A short-form measure of loneliness*. Journal of Personality Assessment. Vol. 51, No.1, pp.69-81.
- Hector-Taylor L and Adams P (1996) *State versus trait loneliness in elderly New Zealanders*. Psychological Reports. Vol. 78, No.3, Pt.2, pp.1329-30.
- Heinemann AW (2005) *Putting outcome measurement in context: A rehabilitation psychology perspective*. Rehabilitation Psychology. Vol. 50, No.1, pp.6-14.
- Heinonen H, Volin L, Uutela A, et al. (2001a) *Gender-associated differences in the quality of life after allogeneic BMT*. Bone Marrow Transplant. Vol. 28, No.5, pp.503-9.
- Heinonen H, Volin L, Uutela A, et al. (2001b) *Quality of life and factors related to perceived satisfaction with quality of life after allogeneic bone marrow transplantation*. Annals of Hematology. Vol. 80, No.3, pp.137-43.
- Helmes E, Csapo KG and Short J-A (1987) *Standardization and validation of the Multidimensional Observation Scale for Elderly Subjects (MOSES)*. Journal of Gerontology. Vol. 42, pp.395-405.
- Helmstadter GC (1964) *Principles of Psychological Measurement*. Appleton-Century-Crofts, New York.
- Henderson AS, Jorm AF, Mackinnon A, Christensen H, Scott LR, Korten AE, et al. (1994) *A survey of dementia in the Canberra population: experience with ICD-10 and DSM-III-R criteria*. Psychological Medicine. Vol. 24, No.2, pp.473-482.
- Henderson C, Hales H and Ruggeri M (2003) *Cross-cultural differences in the conceptualisation of patients' satisfaction with psychiatric services--content validity of the English version of the Verona Service Satisfaction Scale*. Social Psychiatry and Psychiatric Epidemiology. Vol. 38, No.3, pp.142-8.
- Hendriks AA, Vrieling MR, Smets EM, et al. (2001) *Improving the assessment of (in)patients' satisfaction with hospital care*. Medical Care. Vol. 39, No.3, pp.270-83.
- Herrmann M, Ebert AD, Tober D, Hann J and Huth C (1999) *A contrastive analysis of release patterns of biochemical markers of brain damage after coronary artery bypass grafting and valve replacement and their association with the neurobehavioral outcome after cardiac surgery*. European Journal of Cardio-Thoracic Surgery. Vol. 16, No.5, pp.513-518.
- Herrmann W, Kurt S, Gaede K, Apeceche M (1997) *A multi-center Randomized Double-Blind Study on the Efficacy and Safety of Nicergoline in Patients with Multi-Infarct Dementia*. Dement Geriatr Cogn Disord. Vol. 8, pp.9-17.
- Herrman H, Hawthorne G and Thomas R (2002) *Quality of life assessment in people living with psychosis*. Social Psychiatry and Psychiatric Epidemiology. Vol. 37, No.11, pp.510-518.
- Herzog AR and Rodgers WL (1999) *Cognitive performance measures in survey research on older adults*. In Schwarz N, Park DC, Knauper B, Sudman S, (eds). Cognition, ageing, and self reports. Psychology Press, Taylor and Francis Group, Philadelphia.
- Heyman A, Wilkinson WE, Hurwitz BJ, et al. (1987) *Early-onset Alzheimer's disease: clinical predictors of institutionalization and death*. Neurology. Vol. 37, No.6, pp.980-4.  
<http://drugabusestatistics.samhsa.gov/SROS/sros8025.htm> provides method of administering DSRS.
- Hicks-Moore SL (2005) *Relaxing music at mealtime in nursing homes: effects on agitated patients with dementia*. Journal of Gerontological Nursing. Vol. 31, No.12, pp.26-32.
- Higgins M, Koch K, Hynan LS, Carr S, et al. (2005) *Impact of an activities-based adult dementia care program*. Neuropsychiatric Disease And Treatment. Vol. 1, No.2, pp.165-169.

- Hilton BA, Budgen C, Molzahn AE, et al. (2001) *Developing and testing instruments to measure client outcomes at the Comox Valley Nursing Center*. Public Health Nursing. Vol. 18, No.5, pp.327-39.
- Hindmarch I, Lehfeld H, de Jongh P, et al. (1998) *The Bayer Activities of Daily Living Scale (B-ADL)*. Dementia and Geriatric Cognitive Disorders. Vol. 9, No.Suppl 2, pp.20-26.
- Hirono M, Mori E, Ikejiri Y, et al. (1997) *Japanese version of the Neuropsychiatric Inventory- a scoring system for neuropsychiatric disturbance in dementia patients*. No To Shinkei. Vol. 49, No.3, pp.266-271.
- Hjermstad MJ, Fayers PM, Bjordal K, et al. (1998) *Health-related quality of life in the general Norwegian population assessed by the European Organization for Research and Treatment of Cancer Core Quality-of-Life Questionnaire: the QLQ=C30 (+ 3)*. Journal of Clinical Oncology. Vol. 16, No.3, pp.1188-96.
- Hobfoll SE and Walfisch S (1984) *Coping with a threat to life: a longitudinal study of self-concept, social support, and psychological distress*. American Journal of Community Psychology. Vol. 12, No.1, pp.87-100.
- Hodkinson M (1972) *Evaluation of a mental test score for assessment of mental impairment in the elderly*. Age and Ageing. Vol. 1, pp.233-238.
- Hoe J, Hancock G, Livingston G, et al. (2006) *Quality of life of people with dementia in residential care homes*. British Journal of Psychiatry. Vol. 188, pp.460-4.
- Hoe J, Katona C, Roch B, et al. (2005) *Use of the QOL-AD for measuring quality of life in people with severe dementia--the LASER-AD study*. Age and Ageing. Vol. 34, No.2, pp.130-5.
- Hogan A, Hawthorne G, Kethel L, Giles E, White K, Stewart M, Plath B and Code C (2001) *Health-related quality-of-life outcomes from adult cochlear implantation: a cross-sectional study*. Cochlear Implants International. Vol. 2, No.2, pp.115-128.
- Holcomb WR, Parker JC, Leong GB, et al. (1998) *Customer satisfaction and self-reported treatment outcomes among psychiatric inpatients*. Psychiatric Services. Vol. 49, No.7, pp.929-34.
- Holland R, Smith R, Harvey I, Swift L and Lenaghan E (2004) *Assessing quality of life in the elderly: a direct comparison of the EQ-5D and AQL*. Health Economics. Vol. 13, No.8, pp.793-805.
- Hollingworth W, Mackenzie R, Todd CJ and Dixon AK (1995) *Measuring changes in quality of life following magnetic resonance imaging of the knee: SF-36, EuroQol or Rosser index?* Quality of Life Research. Vol. 4, No.4, pp.325-34.
- Holford NH and Peace KE (1992) *Results and validation of a population pharmacodynamic model for cognitive effects in Alzheimer patients treated with tacrine*. Proceedings of the National Academy of Sciences of the United States of America. Vol. 89, No.23, pp.11471-5.
- Holmen K, Ericsson K and Winblad B (2000) *Social and emotional loneliness among non-demented and demented elderly people*. Archives of Gerontology and Geriatrics. Vol. 31, No.3, pp.177-192.
- Holmes C, Hopkins V, Hensford C, MacLaughlin V, et al. (2002) *Lavender oil as a treatment for agitated behaviour in severe dementia: a placebo controlled study*. International Journal of Geriatric Psychiatry. Vol. 17, No.4, pp.305-8.
- House J, Robbins C and Metzner H (1982) *The association of social relationships and activities with mortality: prospective evidence from the Tecumseh Community Health Study*. American Journal of Epidemiology. Vol. 116, No.1, pp.123-140.
- Howat A, Veitch C and Cairns W (2006) *A descriptive study comparing health attitudes of urban and rural oncology patients*. Rural and Remote Health. Vol. 6, No.4, pp.563.
- Howe AL (2006) *Cultural diversity, ageing and HACC: Trends in Victoria in the next 15 years*. Department of Human Services Victoria.

- 
- Hsu LR, Hailey BJ and Range LM (1987) *Cultural and emotional components of loneliness and depression*. Journal of Psychology. Vol. 121, No.1, pp.61-70.
- Hudak PL, Hogg-Johnson S, Bombardier C, et al. (2004) *Testing a new theory of patient satisfaction with treatment outcome*. Medical Care. Vol. 42, No.8, pp.726-39.
- Hughes CP, Berg L, Danziger WL, et al. (1982) *A new clinical scale for the staging of dementia*. British Journal of Psychiatry. Vol. 140, pp.566-72.
- Hughes ME, Waite LJ, Hawkey LC, et al. (2004) *A short scale for measuring loneliness in large surveys*. Research on Aging. Vol. 26, No.6, pp.655-672.
- Hunt CK (2003) *Concepts in Caregiver Research*. Journal of Nursing Scholarship. Vol. 35, No.1, pp.27-32.
- Hunt SM, McKenna SP, McEwen J, et al. (1981) *The Nottingham Health Profile: subjective health status and medical consultations*. Social Science and Medicine. Vol. 15A, No.3, Pt.1, pp.221-9.
- Hunter, E (2003) *Mental Health* in N Thomson (ed), The Health of Indigenous Australians, Oxford University Press, Melbourne, pp127-157.
- Hurtado B, Koepf MJ, Sander JW and Thompson PJ (2006) *The impact of levetiracetam on challenging behavior*. Epilepsy and Behavior. Vol. 8, No.3, pp.588-92.
- Hwang SS, Chang VT, Alejandro Y, et al. (2003) *Caregiver unmet needs, burden, and satisfaction in symptomatic advanced cancer patients at a Veterans Affairs (VA) medical center*. Palliative Support Care. Vol. 1, No.4, pp.319-29.
- Iezzoni LI (2002) *Toward universal design in assessing health care experiences*. Medical Care. Vol. 40, No.9, pp.725-728.
- Ihl R, Frolich L, Dierks T, et al. (1992) *Differential validity of psychometric tests in dementia of the Alzheimer type* Psychiatry Research. Vol. 44, No.2, pp.93-106.
- Ihl R, Grass-Kapanke B, Janner M, et al. (1999) *Neuropsychometric tests in cross sectional and longitudinal studies - a regression analysis of ADAS - cog, SKT and MMSE*. Pharmacopsychiatry. Vol. 32, No.6, pp.248-54.
- Ikeda M, Fukuhara R, Shigendobu K, Hokoishi NM, Nebu A, Komori K and Tanabe H (2007) *Dementia associated mental and behavioural disturbances in elderly people in the community: findings from the first Nakayama study*. Journal Neral. Neurosurg. Psychiatry. Vol. 200, No.75, pp.146-148, available at [www.jnmp.com](http://www.jnmp.com), accessed 8 March, 2007.
- Ikeda S, Niwata S and Igarashi Y (2001) *Evaluating QoL of elderly people with dementia*. Japanese Journal of Pharmacoepidemiology. Vol. 5, pp.99-105.
- Ikeda S, Yamada Y and Ikegami N (2002) *Economic evaluation of donepezil treatment for Alzheimer's disease in Japan*. Dementia and Geriatric Cognitive Disorders. Vol. 13, No.1, pp.33-9.
- Iliffe S and Manthorpe J (2004) *The hazards of early recognition of dementia: a risk assessment*. Ageing and Mental Health. Vol. 8, No.2, pp.99-105.
- Iliffe S and Manthorpe J (2004) *The debate on ethnicity and dementia: from category fallacy to person-centred care?* Ageing and Mental Health. Vol. 8, No.4, pp.283-292.
- Iliffe S, Haines A, Stein A, et al. (1991) *Assessment of elderly people in general practice. 3. Confiding relationships*. British Journal of General Practice. Vol. 41, No.352, pp.459-61.
- Imbimbo BP, Troetel WM, Martell IP, et al. (2000) *A 6-month, double-blind, placebo-controlled trial of eptastigmine in Alzheimer's disease*. Dementia and Geriatric Cognitive Disorders. Vol. 11, No.1, pp.17-24.
-

---

Ingleton C, Morgan J, Hughes P, et al. (2004) *Carer satisfaction with end-of-life care in Powys, Wales: a cross-sectional survey*. Health and Social Care in the Community. Vol. 12, No.1, pp.43-52.

Inouye S (2003) *The Confusion Assessment Method (CAM): Training Manual and Coding Guide*, University School of Medicine, Yale.

Inouye SK, Foreman MD, Mion LC, Katz KH and Cooney LM Jr (2001) *Nurses' recognition of delirium and its symptoms: comparison of nurse and researcher ratings*. Archives of Internal Medicine. Vol. 161, No.20, pp.2467-2473.

Inouye SK, van Dyck CH, Alessi CA, Balkin S, Siegel AP and Horwitz RI (1990) *Clarifying confusion: the confusion assessment method. A new method for detection of delirium*. Annals of Internal Medicine. Vol. 113, No.12, pp.941-948.

IPA (International Psychogeriatric Association) (2003) *The IPA Behavioural and Psychological Symptoms of Dementia*, available at <http://www.ipa-online.net/pdfs/bpsd.asp>, accessed 1 June 2006.

Iverson GL, Hopp GA, de Wolfe K and Solomons K (2002) *Measuring change in psychiatric symptoms using the Neuropsychiatric Inventory: nursing home version*. Intern J Geriatric Psychiatry. Vol. 17, pp.438-443.

Jackson JL and Kroenke K (2001) *The effect of unmet expectations among adults presenting with physical symptoms*. Annals of Internal Medicine. Vol. 134, No.9 Pt 2, pp.889-97.

Jacobs D, Sano M, Marder K, Bell K, Bylsma F, Laflesce G, Albert M, Brandt J and Stern Y (1994) *Age at onset of Alzheimer's disease: relation to pattern of cognitive dysfunction and rate of decline*. Neurology. Vol. 44, No.7, pp.1215-20.

Jacobs JW, Bernhard MR, Delgado A, et al. (1977) *Screening for organic mental syndrome in the medically ill*. Annals of Internal Medicine. Vol. 86, pp.40-46.

Jacobs MR, Strauss ME, Patterson MB and Mack JL (1998) *Characterization of depression in Alzheimer's disease by the CERAD Behavior Rating Scale for Dementia (BRSD)*. American Journal of Geriatric Psychiatry. Vol. 6, No.1, pp.53-8.

Jacoby A, Lecouturier J, Bradshaw C, et al. (1999) *Feasibility of using postal questionnaires to examine carer satisfaction with palliative care: a methodological assessment*. Palliative Medicine. Vol. 13, No.4, pp.285-98.

Jaglal S, Cockerill R, Lemieux-Charles L, et al. (2007) *Perceptions of the process of care among caregivers and care recipients in dementia care networks*. American Journal of Alzheimer's Disease and Other Dementias. Vol. 22, No.2, pp.103-11.

James G (1967) *The medical school of the future and its role in the community*. JAMA. Vol. 202, No.5, pp.415-8.

Jarvis H, Burge FI and Scott CA (1996) *Evaluating a palliative care program: methodology and limitations*. Journal of Palliative Care. Vol. 12, No.2, pp.23-33.

Jenkins C (1992) *Assessment of outcomes of health intervention*. Social Science and Medicine. Vol. 35, pp.367-375.

Jette AM (1980) *Functional capacity evaluation: an empirical approach*. Archives of Physical Medicine and Rehabilitation. Vol. 61, pp.85-89.

Jette AM, Keysor J, Coster WJ, et al. (2005) *Beyond function: predicting participation in a rehabilitation cohort*. Archives of Physical Medicine and Rehabilitation. Vol. 86, pp.2087-2094.

Johannsson HO, Graf W and Pahlman L (2002) *Long-term results of haemorrhoidectomy*. European Journal of Surgery. Vol. 168, No.8-9, pp.485-9.

- Johnson JA, Luo N, Shaw JW, Kind P and Coons SJ (2005) *Valuations of EQ-5D health states: are the United States and United Kingdom different?* Medical Care. Vol. 43, No.3, pp.221-8.
- Johnson JE (1987) *Selecting nursing activities for hospitalized clients*. Journal of Gerontological Nursing. Vol. 13, No.10, pp.29-33.
- Johnson N, Barion A, Rademaker A, et al. (2004) *The Activities of Daily Living Questionnaire: A validation study in patients with dementia*. Alzheimer Disease and Associated Disorders. Vol. 18, No.4, pp.223-230.
- Jones TG, Schinka JA, Vanderploeg RD, Small BJ, Graves AB and Mortimer JA (2002) *3MS normative data for the elderly*. Archives of Clinical Neuropsychology. Vol. 17, pp.171-177.
- Jones RN, Yang FM, Zhang Y, Kiely DK, Marcantonio ER and Inouye SK (2006) *Does educational attainment contribute to risk for delirium? A potential role for cognitive reserve*. Journals of Gerontology Series A-Biological Sciences and Medical Sciences. Vol. 61, No.12, pp.1307-1311.
- Jones RW, Soininen H, Hager K, et al. (2004) *A multinational, randomised, 12-week study comparing the effects of donepezil and galantamine in patients with mild to moderate Alzheimer's disease*. International Journal of Geriatric Psychiatry. Vol. 19, No.1, pp.58-67.
- Jongenelis K, Pot AM, Eisses AMH, et al. (2004) *Prevalence and risk indicators of depression in elderly nursing home patients: The AGED study*. Journal of Affective Disorders. Vol. 83, No.2-3, pp.135-142.
- Jonsen AR (1986) *Bentham in a box: technology assessment and health care allocation*. Law, Medicine and Health Care, Vol. 14, No.3-4, pp.172-4.
- Jonsson L (2003) *Pharmacoeconomics of cholinesterase inhibitors in the treatment of Alzheimer's disease*. Pharmacoeconomics. Vol. 21, No.14, pp.1025-37.
- Jonsson L (2005) *Cost-effectiveness of memantine for moderate to severe Alzheimer's disease in Sweden*. American Journal of Geriatric Pharmacotherapy. Vol. 3, No.2, pp.77-86.
- Jonsson L, Andreasen N, Kilander L, Soininen H, Waldemar G, Nygaard H, Winblad B, Jonhagen ME, Hallikainen M and Wimo A (2006) *Patient- and proxy-reported utility in Alzheimer disease using the EuroQoL*. Alzheimer Disease and Associated Disorders. Vol. 20, No.1, pp.49-55.
- Jorm AF (2004) *The Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE): A review*. International Psychogeriatrics. Vol. 16, No.3, pp.275-293.
- Jorm AF and Jacomb PA (1989) *An Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE): Socio-demographic correlates, reliability, validity and some norms*. Psychological Medicine. Vol. 19, pp.1015-1022.
- Kahn RL (1979) *Aging and social support*. In Riley MW (ed.) Aging from Birth to Death: Interdisciplinary Perspectives. Boulder: Westview Press. pp.77-91.
- Kalisvaart KJ, de Jonghe JFM, Bogaards MJ, Vreeswijk R, Egberts TCG, Burger BJ, et al. (2005) *Haloperidol prophylaxis for elderly hip-surgery patients at risk for delirium: a randomized placebo-controlled study*. Journal of the American Geriatrics Society. Vol. 53, No.10, pp.1658-1666.
- Kane RL and Kane RA (2000) *Assessing older persons: measures, meaning and practical applications*. Oxford University Press, USA.
- Kane RL, Maciejewski M and Finch M (1997) *The relationship of patient satisfaction with care and clinical outcomes*. Medical Care. Vol. 35, No.7, pp.714-30.
- Kaplan R, Anderson J and Ganiats T (1993) *The Quality of Well-being Scale: rationale for a single quality of life index*. In Walker, S and Rosser, R (eds.) (1993). Quality of Life Assessment: Key Issues in the 1990s. Kluwer Academic Publishers, Dordrecht.

- Kaplan R, Bush J and Berry C (1976) *Health status: types of validity and the Index of Well-being*. Health Services Research. Vol. 11, No.4, pp.478-507.
- Kaplan R, Bush J and Berry C (1976) *The reliability, stability and generalizability of a health status index*. Proceedings of the Social Statistics Section. pp.704-709.
- Kaplan R, Ganiats T, Sieber W and Anderson J (1996a) *The Quality of Well-being Scale*. Medical Outcomes Trust Bulletin. Vol. 4, pp. 2-3.
- Kaplan RM (1993) *Quality of life assessment for cost/utility studies in cancer*. Cancer Treat Rev. Vol. 19, Suppl A, pp.85-96.
- Kaplan RM, Alcaraz JE, Anderson JP and Weisman M (1996b) *Quality-adjusted life years lost to arthritis: effects of gender, race, and social class*. Arthritis Care Res. Vol. 9, No.6, pp.473-82.
- Kaplan RM and Anderson JP (1988) *A general health policy model: update and applications*. Health Services Research. Vol. 23, No.2, pp.203-35.
- Kaplan RM, Anderson JP, Patterson TL, McCutchan JA, Weinrich JD, Heaton RK, Atkinson J H, Thal L, Chandler J, Grant I and HNRC Group (1995) *Validity of the Quality of Well-Being Scale for persons with human immunodeficiency virus infection*. Psychosomatic Medicine. Vol. 57, No.2, pp.138-47.
- Kaplan RM, Ganiats TG, Sieber WJ and Anderson JP (1998) *The Quality of Well-Being Scale: critical similarities and differences with SF-36*. International Journal for Quality in Health Care. Vol. 10, No.6, pp.509-20.
- Karagiozis H, Gray S, Sacco J, et al. (1998) *The Direct Assessment of Functional Abilities (DAFA): A comparison to an indirect measure of instrumental activities of daily living*. Gerontologist. Vol. 38, No.1, pp.113-121.
- Karlawish J, Casarett D, et al. (2003) *Why Would Caregivers not want to treat their relative's alzheimer's disease?* Journal of the American Geriatrics Society. Vol. 51, pp.1391-1397.
- Karlidag R, Unal S, Sezer OH, Bay Karabulut A, Battalolu B, But A, et al. (2006) *The role of oxidative stress in postoperative delirium*. General Hospital Psychiatry. Vol. 28, No.5, pp.418-423.
- Kastango KB, Kim Y, et al. (2002) *Verification of scale sub-domains in elderly patients with dementia: a confirmatory factor-analytic approach*. American Journal of Geriatric Psychiatry. Vol. 10, No.6, pp.706-14.
- Kat MG and de Jongle IF (2002) *Neuropsychiatric symptoms of dementia: psychometric properties of the Dutch Neuropsychiatric Inventory (NPI)*. Tijdshef Gerontolog Geriatric. Vol. 33, No.4, pp.150-155.
- Katona C, Livingston G, Cooper C, Ames D, Brodaty H and Chui E (2007) *International Psychogeriatric Association statement on defining and measuring treatment benefits in dementia*. International Psychogeriatrics. Doi: 10.1017/S1041610207005145.
- Katona CL and Aldridge CR (1985) *The dexamethasone suppression test and depressive signs in dementia*. Journal of Affective Disorders. Vol. 8, No.1, pp.83-9.
- Katz S, Ford AB, Moskowitz RW, et al. (1963) *Studies of illness in the aged. The index of ADL: A standardized measure of biological and psychological function*. JAMA. Vol. 185, No.12, pp.914-919.
- Katzman R (1986) *Differential diagnosis of dementing illness*. Neurol Clin. Vol. 4, pp.329-340.
- Kaufers DL, Cummings JL and Christine D (1998) *Differential neuropsychiatric symptom response in Tacrine in Alzheimer's disease: relationship to symptom severity*. J Neuropsychiatry Clin Neurosci. Vol. 10, pp.55-63.
- Kaufers DL, Cummings JL, Christine D, Bray T, Castellon S, Masterman D, MacMillan A, Ketchel P and Dekovsky ST (1998) *Assessing the impact of neuropsychiatric symptoms in Alzheimer's disease; the*

- Neuropsychiatric Inventory Caregiver Distress Scale*. Journal of the American Geriatric Society. Vol. 46, pp.210-215.
- Kawachi I, Colditz G, Ascherio A, et al. (1996) *A prospective study of social networks in relation to total mortality and cardiovascular disease in men in the USA*. Journal of Epidemiology and Community Health. Vol. 50, No.3, pp.245-251.
- Kaye JA (1998) *Diagnostic challenges in dementia*. Neurology. Vol. 51, No.1, Suppl.1, pp.S45-52; discussion pp.S65-47.
- Kealey P and McIntyre I (2005) *An evaluation of the domiciliary occupational therapy service in palliative cancer care in a community trust: a patient and carers perspective*. European Journal of Cancer Care (English Language Edition). Vol. 14, No.3, pp.232-43.
- Kef S (2002) *Psychosocial adjustment and the meaning of social support for visually impaired adolescents*. Journal of Visual Impairment and Blindness. Vol. 96, No.1, pp.22-37.
- Kelly G, Todd J, et al. (2006) *The Overt Behaviour Scale (OBS): a tool for measuring challenging behaviours following ABI in community settings*. Brain Injury. Vol. 20, No.3, pp.307-19.
- Kelleher M and Mannix M (2001) *Carers' views on services for the elderly mentally ill*. Irish Journal of Medical Science. Vol. 170, No.4, pp.239-43.
- Kemp NM, Brodaty H, Pond D, et al. (2002) *Diagnosing dementia in primary care: The accuracy of informant reports*. Alzheimer Disease and Associated Disorders. Vol. 16, No.3, pp.171-176.
- Kerner DN, Patterson TL, Grant I and Kaplan RM (1998) *Validity of the Quality of Well-Being Scale for patients with Alzheimer's disease*. Journal of Ageing and Health. Vol. 10, No.1, pp.44-61.
- Kertesz A, Davidson W and Fox H (1997) *Frontal behavioral inventory: diagnostic criteria for frontal lobe dementia*. Can J Neurol Sci. Vol. 24, pp.29-36.
- Kessler D, Peters TJ, Lee L, et al. (2005) *Social class and access to specialist palliative care services*. Palliative Medicine. Vol. 19, No.2, pp.105-10.
- Khachaturian AS, Gallo JJ and Breitner JC (2000) *Performance characteristics of a two-stage dementia screen in a population sample*. Journal of Clinical Epidemiology. Vol. 53, No.5, pp.531-540.
- Kiely DK, Volicer L, Teno J, et al. (2006) *The validity and reliability of scales for the evaluation of end-of-life care in advanced dementia*. Alzheimer Disease and Associated Disorders. Vol. 20, No.3, pp.176-181.
- Kikano GE, Goodwin MA and Stange KC (1998) *Physician employment status and practice patterns*. Journal of Family Practice. Vol. 46, No.6, pp.499-505.
- Kim O (1999a) *Mediation effect of social support between ethnic attachment and loneliness in older Korean immigrants*. Research in Nursing and Health. Vol. 22, No.2, pp.169-75.
- Kim O (1999b) *Predictors of loneliness in elderly Korean immigrant women living in the United States of America*. Journal of Advanced Nursing. Vol. 29, No.5, pp.1082-8.
- Kim YS, Nibbelink DW and Overall JE (1994) *Factor structure and reliability of the Alzheimer's disease Assessment Scale in a multicenter trial with linopirdine*. Journal of Geriatric Psychiatry and Neurology. Vol. 7, No.2, pp.74-83.
- Kind P (1996) *The EuroQol Instrument: An index of health-related quality of life*. In Spilker, B (ed.) Quality of Life and Pharmacoeconomics in Clinical Trials. Lippincott-Raven Publishers, Philadelphia, pp.191-201.
- Kind P and Gudex C (1994) *Measuring health status in the community: a comparison of methods*. Journal of Epidemiology and Community Health. Vol. 48, pp.86-91.

