Towards a Measure of Function for Home and Community Care Services in Australia: Part 2 – Evaluation of the Screening Tool and Assessment Instruments

Janette Green, Kathy Eagar, Alan Owen, Rob Gordon and Karen Quinsey

Centre for Health Service Development, University of Wollongong

The National Home and Community Care (HACC) Dependency Data Items Project was established to recommend, for national use, validated and reliable instruments for measuring the dependency of people eligible for HACC services. In Stage 2 of the project—reported here—a screening tool and assessment instruments selected in Stage 1 of the project were field-tested in a range of HACC agencies. The performance of the screen and the associated assessments was evaluated, as was their acceptability to HACC staff and clients. The results suggest that all of the five instruments are suitable for the target population of HACC services and they can be used effectively across a broad range of service types and by providers of diverse professional backgrounds.

Key words: *Assessment, Community care, Functional dependency, Activities of daily living*

The National Home and Community Care (HACC) Dependency Data Items Project was established to recommend for national use, validated and reliable instruments for measuring the dependency of people eligible for Home and Community Care (HACC) services. The broad aims of the project were twofold. The first was to determine whether or not screening and assessment instruments chosen after a review of the literature were appropriate for the HACC and aged care target population when administered by HACC- funded staff. The second was to determine how easily the new screening and assessment procedures could be adopted into current assessment practices, identifying potential barriers where possible.

The project was split into two stages. In the first stage, reported elsewhere in this Journal (Eagar, Owen, Marosszeky, & Poulos, 2006), a review was undertaken of existing instruments designed to measure dependency in a variety of populations. A measure of dependency was defined as an instrument that identifies areas in which a person requires assistance with daily living, and that quantifies the extent to which that person has to rely on someone else to help them carry out normal activities in their home and community.

Published scales for the assessment of physical function and daily living tasks largely consider activities of daily living (ADLs) or instrumental activities of daily living (IADLs) as separate entities. ADLs represent the basic, physical functions that underlie normal living such as toileting, dressing, feeding and walking. IADLs (also known as extended or domestic activities of daily living) are those activities needed for continued community residence such as shopping, preparing meals, cleaning the house, managing personal finances, reading and using the telephone.

When the original ADL concept was developed by Katz and colleagues in 1963, it was noted that these activities of daily living form a hierarchy, with older people losing their abilities in the reverse order to which they learned them in childhood. "Early loss" activities, such as housework and managing finances, are gained last and lost first; "late loss" activities, such as dressing and feeding, are gained first and lost last (Katz et al., 1963).

This concept of a functional hierarchy provided the rationale for the adoption of a two-tiered assessment approach. Given that people acquire and lose abilities in a predictable order, a screening instrument could be used to identify where a person sits on the functional hierarchy. It was envisaged that, as a first tier, a simple screening tool could be applied to all current and potential consumers by staff in any service provider agency, with minimal training, irrespective of their background or qualifications.

This screen could be applied on a routine basis and could be developed into a minimum data set (MDS). The initial screen was not considered to be an assessment of client dependency. Instead, the results of the screen were planned to trigger a more in-depth assessment of clients using one or more recommended instruments, namely:

- The Barthel Index (Collin, Wade, Davies, & Horne, 1988) or Functional Independence Measure (FIM, [Granger et al., 1993]) to assess self-care
- Lawton's Instrumental Activities of Daily Living (IADL) Scale, with modifications to make it suitable for the HACC program, to assess domestic functioning
- Behaviour items from the Australian Resident Classification Scale (RCS), with modifications to make them suitable for the HACC program, to assess behaviour
- The Folstein Mini-Mental State Examination (MMSE, [30-point]) to assess cognition.

These assessment instruments would form the second tier in the process. This second tier assessment would not be provided to all clients but only to those with identified needs and functional limitations. Moreover, the functional assessment might be targeted to specific domains, based on the screen results. The second tier was designed specifically to be a functional assessment, which could occur as part of a more comprehensive assessment, or it could be undertaken separately as a specialist assessment for the purpose of a deeper identification of functional limitations and their resulting burdens.

In the second stage of the project—reported here—the screening tool and the assessment instruments were field-tested in a range of HACC agencies (Eagar et al., 2001). The field test focused on assessing the performance of the tools and their acceptability to HACC staff and clients. It also aimed to explore, where possible, the validity and reliability of the screen when used with the target population.

Method

Field testing

A short pilot test was conducted in one HACC agency and one Aged Care Assessment Team (ACAT) in New South Wales, and training and resource manuals were provided to participating agencies before the field test began. The field test spanned a four-week period and involved 40 HACC agencies in the Hunter region of New South Wales, the Kingston/Bayside area of Melbourne, Victoria, Alice Springs in the Northern Territory and Southern Adelaide in South Australia. Agencies were selected to ensure the sample was representative and:

- Covered the broad range of ages that exist within the HACC target group
- Covered a mix of agency sizes
- Included staff who had a mix of skills
- Covered ethno-specific agencies; for example, for Aboriginal and Torres Strait Islander (ATSI) or non-English Speaking Background (NESB) clients.

It was assumed that, in general, the screen would be applied in routine practice whenever contact was made with a service provider for