---

Kinnersley P, Stott N, Peters T, et al. (1996) *A comparison of methods for measuring patient satisfaction with consultations in primary care.* Family Practice. Vol. 13, No.1, pp.41-51.

Knapp MJ, Knopman DS, Solomon PR, et al. (1994) *A 30-week randomized controlled trial of high-dose tacrine in patients with Alzheimer's disease. The Tacrine Study Group.* JAMA. Vol. 272, pp.985-991.

Klimidis S, Reddy P, Minas IH, et al. (2004) *Brief functional English proficiency measure for health survey research.* Australian Psychologist. Vol. 39, No.2, pp.154-165.

Knopman DS, Berg JD, Thomas R, Grundman M, et al. (1999) *Nursing home placement is related to dementia progression: experience from a clinical trial. Alzheimer's disease Cooperative Study.* Neurology. Vol. 52, No.4, pp.714-8.

Knopman DS, DeKosky ST, Cummings JL, Chui HC, Corey-Bloom J, Relkin N et al. (2001) *Practice parameter: Diagnosis of dementia (An evidence-based review). Report of the Quality Standards Subcommittee of the American Academy of Neurology.* Neurology. Vol. 56, pp.1143-1153.

Koffman J and Higginson IJ (2001) *Accounts of carers' satisfaction with health care at the end of life: a comparison of first generation black Caribbeans and white patients with advanced disease.* Palliative Medicine. Vol. 15, No.4, pp.337-45.

Kolibas E, Korinkova V, Novotny V, et al. (2000) *ADAS-cog (Alzheimer's disease Assessment Scale-cognitive subscale)--validation of the Slovak version.* Bratislavske Lekarske Listy. Vol. 101, No.11, pp.598-602.

Kopec JA and Willison KD (2003) *A comparative review of four preference-weighted measures of health-related quality of life.* Journal of Clinical Epidemiology. Vol. 56, No.4, pp.317-25.

Kornblith AB, Herndon JE, 2nd, Zuckerman E, et al. (2001) *Social support as a buffer to the psychological impact of stressful life events in women with breast cancer.* Cancer. Vol. 91, No.2, pp.443-54.

Koss E, Weiner M, Ernesto C, Cohen-Mansfield J, et al. (1997) *Assessing patterns of agitation in Alzheimer's disease patients with the Cohen-Mansfield Agitation Inventory. The Alzheimer's disease Cooperative Study.* Alzheimer Disease and Associated Disorders. Vol. 11 Suppl.2, pp.S45-50.

Kremers IP, Steverink N, Albersnagel FA, et al. (2006) *Improved self-management ability and well-being in older women after a short group intervention.* Aging and Mental Health. Vol. 10, No.5, pp.476-484.

Kring A, Kerr S, Smith D and Neale J (1993) *Flat affect in schizophrenia does not reflect diminished subjective experience of emotion.* Journal of Abnormal Psychology. Vol. 102, pp.507-517

Kristjansson B, Breithaupt K and McDowell I (2001) *Development and validation of an indicator of support for community-residing older Canadians.* Int Psychogeriatr. Vol. 13 Suppl.1, pp.125-35.

Kristjanson LJ (1989) *Quality of terminal care: salient indicators identified by families.* Journal of Palliative Care. Vol. 5, No.1, pp.21-30.

Kristjanson LJ (1993) *Validity and reliability testing of the FAMCARE Scale: measuring family satisfaction with advanced cancer care.* Social Science and Medicine. Vol. 36, No.5, pp.693-701.

Kristjanson LJ, Leis A, Koop PM, et al. (1997) *Family members' care expectations, care perceptions, and satisfaction with advanced cancer care: results of a multi-site pilot study.* Journal of Palliative Care. Vol. 13, No.4, pp.5-13.

Kunik ME, Graham DP, Snow-Turek AL, Molinari VA, et al. (1998) *Contribution of cognitive impairment, depression, and psychosis to the outcome of agitated geropsychiatric inpatients with dementia.* Journal of Nervous and Mental Disease. Vol. 186, No.5, pp.299-303.

Kuriansky JB and Gurland B (1976) *Performance test of activities of daily living.* International Journal of Aging and Human Development. Vol. 7, pp.343-352.

---

- Kurita A, Katayama K, et al. (1997) *Relationship between cognitive deficits, behavioral disturbances and falls in patients with dementia*. Nippon Ronen Igakkai Zasshi. Vol. 34, No.8, pp.662-7.
- La Monica EL, Oberst MT, Madea AR, et al. (1986) *Development of a patient satisfaction scale*. Research in Nursing and Health. Vol. 9, No.1, pp.43-50.
- Labarere J, Fourny M, Jean-Phillippe V, et al. (2004) *Refinement and validation of a French in-patient experience questionnaire*. International Journal of Health Care Quality Assurance. Vol. 17, No.1, pp.17-25.
- Lam LC, Lui VW, Tam CW, et al. (2005) *Subjective memory complaints in Chinese subjects with mild cognitive impairment and early Alzheimer's disease*. International Journal of Geriatric Psychiatry. Vol. 20, No.9, pp.876-82.
- Lam CL, Chan WC, Mok CC, Li SW, et al. (2006) *Validation of the Chinese Challenging Behaviour Scale: clinical correlates of challenging behaviours in nursing home residents with dementia*. International journal of geriatric psychiatry. Vol. 21, No.8, pp.792-9.
- Lam LC, Chui HF and Ng J (1997) *Aggressive behaviour in the Chinese elderly-validation of the Chinese version of the Rating Scale for Aggressive Behavior in the Elderly (RAGE) in hospital and nursing home settings*. International Journal of Geriatric Psychiatry. Vol. 12, No.6, pp.678-681.
- Lam LC, Chiu HF, Ng KO, et al. (1998) *Clock-face drawing, reading and setting tests in the screening of dementia in Chinese elderly adults*. Journals of Gerontology Series B-Psychological Sciences and Social Sciences. Vol. 53, No.6, pp.P353-7.
- Lampléy-Dallas VT (2002) *Research issues for minority dementia patients and their caregivers: what are the gaps in our knowledge base?* Alzheimer Disease and Associated Disorders. Vol. 16 Suppl 2, pp.S46-9.
- Landes A, Sperry S, Strauss M (2005) *Prevalence of Apathy, Dysphoria and Depression in Relation to Dementia Severity in Alzheimer's disease*. J Neuropsychiatry Clin Neurosci. Vol.17, pp.342-349.
- Lang AJ and Stein MB (2001) *Anxiety Disorders: How to recognize and treat the medical symptoms of emotional illness*. Geriatrics. Vol. 56, No.5, pp.24-34.
- Langa KM, Chernew ME, Kabeto MU, Herzog AR, et al. (2001) *National estimates of the quantity and cost of informal caregiving for the elderly with dementia*. Journal of General Internal Medicine. Vol. 16, No.11, pp.770-8.
- Lange RT, Hopp G and Kang N (2004) *Psychometric properties and factors structure of the Neuropsychiatric Inventory Nursing Home version in an elderly neuropsychiatric population*. International Journal of Geriatric Psychiatry. Vol. 19, pp.440-448.
- Larsen D and Rootman I (1976) *Physician role performance and patient satisfaction*. Social Science and Medicine. Vol. 10, pp.29-32.
- Larsen DL, Attkisson CC, Hargreaves WA, et al. (1979) *Assessment of client/patient satisfaction: development of a general scale*. Evaluation and Program Planning. Vol. 2, No.3, pp.197-207.
- Lauderdale SA and Sheikh JI (2003) *Anxiety disorders in older adults*. Clinics in Geriatric Medicine. Vol. 19, No.4, pp.721-741.
- Laurila JV, Pitkala KH, Strandberg TE and Tilvis RS (2002) *Confusion assessment method in the diagnostics of delirium among aged hospital patients: would it serve better in screening than as a diagnostic instrument?* International Journal of Geriatric Psychiatry. Vol. 17, No.12, pp.1112-1119.
- Laurila JV, Pitkala KH, Strandberg TE and Tilvis RS (2004) *Delirium among patients with and without dementia: does the diagnosis according to the DSM-IV differ from previous classifications?* International Journal of Geriatric Psychiatry. Vol. 19, pp.271-277.
- Lawton MP (1994) *Quality of life in Alzheimer disease*. Alzheimer Disease and Associated Disorders. Vol. 8 Suppl.3, pp.138-50.

- Lawton MP, Kleban MH, Moss M, Rovine M, et al. (1989) *Measuring caregiving appraisal*. Journal of Gerontology. Vol. 44, No.3, pp.61-71.
- Lawton M and Brody E (1969) *Assessment of older people: Self-maintaining and instrumental activities in daily living*. Gerontologist. Vol. 9, pp.179-186.
- Lechowski L, De Stampa M, Tortrat D, et al. (2005) *Predictive factors of rate of loss of autonomy in Alzheimer's disease patients. A prospective study of the REAL.FR cohort*. Journal of Nutrition, Health and Ageing. Vol. 9, No.2, pp.100-104.
- Lee DY, Choo IH, Kim KW, Jhoo JH, et al. (2006) *White matter changes associated with psychotic symptoms in Alzheimer's disease patients*. Journal of Neuropsychiatry and Clinical Neurosciences. Vol. 18, No.2, pp.191-8.
- Lee KU, Won WY, Lee HK, Kweon YS, Lee CT, Pae CU, et al. (2005) *Amisulpride versus quetiapine for the treatment of delirium: a randomized, open prospective study*. International Clinical Psychopharmacology. Vol. 20, No. 6, pp.311-314.
- Lehto-Jarnstedt US, Ojanen M and Kellokumpu-Lehtinen P (2004) *Cancer-specific social support received by newly diagnosed cancer patients: validating the new Structural-Functional Social Support Scale (SFSS) measurement tool*. Support Care Cancer. Vol. 12, No.5, pp.326-37.
- Lemiengre J, Nelis T, Joosten E, Braes T, Foreman M, Gastmans C, et al. (2006) *Detection of delirium by bedside nurses using the confusion assessment method*. Journal of the American Geriatrics Society. Vol. 54, No.4, pp.685-689.
- Lester D and Yang B (1992) *Social and economic correlates of the elderly suicide rate*. Suicide and Life-Threatening Behavior. Vol. 22, No.1, pp.36-47.
- Levin C (2000) *Social functioning*. In Kane RL, Kane RA (eds.) Assessing Older Persons. Oxford University Press, Oxford. pp.170-199.
- Levin HS, High WM, et al. (1987) *The neurobehavioural rating scale: assessment of the behavioural sequelae of head injury by the clinician*. Journal of Neurology, Neurosurgery and Psychiatry. Vol. 50, No.2, pp. 183-93.
- Levy MH, Miller BL, Cummings JL, Fairbanks LAS and Craig A (1996) *Alzheimer's Disease and frontotemporal dementias: behavioural distinctions*. Arch. Neurol. Vol. 53, pp.687-690.
- Lewis LM, Miller DK, Morley JE, Nork MJ and Lasater LC (1995) *Unrecognized delirium in ED geriatric patients*. American Journal of Emergency Medicine. Vol. 13, No.2, pp.142-145.
- Lezak MD (2004) *Neuropsychological Assessment* (4<sup>th</sup> ed). Oxford University Press, New York.
- Li M, Ng TP, Kua EH, et al. (2006) *Brief Informant Screening Test for Mild Cognitive Impairment and Early Alzheimer's Disease*. Dementia and Geriatric Cognitive Disorders. Vol. 21, No.5-6, pp.392-402.
- Likert R (1932) *A technique for the measurement of attitudes*. Archives of Psychology. Vol. 140, pp.3-55.
- Likert R, Roslow S and Murphy G (1934) *A simple and reliable method of scoring the Thurstone attitude scales*. Journal of Social Psychology. Vol. 5, pp.228-238.
- Lincoln Centre for Ageing and Community Care Research (2004) *The review and identification of an existing, validated, comprehensive assessment tool*. Australian Institute for Primary Care, La Trobe University.
- Lincoln NB, Francis VM, Lilley SA, et al. (2003) *Evaluation of a stroke family support organiser: a randomized controlled trial*. Stroke. Vol. 34, No.1, pp.116-21.

- Lincoln NB, Walker MF, Dixon A, et al. (2004) *Evaluation of a multiprofessional community stroke team: a randomized controlled trial*. Clinical Rehabilitation. Vol. 18, No.1, pp.40-7.
- Lindeboom R, Vermeulen M, Holman R, et al. (2003) *Activities of daily living instruments: Optimizing scales for neurologic assessments*. Neurology. Vol. 60, pp.738-742.
- Linder-Pelz S (1982) *Toward a theory of patient satisfaction*. Social Science and Medicine. Vol. 16, pp.577-582.
- Linder-Pelz S and Struening EL (1985) *The multidimensionality of patient satisfaction with a clinic visit*. Journal of Community Health. Vol. 10, No.1, pp.42-54.
- Lindsey I, Smilgin-Humphreys MM, Cunningham C, et al. (2002) *A randomized, controlled trial of fibrin glue vs. conventional treatment for anal fistula*. Dis Colon Rectum. Vol. 45, No.12, pp.1608-15.
- Linn M and Linn B (1982) *The Rapid Disability Rating Scale - 2*. JAGS. Vol. 30, No.6, pp.378-382.
- Lippert-Gruner M, Kuchta K, et al. (2006) *Neurobehavioural deficits after severe traumatic brain injury (TBI)*. Brain Injury. Vol. 20, No.6, pp.569-74.
- Lister S and Benson C (2006) *Comparative analysis of dementia and ethnicity in the New South Wales Aged Care Assessment Program: 1996 and 2001*. Australasian Journal on Ageing. Vol. 25, No.1, pp.24-30.
- Litvan I, Mega MS, Cummings JL and Fairbanks L (1996) *Neuropsychiatric aspects of progressive supranuclear palsy*. Neurology. Vol. 47, pp.1184-1189.
- Liu HC, Teng EL, Chuang YY, et al. (2002) *The Alzheimer's disease Assessment Scale: findings from a low-education population*. Dementia and Geriatric Cognitive Disorders. Vol. 13, No.1, pp.21-6.
- Locker D and Dunt D (1978) *Theoretical and methodological issues in sociological studies of consumer satisfaction with medical care*. Social Science and Medicine. Vol. 12, pp.283-292.
- Loewenstein DA, Amigo E, Duara R, et al. (1989) *A new scale for the assessment of functional status in Alzheimer's disease and related disorders*. Journal of Gerontology. Vol. 44, No.4, pp.114-121.
- LoGiudice D, Hassett A, Cook R, et al. (2001) *Equity of access to a memory clinic in Melbourne? Non-English speaking background attenders are more severely demented and have increased rates of psychiatric disorders*. International Journal of Geriatric Psychiatry. Vol. 16, No.3, pp.327-34.
- LoGiudice D, Smith K, Thomas J, et al. (2005) *Kimberley Indigenous Cognitive Assessment tool (KICA): Development of a cognitive assessment tool for older indigenous Australians*. International Psychogeriatrics. Vol. 18, No.2, pp.269-280.
- Logsdon RG, Gibbons LE, McCurry SM, et al. (1999) *Quality of life in Alzheimer's disease: Patient and caregiver reports*. Journal of Mental Health and Aging. Vol. 5, No.1, pp.21-32.
- Logsdon RG, Gibbons LE, McCurry SM, et al. (2002) *Assessing quality of life in older adults with cognitive impairment*. Psychosomatic Medicine. Vol. 64, No.3, pp.510-9.
- Logsdon RG, Teri L, Weiner MF, Gibbons LE, et al. (1999) *Assessment of agitation in Alzheimer's disease: the agitated behavior in dementia scale*. Alzheimer's disease Cooperative Study. Journal of the American Geriatrics Society. Vol. 47, No.11, pp.1354-8.
- Looi JC and Sachdev PS (1999) *Differentiation of vascular dementia from AD on neuropsychological tests*. Neurology. Vol. 53 No.4, pp.670-678.
- Lopez OL, Kuller LH, Becker JT, Jagust WJ, DeKosky ST, Fitzpatrick A, et al. (2005) *Classification of vascular dementia in the Cardiovascular Health Study Cognition Study*. Neurology. Vol. 64, No.9, pp.1539-1547.

- 
- Lopez OL, Becker JT and Sweet RA (2005) *Non-cognitive symptoms in mild cognitive impairment subjects*. Neurocase. Vol. 11, No.1, pp.65-71.
- Lord SR and Menz HB (2002) *Physiologic, psychologic, and health predictors of 6-minute walk performance in older people*. Archives of Physical Medicine and Rehabilitation. Vol. 83, No.7, pp.907-911.
- Lorentz WJ, Scanlan JM and Borson S (2002) *Brief screening tests for dementia*. Canadian Journal of Psychiatry - Revue Canadienne de Psychiatrie. Vol. 47, No.8, pp.723-33.
- Lowe PA and Reynolds CR (2000) *Exploratory analyses of the latent structure of anxiety among older adults*. Educational and Psychological Measurement. Vol. 60, pp.100-116.
- Loy C and Schneider LS (2006) *Galantamine for Alzheimer's disease and mild cognitive impairment*. Cochrane Database of Systematic Reviews. Vol. 1, CD001747.
- Lubart E, Leibovitz A, Shapira A, et al. (2004) *Satisfaction with the care of institutionalized psychogeriatric patients, as reflected by a survey of their relatives*. Archives of Gerontology and Geriatrics. Vol. 38, No.1, pp.45-50.
- Luoma M-L, Vaarama M and Hertto P (2005) *Quality of life (QoL) among demented nursing home residents - Do different instruments provide different results? 18th Congress of the International Association of Gerontology*, Rio de Janeiro, Brazil.
- Luscombe G, Brodaty H and Freeth S (1998) *Younger people with dementia: diagnostic issues, effects on carers and use of services*. International Journal of Geriatric Psychiatry. Vol. 13, No.5, pp.323-30.
- Lyketsos CG, Gonzales-Salvador T, Chin JJ, Baker A et al. (2003) *A follow-up study of change in quality of life among persons with dementia residing in a long-term care facility*. International Journal of Geriatric Psychiatry. Vol. 18, No.4, pp.275-81.
- Lyketsos CG (2007) *Neuropsychiatric symptoms (behavioural and psychological symptoms of dementia) and the development of dementia treatments. Consensus Conference Paper. International Psychogeriatric Association*. International Psychogeriatrics. pp.1-12.
- Lyketsos CG, Breitner JCS and Rabins PV (2001) *An evidence based proposal for the classification of neuropsychiatric disturbance in Alzheimer's disease*. International Journal of Geriatric Psychiatry. Vol. 16, pp.1037-1042.
- Lyketsos CG, Steinberg M, Tschanz JY, Norton MC, Steffens DC and Breiner JC (2000) *Mental and behavioural disturbances in dementia: findings for the Cache County study on memory in aging*. Am J Psychiatry. Vol. 157, pp.708-714.
- MacDonald L, Sibbald B and Hoare C (1988) *Measuring patient satisfaction with life in a long-stay psychiatric hospital*. International Journal of Social Psychiatry. Vol. 34, No.4, pp.292-304.
- Mackenzie AE and Chang AM (2002) *Predictors of quality of life following stroke*. Disability and Rehabilitation. Vol. 24, No.5, pp.259-65.
- Mackenzie A, Holroyd E and Lui MH (1998) *Community nurses assessment of the needs of Hong Kong family carers who are looking after stroke patients*. International Journal of Nursing Studies. Vol. 35, No.1, pp.132-140.
- Mack JL and Patterson MB (2006) *An empirical basis for domains in the analysis of dependency in the activities of daily living (ADL): results of a confirmatory factor analysis of the Cleveland Scale for Activities of Daily Living (CSADL)*. Clinical Neuropsychologist. Vol. 20, No.4, pp.662-677.
- Mack JL, Patterson MB and Tariot PN (1999) *Behavior Rating Scale for Dementia: development of test scales and presentation of data for 555 individuals with Alzheimer's disease*. Journal of Geriatric Psychiatry and Neurology. Vol. 12, No.4, pp.211-23.
-

- 
- Mackin RS, Depp C, Arian P, et al. (2006) *Overview of psychiatric assessment with dementia patients*. In Yeo G, Gallagher-Thompson D, (eds). Ethnicity and the Dementias. 2nd ed. Routledge, Taylor and Francis Group, New York, p 71-86.
- MacNeil SE and Lichtenberg PA (1997) *Home alone: the role of cognition in return to independent living*. Archives of Physical Medicine and Rehabilitation. Vol. 78, No.7, pp.755-758.
- Mador J, Hecker J and Clark M (2003) *Evaluation of donepezil in Alzheimer's disease*. Australasian Journal on Ageing. Vol. 22, No.3.
- Mador J, Hecker J and Clark M (2003) *Evaluation of donepezil in Alzheimer's disease*. Australasian Journal on Ageing. Vol. 22, No.3.
- Magaziner J (1997) *Use of proxies to measure health and functional outcomes in effectiveness research in persons with Alzheimer disease and related disorders*. Alzheimer Disease and Associated Disorders. Vol. 11, Suppl.6, pp.168-74.
- Mahon NE and Yarcheski A (1990) *The dimensionality of the UCLA Loneliness Scale in early adolescents*. Research in Nursing and Health. Vol. 13, No.1, pp.45-52.
- Mahoney AE (2003) *Age- or stage-appropriate? Recreation and the relevance of Piaget's theory in dementia care*. American Journal of Alzheimer's Disease and Other Dementias. Vol. 18, No.1, pp.24-30.
- Mahoney FI and Barthel DW (1965) *Functional evaluation: the Barthel Index*. Maryland State Medical Journal. Vol. 14, pp.61-65.
- Mahurin RK, DeBettignies BH and Pirozzolo FJ (1991) *Structured assessment of independent living skills: preliminary report of a performance measure of functional abilities in dementia*. Journal of Gerontology. Vol. 46, No.2, pp.P58-P66.
- Malouf R and Birks J (2004) *Donepezil for vascular cognitive impairment*. Cochrane Database Systematic Review. Vol. 1, No.CD004395.
- Malloy PF and Boyle PA (2005) *Apathy and Its Treatment in Alzheimer's disease and Other Dementias*. Psychiatric Times. Vol. 22, No.13.
- Manning WG, Leibowitz A, Goldberg GA, et al. (1984) *A controlled trial of the effect of a prepaid group practice on use of services*. New England Journal of Medicine. Vol. 310, No.23, pp.1505-10.
- Mant J, Carter J, Wade DT, et al. (2000) *Family support for stroke: a randomised controlled trial*. Lancet. Vol. 356, No.9232, pp.808-13.
- Manzano JM, Llorca G, Ledesma A, et al. (1994) *Spanish adaptation of the Alzheimer's disease assessment scale (ADAS)*. Actas Luso Esp Neurol Psiquiatr Cienc Afines. Vol. 22, No.6, pp.64-70.
- Marangoni C and Ickes W (1989) *Loneliness: a theoretical review with implications for measurement*. Journal of Social and Personal Relationships. Vol. 6, pp.93-128.
- Marcantonio ER, Michaels M and Resnick NM (1998) *Diagnosing delirium by telephone*. Journal of General Internal Medicine. Vol. 13, No.9, pp.621-623.
- Margolis SA, Al-Marzouq S, Revel T, et al. (2003) *Patient satisfaction with primary health care services in the United Arab Emirates*. International Journal for Quality in Health Care. Vol. 15, No.3, pp.241-9.
- Marin RS (1990) *Differential diagnosis and classification of apathy*. American Journal of Psychiatry. Vol. 147, pp.22-30.
- Marin RS, Biedrzycki RC and Firinciogullari S (1991) *Reliability and validity of the Apathy Evaluation Scale*. Psychiatry Research. Vol. 38, No.2, pp.143-162.
-

---

Marin RS, Butters MA, Mulsant BH, et al. (2003) *Apathy and executive function in depressed elderly*. Journal of Geriatric Psychiatry and Neurology. Vol.16, No.2, pp.112-116.

Marin DB, Flynn S, Mare M, et al. (2001) *Reliability and validity of a chronic care facility adaptation of the Clinical Dementia Rating scale*. International Journal of Geriatric Psychiatry. Vol. 16, No.8, pp.745-50.

Marks IM and Mathews AM (1979) *Brief standard self-rating for phobic patients*. Behaviour Research and Therapy. Vol. 17, pp.263-267.

Marquis KH (1970) *Effects of social reinforcement on health reporting in the household interview*. Sociometry. Vol. 33, pp.205-215.

Marra CA, Woolcott JC, Kopec JA, Shojania K, Offer R, Brazier JE, Esdaile JM and Anis AH (2005) *A comparison of generic, indirect utility measures (the HUI2, HUI3, SF-6D, and the EQ-5D) and disease-specific instruments (the RAQoL and the HAQ) in rheumatoid arthritis*. Social Science and Medicine. Vol. 60, No.7, pp.1571-82.

Martin C, McKenzie S and Ames D (1994) *Disturbed behaviour in dementia sufferers: A comparison of three nursing home settings*. International Journal of Geriatric Psychiatry. Vol. 9, No.5, pp.393-398.

Martin A, Glasziou P and Simes R (1996) *A Utility-Based Quality of Life Questionnaire for Cardiovascular Patients: Reliability and Validity of the UBQ-H(ear) Items*. NHMRC Clinical Trials Centre, University of Sydney, Sydney.

Martina CMS and Stevens NL (2006) *Breaking the cycle of loneliness? Psychological effects of a friendship enrichment program for older women*. Aging and Mental Health. Vol. 10, No.5, pp.467-475.

Martin-Cook K, Hynan LS, Rice-Koch K, et al. (2005) *Responsiveness of the quality of life in late-stage dementia scale to psychotropic drug treatment in late-stage dementia*. Dementia and Geriatric Cognitive Disorders. Vol. 19, No.2-3, pp.82-5.

Martinon-Torres G, Fioravanti M and Grimley EJ (2004) *Trazodone for agitation in dementia*. Cochrane Database of Systematic Reviews. No.4, pp.CD004990.

Mason A, Weatherly H, Spilsbury K, et al. (2007) *A systematic review of the effectiveness and cost-effectiveness of different models of community-based respite care for frail older people and their carers*. Health Technology Assessment. Vol. 11, No.15, pp.1-176.

Mast BT, Fitzgerald J, Steinberg J, et al. (2001) *Effective screening for Alzheimer's disease among older African Americans*. Clinical Neuropsychologist. Vol. 15, No.2, pp.196-202.

Mathers C, Vos T and Stevenson C (1999) *The Burden of Disease and Injury in Australia*, Australian Institute of Health and Welfare, Canberra.

Mathias JL (2003) *Neurobehavioral functioning of persons with Parkinson's disease*. Applied Neuropsychology. Vol. 10, No.2, pp.57-68.

Mathuranath PS, Nestor PJ, Berrios GE, et al. (2000) *A brief cognitive test battery to differentiate Alzheimer's disease and frontotemporal dementia*. Neurology. Vol. 55, No.11, pp.1613-1620.

Mattis S (1976) *Mental status examination for organic mental syndrome in the elderly patient*. In Bellak L, Karasu T, (eds). Geriatric psychiatry: A handbook for psychiatrists and primary care physicians. Grune & Stratton, New York, pp.77-101.

May KM (1992) *Middle-Eastern immigrant parents' social networks and help-seeking for child health care*. Journal of Advanced Nursing. Vol. 17, No.8, pp.905-12.

Mayers AM and Svartberg M (2001) *Existential loneliness: a review of the concept, its psychosocial precipitants and psychotherapeutic implications for HIV-infected women*. British Journal of Medical Psychology. Vol. 74, No.Pt 4, pp.539-53.

- Mazaux JM, Masson F, et al. (1997) *Long-term neuropsychological outcome and loss of social autonomy after traumatic brain injury*. Archives of Physical Medicine and Rehabilitation. Vol. 78, No.12, pp.1316-20.
- McCulla MM, Coats M, Van Fleet N, et al. (1989) *Reliability of clinical nurse specialists in the staging of dementia*. Archives of Neurology. Vol. 46, No.11, pp.1210-1.
- McCauley SR, Levin HS, et al. (2001) *The neurobehavioural rating scale-revised: sensitivity and validity in closed head injury assessment*. Journal of Neurology, Neurosurgery and Psychiatry. Vol. 71, No.5, pp.643-51.
- McCracken CF, Boneham MA, Copeland JR, et al. (1997) *Prevalence of dementia and depression among elderly people in black and ethnic minorities*. British Journal of Psychiatry. Vol. 171, pp.269-73.
- McCusker J (1984) *Development of scales to measure satisfaction and preferences regarding long-term and terminal care*. Medical Care. Vol. 22, No.5, pp.476-93.
- McCusker J, Cole M, Bellavance F and Primeau F (1998) *Reliability and validity of a new measure of severity of delirium*. International Psychogeriatrics. Vol. 10, No.4, pp.421-433.
- McDowell I (2006) *Measuring Health: A Guide to Rating Scales and Questionnaires (3rd ed.)* Oxford University Press, Oxford.
- McDowell I and Newell C (eds.) (1987) *Measuring health: a guide to rating scales and questionnaires*. Oxford University Press, New York.
- McGee S, Orengo C, Kunik ME, Molinari V, et al. (1997) *Delirium in geropsychiatric patients: Patient characteristics and treatment outcomes*. Journal of Geriatric Psychiatry and Neurology. Vol. 10, pp.7-10.
- McGrath P, Moore A, McNaught M, et al. (2005) *Another form to fill in! Clients' reflections on the hospice use of questionnaires*. Support Care Cancer. Vol. 13, No.9, pp.691-701.
- McHorney CA (1996) *Measuring and monitoring general health status in elderly persons: Practical and methodological issues in using the SF-36 health survey*. The Gerontologist. Vol. 36, No.5, Health Module, pp. 571.
- McHorney CA and Fleishman JA (2006) *Assessing and understanding measurement equivalence in health outcome measures*. Medical Care. Vol. 44, No.11 Suppl 3, pp.S205-S210.
- McKee MD, Cunningham M, Jankowski KR, et al. (2001) *Health-related functional status in pregnancy: relationship to depression and social support in a multi-ethnic population*. Obstetrics and Gynecology. Vol. 97, No.6, pp.988-93.
- McKeith I, Cummings J, Lovestone S, Harvey R and Wilkinson D (1999) *Outcome measures in Alzheimer's Disease*. Martin Dunitz, London.
- McKeith IG, Dickson DW, Lowe J, Emre M, O'Brien JT, Feldman H, et al. (2005) *Diagnosis and management of dementia with Lewy bodies: third report of the DLB Consortium. [erratum appears in Neurology. 2005 Dec 27; 65(12):1992] [summary for patients in Neurology. 2005 Dec 27; 65(12):E26-7; PMID: 16380603]*. Neurology. Vol. 65, No.12, pp.1863-1872.
- McKeith IG, Galasko D, Kosaka K, Perry EK, Dickson DW, Hansen LA, et al. (1996) *Consensus guidelines for the clinical and pathologic diagnosis of dementia with Lewy bodies (DLB): report of the consortium on DLB international workshop*. Neurology. Vol. 47, No.5, pp.1113-1124.
- McKeith IG, Perry EK and Perry RH (1999) *Report of the second dementia with Lewy body international workshop: diagnosis and treatment*. Consortium on Dementia with Lewy Bodies. Neurology. Vol. 53, No.5, pp.902-905.
- McKhann G, Drachman D, Folstein M, Katzman R, Price D and Stadlan EM (1984) *Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease*. Neurology. Vol. 34, No.7, pp.939-944.

- McKie J and Richardson J (2003) *The rule of rescue*. Social Science and Medicine. Vol. 56, No.12, pp.2407-19.
- McKinley RK, Fraser RC, Baker RH, et al. (2004) *The relationship between measures of patient satisfaction and enablement and professional assessments of consultation competence*. Medical Teacher. Vol. 26, No.3, pp.223-8.
- McMahon PM, Araki SS, Neumann, PJ, Harris GJ and Gazelle GS (2000) *Cost-effectiveness of functional imaging tests in the diagnosis of Alzheimer disease*. Radiology. Vol. 217, No.1, pp.58-68.
- McNamee P, Bond J and Buck D (2001) *Costs of dementia in England and Wales in the 21st century*. British Journal of Psychiatry. Vol. 179, pp.261-6.
- McNicoll L, Pisani MA, Ely EW, Gifford D and Inouye SK (2005) *Detection of delirium in the intensive care unit: comparison of confusion assessment method for the intensive care unit with confusion assessment method ratings*. Journal of the American Geriatrics Society. Vol. 53, No.3, pp.495-500.
- McQuellon RP, Russell GB, Rambo TD, et al. (1998) *Quality of life and psychological distress of bone marrow transplant recipients: the 'time trajectory' to recovery over the first year*. Bone Marrow Transplant. Vol. 21, No.5, pp.477-86.
- McShane R, Birks J, Hermans D and Hicks K (2006) *Dementia and Cognitive Improvement Group*. *About The Cochrane Collaboration* (Cochrane Review Groups (CRGs) 2006, Issue 2, Art No.DEMENTIA.
- Meagher DJ, Moran M, Raju B, Gibbons D, Donnelly S, Saunders J, et al. (2007) *Phenomenology of delirium. Assessment of 100 adult cases using standardised measures*. Br J Psychiatry. Vol. 190, pp.135-141.
- Meakin R and Weinman J (2002) *The 'Medical Interview Satisfaction Scale' (MISS-21) adapted for British general practice*. Family Practice. Vol. 19, No.3, pp.257-63.
- Meehan T, Bergen H and Stedman T (2002) *Monitoring consumer satisfaction with inpatient service delivery: the Inpatient Evaluation of Service Questionnaire*. Australian and New Zealand Journal of Psychiatry. Vol. 36, No.6, pp.807-11.
- Meek PD, McKeithan K and Schumock GT (1998) *Economic considerations in Alzheimer's disease*. Pharmacotherapy. Vol. 18, No.2, Pt.2, pp.68-73, discussion pp.79-82.
- Mega MS, Cummings JL, Fiorello R and Gornbein J (1996) *The spectrum of behavioural change in Alzheimer's Disease*. Neurology. Vol. 46, pp.130-135.
- Melzer D, Bedford S, Denning T, et al. (1996) *Carers and the monitoring of psychogeriatric community teams*. International Journal of Geriatric Psychiatry. Vol. 11, No.12, pp.1057-1-61.
- Mendez MF, Perryman KM, et al. (1998) *Behavioral differences between frontotemporal dementia and Alzheimer's disease: a comparison on the BEHAVE-AD rating scale*. International Psychogeriatrics. Vol. 10, No.2, pp.155-62.
- Mendes de Leon CF, Glass TA and Berkman LF (2003) *Social engagement and disability in a community population of older adults: the New Haven EPESE*. American Journal of Epidemiology. Vol. 157, No.7, pp.633-42.
- Merritt BK and Fisher AG (2003) *Differences between men and women in ADL performance*. Archives of Physical Medicine and Rehabilitation. Vol. 84, pp.1872-1877.
- Meyer TJ, Miller ML, Metzger RL, et al. (1990) *Development and validation of the Penn State Worry Questionnaire*. Behaviour Research and Therapy. Vol. 28, pp.487-495.

---

Michelson H, Bolund C, Nilsson B, et al. (2000) *Health-related quality of life measured by the EORTC QLQ-C30--reference values from a large sample of Swedish population*. Acta Oncologica. Vol. 39, No.4, pp.477-84.

Midlov P, Bondesson A, et al. (2002) *Descriptive study and pharmacotherapeutic intervention in patients with epilepsy or Parkinson's disease at nursing homes in southern Sweden*. European Journal of Clinical Pharmacology. Vol. 57, No.12, pp.903-10.

Miller RJ, Snowdon J and Vaughan R (1995) *The use of the Cohen-Mansfield Agitation Inventory in the assessment of behavioral disorders in nursing homes*. Journal of the American Geriatrics Society. Vol. 43, No.5, pp.546-9.

Miller PA (1997) *Support systems of nurses recovering from chemical dependency: a pilot study*. Holistic Nursing Practice. Vol. 11, No.4, pp.56-70.

Miller TR and Cleary TA (1993) *Direction of wording effects in balanced scales*. Education and Psychological Measurement. Vol. 53, No.1, pp.51-60.

Minicuci N, Maggi S, Noale M, et al. (2003) *Predicting mortality in older patients: The Velca Study*. Aging-Clinical and Experimental Research. Vol. 15, No.4, pp.328-335.

Mizoguchi T, Iijima S, et al. (1993) *Reliability and validity of a Japanese version of the Dementia Behavior Disturbance Scale*. Nippon Ronen Igakkai Zasshi. Vol. 30, No.10, pp.835-40.

Mohs RC and Cohen L (1988) *Alzheimer's Disease Assessment Scale (ADAS)*. Psychopharmacology Bulletin. Vol. 24, No.4, pp.627-8.

Mohs RC, Doody RS, Morris JC, et al (2001) *A 1-year, placebo-controlled preservation of function survival study of donepezil in AD patients*. Neurology. Vol. 57, No.3, pp.481-488.

Molloy DW, Alemayehu E and Roberts R (1991) *Reliability of a Standardized Mini-Mental State Examination compared with the traditional Mini-Mental State Examination*. American Journal of Psychiatry. Vol. 148, No.1, pp.102-105.

Moniz-Cook E, Woods R, Gardiner E, Silver M, et al. (2001) *The Challenging Behaviour Scale (CBS): development of a scale for staff caring for older people in residential and nursing homes*. British Journal of Clinical Psychology. Vol. 40, No.3, pp.309-22.

Montalto M (1996) *Patients' and carers' satisfaction with hospital-in-the-home care*. International Journal for Quality in Health Care. Vol. 8, No.3, pp.243-51.

Monteiro IM, Boksay I, et al. (2001) *Addition of a frequency-weighted score to the Behavioral Pathology in Alzheimer's Disease Rating Scale: the BEHAVE-AD-FW: methodology and reliability*. European Psychiatry: The Journal of the Association of European Psychiatrists. Vol. 16, Suppl.1, pp.5s-24s.

Montgomery SA and Asberg M (1979) *A new depression scale designed to be sensitive to change*. British Journal of Psychiatry. Vol. 134, pp.382-9.

Moore DJ, Palmer BW, Patterson TL, et al. (2007) *A review of performance-based measures of functional living skills*. Journal of Psychiatric Research. Vol. 41, pp.97-118.

Moore JT, Bobula JA, Short TB, et al. (1983) *A functional dementia scale*. The Journal of Family Practice. Vol. 16, pp.499-503.

Moorer P and Suurmeijer TPBM (1993) *Unidimensionality and cumulativeness of the Loneliness Scale using Mokken Scale Analysis for polychotomous items*. Psychological Reports. Vol. 73, No.3, Pt.2, pp.1324-1326.

Morales LS, Flowers C, Gutierrez P, et al. (2006) *Item and scale differential functioning of the Mini-Mental State Exam assessed using the Differential Item and Test Functioning (DFIT) Framework*. Medical Care. Vol. 44, No.11 Suppl 3, pp.S143-51.

- 
- Morano CD, Cisler RA and Lemerond J (1993) *Risk factors for adolescent suicidal behavior: loss, insufficient familial support, and hopelessness*. Adolescence. Vol. 28, No.112, pp.851-65.
- Morrell CJ, Spiby H, Stewart P, et al. (2000) *Costs and benefits of community postnatal support workers: a randomised controlled trial*. Health Technology Assessment. Vol. 4, No.6, pp.1-100.
- Morris JC (1993) *The Clinical Dementia Rating (CDR): Current version and scoring rules*. Neurology. Vol. 43, No.11, pp.2412-2414.
- Morris JC (1997) *Clinical dementia rating: a reliable and valid diagnostic and staging measure for dementia of the Alzheimer type*. International Psychogeriatrics. Vol. 9 Suppl 1, pp.173-6; discussion 177-8.
- Morris JN, Fries BE, Mehr DR, et al. (1994) *MDS Cognitive Performance Scale*. Journal of Gerontology. Vol. 49, No.4, pp.M174-82.
- Morris JC, Heyman A, Mohs RC, et al. (1989) *The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's Disease*. Neurology. Vol. 39, No.9, pp.1159-65.
- Morris JC, McKeel DW, Jr., Fulling K, et al. (1988) *Validation of clinical diagnostic criteria for Alzheimer's disease*. Annals of Neurology. Vol. 24, No.1, pp.17-22.
- Morris JC, McKeel DWJ, Storandt M, et al. (1991) *Very mild Alzheimer's disease: informant-based clinical, psychometric, and pathologic distinction from normal aging*. Neurology. Vol. 41, No.4, pp.469-78.
- Morris JC, Storandt M, McKeel DW, Jr., et al. (1996) *Cerebral amyloid deposition and diffuse plaques in "normal" aging: Evidence for presymptomatic and very mild Alzheimer's disease*. Neurology. Vol. 46, No.3, pp.707-19.
- Morris JC, Storandt M, Miller JP, et al. (2001) *Mild cognitive impairment represents early-stage Alzheimer disease*. Archives of Neurology. Vol. 58, No.3, pp.397-405.
- Mozley CG, Huxley P, Sutcliffe C, et al. (1999) *'Not knowing where I am doesn't mean I don't know what I like': cognitive impairment and quality of life responses in elderly people*. International Journal of Geriatric Psychiatry. Vol. 14, No.9, pp.776-83.
- Muller RT and Lemieux KE (2000) *Social support, attachment, and psychopathology in high risk formerly maltreated adults*. Child Abuse and Neglect. Vol. 24, No.7, pp.883-900.
- Mulkay M, Ashmore M and Pinch T (1987) *Measuring the quality of life: a sociological intervention concerning the application of economics to health care*. Sociology. Vol. 21, pp.541-564.
- Mungas D (2006) *Neuropsychological assessment of hispanics elders*. In Yeo G, Gallagher-Thompson D, (eds.) Ethnicity and the Dementias. 2nd ed. Routledge, Taylor and Francis Group, New York, p 71-86.
- Mungas D, Weller P, Franzi C and Henry R (1989) *Assessment of disruptive behaviour associated with dementia: the Disruptive Behaviour Rating Scale*. Journal of Geriatric Psychiatry. Vol. 2, pp.196-202.
- Munro BH, Jacobsen BS and Brooten DA (1994) *Re-examination of the psychometric characteristics of the La Monica-Oberst Patient Satisfaction Scale*. Research in Nursing and Health. Vol. 17, No.2, pp.119-25.
- Munsey C (2006) *Testing for impairment: Researchers are seeking culturally valid methods to assess cognitive functioning in the minority elderly*. Monitor on Psychology (APA). Vol. 37, No.10, pp.54.
- Muo R, Schindler A, Venero I, Schindler O, Ferrario E and Frisoni GB (2005) *Alzheimer's disease-associated disability: an ICF approach*. Disability and Rehabilitation. Vol. 27, No.23, pp.1405-1413.
- Murphy PA, Prewitt TE, Bote E, et al. (2001) *Internal locus of control and social support associated with some dietary changes by elderly participants in a diet intervention trial*. Journal of the American Dietetic Association. Vol. 101, No.2, pp.203-8.
-

- MVHGroup (1995) *The Measurement and Valuation of Health: Final Modelling of Valuation Tariffs*, Centre for Health Economics, York.
- Nagels G, Engelborghs S, Vloeberghs E, Van Dam D, et al. (2006) *Actigraphic measurement of agitated behaviour in dementia*. International Journal of Geriatric Psychiatry. Vol. 21, No.4, pp.388-93.
- Naglie G, Tomlinson G, Tansey C, Irvine J, Ritvo P, Black SE, Freedman M, Silberfeld M and Krahn M (2006) *Utility-based quality of life measures in Alzheimer's disease*. Quality of Life Research. Vol. 15, No.4, pp.631-43.
- National Centre for Classification in Health (2002a) *ICD-10-AM Mental Health Manual. An integrated classification and diagnostic tool for community-based mental health services (1st ed.)* University of Sydney, Sydney.
- National Centre for Classification in Health (2002b) *The international statistical classification of diseases and related health problems, 10th revision, Australian modification (ICD-10-AM) (3rd ed.)* University of Sydney, Sydney.
- National Institute for Health and Clinical Excellence (NICE) (2007) *TA111 Alzheimer's disease - donepezil, galantamine, rivastigmine (review) and memantine: guidance (amended)*. National Institute for Health and Clinical Excellence, National Health Service, London. Available at <http://guidance.nice.org.uk/TA111/guidance/pdf/English/download.dspix>, accessed 27 September 2007.
- Nay R, Scherer S, Pitcher A, Koch S, Browning M, Flicker L, et al. (2003) *Responding to Behaviours of Concern Among People Living in Residential Aged Care*. La Trobe University, Melbourne.
- Nelson EC, Wasson J, Kirk J, et al. (1987) *Assessment of function in routine clinical practice: description of the COOP Chart method and preliminary findings*. J Chronic Dis. Vol. 40, Suppl.1, pp.55S-63S.
- Nelson CJ, Rosenfeld B, Breitbart W, et al. (2002) *Spirituality, religion, and depression in the terminally ill*. Psychosomatics. Vol. 43, No.3, pp.213-20.
- Nelson JE, Tandon N, Mercado AF, Camhi SL, Ely EW and Morrison RS (2006) *Brain dysfunction: another burden for the chronically critically ill*. Archives of Internal Medicine. Vol. 166, No.18, pp.1993-1999.
- Nelson PB (1989) *Social support, self-esteem, and depression in the institutionalized elderly*. Issues in Mental Health Nursing. Vol. 10, No.1, pp.55-68.
- Neugebauer A, Katz PP and Pasch LA (2003) *Effect of valued activity disability, social comparisons, and satisfaction with ability on depressive symptoms in rheumatoid arthritis*. Health Psychology. Vol. 22, No.3, pp.253-262.
- Neumann PJ, Araki SS and Gutterman EM (2000) *The use of proxy respondents in studies of older adults: Lessons, challenges and opportunities*. JAGS. Vol. 48, pp.1646-1654.
- Neumann PJ, Hermann RC, Kuntz KM, Araki SS, Duff SB, Leon J, Berenbaum PA, Goldman PA, Williams LW and Weinstein MC (1999a) *Cost-effectiveness of donepezil in the treatment of mild or moderate Alzheimer's disease*. Neurology. Vol. 52, No.6, pp.1138-45.
- Neumann PJ, Kuntz KM, Leon J, Araki SS, Hermann RC, Hsu MA. and Weinstein MC (1999b) *Health utilities in Alzheimer's disease: a cross-sectional study of patients and caregivers*. Medical Care. Vol. 37, No.1, pp.27-32.
- Neumann PJ, Sandberg EA, Araki SS, Kuntz KM, Feeny D and Weinstein MC (2000). *A comparison of HUI2 and HUI3 utility scores in Alzheimer's disease*. Medical Decision Making. Vol. 20, No.4, pp.413-22.
- Neundorfer MM, McClendon MJ, Smyth KA, Stuckey JC, et al. (2001) *A longitudinal study of the relationship between levels of depression among persons with Alzheimer's disease and levels of depression among their family caregivers*. Journals of Gerontology Series B-Psychological Sciences and Social Sciences. Vol. 56, No.5, pp.P301-13.

- 
- Neville CC and Byrne GJ (2002) *Behaviour of older people admitted for residential respite care*. Aust J Adv Nurs. Vol. 20, No.1, pp.8-12.
- Neville CC and Byrne GJ (2007) *Prevalence of disruptive behaviour displayed by older people in community and residential respite care settings*. International Journal of Mental Health Nursing. Vol. 16, No.2, pp.81-5.
- Newberg A, Cotter A, et al. (2003) *A Metabolic Imaging Severity Rating Scale for the Assessment of Cognitive Impairment*. Clinical Nuclear Medicine. Vol. 28, No.7, pp.565-570.
- Newsome PR and Wright GH (1999) *A review of patient satisfaction: 1. Concepts of satisfaction*. British Dental Journal. Vol. 186, No.4 Spec No, pp.161-5.
- Nguyen TD, Attkisson CC and Stegner BL (1983) *Assessment of patient satisfaction: development and refinement of a service evaluation questionnaire*. Evaluation and Program Planning. Vol. 6, No.3-4, pp.299-313.
- Nicoll M, Ashworth M, McNally L, et al. (2002) *Satisfaction with respite care: a pilot study*. Health Soc Care Community. Vol. 10, No.6, pp.479-84.
- Nishikawa K, Nakayama M, Omote K and Namiki A (2004) *Recovery characteristics and post-operative delirium after long-duration laparoscope-assisted surgery in elderly patients: propofol-based vs. sevoflurane-based anaesthesia*. Acta Anaesthesiologica Scandinavica. Vol. 48, No.2, pp.162-168.
- Njegovan V, Hing MM, Mitchell SL, et al. (2001) *The hierarchy of functional loss associated with cognitive decline in older persons*. Journals of Gerontology Series A - Biological Sciences and Medical Sciences. Vol. 56, No.10, pp.M638-M643.
- Nolan M and Grant G (1992) *Regular Respite: An Evaluation of Hospital Rota Bed Scheme for Elderly People*. Age Concern Institute of Gerontology. Report nr Research Paper N 6.
- Nolan MR, Grant G and Ellis NC (1990) *Stress is in the eye of the beholder: reconceptualizing the measurement of carer burden*. Journal of Advanced Nursing. Vol. 15, No.5, pp.544-55.
- Nolan M, Grant G and Keady J (1996) *Understanding family care : a multi-dimensional model of caring and coping*. Open University Press, Philadelphia, PA.
- Nolen-Hoeksema S and Ahrens C (2002) *Age differences and similarities in the correlates of depressive symptoms*. Psychology and Aging. Vol. 17, No.1, pp.116-24.
- Norbeck JS (1984) *The Norbeck Social Support Questionnaire*. Birth Defects Original Article Series. Vol. 20, No.5, pp.45-57.
- Norbeck JS (1995) *Scoring Instructions for the Norbeck Social Support Questionnaire*. University of California.
- Norbeck JS, Lindsey AM and Carrieri VL (1981) *The development of an instrument to measure social support*. Nursing Research. Vol. 30, No.5, pp.264-9.
- Norbeck JS, Lindsey AM and Carrieri VL (1983) *Further development of the Norbeck Social Support Questionnaire: normative data and validity testing*. Nursing Research. Vol. 32, No.1, pp.4-9.
- Nord E (1993) *Unjustified use of the Quality of Well-Being Scale in priority setting in Oregon* Health Policy. Vol. 24, No.1, pp.45-53.
- Nord E (1997) *A Review of Synthetic Health Indicators*, National Institute of Public Health for the OECD Directorate for Education, Employment, Labour and Social Affairs, Oslo.
- Nord E, Richardson J and Macarounas-Kirchmann K (1993) *Social evaluation of health care versus personal evaluation of health states. Evidence on the validity of four health-state scaling instruments using Norwegian and Australian surveys*. International Journal of Technology Assessment in Health Care. Vol. 9, No.4, pp.463-78.
-

Novella JL, Boyer F, Jochum C, et al. (2006). *Agreement between patients' and proxies' reports of quality of life in Alzheimers disease*. Quality of Life Research. Vol. 10, No 5, pp 443-52.

Novella JL, Boyer F, Jochum C, et al. (2006) *Health status in patient's with Alzheimer's disease: An investigation of inter-rater agreement*. Quality of Life Research. Vol. 15, pp.811-819.

Novella JL, Jochum C, Ankri J et al. (2001) *Measuring general health status in dementia: practical and methodological issues in using the SF-36*. Journal of Aging-Clinical and Experimental Research. Vol. 13, pp 362-9.

Nunnally JC and Bernstein IH (1994) *Psychometric Theory*. McGraw-Hill, New York. 752 p.

Nybo T, Sainio M, et al. (2005) *Middle age cognition and vocational outcome of childhood brain injury*. Acta Neurologica Scandinavica. Vol. 112, No.5, pp.338-42.

O'Connell B, et al. (2007) *Continence Assessment Tool for Residential Aged Care (Draft)*. School of Nursing, Deakin University.

O'Connell ME, Tuokko H, Graves RE and Kadlec H (2004) *Correcting the 3MS for bias does not improve accuracy when screening for cognitive impairment or dementia*. Journal of Clinical and Experimental Neuropsychology. Vol. 26, No.7, pp.970-980.

O'Connell B, Young J and Twigg D (1999) *Patient satisfaction with nursing care: a measurement conundrum*. International Journal of Nursing Practice. Vol. 5, No.2, pp.72-7.

O'Connor DW, Blessed G, Cooper B, et al. (1996) *Cross-national interrater reliability of dementia diagnosis in the elderly and factors associated with disagreement*. Neurology. Vol. 47, No.5, pp.1194-9.

O'Connor D, Phinney A, Smith A, et al. (2007) *Personhood in dementia care: Developing a research agenda for broadening the vision*. Dementia. Vol. 6, No.1, pp.121-142.

O'Connor R (1993) *Issues in the Measurement of Health-Related Quality of Life*, National Centre for Health Program Evaluation, Melbourne.

OECD (2006) *OECD Health Data*. Organization for Economically Developed Countries, Geneva.

Ohki T, Matsushima E, Shibuya M and Sunamori M (2006) *An evaluation strategy for the early detection of postoperative delirium*. Psychiatry and Clinical Neurosciences. Vol. 60, No.3, pp.277-282.

Olazaran J, Muniz R, Reisberg B, et al. (2004) *Benefits of cognitive-motor intervention in MCI and mild to moderate Alzheimer disease*. Neurology. Vol. 63, No.12, pp.2348-53.

O'Malley KJ, Orengo CA, Kunik ME, Snow L and Molinari V (2002) *Measuring aggression in older adults: a latent variable modeling approach*. Aging Mental Health. Vol. 6, pp.231-238.

O'Leary D, Jyringi D and Sedler M (2005) *Childhood conduct problems, stages of Alzheimer's disease, and physical aggression against caregivers*. International Journal of Geriatric Psychiatry. Vol. 20, No.5, pp.401-5.

Onega LL (2006) *Assessment of psychoemotional and behavioral status in patients with dementia*. Nursing Clinics of North America. Vol. 41, No.1, pp.23-41.

Opie J, Doyle C and O'Connor DW (2002) *Challenging behaviours in nursing home residents with dementia: a randomized controlled trail of multi-disciplinary interventions*. International Journal of Geriatric Psychiatry. Vol. 17, No.1, pp.6-13.

Orlando Edelen MO, Thissen D, Teresi JA, et al. (2006) *Identification of differential item functioning using item response theory and the likelihood-based model comparison approach. Application to the Mini-Mental State Examination*. Medical Care. Vol. 44, No.11 Suppl 3, pp.S134-42.

- 
- Orrell M, Spector A, Thorgrimsen L, et al. (2005) *A pilot study examining the effectiveness of maintenance Cognitive Stimulation Therapy (MCST) for people with dementia*. International Journal of Geriatric Psychiatry. Vol. 20, No.5, pp.446-51.
- O'Rourke D, Sudman S, Johnson T, et al. (1999) *Cognitive testing of cognitive functioning questions*. In Schwarz N, Park DC, Knauper B, Sudman S, (eds). Cognition, ageing, and self reports. Psychology Press, Taylor and Francis Group, Philadelphia.
- Osborn DPJ, Fletcher AE, Smeeth L, et al. (2002) *Geriatric Depression Scale Scores in a representative sample of 14 545 people aged 75 and over in the United Kingdom: results from the MRC Trial of Assessment and Management of Older People in the Community*. International Journal of Geriatric Psychiatry. Vol.17, No.4, pp.375-382.
- Ott BR, Tate CA, et al. (1996) *Gender differences in the behavioral manifestations of Alzheimer's disease*. J Am Geriatr Soc. Vol. 44, No.5, pp.583-7.
- Owen A, Eagar K and Ramsay L (2004) *Ongoing Needs Identification in Community Care: How to use the Tier 1 Assessment and Referral Tools*. Centre for Health Service Development, University of Wollongong.
- Oye T, Loke YA, Chan MF and Kwok T (2005) *Psychometric properties of the Chinese version of the Cost of Care Index to measure caregiver burden among Chinese informal caregivers*. Asian Journal of Nursing Studies. Vol. 8, No.2, pp.4-13.
- Pachana NA, Byrne GJ, Siddle H, et al. (2007) *Development and validation of the Geriatric Anxiety Inventory*. International Psychogeriatrics. Vol.19, No.1, pp.103-114.
- Pae CU, Lee SJ, Lee CU, Lee C and Paik IH (2004) *A pilot trial of quetiapine for the treatment of patients with delirium*. Human Psychopharmacology. Vol. 19, No.2, pp.125-127.
- Palmstierna T and Wistedt B (1987) *Staff Observation Aggression Scale. Presentation and observation*. Acta Psychiatrica Scandinavia. Vol. 76, pp.657-663.
- Pandav R, Fillenbaum G, Ratcliff G, et al. (2002) *Sensitivity and specificity of cognitive and functional screening instruments for dementia: The Indo-U.S. dementia epidemiology study*. Journal of the American Geriatrics Society. Vol. 50, No.3, pp.554-561.
- Pang MW, Chan LW and Yip SK (2003) *One-year urodynamic outcome and quality of life in patients with concomitant tension-free vaginal tape during pelvic floor reconstruction surgery for genitourinary prolapse and urodynamic stress incontinence*. International Urogynecology Journal and Pelvic Floor Dysfunction. Vol. 14, No.4, pp.256-60; discussion 259-60.
- Pang SK, Ip WY and Chang AM (2001) *Psychosocial correlates of fluid compliance among Chinese haemodialysis patients*. Journal of Advanced Nursing. Vol. 35, No.5, pp.691-8.
- Paradies, Y, (2007) *Racsim*. In Carson B, Dunbar T, Chenhall RD and Bailie R (eds.) The Social Determinants of Indigenous Health. Allen and Unwin, Sydney, pp.65-86.
- Parasuraman A, Zeithaml VA and Berry L (1985) *A conceptual model of service quality and it's implications for future research*. Journal of Marketing. Vol. 49, No.4, pp.41-50.
- Parasuraman A, Zeithaml VA and Berry L (1988) *SERVQUAL: a multiple-item scale for measuring consumer perceptions of service quality*. Journal of Retailing. Vol. 64, No.1, pp.12-37.
- Park DC (1999) *Cognitive aging, processing resources, and self-report*. In Schwarz N, Park DC, Knauper B, Sudman S, (eds.) Cognition, ageing, and self reports. Psychology Press, Taylor and Francis Group, Philadelphia.
- Parker SG, Peet SM, Jagger C, et al (1998) *Measuring health status in older patients. The SF-36 in practice*. Age and Ageing. Vol. 27, No.1, pp.13-18.
-

---

Pascual LF, Saz P, Larumbe R, et al. (1997) *Standardization of the Alzheimer's Disease Assessment Scale in a Spanish population*. Neurologia. Vol. 12, No.6, pp.238-44.

Patel V and Hope RA (1992a) *A rating scale for aggressive behaviour in the elderly: RAGE*. Psychological Medicine. Vol. 22, pp.211-221.

Patel V and Hope RA (1992b) *Aggressive behaviour in elderly psychogeriatric inpatients*. Acta. Psychiatr Scand. Vol. 85, pp.131-135.

Patel V and Hope RA (1992c) *The assessment of behavioural phenomena in dementia*. In A Burns (ed.). Ageing and Dementia: a Methodological Approach. Edward Arnold, UK.

Patel V and Hope RA (1993a) *Aggressive behaviour in elderly people with dementia: A review*. International Journal of Geriatric Psychiatry. Vol. 8, pp.457-472.

Patel V, Hope T, Hall J and Fairburn CG (1995) *Three methodological issues in the development of observer-rated behaviour rating scales: experience from the development of the Rating Scale for Aggressive Behaviour in the Elderly*. International Journal of Methods in Psychiatric Research. Vol. 5, pp.21-27.

Patrick DL and Bergner M (1990) *Measurement of health status in the 1990s*. Annual Review of Public Health. Vol. 11, pp.165-183.

Patterson CJS, Gauthier S, Bergman H, Cohen CA, Feightner JW, Felman H, et al. (1999) *Management of dementia disorders. The recognition, assessment and management of dementing disorders- conclusions. Canadian Consensus Conference on Dementia*. Canadian Medical Association Journal, Vol. 160, No.12 (supplement) (online at [www.cma.ca](http://www.cma.ca))

Patterson MB, Mack JL, Mackell JA, Thomas R, et al. (1997) *A longitudinal study of behavioral pathology across five levels of dementia severity in Alzheimer's disease: the CERAD Behavior Rating Scale for Dementia. The Alzheimer's Disease Cooperative Study*. Alzheimer Disease and Associated Disorders. Vol. 11 Suppl.2, pp.S40-4.

Patterson MB, Mack JL, Neundorfer MM, et al. (1992) *Assessment of functional ability in Alzheimer disease: A review and a preliminary report on the Cleveland Scale for Activities of Daily Living* Alzheimer Disease and Associated Disorders. Vol. 6, pp.145-163.

Patterson MB and Mack JL (2001) *The Cleveland Scale for Activities of Daily Living (CSADL): Its reliability and validity*. Journal of Clinical Geropsychology. Vol. 7, No.1, pp.15-28.

Pavot W, Diener E, Colvin CR, et al. (1991) *Further validation of the Satisfaction with Life Scale: evidence for the cross-method convergence of well-being measures*. Journal of Personality Assessment. Vol. 57, No.1, pp.149-61.

Pearlin LI, Mullan JT, Semple SJ and Skaff MM (1990) *Caregiving and the stress process: an overview of concepts and their measures [see comments]*. Gerontologist. Vol. 30, No.5, pp.583-94.

Pearson VI (2000) *Assessment of function in older adults*. In Kane R, Kane R, (eds.) Assessing older persons: measures, meaning and practical applications. Oxford University Press, USA.

Pedhazur E and Schmelkin L (1991) *Measurement, Design and Analysis: An Integrated Approach*. Lawrence Erlbaum, Hillsdale.

Pena-Casanova J, Aguilar M, Santacruz P, et al. (1997a) *Adaptation and normalization of the Alzheimer's disease Assessment Scale for Spain (NORMACODEM) (II)*. [UI: 9147454] Neurologia. Vol. 12, No.2, pp.69-77.

Pena-Casanova J, Meza M, Bertran-Serra I, et al. (1997b) *Shortened version of the Barcelona test (III): criterion validity with the ADAS-Cog* Neurologia Croatica. Vol. 12, No.3, pp.117-9.

Peplau LA and Perlman D, (eds.) (1982). *Loneliness: A Sourcebook of Current Theory, Research and Therapy*. Wiley, New York.

---

- Perez-Madrinan G, Cook SE, Saxton JA, Miyahara S, et al. (2004) *Alzheimer disease with psychosis: excess cognitive impairment is restricted to the misidentification subtype*. American Journal of Geriatric Psychiatry. Vol. 12, No.5, pp.449-56.
- Perivolaris A, LeClerc CM, Wilkinson K and Buchanan S (2006) *An enhanced dining program for persons with dementia*. Alzheimer's Care Quarterly. Vol. 7, No.4, pp.258-67.
- Perlman D, Gerson AC and Spinner B (1978) *Loneliness among senior citizens: an empirical report*. Essence. Vol. 2, pp.239-248.
- Peskind ER, Potkin SG, Pomara N, et al. (2006) *Memantine treatment in mild to moderate Alzheimer disease: a 24-week randomized, controlled trial*. American Journal of Geriatric Psychiatry. Vol. 14, No.8, pp.704-715.
- Petersen RC (1995) *Normal aging, mild cognitive impairment, and early Alzheimer's disease*. The Neurologist. Vol. 1, pp.326-344.
- Petersen RC, Smith GE, Waring SC, Ivnik RJ, Kokmen E and Tangalos EG (1997) *Aging, Memory, and Mild Cognitive Impairment*. International Psychogeriatrics. Vol. 9, suppl.1, pp.65-69.
- Petersen RC, Stevens JC, Ganguli M, Tangalos EG, Cummings JL and DeKosky ST (2001) *Practice parameter: Early detection of dementia: Mild cognitive impairment (an evidence-based review)*. Neurology. Vol. 56, No.9, pp.1133-1142.
- Petracca G, Teson A, Chemerinski E, et al. (1996) *A double-blind placebo-controlled study of clomipramine in depressed patients with Alzheimer's disease*. Journal of Neuropsychiatry and Clinical Neurosciences. Vol. 8, No.3, pp.270-275.
- Petrak JA, Doyle AM, Smith A, et al. (2001) *Factors associated with self-disclosure of HIV serostatus to significant others*. British Journal of Health Psychology. Vol. 6, No.Pt 1, pp.69-79.
- Pettit T, Livingston G, Manela M, et al. (2001) *Validation and normative data of health status measures in older people: the Islington study*. International Journal of Geriatric Psychiatry. Vol. 16, No.11, pp.1061-70.
- Pfeffer RI, Kurosaki TT, Harrah CH, et al. (1982) *Measurement of functional activities in older adults in the community*. Journal of Gerontology. Vol. 37, pp.323-329.
- Pfeffer E (1975) *A Short Portable Mental Status Questionnaire for the assessment of organic brain deficit in elderly patients*. Journal of the American Geriatrics Society. Vol. 23, pp.433-441.
- Philp I (2000) *EASY-Care: A systematic approach to the assessment of older people*. Geriatric Medicine. Vol. 30, No.5, pp.15-19.
- Pirkis JE, Burgess PM, Kirk PK, Dodson S, et al. (2005) *A review of the psychometric properties of the Health of the Nation Outcome Scales (HoNOS) family of measures*. Health and Quality of Life Outcomes. Vol. 3.
- Pickard AS, Johnson JA and Feeny DH (2005) *Responsiveness of generic health-related quality of life measures in stroke*. Quality of Life Research. Vol. 14, No.1, pp.207-19.
- Pickard AS, Johnson JA and Feeny DH, Shuaib A, Carriere KC and Nasser AM (2004) *Agreement between patient and proxy assessments of health-related quality of life after stroke using the EQ-5D and Health Utilities Index*. Stroke. Vol. 35, No.2, pp.607-12.
- Pioggiosi P, Forti P, Ravaglia G, Berardi D, Ferrari G and De Ronchi D (2004) *Different classification systems yield different dementia occurrence among nonagenarians and centenarians*. Dementia and Geriatric Cognitive Disorders. Vol. 17, No.1-2, pp.35-41.
- Pohjasvaara T, Mantyla R, Ylikoski R, Kaste M and Erkinjuntti T (2000) *Comparison of different clinical criteria (DSM-III, ADDTC, ICD-10, NINDS-AIREN, DSM-IV) for the diagnosis of vascular dementia*. National

---

*Institute of Neurological Disorders and Stroke-Association Internationale pour la Recherche et l'Enseignement en Neurosciences. Stroke. Vol. 31, No.12, pp.2952-2957.*

Politis AM, Mayer LS, Passa M and Maillis A (2004) *Validity and reliability of the newly translated Hellenic Neuropsychiatric Inventory (H-NPI) applied to Greek outpatients with Alzheimer's Disease: a study of the disturbing behaviours among referrals to a memory clinic. International Journal of Geriatric Psychiatry. Vol. 19, pp.203-208.*

Pollock BG, Mulsant BH, et al. (2002) *Comparison of citalopram, perphenazine, and placebo for the acute treatment of psychosis and behavioral disturbances in hospitalized, demented patients. American Journal of Psychiatry. Vol. 159, No.3, pp.460-5.*

Ponce H, Molinari V, Kunik ME, Orengo C, et al. (1998) *Placement predictors of geropsychiatric inpatients: Home versus nursing home. Journal of Gerontological Social Work. Vol. 29, No.4, pp.3-12.*

Portet F, Ousset PJ, Visser PJ, Frisoni GB, Nobili J, Scheltens P, et al. (2006) *Mild cognitive impairment in medical practice: critical review of the concept and new diagnostic procedure. Report of the MCI working group of the European Consortium on Alzheimer's disease (EADC), available at <http://www.jnnp.bmjournals.com>, accessed 2 May 2006.*

Pound P, Gompertz P and Ebrahim S (1993) *Development and results of a questionnaire to measure carer satisfaction after stroke. Journal of Epidemiology and Community Health. Vol. 47, No.6, pp.500-5.*

Pound P, Gompertz P and Ebrahim S (1994) *Patients' satisfaction with stroke services. Clinical Rehabilitation. Vol. 8, pp.7-17.*

Powell AL (2002) *On issues pertinent to Alzheimer disease and cultural diversity. Alzheimer Disease and Associated Disorders. Vol. 16, No.Suppl 2, pp.S43-S45.*

Power M, Quinn K, Schmidt S and the WHOQOL-OLD Group (2005) *Development of the WHOQOL-OLD module. Quality of Life Research. Vol. 14, pp.2197-2214.*

Prince M, Acosta D, Chiu H, et al. (2004) *Effects of education and culture on the validity of the Geriatric Mental State and its AGE-CAT algorithm. British Journal of Psychiatry. Vol. 185, No.5, pp.429-436.*

Pritchard EJ and Dewing J (2001) *A multi-method evaluation of an independent dementia care service and its approach. Aging and Mental Health. Vol. 5, No.1, pp.63-72.*

Procidano ME and Heller K (1983) *Measures of perceived social support from friends and from family: three validation studies. American Journal of Community Psychology. Vol. 11, No.1, pp.1-24.*

Productivity Commission (2005a) *Economic Implications of an Ageing Australia*, Commonwealth of Australia, Canberra.

Productivity Commission (2005b) *Impacts of Medical Technology in Australia*, Productivity Commission, Melbourne.

Purnell LD and Paulanka BJ (2005) *Guide to culturally competent health care*. F.A. Davis and Company, Philadelphia.

Quartararo M, O'Neill TJ, Tang C, et al. (1991) *Assessing the residential care needs of nursing home applicants. Australian Journal of Public Health. Vol. 15, No.3, pp.222-227.*

Quintana JM, Gonzalez N, Bilbao A, et al. (2006) *Predictors of patient satisfaction with hospital health care. BMC Health Services Research. Vol. 6, pp.102.*

Rabins PV and Black BS (2007) *Measuring quality of life in dementia: purposes, goals, challenges and progress. Consensus conference paper, International Psychogeriatrics.*

Rabins PV and Kasper JD (1997) *Measuring quality of life in dementia: conceptual and practical issues. Alzheimer Disease and Associated Disorders. Vol. 11 Suppl.6, pp.100-4.*

---

Rabins PV, Kasper JD, Kleinman JD and Black BS (1999) *Concepts and methods in the development of the ADRQL: An instrument for assessing health-related quality of life in persons with Alzheimer's Disease*. Journal of Mental Health and Aging. Vol. 5, No.1, pp.33-48.

Rabin R and de Charro F (2001) *EQ-5D: a measure of health status from the EuroQol Group*. Annals of Medicine. Vol. 33, pp.337-343.

Ramsay M, Winget C and Higginson I (1995) *Review: Measures to determine the outcome of community services for people with dementia*. Age and Ageing. Vol. 24, No.1, pp.73-83.

Rabinowitz J, Davidson M, De Deyn PP, Katz I, et al. (2005) *Factor analysis of the Cohen-Mansfield Agitation Inventory in three large samples of nursing home patients with dementia and behavioral disturbance*. American Journal of Geriatric Psychiatry. Vol. 13, No.11, pp.991-8.

Rabinowitz J, Katz IR, De Deyn PP, Brodaty H, et al. (2004) *Behavioral and psychological symptoms in patients with dementia as a target for pharmacotherapy with risperidone*. Journal of Clinical Psychiatry. Vol. 65, No.10, pp.1329-34.

Radloff LS and Teri L (1986) *Use of the Center for Epidemiological Studies-Depression Scale with older adults*. Clinical Gerontologist. Vol. 5, No.1-2, pp.119-136.

Rainer MK, Mucke HA, Kruger-Rainer C, Haushofer M, et al. (2004) *Zotepine for behavioural and psychological symptoms in dementia: an open-label study.[see comment]*. CNS Drugs. Vol. 18, No.1, pp.49-55.

Ramadan FH and Naughton BJ (1999) *Paroxetine for verbal agitation*. Annals of Long Term Care. Vol. 7, No.7, pp.282-286.

Ramirez Diaz SP, Gregorio PG, Ribera Casado JM, et al. (2005) *The need for a consensus in the use of assessment tools for Alzheimer's disease: the Feasibility Study (assessment tools for dementia in Alzheimer Centres across Europe), a European Alzheimer's Disease Consortium's (EADC) survey*. International Journal of Geriatric Psychiatry. Vol. 20, No.8, pp.744-8.

Ramirez M, Teresi JA, Holmes D, et al. (2006) *Differential item functioning (DIF) and the Mini-Mental State Examination (MMSE). Overview, sample, and issues of translation*. Medical Care. Vol. 44, No.11 Suppl 3, pp.S95-S106.

Ramsay L, Owen A, Eagar K, Marosszeky N, Samsa P, Rix M, Fildes D, Willmott L (2006) *An Australian Carer Needs Assessment Instrument: The Recommended Approach*. Centre for Health Service Development, University of Wollongong.

Rapley M and Antaki C (1996) *A conversation analysis of the 'acquiescence' of people with learning disabilities*. Journal of Community and Applied Social Psychology. Vol. 6, No.3, pp.207-227.

Rapoport M, McCauley S, et al. (2002) *The role of injury severity in neurobehavioral outcome 3 months after traumatic brain injury*. Neuropsychiatry, Neuropsychology, and Behavioral Neurology. Vol. 15, No.2, pp.123-32.

Rasclé N, Bruchon-Schweitzer M and Sarason IG (2005) *Short form of Sarason's Social Support Questionnaire: French adaptation and validation*. Psychological Reports. Vol. 97, No.1, pp.195-202.

Ray WA, Taylor JA, Lichtenstein MJ and Meador KG (1992) *The Nursing Home Behavior Problem Scale*. Journal of Gerontology. Vol. 47, No.1, pp.M9-16.

Ready RE, Ott BR and Grace J (2004) *Patient versus informant perspectives of Quality of Life in Mild Cognitive Impairment and Alzheimer's disease*. International Journal of Geriatric Psychiatry. Vol. 19, No.3, pp.256-65.

Ready RE, Ott BR, Grace J and Fernandez I (2002) *The Cornell-Brown Scale for Quality of Life in dementia. Alzheimer Disease and Associated Disorders*. Vol. 16, No.2, pp.109-15.

- Reid J and Dhamarrandji M (1978) *Curing not caring: why Aboriginal patients 'abscond'*. New Doctor: The Journal of the Doctors Reform Society. Vol. 8, pp.27-32.
- Reid M, Glazener C, Murray GD, et al. (2002) *A two-centred pragmatic randomised controlled trial of two interventions of postnatal support*. BJOG. Vol. 109, No.10, pp.1164-70.
- Reisberg B (1988) *Functional assessment staging (FAST)*. Psychopharmacology Bulletin. Vol. 24, pp.653-659.
- Reisberg B, Borenstein J, et al. (1987) *BEHAVE-AD: A clinical rating scale for the assessment of pharmacologically remediable behavioral symptomatology in Alzheimer's disease*. In Altman, H.J (ed.) Alzheimer's disease: Problems, prospects, and perspectives. Plenum Press, New York. pp.1-16.
- Reisberg B, Borenstein J, Salob SP, Ferris SH, Franssen E and Georgotas A (1987) *Behavioural symptoms in Alzheimer's disease: phenomenology and treatment*. Journal of Clinical Psychiatry. Vol. 48, Suppl, pp.9-15.
- Reisberg B and Ferris SH (1988) *Brief Cognitive Rating Scale (BCRS)*. Psychopharmacology Bulletin. Vol. 24, pp.629-636.
- Reisberg B, Ferris SH, de Leon MJ and Crook T (1982) *The Global Deterioration Scale for assessment of primary degenerative dementia*. American Journal of Psychiatry. Vol. 139, No.9, pp.1136-1164.
- Reisberg B, Finkel S, Overall J, et al. (2001) *The Alzheimer's Disease Activities of Daily Living International Scale (ADL-IS)*. International Psychogeriatrics. Vol. 13, No.2, pp.163-181.
- Reisberg B, Monteiro I, et al. (2000) *Do many of the behavioral and psychological symptoms of dementia constitute a distinct clinical syndrome? Current evidence using the BEHAVE-AD*. International Psychogeriatrics. Vol. 12, Suppl.1, pp.155-164.
- Retsas AP and Crabbe H (1998) *Use of physical restraints in aged care facilities in New South Wales*. International Journal of Nursing Studies. Vol. 35, pp.177-183.
- Retsinas J and Garrity P (1985) *Nursing home friendships*. Gerontologist. Vol. 25, pp.376-381.
- Reuben DB and Siu AL (1990) *An objective measure of physical function of elderly outpatients: The Physical Performance Test*. JAGS. Vol. 38, No.10, pp.1105-1112.
- Reuben DB, Valle LA, Hays RD, et al. (1995) *Measuring physical function in community dwelling older persons: A comparison of self-administered, interview-administered, and performance-based measures*. JAGS. Vol. 43, No.1, pp.17-23.
- Revenson TA and Johnson JL (1984) *Social and demographic correlates of loneliness in late life*. American Journal of Community Psychology. Vol. 12, No.1, pp.71-85.
- Richardson C and Zumbo B (2000) *A statistical examination of the Health Utility Index-Mark III as a summary measure of health status for a general population survey*. Social Indicators Research. Vol. 51, No.2, pp.171-191.
- Richardson J (1994) *Cost utility analysis: what should be measured?* Social Science and Medicine. Vol. 39, No.1, pp.7-21.
- Richardson J and Hawthorne G (2000) *Negative Utilities and the Evaluation of Complex Health States: Issues Arising from the Scaling of a Multiattribute Utility Instrument*. Centre for Health Program Evaluation, Melbourne.
- Richeson NE (2003) *Effects of animal-assisted therapy on agitated behaviors and social interactions of older adults with dementia: an evidence-based therapeutic recreation intervention*. American Journal of Recreation Therapy. Vol. 2, No.4, pp.9-16.

- Richeson NE and Neill DJ (2004) *Therapeutic recreation music intervention to decrease mealtime agitation and increase food intake in older adults with dementia*. American Journal of Recreation Therapy. Vol. 3, No.1, pp.37-41.
- Riedel-Heller SG, Busse A, Aurich C, Matschinger H and Angermeyer MC (2001) *Prevalence of dementia according to DSM-III-R and ICD-10: results of the Leipzig Longitudinal Study of the Aged (LEILA75+) Part 1*. British Journal of Psychiatry. Vol. 179, pp.250-4.
- Riepe MW, Adler G, Ibach B, et al. (2006) *Adding memantine to rivastigmine therapy in patients with mild-to-moderate Alzheimer's disease: Results of a 12-week, open-label pilot study*. Primary Care Companion to the Journal of Clinical Psychiatry. Vol. 8, No.5, pp.258-263.
- Riemsma RP, Forbes CA, Glanville JM, et al. (2001) *General health status measures for people with cognitive impairment: learning disability and acquired brain injury*. Health Technology Assessment. Vol. 5, No.6, pp.1-100.
- Ringdal GI, Jordhoy MS and Kaasa S (2002) *Family satisfaction with end-of-life care for cancer patients in a cluster randomized trial*. Journal of Pain and Symptom Management. Vol. 24, No.1, pp.53-63.
- Ringdal GI, Jordhoy MS and Kaasa S (2003) *Measuring quality of palliative care: psychometric properties of the FAMCARE Scale*. Quality of Life Research. Vol. 12, No.2, pp.167-76.
- Robert PH, Clairet S, Benoit M, et al. (2002) *The apathy inventory: assessment of apathy and awareness in Alzheimer's disease, Parkinson's disease and mild cognitive impairment*. International Journal of Geriatric Psychiatry. Vol. 17, No.12, pp.1099-105.
- Roberts RE, Pascoe GC and Attkisson CC (1983) *Relationship of service satisfaction to life satisfaction and perceived well-being*. Evaluation and Program Planning. Vol. 6, No.3-4, pp.373-83.
- Robinson A, Dolan P and Williams A (1997) *Valuing health states using VAS and TTO: what lies behind the numbers?* Social Science and Medicine. Vol. 45, No.8, pp.1289-1297.
- Robinson A, Loomes G and Jones-Lee M (2001) *Visual analog scales, standard gambles and relative risk aversion*. Medical Decision Making. Vol. 21, pp.17-27.
- Rockwood K (2004) *Size of the treatment effect on cognition of cholinesterase inhibition in Alzheimer's disease*. Journal of Neurology, Neurosurgery and Psychiatry. Vol. 75, No.5, pp.677-85.
- Rockwood K (2007) *The measuring, meaning and importance of activities of daily living (ADLs) as an outcome*. International Psychogeriatrics. Vol. 19, No.3, pp.467-482.
- Rockwood K, Goodman J, Flynn M and Stolee P (1996) *Cross-validation of the Delirium Rating Scale in older patients*. Journal of the American Geriatrics Society. Vol. 44, No.7, pp.839-842.
- Rockwood K, Strang D, MacKnight C, et al. (2000) *Interrater reliability of the Clinical Dementia Rating in a multicenter trial*. Journal of the American Geriatrics Society. Vol. 48, No.5, pp.558-9.
- Rodgers H, Atkinson C, Bond S, et al. (1999) *Randomized controlled trial of a comprehensive stroke education program for patients and caregivers*. Stroke. Vol. 30, No.12, pp.2585-91.
- Rogers JC, Holm MB, Goldstein G, et al. (1994) *Stability and change in functional assessment of patients with geropsychiatric disorders*. American Journal of Occupational Therapy. Vol. 48, No.10, pp.914-918.
- Rogers SL and Friedhoff LT (1996) *The efficacy and safety of donepezil in patients with Alzheimer's disease: results of a US Multicentre, Randomized, Double-Blind, Placebo-Controlled Trial. The Donepezil Study Group*. Dementia. Vol. 7, No.6, pp.293-303.
- Rogers SL, Farlow MR, Doody RS, et al. (1998) *A 24-week, double-blind, placebo-controlled trial of donepezil in patients with Alzheimer's disease. Donepezil Study Group. [see comment]*. Neurology. Vol. 50, No.1, pp.136-45.

- 
- Rokach A (2000) *Loneliness and the life cycle*. Psychological Reports. Vol. 86, pp.629-642.
- Rolfson DB, McElhaney JE, Jhangri GS and Rockwood K (1999) *Validity of the confusion assessment method in detecting postoperative delirium in the elderly*. International Psychogeriatrics. Vol. 11, No.4, pp.431-438.
- Roman GC, Wilkinson DG, Doody RS, et al. (2005) *Donepezil in vascular dementia: combined analysis of two large-scale clinical trials*. Dementia and Geriatric Cognitive Disorders. Vol. 20, No.6, pp.338-44.
- Roman GC, Tatemichi TK, Erkinjuntti T, Cummings JL, Masdeu JC, Garcia JH, et al. (1993) *Vascular dementia: diagnostic criteria for research studies. Report of the NINDS-AIREN International Workshop. [see comment]*. Neurology. Vol. 43, No.2, pp.250-260.
- Robert PH, Verhey FRJ, Byrne EJ, Hurt C, De Deyn PP, Nobili F, et al. (2005) *Grouping for behavioural and psychological symptoms in dementia: clinical and biological aspects: Consensus paper on the European Alzheimer's disease consortium*. European Psychiatry. Vol. 20, pp.490-496.
- Rosen J, Bobys PD, Mazumdar S, Mulsant BH, et al. (1999) *OBRA regulations and neuroleptic use: defining agitation using the Pittsburgh Agitation Scale and the Neurobehavioral Rating Scale*. Annals of Long-Term Care. Vol. 7, No.12, pp.429-36.
- Rosen J, Burgio L, Kollar M, Cain M, et al. (1994) *The Pittsburgh Agitation Scale: a user friendly instrument for rating agitation in dementia patients*. American Journal of Geriatric Psychiatry. Vol. 2, pp.52-9.
- Rosen J, Sweet RA, Mulsant BH, Rifai AH, Pasternak R and Zubenko GS (1994) *The Delirium Rating Scale in a psychogeriatric inpatient setting*. Journal of Neuropsychiatry and Clinical Neurosciences. Vol. 6, No.1, pp.30-35.
- Rosen WG, Mohs R and Davis KL (1986) *Longitudinal changes: Cognitive, behavioral and affective patterns in Alzheimer's disease*. In Poon L, editor. Handbook for clinical memory assessment of older adults. Washington, DC: American Psychological Association. pp.294-301.
- Rosen WG, Mohs RC and Davis KL (1984) *A new rating scale for Alzheimer's disease* American Journal of Psychiatry. Vol. 141, No.11, pp.1356-64.
- Rosenberg PR, Meikle MM and Lyketsos CG (2005) *Caregiver assessment of patient's depression in Alzheimer's Disease: longitudinal analysis in a drug treatment study*. American Journal of Geriatric Psychiatry. Vol. 13, pp.822-826.
- Ross C, Steward C and Sinacore J (1993) *The importance of patient preferences in the measurement of health care satisfaction*. Medical Care. Vol. 31, No.12, pp.1138-1149.
- Ross CK, Frommelt G, Hazelwood L, et al. (1987) *The role of expectations in patient satisfaction with medical care*. J Health Care Mark. Vol. 7, No.4, pp.16-26.
- Rosser R (1993) *A health index and output measure*. In Walker S and Rosser R (eds.) Quality of Life Assessment: Key Issues in the 1990s. Kluwer Academic Publishers, Dordrecht.
- Roth M (1981) *The diagnosis of dementia in late and middle life*. In Mortimer JA, Schuman LM, (eds.) The epidemiology of dementia. Oxford University Press, New York. pp.24-61.
- Roth M, Tym E, Mountjoy CQ, et al. (1986) *CAMDEX: A standardised instrument for the diagnosis of mental disorder in the elderly with special reference to the early detection of dementia*. British Journal of Psychiatry. Vol. 149, pp.698-709.
- Rothenhausler HB, Grieser B, Nollert G, Reichart B, Schelling G and Kapfhammer HP (2005) *Psychiatric and psychosocial outcome of cardiac surgery with cardiopulmonary bypass: a prospective 12-month follow-up study*. General Hospital Psychiatry. Vol. 27, No.1, pp.18-28.
- Rouse MO (1967) *Organizing and delivering health care*. 3. JAMA. Vol. 202, No.4, pp.305.
-

- Routasalo PE, Savikko N, Tilvis RS, et al. (2006) *Social contacts and their relationship to loneliness among aged people - a population-based study*. Gerontology. Vol. 52, No.3, pp.181-7.
- Rowbotham MC (1953) *Centralized or decentralized service; Faulkner Hospital combines them for greater efficiency and patient satisfaction*. Mod Hosp. Vol. 80, No.6, pp.114-6.
- Rowland JT, Basic D, Storey JE, et al. (2006) *The Rowland Universal Dementia Assessment Scale (RUDAS) and the Folstein MMSE in a multicultural cohort of elderly persons*. International Psychogeriatrics. Vol. 18, No.1, pp.111-20.
- Royal Australian College of General Practitioners (RACGP) (2005) *Medical care of older persons in residential aged care facilities (Silver book)* (4th Ed) Royal Australian College of General Practitioners.
- Ruberman W, Weinblatt E, Goldberg J, et al. (1984) *Psychosocial influences on mortality after myocardial infarction*. New England Journal of Medicine. Vol. 311, pp.552-559.
- Rubin HR, Gandek B, Rogers WH, et al. (1993) *Patients' ratings of outpatient visits in different practice settings. Results from the Medical Outcomes Study*. Journal of the American Medical Association. Vol. 270, No.7, pp.835-40.
- Ruggeri M and Dall'Agnola R (1993) *The development and use of the Verona Expectations for Care Scale (VECS) and the Verona Service Satisfaction Scale (VSSS) for measuring expectations and satisfaction with community-based psychiatric services in patients, relatives and professionals*. Psychological Medicine. Vol. 23, pp.511-523.
- Ruchinkas RA, Singer HK and Repetz NK (2001) *Clock drawing, clock copying, and physical abilities in geriatric rehabilitation*. Archives of Physical Medicine and Rehabilitation. Vol. 82, No.7, pp.920-924.
- Ruiz Perez I, Rodriguez Bano J, Lopez Ruz MA, et al. (2005) *Health-related quality of life of patients with HIV: impact of sociodemographic, clinical and psychosocial factors*. Quality of Life Research. Vol. 14, No.5, pp.1301-10.
- Russell D (1982) *The measurement of loneliness*. In Peplau L, Perlman D, (eds.) Loneliness: A Sourcebook of Current Theory, Research and Therapy. John Wiley, New York. pp.81-104.
- Russell D, Peplau L and Cutrona C (1980) *The revised UCLA Loneliness Scale: concurrent and discriminant validity evidence*. Journal of Personality and Social Psychology. Vol. 39, No.3, pp.472-480.
- Russell D, Peplau LA and Ferguson ML (1978) *Developing a measure of loneliness*. Journal of Personality Assessment. Vol. 42, No.3, pp.290-4.
- Russell DW (1996) *UCLA Loneliness Scale (Version 3): reliability, validity, and factor structure*. Journal of Personality Assessment. Vol. 66, No.1, pp.20-40.
- Ryden M (1988) *Aggressive behaviour in persons with dementia who live in the community*. Alzheimer Disease and Associated Disorders. Vol. 2, pp.342-355.
- Sackett D and Torrance G (1978) *The utility of different health states as perceived by the general public*. Journal of Chronic Diseases. Vol. 31, pp.697-704.
- Sadler EA, Braam AW, Broese van Groenou MI, et al. (2006) *Cosmic transcendence, loneliness, and exchange of emotional support with adult children: a study among older parents in The Netherlands*. European Journal of Ageing. Vol. 3, No.3, pp.146.
- Sainfort F, Becker M and Diamond R (1996) *Judgements of quality of life of individuals with severe mental disorders: patient self-report versus provider perspectives*. American Journal of Psychiatry. Vol. 153, pp.497-502.
- Samus QM, Rosenblatt A, Steele C, Baker A, et al. (2005) *The association of neuropsychiatric symptoms and environment with quality of life in assisted living residents with dementia*. Gerontologist. Vol. 45 Spec No 1, No.1, pp.19-26.

- Sanavio E (1988) *Obsessions and compulsions: the Padua Inventory*. Behaviour Research and Therapy. Vol. 26, pp.167-177.
- Sand L and Strang P (2006) *Existential loneliness in a palliative home care setting*. Journal of Palliative Medicine. Vol. 9, No.6, pp.1376-87.
- Sands LP, Ferreira P, Stewart AL, et al. (2004) *What explains differences between dementia patients' and their caregivers' ratings of patients' quality of life*. Am J Geriatric Psychiatry. Vol. 12, No.3, pp.272-280.
- Sansoni J, Marosszeky N, Sansoni E, et al. (2006) *Refining Continence Measurement Tools*. Centre for Health Service Development, University of Wollongong and the Department of Psychiatry, The University of Melbourne.
- Sansoni, J and Senior K (1998) *Health outcomes measurement and the coordinated care trials for Aboriginal and Torres Straits islander peoples*. Health Outcomes. Bulletin 9.
- Sarason I, Sarason B, Shearin E, et al. (1987) *A brief measure of social support: practical and theoretical implications*. Journal of Social and Personal Relationships. Vol. 4, pp.497-510.
- Sarason IG, Levine H, Basham RB, et al. (1983) *Assessing social support: the Social Support Questionnaire*. Journal of Personality and Social Psychology. Vol. 44, No.1, pp.127-139.
- Saxton J, McGonigle-Gibson K, Swihart A, et al. (1990) *Assessment of severely impaired patients: Description and validation of a new neuropsychological test battery*. Psychological Assessment. Vol. 2, No.298-303.
- Schafer KA, Tractenberg RE, Sano M, et al. (2004) *Reliability of monitoring the clinical dementia rating in multicenter clinical trials*. Alzheimer Disease and Associated Disorders. Vol. 18, No.4, pp.219-22.
- Schaffer MA and Lia-Hoagberg B (1997) *Effects of social support on prenatal care and health behaviors of low-income women*. Journal of Obstetric, Gynecologic, and Neonatal Nursing. Vol. 26, No.4, pp.433-40.
- Schechter S, Beatty P and Willis GB (1999) *Asking survey respondents about health status: Judgment and response issues*. In Schwarz N, Park DC, Knauper B, Sudman S, (eds.) Cognition, ageing, and self reports. Psychology Press, Taylor and Francis Group, Philadelphia.
- Schipper H, Clinch JJ and Olweny CLM (1996) *Quality of life studies: definitions and conceptual issues*. In Spilker B (ed.) Quality of Life and Pharmacoeconomics in Clinical Trials. 2 ed. Lippincott-Raven, New York. pp.11-23.
- Schmeidler J, Mohs RC and Aryan M (1998) *Relationship of disease severity to decline on specific cognitive and functional measures in Alzheimer disease*. Alzheimer Disease and Associated Disorders. Vol. 12, No.3, pp.146-51.
- Schneider F, Bohner H, Habel U, Salloum JB, Stierstorfer A, Hummel TC, et al. (2002) *Risk factors for postoperative delirium in vascular surgery*. General Hospital Psychiatry. Vol. 24, No.1, pp.28-34.
- Schneider J, Murray J, Banerjee S, et al. (1999) *EURO CARE: a cross-national study of co-resident spouse carers for people with Alzheimer's disease: I--Factors associated with carer burden*. International Journal of Geriatric Psychiatry. Vol. 14, No.8, pp.651-61.
- Schneider LS (2001) *Assessing outcomes in Alzheimer Disease*. Alzheimer Disease and Associated Disorders. Vol. 15, No.Suppl 1, pp.S8-S18.
- Schreiner AS, Yamamoto E and Shiotani H (2000) *Agitated behavior in elderly nursing home residents with dementia in Japan*. Journals of Gerontology Series B-Psychological Sciences and Social Sciences. Vol. 55, No.3, pp.P180-6.
- Schreinzer D, Ballaban T, Brannath W, Lang T, et al. (2005) *Components of behavioral pathology in dementia*. International Journal of Geriatric Psychiatry. Vol. 20, No.2, pp.137-45.

- Schulz R, Belle SH, Czaja SJ, McGinnis KA, et al. (2004) *Long-term care placement of dementia patients and caregiver health and well-being*. Journal of the American Medical Association. Vol. 292, No.8, pp.961-967.
- Schulz R, O'Brien A, Czaja S, Ory M, et al. (2002) *Dementia caregiver intervention research: In search of clinical significance*. The Gerontologist. Vol. 42, No.5, pp.589-602.
- Schulz R and Williamson GM (1997) *The measurement of caregiver outcomes in Alzheimer disease research*. Alzheimer Disease and Associated Disorders. Vol. 11, Suppl 6, pp.117-24.
- Schultz RR, Siviero MO and Bertolucci PH (2001) *The cognitive subscale of the "Alzheimer's Disease Assessment Scale" in a Brazilian sample*. Brazilian Journal of Medical and Biological Research. Vol. 34, No.10, pp.1295-302.
- Schwarz R and Hinze A (2001) *Reference data for the quality of life questionnaire EORTC QLQ-C30 in the general German population*. European Journal of Cancer. Vol. 37, No.11, pp.1345-51.
- Sclan S, Reisberg B (1992) *FAST in Alzheimer's disease: reliability, validity and ordinality*. International Psychogeriatrics. Vol. 4 (suppl. 1), pp. 55-69.
- Sclan SG, Saillon A, et al. (1996) *The Behavior Pathology in Alzheimer's Disease Rating Scale (BEHAVE-AD): Reliability and analysis of symptom category scores*. International Journal of Geriatric Psychiatry. Vol. 11, No.9, pp.819-830.
- Scocco P, Fantoni G and Caon F (2005) *Role of depressive and cognitive status in self-reported evaluation of quality of life in older people: comparing proxy and physician perspectives*. Age Ageing. Vol. 35, No.2, pp.166-71.
- Scottish Intercollegiate Guidelines Network (SIGN) (2006) *Management of patients with dementia: A national clinical guideline*. SIGN, Edinburgh.
- Seale CF (1989) *What happens in hospices: a review of research evidence*. Social Science and Medicine. Vol. 28, No.6, pp.551-9.
- Seckel MM and Birney MH (1996) *Social support, stress, and age in women undergoing breast biopsies*. Clinical Nurse Specialist. Vol. 10, No.3, pp.137-43.
- Selwood A, Thorgrimsen L and Orrell M (2005) *Quality of life in dementia--a one-year follow-up study*. International Journal of Geriatric Psychiatry. Vol. 20, No.3, pp.232-7.
- Senior K (2003) *A gudbala Laif? Health and well being in a remote Aboriginal community-what are the problems and where lies responsibility?* PhD thesis, Australian National University, Canberra.
- Serra L, Perri R, Carlesimo GA, et al. (2004) *ADAS-CoG and mental deterioration battery: Different instruments for different aspects of cognitive decline in Alzheimer's disease*. Neuroscience Research Communications. Vol. 35, No.3, pp.184-192.
- Seymour DG, Ball AE, Russell EM, et al. (2001) *Problems in using health survey questionnaires in older patients with physical disabilities. The reliability and validity of the SF-36 and the effect of cognitive impairment*. Journal of Evaluation in Clinical Practice. Vol. 7, pp. 411-8.
- Shader R, Harmatz J, Salzman C (1974) *A New Scale for Clinical Assessment in Geriatric Populations: Sandoz Clinical Assessment – Geriatric (SCAG)*. JAGS. Vol.12, No.3, pp.107-113.
- Shah A and Allen H (1999) *Editorial: Is Improvement possible in the measurement of behaviour disturbance in dementia?* International Journal of Geriatric Psychiatry. Vol. 14, pp.512-519.
- Shah A (1999a) *Some methodological issues in using aggression rating scales in intervention studies among institutionalised elderly*. Int. Psychogeriatric. Vol. 11, No.4, pp.439-444.

- Shah A (1999b) *Aggressive behaviour in the elderly*. Int J. Psychiat.Clin. Pract. Vol. 3, pp.85-103.
- Shah A, Chiu E and Ames D (2000) *Aggressive behaviour and environmental characteristics in Australian nursing homes*. Ageing and Mental Health. Vol. 4, No.1, pp.43-47.
- Shah A, Chiu E, Ames D, Harrigan S and McKenzie D (2000) *Characteristics of aggressive subjects in Australian (Melbourne) nursing homes*. International Psychogeriatrics. Vol. 12, No.2, pp.145-161.
- Shah A, Dalvi M and Thompson T (2005) *Behavioural and psychological signs and symptoms of dementia across cultures: current status and the future*. International Journal of Geriatric Psychiatry. Vol. 20, No.12, pp.1187-95.
- Shah A and De T (1997) *The relationship between two scales measuring aggressive behaviour among continuing-care psychogeriatric inpatients*. International Psychogeriatrics. Vol. 9, pp.471-477.
- Shah A, Ellanchenny N, et al. (2005) *A comparative study of behavioral and psychological symptoms of dementia in patients with Alzheimer's disease and vascular dementia referred to psychogeriatric services in Korea and the U.K.* Int Psychogeriatr. Vol. 17, No.2, pp.207-19.
- Shah A, Evans H and Parkash N (1998) *Evaluation of three aggression/agitation behaviour rating scales for use on an acute admission and assessment psychogeriatric ward*. International Journal of Geriatric Psychiatry. Vol. 13, No.6, pp.415-420.
- Shalek M, Richeson NE and Buettner LL (2004) *Air mat therapy for treatment of agitated wandering: an evidence-based recreational therapy intervention*. American Journal of Recreation Therapy. Vol. 3, No.3, pp.18-26.
- Shankar KK, Walker M, Frost D, et al. (1999) *The development of a valid and reliable scale for rating anxiety in dementia (RAID)*. Ageing and Mental Health Vol.3, No.1, pp.39-49.
- Shaughnessy PW, Crisler KS and Schlenker RE (1997) *Medicare's OASIS: Standardized Outcome and Assessment Information Set for Home Health Care, OASIS-B*. Center for Health Services and Policy Research, University of Colorado Health Sciences Center.
- Shaw JW, Johnson JA and Coons SJ (2005) *US valuation of the EQ-5D health states: development and testing of the D1 valuation model*. Medical Care. Vol. 43, No.3, pp.203-20.
- Shepperd S and Iliffe S (1998) *The effectiveness of hospital at home compared with in-patient hospital care: a systematic review*. Journal of Public Health Medicine. Vol. 20, No.3, pp.344-350.
- Shepperd S and Iliffe S (2005) *Hospital-at-home versus in-patient hospital care*. Cochrane Database of Systematic Reviews. No.3, pp.CD000356.
- Sherbourne CD and Meredith L (1992) *Quality of self-report data: A comparison of older and younger chronically ill patients*. Journal of Gerontology. Vol. 47, pp.S204-S211.
- Sherbourne C and Stewart A (1991) *The MOS social support survey*. Social Science and Medicine. Vol. 32, pp.705-714.
- Shergill S, Shankar K, Seneviratna K and Orrell M (1999) *The validity and reliability of the Health of the Nation Outcome Scales (HoNOS) in the elderly*. Journal of Mental Health (UK). Vol. 8, pp.511-521.
- Shin IS, Carter M, Masterman D, et al. (2005) *Neuropsychiatric symptoms and quality of life in Alzheimer disease*. American Journal of Geriatric Psychiatry. Vol. 13, No.6, pp.469-74.
- Sigelman CK, Budd EC, Spanhel CL and Schoenrock CJ (1981a) *Asking questions of retarded persons: a comparison of yes-no and either-or formats*. Applied Research in Mental Retardation. Vol. 2, No.4, pp.347-57.
- Sigelman CK, Budd EC, Spanhel CL and Schoenrock CJ (1981b) *When in doubt, say yes: acquiescence in interviews with mentally retarded persons*. Mental Retardation. Vol. 19, No.2, pp.53-8.

- Sinoff G, Ore L, Zlotogorsky D, et al. (1999) *Short Anxiety Screening Test - a brief instrument for detecting anxiety in the elderly*. International Journal of Geriatric Psychiatry. Vol. 14, pp.1062-1071.
- Shore D, Overman CA and Wyatt RJ (1983) *Improving accuracy in the diagnosis of dementia*. Journal of Clinical Psychiatry. Vol. 44, pp.207-212.
- Shyu YI, Tang WR, Liang J, et al. (2006) *Psychometric testing of the social support survey on a Taiwanese sample*. Nursing Research. Vol. 55, No.6, pp.411-7.
- Sigelman CK, Budd EC, Spanhel CL and Schoenrock CJ (1981a) *Asking questions of retarded persons: a comparison of yes-no and either-or formats*. Applied Research in Mental Retardation. Vol. 2, No.4, pp.347-57.
- Sigelman CK, Budd EC, Spanhel CL and Schoenrock CJ (1981b) *When in doubt, say yes: acquiescence in interviews with mentally retarded persons*. Mental Retardation. Vol. 19, No.2, pp.53-8.
- Silver MH, Jilinskaia E and Perls TT (2001) *Cognitive functional status of age-confirmed centenarians in a population-based study*. Journals of Gerontology Series B - Psychological Sciences and Social Sciences. Vol. 56, No.3, pp.P134-P140.
- Silvestrini M, Pasqualetti P, Baruffaldi R, et al. (2006) *Cerebrovascular reactivity and cognitive decline in patients with Alzheimer disease*. Stroke. Vol. 37, No.4, pp.1010-5.
- Simon C, Little P, Birtwistle J, et al. (2003) *A questionnaire to measure satisfaction with community services for informal carers of stroke patients: construction and initial piloting*. Health and Social Care in the Community. Vol. 11, No.2, pp.129-37.
- Simpson RG (1997) *Carers as equal partners in care planning*. Journal of Psychiatric and Mental Health Nursing. Vol. 4, No.5, pp.345-54.
- Simpson RG, Scothern G and Vincent M (1995) *Survey of carer satisfaction with the quality of care delivered to in-patients suffering from dementia*. Journal of Advanced Nursing. Vol. 22, No.3, pp.517-27.
- Sin MK, Kang DH and Weaver M (2005) *Relationships of asthma knowledge, self-management, and social support in African American adolescents with asthma*. International Journal of Nursing Studies. Vol. 42, No.3, pp.307-13.
- Sinha D, Zemlan FP, Nelson S, Bienenfeld D, Thienhaus O, Ramaswamy G, et al. (1992) *A new scale for assessing behavioral agitation in dementia*. Psychiatry Research. Vol. 41, No.1, pp.73-88.
- Sinha SP and Nayyar P (2002) *Social support and self-control as variables in attitude toward life and perceived control among older people in India*. Journal of Social Psychology. Vol. 142, No.4, pp.527-40.
- Sintonen H (1994) *The 15D measure of health-related quality of life: feasibility, reliability and validity of its valuation system*, National Centre for Health Program Evaluation, Melbourne.
- Sintonen H (2001) *The 15D instrument of health-related quality of life: properties and applications*. Annals of Medicine. Vol. 33, pp.328-336.
- Sintonen H and Pekurinen M (1993) *A fifteen-dimensional measure of health-related quality of life (15D) and its applications*. In Walker S and Rosser R (eds.) Quality of Life Assessment: Key Issues in the 1990s. Kluwer Academic Publishers, Dordrecht.
- Sitzia J (1999) *How valid and reliable are patient satisfaction data? An analysis of 195 studies*. International Journal for Quality in Health Care. Vol. 11, No.4, pp.319-28.
- Sitzia J and Wood N (1997) *Patient satisfaction: a review of issues and concepts*. Social Science and Medicine. Vol. 45, No.12, pp.1829-1843.

- Skolnick AA (1998) *MediCaring Project to demonstrate, evaluate innovative end-of-life program for chronically ill*. JAMA. Vol. 279, No.19, pp.1511-2.
- Skjerve A, Holsten F, Aarsland D, Bjorvatn B, et al. (2004) *Improvement in behavioral symptoms and advance of activity acrophase after short-term bright light treatment in severe dementia*. Psychiatry and Clinical Neurosciences. Vol. 58, No.4, pp.343-7.
- Slevin ML, Plant H, Lynch D, et al. (1988) *Who should measure quality of life, the doctor or the patient?* British Journal of Cancer. Vol. 57, No.1, pp.109-12.
- Sloane PD, Zimmerman S, Williams CS, et al. (2005) *Evaluating the quality of life of long-term care residents with dementia*. Gerontologist. Vol. 45, Spec No 1, No.1, pp.37-49.
- Smith SC, Lamping DL, Banerjee S, et al. (2005) *Measurement of health-related quality of life for people with dementia: development of a new instrument (DEMQOL) and an evaluation of current methodology*. Health Technology Assessment (Winchester, England). Vol. 9, No.10, pp.1-93.
- Smith SC, Murray J, Banerjee S, et al. (2005b) *What constitutes health-related quality of life in dementia? Development of a conceptual framework for people with dementia and their carers*. International Journal of Geriatric Psychiatry. Vol. 20, No.9, pp.889-95.
- Snellgrove CA and Hecker JR (2005) *Brief checklist for non-cognitive symptoms of dementia*. Australasian Journal on Ageing. Vol. 24, No.2, pp.88-93.
- Snow AL, Cook KF, Lin P-S, et al. (2005a) *Proxies and other external raters: Methodological considerations*. Health Services Research. Vol. 40, No.5, pp.1676-1693.
- Snow AL, Kunik ME, Molinari VA, et al. (2005b) *Accuracy of self-reported depression in persons with dementia*. JAGS. Vol. 53, pp.389-396.
- Snyder M, Ryden MB, Shaver P, et al. (1998) *The Apparent Emotion Rating Instrument: assessing affect in cognitively impaired elders*. Clinical Gerontologist. Vol.18, No.4, pp.17-29.
- Sockeel P, Dujardin K, Devos D, et al. (2006) *The Lille Apathy Rating Scale (LARS): a new instrument for detecting and quantifying apathy: Validation in Parkinson's disease*. Journal of Neurology, Neurosurgery and Psychiatry. Vol.77, No.5, pp.579-584.
- Solano CH (1980) *Two measures of loneliness: a comparison*. Psychological Reports. Vol. 46, No.1, pp.23-8.
- Soliman AA (1992) *Assessing the quality of health care: a consumerist approach*. Health Marketing Quarterly. Vol. 10, No.1-2, pp.121-41.
- Solomon P, Adams F, Groccia M, et al. (1999) *Correlational Analysis of Five Commonly Used Measures of Mental Status/Functional Abilities in Patients with Alzheimer Disease*. Alz Dis and Ass Disorders. Vol.13, No.3 pp.147-150.
- Solomon PR, Hirschhoff A, Kelly B, et al. (1998) *A 7 minute neurocognitive screening battery highly sensitive to Alzheimer's disease*. Archives of Neurology. Vol. 55, pp.349-355.
- Sorensen S, Pinquart M and Duberstein P (2002) *How effective are interventions with caregivers? An updated meta-analysis*. Gerontologist. Vol. 42, No.3, pp.356-72.
- Spark R Donovan RJ and Binns C (1992) *Aboriginal peoples' perceptions of their own and their community's health*. Health Promotion Journal of Australia. pp.60-61.
- Spear J (2003) *A new measure of consumer expectations, perceptions and satisfaction for patients and carers of older people with mental health problems*. Australasian Psychiatry. Vol. 11, No.3, pp.330.

- Spear J, Chawla S, O'Reilly M and Rock D (2002) *Does the HoNOS 65+ meet the criteria for a clinical outcome indicator for mental health services for older people?* International Journal of Geriatric Psychiatry. Vol. 17, pp.226-230.
- Spector A and Orrell M (2006) *Quality of Life (QoL) in Dementia: A Comparison of the Perceptions of People With Dementia and Care Staff in Residential Homes*. Alzheimer Disease and Associated Disorders. Vol. 20, No.3, pp.160-5.
- Spector A, Thorgrimsen L, Woods B, et al. (2003) *Efficacy of an evidence-based cognitive stimulation therapy programme for people with dementia: randomised controlled trial.[see comment]*. British Journal of Psychiatry. Vol. 183, pp.248-54.
- Spector WD (1997) *Measuring functioning in daily activities for persons with dementia*. Alzheimer Disease and Associated Disorders. Vol. 11, No.Suppl 6, pp.81-90.
- Spielberger CD, Gorsuch R and Lushene RE (1970) *Manual for the State-Trait Anxiety Inventory*. Consulting Psychologists Press, Palo Alto, CA.
- Spiegel R, Brunner C, Ermini-Funfschilling D, et al. (1991) *A new behavioural assessment scale for geriatric out- and in-patients: the NOSGER (Nurses' Observation Scale for Geriatric Patients)*. JAGS. Vol. 39, pp.339-347.
- Spitzer WO (1987) *State of science 1986: Quality of life and functional status as target variables for research*. Journal of Chronic Diseases. Vol. 40, No.6, pp.465-471.
- Sprangers M and Aaronson N (1992) *The role of health care providers and significant others in evaluating the quality of life of patients with chronic disease: a review*. Journal of Clinical Epidemiology. Vol. 45, No.7, pp.743-760.
- Springer MB, Newman A, Weaver AJ, et al. (2003) *Spirituality, depression, and loneliness among Jewish seniors residing in New York City*. Journal of Pastoral Care and Counseling. Vol. 57, No.3, pp.305-18.
- Starkstein SE, Ingram L, Garau ML, et al. (2005) *On the overlap between apathy and depression in dementia*. Journal of Neurology, Neurosurgery and Psychiatry. Vol. 76, No.8, pp.1070-1074.
- Starkstein SE, Mayberg HS, Preziosi TJ, et al. (1992) *Reliability, validity, and clinical correlates of apathy in Parkinson's disease*. Journal of Neuropsychiatry and Clinical Neurosciences. Vol. 4, No.2, pp.134-9.
- Stavem K (1999) *Reliability, validity and responsiveness of two multiattribute utility measures in patients with chronic obstructive pulmonary disease*. Quality of Life Research. Vol. 8, No.1-2, pp.45-54.
- Steffens DC, Otey E, Alexopoulos GS, et al. (2006) *Perspectives on depression, mild cognitive impairment, and cognitive decline*. Archives of General Psychiatry. Vol. 63, No.2, pp.130-8.
- Steinberg, M., Tschanz, JT., Norton, M., Breitner, JCS. and Liketsos, CG. (2006) *Point and five-year period prevalence of neuropsychiatric symptoms in dementia: The Catch County Study*. Poster, International Conference on Alzheimer's Disease, Madrid, July 2006.
- Stern Y, Hesdorffer D, Sano M and Mayeux R (1990) *Measurement and prediction of functional capacity in Alzheimer's disease*. Neurology. Vol. 40, No. 1, pp.8-14
- Stephenson C, Wilson S and Gladman JR (1995) *Patient and carer satisfaction in geriatric day hospitals*. Disability and Rehabilitation. Vol. 17, No.5, pp.252-5.
- Steptoe A, Owen N, Kunz-Ebrecht SR, et al. (2004) *Loneliness and neuroendocrine, cardiovascular, and inflammatory stress responses in middle-aged men and women*. Psychoneuroendocrinology. Vol. 29, No.5, pp.593-611.
- Stern Y (1994) *Assessing Patient Dependence in Alzheimer's disease*. Journal of Gerontology. Vol. 49, No.5, pp.216-225.

- 
- Stern Y, Hesdorffer D, Sano M and Mayeux R (1991) *Measurement and prediction of functional capacity in Alzheimer's disease*. Neurology. Vol. 40, No.1, pp.8-14.
- Stevens N and Westerhof GJ (2006) *Marriage, social integration, and loneliness in the second half of life: a comparison of Dutch and German men and women*. Research on Aging. Vol. 28, No.6, pp.713.
- Stewart A, Phillips R and Dempsey G (1998) *Pharmacotherapy for people with Alzheimer's disease: a Markov-cycle evaluation of five years' therapy using donepezil*. International Journal of Geriatric Psychiatry. Vol. 13, No.7, pp.445-53.
- Stewart MJ, Hirth AM, Klassen G, et al. (1997) *Stress, coping, and social support as psychosocial factors in readmissions for ischaemic heart disease*. International Journal of Nursing Studies. Vol. 34, No.2, pp.151-63.
- Stokes JP and Wilson DG (1984) *The inventory of socially supportive behaviors: dimensionality, prediction, and gender differences*. American Journal of Community Psychology. Vol. 12, No.1, pp.53-69.
- Stone SP, Ali B, Auberleek I, et al. (1994) *The Barthel Index in clinical practice: use on a rehabilitation ward for elderly people*. Journal of the Royal College of Physicians of London. Vol. 28, No.5, pp.419-423.
- Storch EA, Roth DA, Coles ME, et al. (2004) *The measurement and impact of childhood teasing in a sample of young adults*. Journal of Anxiety Disorders. Vol. 18, No.5, pp.681-94.
- Storey JE, Rowland JT, Basic D, et al. (2004) *The Rowland Universal Dementia Assessment Scale (RUDAS): a multicultural cognitive assessment scale*. [erratum appears in *Int Psychogeriatry*. 2004 Jun;16(2):218]. International Psychogeriatrics. Vol. 16, No.1, pp.13-31.
- Storey JE, Rowland JT, Basic D, et al. (2002) *Accuracy of the clock drawing test for detecting dementia in a multicultural sample of elderly Australian patients*. International Psychogeriatrics. Vol. 14, No.3, pp.259-71.
- Straker DA, Shapiro PA and Muskin PR (2006) *Aripiprazole in the treatment of delirium*. Psychosomatics. Vol. 47, No.5, pp.385-391.
- Strauss ME and Sperry SD (2002) *An informant-based assessment of apathy in Alzheimer disease*. Neuropsychiatry, Neuropsychology, and Behavioral Neurology. Vol.15, No.3, pp.176-183.
- Streiner DL and Norman GR (2003) *Health measurement scales: A practical guide to their development and use*. (3rd ed.) Oxford University Press, Oxford.
- Streiner D and Norman G (2006) *Health Measurement Scales: A Practical Guide to their Development and Use*. Oxford Medical Publications, Oxford.
- Stull DE, Kosloski K and Kercher K (1994) *Caregiver burden and generic well-being: opposite sides of the same coin?*[see comment]. Gerontologist. Vol. 34, No.1, pp.88-94.
- Sturm JW, Donnan GA, Dewey HM, Macdonell RA, Gilligan AK, Srikanth V and Thrift AG (2004) *Quality of life after stroke: the North East Melbourne Stroke Incidence Study (NEMESIS)*. Stroke. Vol. 35, No.10, pp.2340-5.
- Sturm JW, Osborne RH, Dewey HM, Donnan GA, Macdonell RA and Thrift AG (2002) *Brief comprehensive quality of life assessment after stroke: the Assessment of Quality of Life instrument in the north East Melbourne stroke incidence study (NEMESIS)*. Stroke. Vol. 33, No.12, pp.2888-94.
- Stuss DT, van Reekum R, Murphy KJ (2000) *Differentiation of states and causes of apathy*. In Borod J (ed.) The Neuropsychology of Emotion. Oxford University Press, New York. pp 340-363.
- Suchman EA (1965) *Social patterns of illness and medical care*. Journal of Health and Human Behavior. Vol. 106, pp.2-16.
-

- Suh GH (2004) *Agitated behaviours among the institutionalized elderly with dementia: validation of the Korean version of the Cohen-Mansfield Agitation Inventory*. International Journal of Geriatric Psychiatry. Vol. 19, No.4, pp.378-85.
- Suh GH and Kim SK (2004) *Behavioral and Psychological Signs and Symptoms of Dementia (BPSD) in antipsychotic-naive Alzheimer's Disease patients*. International Psychogeriatrics. Vol. 16, No.3, pp.337-50.
- Suh GH, Greenspan AJ and Choi SK (2006) *Comparative efficacy of risperidone versus haloperidol on behavioural and psychological symptoms of dementia*. International journal of geriatric psychiatry. Vol. 21, No.7, pp.654-60.
- Suh GH, Ju YS, Yeon BK, et al. (2004) *A longitudinal study of Alzheimer's disease: rates of cognitive and functional decline*. International Journal of Geriatric Psychiatry. Vol. 19, No.9, pp.817-24.
- Suh GH, Son HG, Ju YS, Jcho KH, et al. (2004) *A randomized, double-blind, crossover comparison of risperidone and haloperidol in Korean dementia patients with behavioral disturbances*. American Journal of Geriatric Psychiatry. Vol. 12, No.5, pp.509-16.
- Sullivan M, Karlsson J and Ware J (1995) *The Swedish SF-36 Health Survey--I. Evaluation of data quality, scaling assumptions, reliability and construct validity across general populations in Sweden*. Social Science and Medicine. Vol. 41, No.10, pp.1349-58.
- Sultzer DL, Berisford MA, et al. (1995) *The Neurobehavioral Rating Scale: reliability in patients with dementia*. Journal of Psychiatric Research. Vol. 29, No 3, pp.185-91.
- Sultzer DL, Levin HS, et al. (1992) *Assessment of cognitive, psychiatric, and behavioral disturbances in patients with dementia: the Neurobehavioral Rating Scale*. Journal of the American Geriatrics Society. Vol. 40, No.6, pp.549-55.
- Sultzer DL, Levin HS, et al. (1993) *A comparison of psychiatric symptoms in vascular dementia and Alzheimer's disease*. American Journal of Psychiatry. Vol. 150, No.12, pp.1806-12.
- Sultzer DL, Mahler ME, et al. (1995) *The relationship between psychiatric symptoms and regional cortical metabolism in Alzheimer's disease*. Journal of Neuropsychiatry and Clinical Neurosciences. Vol. 7, No.4, pp.476-84.
- Summers WK, DeBoynton V, Marsh GM, et al. (1990) *Comparison of seven psychometric instruments used for evaluation of treatment effect in Alzheimer's dementia*. Neuroepidemiology. Vol. 9, No.4, pp.193-207.
- Sunderland T, Hill JL, Mellow AM, et al. (1989) *Clock drawing in Alzheimer's disease*. Journal of the American Geriatrics Society. Vol. 37, pp.725-729.
- Sunderland T, Hill JL, Lawlor BA, et al. (1988) *NIMH Dementia Mood Assessment Scale (DMAS)*. Psychopharmacology Bulletin. Vol. 24, No.4, pp.747-53.
- Swami V, Chamorro-Premuzic T, Sinniah D, et al. (2007) *General health mediates the relationship between loneliness, life satisfaction and depression : A study with Malaysian medical students*. Social Psychiatry and Psychiatric Epidemiology. Vol. 42, No.2, pp.161-6.
- Swearer JM, Hoople NE, Kane KJ and Drachman DA (1996) *Predicting aberrant behaviour in Alzheimer's Disease*. Neuropsychiatry, Neuropsychology and Behavioural Neurology. Vol. 9, No.3, pp.162-170.
- Swindells S, Mohr J, Justis JC, et al. (1999) *Quality of life in patients with human immunodeficiency virus infection: impact of social support, coping style and hopelessness*. International Journal of STD and AIDS. Vol. 10, No.6, pp.383-91.
- Tabert MH, Albert SM, Borukhova-Milov L, et al. (2002) *Functional deficits in patients with mild cognitive impairment: Prediction of AD*. Neurology. Vol. 58, pp.758-764.
- Takeuchi T, Furuta K, Hirasawa T, Masaki H, Yukizane T, Atsuta H, et al. (2007) *Perospirone in the treatment of patients with delirium*. Psychiatry and Clinical Neurosciences. Vol. 61, No.1, pp.67-70.

- Talwalker S, Overall JE, Srirama MK, et al. (1996) *Cardinal features of cognitive dysfunction in Alzheimer's disease: A factor analytic study of the Alzheimer's Disease Assessment Scale*. Journal of Geriatric Psychiatry and Neurology. Vol. 9, pp.39-46.
- Tamim H, McCusker J and Dendukuri N (2002) *Proxy reporting of quality of life using the EQ-5D*. Medical Care. Vol. 40, No.12, pp.1186-95.
- Tariot PN (1996) *CERAD Behavior Rating Scale for Dementia*. International Psychogeriatrics. Vol. 8, Suppl.3, pp.514-515.
- Tariot PN, Cummings JL, Katz IR, et al. (2001) *A randomized, double-blind, placebo-controlled study of the efficacy and safety of donepezil in patients with Alzheimer's disease in the nursing home setting*. Journal of the American Geriatrics Society. Vol. 49, No.12, pp.1590-9.
- Tariot PN, Mack JL, Patterson MB, Edland SD, Weiner MF, Fillenbaum G, et al. (1995) *The Behavior Rating Scale for Dementia of the Consortium to Establish a Registry for Alzheimer's Disease*. The Behavioral Pathology Committee of the Consortium to Establish a Registry for Alzheimer's Disease. American Journal of Psychiatry. Vol. 152, No.9, pp.1349-1357.
- Tariot PN, Schneider LS, Mintzer JE, Cutler AJ, et al. (2001) *Safety and tolerability of divalproex sodium in the treatment of signs and symptoms of mania in elderly patients with dementia: Results of a double-blind, placebo-controlled trial*. Current Therapeutic Research, Clinical and Experimental. Vol. 62, No.1, pp.51-67.
- Tedesco M, Ratti GS, Mennella S, et al. (1999) *Comparison of losartan and hydrochlorothiazide on cognitive function and quality of life in hypertensive patients*. American Journal of Hypertension. Vol. 12, No.11, pp.1130-1134.
- Teng EL and Chui HC (1987) *The Modified Mini-Mental State (3MS) Examination*. Journal of Clinical Psychiatry. Vol. 48, pp.314-318.
- Teng EL, Hasegawa K, Homma A, et al. (1994) *The Cognitive Abilities Screening Instrument (CASI): A practical test for cross-cultural epidemiological studies of dementia*. International Psychogeriatrics. Vol. 6, pp.45-58.
- Teng EL and Manly JJ (2005) *Neuropsychological testing: helpful or harmful?* Alzheimer Disease and Associated Disorders. Vol. 19, No.4, pp.267-271.
- Teno JM, Landrum K and Lynn J (1997) *Defining and measuring outcomes in end-stage dementia*. Alzheimer Disease and Associated Disorders. Vol. 11, Suppl 6, pp.25-9.
- Teri L, Logsdon RG, Peskind E, Raskind M, et al. (2000) *Treatment of agitation in AD: a randomized, placebo-controlled clinical trial.[see comment][erratum appears in Neurology 2001 Feb 13;56(3):426]*. Neurology. Vol. 55, No.9, pp.1271-8.
- Teri L, McCurry SM, Logsdon R, et al. (2005) *Training community consultants to help family members improve dementia care: a randomized controlled trial*. Gerontologist. Vol. 45, No.6, pp.802-11.
- Teri L, Rabins P, Whitehouse P, Berg L, Reisberg B, Sunderland T, Eicheleman B and Phelps C (1992) *Management of behaviours disturbance in Alzheimer's disease: Current knowledge and future directions*. Alzheimer's Disease and Associated Disorders. Vol. 6, No.2, pp.77-88.
- Teri L, Truax P, Logsdon R, Uomoto J, Zarit S and Vitaliano PP (1992) *Assessment of behavioral problems in dementia: the revised memory and behavior problems checklist*. Psychology and Aging. Vol. 7, No.4, pp.622-631.
- Teunisse S, Derix MMA and van Crevel H (1991) *Assessing the severity of dementia: patient and caregiver*. Archive of Neurology. Vol. 48, pp.274-277.
- Thoits P (1982) *Conceptual, methodological and theoretical problems in studying social support as a buffer against life stress*. Journal of Health and Social Behavior. Vol. 23, pp.145-159.

- Thomas P, Lalloue F, Preux PM, et al. (2006) *Dementia patients' caregiver' quality of life: the PIXEL study*. International Journal of Geriatric Psychiatry. Vol. 21, No.1, pp.50-6.
- Thomas S, Nay R, Fonda D, et al. (2006) *Continence Outcomes Measurement Suite Project: Final Report*, Australian Government Department of Health and Ageing, Canberra.
- Thomas VS, Rockwood K and McDowell I (1998) *Multidimensionality in Instrumental and Basic Activities of Daily Living*. J Clin Epidemiol. Vol. 51, No.4, pp.315-321.
- Thompson AG and Sunol R (1995) *Expectations as determinants of patient satisfaction: concepts, theory and evidence*. International Journal for Quality in Health Care. Vol. 7, No.2, pp.127-41.
- Thorgrimsen L, Selwood A, Spector A, et al. (2003) *Whose quality of life is it anyway? The validity and reliability of the Quality of Life-Alzheimer's Disease (QoL-AD) scale.[see comment]*. Alzheimer Disease and Associated Disorders. Vol. 17, No.4, pp.201-8.
- Thurstone L (1947) *Multiple Factor Analysis*. University of Chicago Press, Chicago.
- Tilling K, Coshall C, McKeivitt C, et al (2005) *A family support organiser for stroke patients and their carers: a randomised controlled trial*. Cerebrovascular Diseases. Vol. 20, No.2, pp.85-91.
- Tierney MC, Herrmann N, Geslani DM, et al. (2003) *Contribution of Informant and Patient Ratings to the Accuracy of the Mini-Mental State Examination in Predicting Probable Alzheimer's Disease*. Journal of the American Geriatrics Society. Vol. 51, No.6, pp.813-818.
- Tomaka J, Thompson S and Palacios R (2006) *The relation of social isolation, loneliness, and social support to disease outcomes among the elderly*. Journal of Aging and Health. Vol. 18, No.3, pp.359-84.
- Tombaugh TN, McDowell I, Kristjansson B and Hubley AM (1996) *Mini-Mental State Examination (MMSE) and the Modified MMSE (3MS): A psychometric comparison and normative data*. Psychological Assessment. Vol. 8, No.1, pp.48-59.
- Torrance G (1986) *Measurement of health state utilities for economic appraisal: a review*. Journal of Health Economics. Vol. 5, pp.1-30.
- Torrance G, Boyle M and Horwood S (1982) *Application of multi-attribute theory to measure social preferences for health states*. Operations Research. Vol. 30, pp.1043-1069.
- Torrance GW, Feeny D and Furlong W (2001) *Visual analog scales: do they have a role in the measurement of preferences for health states?* Medical Decision Making. Vol. 21, No.4, pp.329-34.
- Torrance GW, Furlong W, Feeny D and Boyle M (1995) *Multi-attribute preference functions*. Health Utilities Index. Pharmacoeconomics. Vol. 7, No.6, pp.503-20.
- Torti FM Jr., Gwyther LP, Reed SD, et al. (2004) *A multinational review of recent trends and reports in dementia caregiver burden*. Alzheimer Disease and Associated Disorders. Vol. 18, No.2, pp.99-109.
- Townsend P (1963) *The Family Life of Old People*. Penguin Books, London.
- Townsend P (1973) *Isolation and loneliness in the aged*. In Weiss RS (ed.) Loneliness: The Experience of Emotional and Social Isolation. Cambridge, MA: Massachusetts Institute of Technology. pp.175-188.
- The American Heritage Stedman's Medical Dictionary (2004) *The American Heritage Stedman's Medical Dictionary (2nd ed.)*. Houghton Mifflin Company, USA.
- Tom Mesuer PhD, Associate Prof, Director of Education and Rural Outreach Alzheimer Disorder Research Centre, Washington University (2007) [Personal communication]. [meusert@neuro.wustl.edu](mailto:meusert@neuro.wustl.edu). The Washington University is the most important resource for the GDS.
- The Lund and Manchester Groups (1994) *Clinical and neuropathological criteria for frontotemporal dementia*. Journal of Neurology, Neurosurgery and Psychiatry. Vol. 57, pp.416-18.

Tractenberg RE, Gamst A, Thomas RG, Patterson M, et al. (2002) *Investigating emergent symptomatology as an outcome measure in a behavioral study of Alzheimer's disease.* Journal of Neuropsychiatry and Clinical Neurosciences. Vol. 14, No.3, pp.303-10.

Tractenberg RE, Patterson M, Weiner MF, Teri L, et al. (2000) *Prevalence of symptoms on the CERAD behavior rating scale for dementia in normal elderly subjects and Alzheimer's disease patients.[erratum appears in J Neuropsychiatry Clin Neurosci 2001 Winter;13(1):95].* Journal of Neuropsychiatry and Clinical Neurosciences. Vol. 12, No.4, pp.472-9.

Tractenberg RE, Schafer K and Morris JC (2001) *Interobserver disagreements on clinical dementia rating assessment: interpretation and implications for training.* Alzheimer Disease and Associated Disorders. Vol. 15, No.3, pp.155-61.

Tractenberg RE, Weiner MF, Patterson MB, Gamst A, et al. (2002) *Emergent psychopathology in Alzheimer's disease patients over 12 months associated with functional, not cognitive, changes.* Journal of Geriatric Psychiatry and Neurology. Vol. 15, No.2, pp.110-7.

Tractenberg RE, Weiner MF, Patterson MB, Teri L, et al. (2003) *Comorbidity of psychopathological domains in community-dwelling persons with Alzheimer's disease.* Journal of Geriatric Psychiatry and Neurology. Vol. 16, No.2, pp.94-9.

Trzepacz PT (1999) *The Delirium Rating Scale. Its use in consultation-liaison research.* Psychosomatics. Vol. 40, No.3, pp.193-204.

Trzepacz PT, Baker RW and Greenhouse J (1988) *A symptom rating scale for delirium.* Psychiatry Research. Vol. 23, No.1, pp.89-97.

Trzepacz PT, Mittal D, Torres R, Canary K, Norton J and Jimerson N (2001) *Validation of the Delirium Rating Scale-revised-98: comparison with the delirium rating scale and the cognitive test for delirium.[erratum appears in J Neuropsychiatry Clin Neurosci 2001 Summer;13(3):433].* Journal of Neuropsychiatry and Clinical Neurosciences. Vol. 13, No.2, pp.229-242.

Trzepacz PT, Mulsant BH, Dew MA, Pasternak R, Sweet RA and Zubenko GS (1998) *Is delirium different when it occurs in dementia? A study using the delirium rating scale.* J Neuropsychiatry Clin Neurosci. Vol. 10, No.2, pp.199-204.

Tschanz JT, Welsh-Bohmer KA, Plassman BL, Norton MC, Wyse BW, Breitner JCS, et al. (2002) *An adaptation of the modified mini-mental state examination: analysis of demographic influences and normative data: the cache county study.* Neuropsychiatry, Neuropsychology, and Behavioral Neurology. Vol. 15, No.1, pp.28-38.

Tsolaki M, Fountoulakis K, Nakopoulou E, et al. (1997) *Alzheimer's Disease Assessment Scale: the validation of the scale in Greece in elderly demented patients and normal subjects.* Dementia and Geriatric Cognitive Disorders. Vol. 8, No.5, pp.273-80.

Turner RJ (1981) *Social support as a contingency in psychological wellbeing.* Journal of Health and Social Behavior. Vol. 22, pp.357-367.

Turner S (2004) *Are the Health of the Nation Outcome Scales (HoNOS) useful for measuring outcomes in older people's mental health services?* Ageing and Mental Health. Vol. 8, No. 5, pp. 387-396.

USPSTF (U.S. Preventive Service Task Force) (2003) *Recommendations and Rationale: Screening for Dementia*, available at <http://www.ahcpr.gov/clinic/3rduspstf/dementia/dementrr.htm>, accessed 21 June 2006.

Vahey DC, Aiken LH, Sloane DM, et al. (2004) *Nurse burnout and patient satisfaction.* Medical Care. Vol. 42, No.2 Suppl, pp.II57-66.

- Valvanne J, Luoma M-L, Ylonen L, et al. 2005. *Quality of life among clients in home care - Are there differences between demented and non-demented clients*. 18th Congress of the International Association of Gerontology. Rio de Janeiro, Brazil.
- van Agt HM, Essink-Bot ML, Krabbe PF and Bonsel GJ (1994) *Test-retest reliability of health state valuations collected with the EuroQol questionnaire*. Social Science and Medicine. Vol. 39, No.11, pp.1537-1544.
- van Baalen B, Odding E, et al. (2006) *Reliability and sensitivity to change of measurement instruments used in a traumatic brain injury population*. Clinical Rehabilitation. Vol. 20, No.8, pp.686-700.
- van Baarsen B, Smit JH, Snijders TAB, et al. (1999) *Do personal conditions and circumstances surrounding partner loss explain loneliness in newly bereaved older adults*. Ageing and Society. Vol. 19, No.4, pp.441-469.
- van Baarsen B, Snijders TAB, Smit JH, et al. (2001) *Lonely but not alone: emotional isolation and social isolation as two distinct dimensions of loneliness in older people*. Educational and Psychological Measurement. Vol. 61, No.1, pp.119-135.
- van de Water MS, Kutner M, Parmelee PA, et al. (2003) *Which long-term care residents should be asked to complete a customer satisfaction survey?* Journal of the American Medical Directors Association. Vol. 4, No.5, pp.257-63.
- van Diepen E, Baillon SF, Redman J, Rooke N, et al. (2006) *A pilot study of the physiological and behavioural effects of Snoezelen in dementia*. British Journal of Occupational Therapy. Vol. 65, No.2, pp.61-66.
- van Reekum R, Stuss DT and Ostrander L (2005) *Apathy: Why care?* J Neuropsychiatry Clin Neurosci. Vol. 17, No.1, pp.7-19.
- van Tilburg T, Havens B and de Jong Gierveld J (2004) *Loneliness among older adults in the Netherlands, Italy, and Canada: a multifaceted comparison*. Canadian Journal on Aging. Vol. 23, No.2, pp.169-180.
- Vanier M, Mazaux JM, et al. (2000) *Assessment of neuropsychologic impairments after head injury: interrater reliability and factorial and criterion validity of the Neurobehavioral Rating Scale-Revised*. Archives of Physical Medicine and Rehabilitation. Vol. 81, No.6, pp.796-806.
- Venn, RD (1983) *The Sandoz Clinical Assessment Geriatric (SCAG) Scale. A general purpose psychogeriatric rating scale*. Gerontology. Vol. 29, pp.185-198.
- Verkaik R, van Weert JC and Francke AL (2005) *The effects of psychosocial methods on depressed, aggressive and apathetic behaviors of people with dementia: a systematic review*. International Journal of Geriatric Psychiatry. Vol.20, No.4, pp.301-314.
- Vespa A, Gori G and Spazzafumo L (2002) *Evaluation of non-pharmacological intervention on antisocial behavior in patients suffering from Alzheimer's disease in a day care center*. Archives of gerontology and geriatrics. Vol. 34, No.1, pp.1-8.
- Vespa A, Gori S, Bonaiuto S, Cruciani G, et al. (2002) *Validation of the Cohen-Mansfield agitation inventory (CMAI-long form) on a sample of Italian demented patients*. Archives of gerontology and geriatrics. Vol. 8, pp.377-83.
- Vetter PH, Krauss S, et al. (1999) *Vascular dementia versus dementia of Alzheimer's type: do they have differential effects on caregivers' burden?* Journals of Gerontology Series B-Psychological Sciences and Social Sciences. Vol. 54, No.2, pp.S93-8.
- Victor C, Scambler S, Bond J, et al. (2000) *Being alone in later life: loneliness, social isolation and living alone*. Reviews in Clinical Gerontology. Vol. 10, No.4, pp.407-417.
- Victorian Transcultural Psychiatry Unit (2006) *Guidelines for working effectively with interpreters in mental health settings*. Victorian Transcultural Psychiatry Unit.

- Villardita LC and Lomeo (1992) *Alzheimer's disease: Correlational analysis of three screening tests and three behaviour scales*. Acta Neurologica Scandinavica. Vol. 80, No.6, pp.603-8.
- Vilkki J, Ahola K, et al. (1994) *Prediction of psychosocial recovery after head injury with cognitive tests and neurobehavioral ratings*. Journal of Clinical and Experimental Neuropsychology: Official Journal of the International Neuropsychological Society. Vol. 16, No.3, pp.325-38.
- Villanueva MR, Smith TL, Erickson JS, Lee AC, et al. (2003) *Pain Assessment for the Dementing Elderly (PADE): reliability and validity of a new measure.[see comment]*. Journal of the American Medical Directors Association. Vol. 4, No.1, pp.1-8.
- Visser M, Fletcher A, Parr G, Simpson A and Bulpitt C (1994) *A comparison of three quality of life instruments in subjects with angina pectoris: the Sickness Impact Profile, the Nottingham Health Profile, and the Quality of Well Being Scale*. Journal of Clinical Epidemiology. Vol. 47, No.2, pp.157-163.
- Vitaliano PP, Young HM and Russo J (1991) *Burden: a review of measures used among caregivers of individuals with dementia*. Gerontologist. Vol. 31, No.1, pp.67-75.
- Volicer L, Camberg L, Hurley AC, Ashley J, et al. (1999) *Dimensions of decreased psychological well-being in advanced dementia*. Alzheimer Disease and Associated Disorders. Vol. 13, No.4, pp.192-201.
- Volicer L, Hurley AC and Blasi ZV (2001) *Scales for evaluation of end-of-life care in dementia*. Alzheimer Disease and Associated Disorders. Vol. 15, No.4, pp.194-200.
- von Winterfeldt D and Edwards W (1986) *Decision Analysis and Behavioural Research*, Cambridge University Press, Cambridge.
- Voruganti L, Heslegrave R, Awad A and Seeman M (1998) *Quality of life measurement in schizophrenia: reconciling the quest for subjectivity with the question of reliability*. Psychological Medicine. Vol. 28, pp.165-172.
- Voyer P, McCusker J, Cole M and Khomenko L (2006) *Influence of prior cognitive impairment on the severity of delirium symptoms among older patients*. Journal of Neuroscience Nursing. Vol. 38, No.2, pp.90-101.
- Wagle A, Komorita NI and Lu ZJ (1997) *Social support and breast self-examination*. Cancer Nursing. Vol. 20, No.1, pp.42-8.
- Waite L, Grayson D, Jorm AF, et al. (1999) *Informant-based staging of dementia using the clinical dementia rating*. Alzheimer Disease and Associated Disorders. Vol. 13, No.1, pp.34-7.
- Wall JR, Deshpande SA, MacNeill SE, et al. (1998) *The Fuld Object Memory Evaluation, a useful tool in the assessment of urban geriatric patients*. Clinical Gerontologist. Vol. 19, No.1, pp.39-49.
- Walker SA, Naji SA, Howie FL, et al. (2001) *The effectiveness of psychiatric discharge coordinators*. Primary Care Psychiatry. Vol. 6, No.4, pp.149-153.
- Wang H, Yu X, Li S, et al. (2004) *The cognitive subscale of Alzheimer's Disease Assessment Scale, Chinese version in staging of Alzheimer disease*. Alzheimer Disease and Associated Disorders. Vol. 18, No.4, pp.231-5.
- Wang L, Swank JS, Glick IE, et al. (2003) *Changes in hippocampal volume and shape across time distinguish dementia of the Alzheimer type from healthy aging*. Neuroimage. Vol. 20, No.2, pp.667-82.
- Ware J (2003) *Standardizing health metrics: the SF-36 Health Survey and beyond*. In Sansoni J and Tilley L (eds.) Conference Proceedings, Health Outcomes 2003: The Quest for Practice Improvement. Canberra, 20-21 August 2003.
- Ware JE (1987) *Standards for validating health measures: definition and content*. Journal of Chronic Diseases. Vol. 40, No.6, pp.473-480.

- Ware JE, Jr. and Hays RD (1988) *Methods for measuring patient satisfaction with specific medical encounters*. Medical Care. Vol. 26, No.4, pp.393-402.
- Ware JE, Kosinski M and Dewey JE (2001) *How to Score Version Two of the SF-36 Health Survey* (2<sup>nd</sup> ed.) QualityMetric Incorporated, Lincoln RI.
- Ware JE, Kosinski M, Turner-Bowker DM and Gandek B (2002) *How to Score Version 2 of the SF-12 Health Survey (With a Supplement Documenting Version 1)*. QualityMetric Incorporated, Lincoln RI
- Ware JE, Snyder MK, Wright WR, et al. (1983) *Defining and measuring patient satisfaction with medical care*. Evaluation and Program Planning. Vol. 6, pp.247-263.
- Warner SC and Williams JI (1987) *The Meaning in Life Scale: determining the reliability and validity of a measure*. Journal of Chronic Diseases. Vol. 40, No.6, pp.503-12.
- Watson LC, Lewis CL and Fillenbaum GG (2004) *Asking family about memory loss: Is it helpful ?* J General Internal Medicine. Vol. 20, pp.28-32.
- Watson LC and Pignone MP (2003) *Screening accuracy for late-life depression in primary care: a systematic review*. Journal of Family Practice. Vol. 52, No.12, pp.956-964.
- Watt S, Sword W, Krueger P, et al. (2002) *A cross-sectional study of early identification of postpartum depression: implications for primary care providers from The Ontario Mother and Infant Survey*. BMC Family Practice. Vol. 3, pp.5.
- Weiner DK, Herr K and Rudy TE (eds.) (2002) *Persistent pain in older adults*. Springer Publishing Company, New York.
- Weiner MF, Koss E, Patterson M, Jin S, et al. (1998) *A comparison of the Cohen-Mansfield agitation inventory with the CERAD behavioral rating scale for dementia in community-dwelling persons with Alzheimer's disease*. Journal of Psychiatric Research. Vol. 32, No.6, pp.347-51.
- Weiner MF, Martin-Cook K, Foster BM, Saine K, et al. (2000) *Effects of donepezil on emotional/behavioral symptoms in Alzheimer's disease patients*. Journal of Clinical Psychiatry. Vol. 61, No.7, pp.487-92.
- Weiner MF, Martin-Cook K, Svetlik DA, et al. (2000) *The quality of life in late-stage dementia (QUALID) scale*. Journal of the American Medical Directors Association. Vol. 1, No.3, pp.114-116.
- Weiner MF, Tractenberg R, Teri L, Logsdon R, et al. (2000) *Quantifying behavioral disturbance in Alzheimer's disease patients*. Journal of Psychiatric Research. Vol. 34, No.2, pp.163-7.
- Weiner MF, Tractenberg RE, Jin S, Gamst A, et al. (2002) *Assessing Alzheimer's disease patients with the Cohen-Mansfield Agitation Inventory: scoring and clinical implications*. Journal of Psychiatric Research. Vol. 36, No.1, pp.19-25.
- Weiner MF, Williams B and Risser RC (1997) *Assessment of behavioral symptoms in community-dwelling dementia patients*. American Journal of Geriatric Psychiatry. Vol. 5, No.1, pp.26-30.
- Weiner MF, Vobach S, Svetlik D, et al. (1993) *Cortisol secretion and Alzheimer's disease progression: a preliminary report*. Biological Psychiatry. Vol. 34, No.3, pp.158-61.
- Weiss R (1974) *The provisions of social relationships*. In Rubin Z, (ed.) Doing Unto Others: Joining, Molding, Conforming, Helping, Loving. Prentice Hall, Englewood Cliffs. pp.17-26.
- Weiss RS (1973) *Loneliness: The Experience of Emotional and Social Isolation*. MIT Press, Cambridge, MA.
- Welsh KA, Butters N, Hughes CP, et al. (1991) *Detection of abnormal memory decline in mild cases of Alzheimer's disease using CERAD neuropsychological measures*. Archives of Neurology. Vol. 48, pp.278-282.

- Wells DL, Dawson P, Sidani S, Craig D, et al. (2000) *Effects of an abilities-focused program of morning care on residents who have dementia and on caregivers*. Journal of the American Geriatrics Society. Vol. 48, No.4, pp.442-9.
- Wellwood I, Dennis M and Warlow C (1995) *Patients' and carers' satisfaction with acute stroke management*. Age and Ageing. Vol. 24, No.6, pp.519-24.
- Wenger C (1983) *Loneliness: a problem of measurement*. In Jerrome D, (ed.) Aging in Modern Society: Contemporary Approaches. Croom Helm, London. pp.145-167.
- Wenger GC and Burholt V (2004) *Changes in levels of social isolation and loneliness among older people in a rural area: a twenty-year longitudinal study*. Canadian Journal on Aging. Vol. 23, No.2, pp.115-27.
- Werner P, Cohen-Mansfield J, Braun J and Marx MS (1989) *Physical restraint and agitation in nursing home residents*. Journal of the American Geriatric Society. Vol. 37, pp.1122-1126.
- Werner P, Cohen-Mansfield J, Koroknay V and Braun J (1994) *The impact of a restraint-reduction program on nursing home residents*. Geriatric Nursing. Vol. 15, No.3, pp.142-46.
- Westaway MS, Rheeder P, Van Zyl DG, et al. (2003) *Interpersonal and organizational dimensions of patient satisfaction: the moderating effects of health status*. International Journal for Quality in Health Care. Vol. 15, No.4, pp.337-44.
- Westaway MS, Seager JR, Rheeder P, et al. (2005) *The effects of social support on health, well-being and management of diabetes mellitus: a black South African perspective*. Ethnicity and Health. Vol. 10, No.1, pp.73-89.
- Westhuis D and Thyer BA (1989) *Development and validation of the clinical anxiety scale: a rapid assessment instrument for empirical practice*. Educational and Psychological Measurement. Vol. 49, pp.153-163.
- Wetterling T, Kanitz RD and Borgis KJ (1996) *Comparison of different diagnostic criteria for vascular dementia (ADDC, DSM-IV, ICD-10, NINDS-AIREN)*. Stroke. Vol. 27, No.1, pp.30-36.
- Weyer G, Erzigkeit H, Kanowski S, et al. (1997) *Alzheimer's Disease Assessment Scale: reliability and validity in a multicenter clinical trial*. International Psychogeriatrics. Vol. 9, No.2, pp.123-38.
- Weyerer S, Killmann U, Ames D, et al. (1999) *The Even Briefer Assessment Scale for Depression (EBAS DEP): its suitability for the elderly in geriatric care in English- and German-speaking countries*. International Journal of Geriatric Psychiatry. Vol. 14, No.6, pp.473-80.
- Whall AL, Black MEA, Yankou DJ, Groh CJ, et al (1999) *Nurse aides' identification of onset and level of agitation in late stage dementia patients*. American Journal of Alzheimer's Disease. Vol. 14, No.4, pp.202-206.
- White HK, McConnell ES, Bales CW and Kuchibhatia M (2004) *The 6<sup>th</sup> month study of the relationship between weight loss and behavioural symptoms in institutionalised Alzheimer's disease subjects*. Journal of the American Medical Directors Association. Vol. 5, No.2, pp. 89-97.
- White MB (1972) *Importance of selected nursing activities*. Nursing Research. Vol. 21, No.1, pp.4-14.
- Whitehouse PJ, Patterson MB, Strauss ME, Geldmacher DS, et al. (1996) *Hallucinations*. International Psychogeriatrics. Vol. 8, No.Suppl 3, pp.387-392.
- WHOQoL Group (1993) *Study protocol for the World Health Organization project to develop a quality of life assessment instrument (WHOQOL)*. Quality of Life Research. Vol. 2, pp.153-159
- WHOQOL Group (1998a) *The World Health Organization quality of life assessment (WHOQOL): Development and general psychometric properties*. Social Science and Medicine. Vol. 12, pp.1569-1585.

---

WHOQOL Group (1998b) *Group development of the World Health Organization WHOQOL-BREF quality of life assessment*. Psychological Medicine. Vol. 28, pp.551-558.

WHOQoL Group (1998) *Development of the World Health Organization WHOQOL-BREF Quality of Life Assessment*. Psychological Medicine. Vol. 28, No.3, pp.551-558.

Wilkinson EK, Salisbury C, Bosanquet N, et al. (1999) *Patient and carer preference for, and satisfaction with, specialist models of palliative care: a systematic literature review*. Palliative Medicine. Vol. 13, No.3, pp.197-216.

Wilkosz PA, Miyahara S, Lopez OL, Dekosky ST, et al. (2006) *Prediction of psychosis onset in Alzheimer disease: The role of cognitive impairment, depressive symptoms, and further evidence for psychosis subtypes*. American Journal of Geriatric Psychiatry. Vol. 14, No.4, pp.352-60.

Williams A (1992) *Review article: measuring functioning and well-being*. In Stewart A, Ware J, (eds). The Medical Outcomes Study Approach. Durham: Duke University Press.

Williams AK (2005) *Motivation and dementia*. Topics in Geriatric Rehabilitation. Vol.21, No.2, pp.123-126.

Williams B (1994) *Patient satisfaction: a valid concept?* Social Science and Medicine. Vol. 38, No.4, pp.509-516.

Williams C, Zimmerman S, Sloane P, Reed P (2005) *Characteristics associated with pain in long-term residents with dementia*. The Gerontologist. Vol. 45, pp.68-73.

Williams M, Gaylord S and Gerrity M (1994) *The Timed Manual Performance test as a predictor of hospitalization and death in a community-based elderly population*. JAGS. Vol. 42, No.1, pp.21-27.

Williams PD, Williams AR and Griggs C (1990) *Children at home on mechanical assistive devices and their families: a retrospective study*. Maternal-Child Nursing Journal. Vol. 19, No.4, pp.297-311.

Williamson J, Stokoe IH, Gray S, et al. (1964) *Old people at home. Their unreported needs*. Lancet. Vol. 13, pp.1117-20.

Wilson A, Wynn A and Parker H (2002) *Patient and carer satisfaction with 'hospital at home': quantitative and qualitative results from a randomised controlled trial*. British Journal of General Practice. Vol. 52, No.474, pp.9-13.

Wilson D, Cutts J, Lees I, et al. (1992) *Psychometric properties of the revised UCLA Loneliness Scale and two short-form measures of loneliness in Zimbabwe*. Journal of Personality Assessment. Vol. 59, No.1, pp.72-81.

Wilson W and Goldschmidt P (1995) *Quality Management in Healthcare*. McGraw-Hill Book Company Australia, Sydney.

Wilson RS, Krueger KR, Arnold SE, et al. (2007) *Loneliness and risk of Alzheimer disease*. Archives of General Psychiatry. Vol. 64, No.2, pp.234-40.

Wimo A, Mattson B, Krakau I, Eriksson T, Nelvig A and Karlsson G (1995) *Cost-utility analysis of group living in dementia care*. International Journal of Technology Assessment in Health Care. Vol. 11, No.1, pp.49-65.

Winblad B, Palmer K, Kivipelto M, Jelic V, Fratiglioni L, Wahlund LO, et al. (2004) *Mild cognitive impairment – beyond controversies, towards a consensus: report of the International Working Group on Mild Cognitive Impairment*. Journal of Internal Medicine. Vol. 256, pp.240-246.

Wingfield A (1999) *Comprehending spoken questions: Effects of cognitive and sensory change in adult aging*. In Schwarz N, Park DC, Knauper B, Sudman S, (eds). Cognition, ageing, and self reports. Psychology Press, Taylor and Francis Group, Philadelphia.

Winzelberg GS, Williams CS, Preisser JS, et al. (2005) *Factors associated with nursing assistant quality-of-life ratings for residents with dementia in long-term care facilities.* Gerontologist. Vol. 45, Spec No.1, No.1, pp.106-14.

Wisocki PA (1988) *Worry as a phenomenon relevant to the elderly.* Behavior Therapy. Vol. 19, No.3, pp.369-379.

Wistedt B, Rasmussen A, Pedersen L, Malm U, Tracksman-Bendtz L, Wakelin J 7 Bech P (1990) *The development of an observer-scale for measuring social dysfunction and aggression.* Pharmacopsychiatry. Vol. 23, No.6, pp.249-52.

Wlodarczyk JH, Brodaty H and Hawthorne G (2004) *The relationship between quality of life, Mini-Mental State Examination, and the Instrumental Activities of Daily Living in patients with Alzheimer's Disease.* Archives of Gerontology and Geriatrics. Vol. 39, No.1, pp.25-33.

Wolf MH, Putnam SM, James SA, et al. (1978) *The Medical Interview Satisfaction Scale: development of a scale to measure patient perceptions of physician behavior.* Journal of Behavioral Medicine. Vol. 1, No.4, pp.391-401.

Wolf MH and Stiles WB (1981) *Further development of the Medical Interview Satisfaction Scale.* American Psychological Association Convention, Los Angeles.

Wolfe N (2002) *Cross-cultural neuropsychology of aging and dementia: An update.* In Ferraro, Richard (ed.) Minority and Cross-Cultural Aspects of Neuropsychological Assessment, Taylor and Francis.

Wolstenholme J, Fenn P, Gray A, et al. (2002) *Estimating the relationship between disease progression and cost of care in dementia.* British Journal of Psychiatry. Vol. 181, pp.36-42.

Wood S, Cummings JL, Hsu MA, et al. (2000) *The use of the Neuropsychiatric Inventory in nursing home residents: characteristics and measurement.* American Journal of Geriatric Psychiatry. Vol. 8, pp.75-83.

Wood-Dauphinee S (1999) *Assessing quality of life in clinical research: from where have we come and where are we going?* Journal of Clinical Epidemiology. Vol. 52, No.4, pp.355-63.

Wood-Dauphinee S and Williams JI (1987) *Reintegration to Normal Living as a proxy to quality of life.* Journal of Chronic Diseases. Vol. 40, No.6, pp.491-502.

Wood J, Taylor S and Lichtman R (1985) *Social comparison in adjustment to breast cancer.* Journal of Personality and Social Psychology. Vol. 49, pp.1169-1183.

Woods B, Thorgrimsen L, Spector A, et al. (2006) *Improved quality of life and cognitive stimulation therapy in dementia.* Aging and Mental Health. Vol. 10, No.3, pp.219-226.

World Health Organisation (1992) *The ICD-10 Classification of Mental and Behavioural Disorders: Clinical descriptions and diagnostic guidelines.* WHO, Geneva.

World Health Organisation (1992) *The ICD-10 Classification of Mental and Behavioural Disorders: Clinical descriptions and diagnostic guidelines.* WHO, Geneva.

World Health Organisation (1993) *The ICD-10 Classification of Mental and Behavioural Disorders: Diagnostic criteria for research.* WHO, Geneva.

Wu AW, Jacobson DL, Berzon RA, Revicki DA, van der Horst C, Fichtenbaum CJ, Saag MS, Lynn L, Hardy D and Feinberg J (1997) *The effect of mode of administration on medical outcomes study health ratings and EuroQol scores in AIDS.* Quality of Life Research. Vol. 6, No.1, pp.3-10.

Wyrwich KW, Metz SM, Babu AN, et al. (2002) *The reliability of retrospective change assessments.* Quality of Life Research. Vol. 11, No.7, pp.636.

Yasuda N, Zimmerman S, Hawkes WG, et al. (2004) *Concordance of proxy-perceived change and measured change in multiple domains of function in older persons.* JAGS. Vol. 52, pp.1157-1162.

- Yeo G and Gallagher-Thompson D (2006) *Ethnicity and the Dementias*. Routledge, Taylor and Francis Group, New York.
- Yesavage JA, Brink TL, Rose TL, et al. (1983) *Development and validation of a geriatric depression screening scale: a preliminary report*. Journal of Psychiatric Research. Vol. 17, No.1, pp.37-49.
- Yudovsky SC, Silver JM, Jackson W, Endicott J and Williams D (1986) *The Overt Aggression Scale for objective rating of verbal and physical aggression*. American Journal of Psychiatry. Vol. 143, pp.35-39.
- Youn JC, Lee DY, Kim KW, et al. (2002) *Development of the Korean version of Alzheimer's Disease Assessment Scale (ADAS-K)*. International Journal of Geriatric Psychiatry. Vol. 17, No.9, pp.797-803
- Yu DS, Lee DT and Woo J (2004) *Psychometric testing of the Chinese version of the medical outcomes study social support survey (MOS-SSS-C)*. Research in Nursing and Health. Vol. 27, No.2, pp.135-43.
- Zakahi WR and Duran RL (1982) *All the lonely people: the relationship among loneliness, communicative competence and communication anxiety*. Communication Quarterly. Vol. 30, No.3, pp.203-209.
- Zarit SH, Reever KE and Bach-Peterson J (1980) *Relatives of the impaired elderly: correlates of feelings of burden*. Gerontologist. Vol. 20, No.6, pp.649-55.
- Zarit SH, Todd PA and Zarit JM (1986) *Subjective burden of husbands and wives as caregivers: a longitudinal study*. Gerontologist. Vol. 26, No.3, pp.260-6.
- Zebiene E, Razgauskas E, Basys V, et al. (2004) *Meeting patient's expectations in primary care consultations in Lithuania*. International Journal of Quality in Health Care. Vol. 16, No.1, pp.83-9.
- Zec RF, Landreth ES, Vicari SK, et al. (1992) *Alzheimer disease assessment scale: useful for both early detection and staging of dementia of the Alzheimer type*. Alzheimer Disease and Associated Disorders. Vol. 6, No.2, pp.89-102.
- Zemlan FP (1996) *Velnacrine for the treatment of Alzheimer's disease: a double-blind, placebo-controlled trial*. The Mentane Study Group. Journal of Neural Transmission. Vol. 103, No.8-9, pp.1105-16.
- Zemlan FP (1996) *Velnacrine for the treatment of Alzheimer's disease: a double-blind, placebo-controlled trial*. The Mentane Study Group. Journal of Neural Transmission. Vol. 103, No.8-9, pp.1105-16.
- Zieber CG, Hagen B, Armstrong-Esther C and Aho M (2005) *Pain and agitation in long-term care residents with dementia: use of the Pittsburgh Agitation Scale*. International Journal of Palliative Nursing. Vol. 11, No.2, pp.71-8.
- Zigmond AS and Snaith RP (1983) *The Hospital Anxiety and Depression Scale*. Acta Psychiatrica Scandinavica. Vol. 67, No.6, pp.361-370.
- Zillmer EA, Fowler P, Gutnick HN, Becker E (1990) *Comparison of two cognitive bedside screening instruments in nursing home residents: a factor analysis study*. Journal of Gerontology. Vol. 45, No.2, pp.69-74.
- Zuidema SU, de Jonghe JF, Verhey FR and Koopmans RT (2007) *Agitation in Dutch institutionalized patients with dementia: factor analysis of the Dutch version of the Cohen-Mansfield Agitation Inventory*. Dementia and Geriatric Cognitive Disorders. Vol. 23, No.1, pp.35-41.

**Additional internet sites visited to search for dementia staging and descriptive measures:**

[http://www.alz.org/professionals\\_and\\_researchers\\_conducting\\_an\\_assessment.asp](http://www.alz.org/professionals_and_researchers_conducting_an_assessment.asp)

<http://www.neurotransmitter.net/Alzheimerscales.htm>

<http://alz.org/Services/LibraryServices.asp>

[http://www.alz.org/professionals\\_and\\_researchers\\_general\\_resources.asp](http://www.alz.org/professionals_and_researchers_general_resources.asp)

<http://www.alzforum.org/dis/dia/tes/neuropsychological.asp>  
[Greenfield@alz.org](mailto:Greenfield@alz.org)

[www.adrc.wustl.edu/adrc](http://www.adrc.wustl.edu/adrc)

<http://alzheimer.wustl.edu/cdr/default.htm>

<http://www.hnrc.ucsd.edu/publications/index.php>

[www.stroke.com](http://www.stroke.com)

<http://www.Alzheimer-insights.com/insights/vol2no3/vol2no3.htm>

<http://www.medicine.uiowa.edu/igec/tools/default.asp>

<http://library.csus.edu/guides/rogenmoserd/heas/nursing2004.pdf>

<http://libraries.uta.edu/helen/Test&Meas/testmainframe.htm>

<http://www.hartfordign.org/resources/education/tryThis.html>

---

## Glossary of Terms

ANOVA	Analysis of variance.
ARSB	Acquiescent response set bias. This refers to the situation where a respondent provides biased answers to questions because he/she wishes to please the researchers.
Ceiling effect	Refers to scores on an instrument being 'bunched' up at the top end of the scoring range.
Cohen's effect size (d)	This quantifies the size of differences between groups or over time. Cohen classified effect sizes into small effects (<0.20), moderate (~0.50) and large effect sizes (>0.80).
Correlation	Describes the linear relationship between two variables, and is used in psychometrics as a test of validity. The conventional interpretation is that a correlation of <0.60 between two variables would indicate that they measuring different things; between 0.60-0.80 indicates they share something in common, but are not measuring the same thing; and correlations >0.80 imply the two measures are probably measuring the same thing. Correlations of >0.90 are needed before it can be asserted that the two measures are equivalent.
Coverage	Describes how well the descriptive system of a manifest instrument covers the latent construct of interest.
Cronbach $\alpha$	Measure of the reliability of a scale, based on examining the internal consistency of responses to items forming the scale. Cronbach $\alpha$ is based on both the correlations between items and the number of items within an instrument. However, where data distributions are highly skewed, $\alpha$ will represent the lower boundary of reliability rather than an accurate estimate.
Descriptive system	Refers to the actual items of an instrument and how these items are organized within an instrument.
DIF	Differential item functioning. DIF describes the extent to which two or more groups of respondents interpret an item differently (i.e. whether the item has significantly different meaning for the different groups).
Double-counting	Describes where the same issue is counted twice or more within an instrument. If there are redundant items in an instrument, then adding up their scores will produce double-counting.
End aversion	Describes where respondents avoid selecting an extreme option. E.g. a person may wish to avoid stating that they are 'extremely dissatisfied', so they will state that they are 'dissatisfied'.
Guttman scale	Describes a response scale where the responses progressively increase (e.g. <i>none, some, a lot, many</i> ).
Guttman scalogram	Describes an instrument comprising Guttman scales, where respondents order their responses such that $a < b < c < d$ etc.

---

---

Homogenous scale	Describes a scale where all the items in the scale are measuring the same latent construct. Ideally, all scales should be homogenous as this minimizes measurement error. Homogeneity is usually tested using factor analysis, which groups items according to how well they are correlated. It is also tested for with IRT and Mokken analysis.
Internal consistency	Describes the extent to which a scale is reliable. The most common method of testing for internal consistency is Cronbach $\alpha$ .
Instrument	An instrument is the formal language used to describe the descriptive system of a measure. It usually comprises several scales, each of which contains several items.
Item	Is the term used to describe a single question, where the psychometric properties of the question are known. In contrast a 'question' has no formally known measurement properties. Items consist of two parts: the item stem is the question part, and the item response set is the response part.
Item response theory	Referred to as IRT and modern test theory, this postulates that a person's probability of selecting a particular response to an item is conditional upon their ability to select the correct response for him/her, and that abilities and probabilities can be separately described.
IRTC	Item rest of test correlation. Describes how well an item fits a scale based on correlations with the other items in the scale.
Kappa ( $\kappa$ )	A measure of the level of agreement between two observers.
Latent construct	Describes an object that doesn't exist but that is presumed to exist, such as love. For example, in this report a latent construct of interest is patient satisfaction. A latent construct is defined by a theoretical model postulated by the researchers.
Likert scale	Describes a scale where the distance of responses from a mid-point indicates the strength of agreement or disagreement with a statement (e.g. the responses to the question: <i>You are satisfied with your treatment</i> might be: <i>strongly disagree/disagree/neither/agree/strongly agree</i> ).
Manifest instrument	The descriptive system of an instrument that is used to represent a latent construct. It is the instrument that is administered to respondents.
Mokken analysis	A form of item response theory analysis which assesses the unidimensionality of a scale based on the axioms of Guttman scalogram measurement.
Mokken rho ( $\rho$ )	Internal consistency reliability estimate for use with Guttman scales where the data are highly skewed.
Monotonicity	Monotonicity describes where mean scores on an instrument vary in order on a known response set from a criterion. For example, if the criterion is good health and the response set is Excellent/Very good/Good/Fair/Poor, and the scores of interest are walking rate, a monotonic relationship would be where the mean walking rates, for each response set level, were Excellent: 140cm/sec; Very Good: 135 cm/sec; Good: 128 cm/sec; Fair: 113 cm/sec; Poor: 102 cm/sec.

---

---

Nomological net	Because validity is never established, researchers collect a variety of different types of validity evidence relating to an instrument. Where sufficient evidence is collected this is referred to as a nomological net of evidence.
Psychometrics	This is the discipline of measurement, where psychometric refers to the formal measurement properties of an item, scale or instrument
Redundancy	Refers to items that are not needed in a scale, i.e. their presence does not contribute to the scale, and the scale is as reliable and valid with these items removed.
Relative efficiency (RE)	Describes how responsive a scale is when compared with another scale
Reliability	Describes the stability of scale scores. A person who scores $X$ on a scale should also score $X$ on the scale if they complete the scale a second time. Reliability is usually assessed through correlation at test-retest, Cronbach $\alpha$ , or the correlation between half of a scale compared with the other half, administered at the same time (split-half reliability). It is difficult to be precise about the desired reliability levels. For example, longer scales will have higher reliability than shorter scales. The conventions are that for comparison of groups reliability should be within the range 0.70 to 0.90. For individual assessment (e.g. clinical diagnosis) the literature has suggested values in the range of 0.70 to 0.95.
Response scale	Sub-set of response set. Items often use a response scale on which the respondent selects the response that best describes his/her position. E.g. An item may ask <i>Do you leak urine?</i> and the response scale might be <i>Not at all, a little, some, a moderate amount, a lot.</i>
Response set	The set of responses attached to any item, regardless of whether they form a response scale, are multiple selection or other type.
Responsiveness	Describes the sensitivity of a scale to differences in the same respondents in the underlying condition over time.
Scale	Refers to a collection of items that, between them, measure a construct. It is accepted that the items within a scale should be homogenous. Several scales may be included in an instrument.
SD	Standard deviation.
Sensitivity	Describes the responsiveness of a scale to different groups of respondents who have different known conditions, where the comparisons are made at a single point in time.
Then-test	Describes the difference between a participant's current health state and his/her previous health state, where the previous health state is assessed by asking the participant to recall his/her previous health state.
T-score	Standardized scores where the mean score = 50 and the standard deviation = 10.
Validity	Refers to evidence that suggests an instrument (or scale) measures what it is claimed to measure. Since validity is made up of two components – the

---

properties of the descriptive system and the ability of the respondents – validity varies from sample to sample. Researchers therefore collect different types of validation evidence about an instrument; hence the ‘nomological net of evidence’ . Because respondents vary in their ability to answer questions (e.g. consider those who are continent compared with those who are incontinent), an instrument that has validity in one population sample may not be valid in another sample. Therefore validation exercises should be undertaken each time an instrument is used in a new study.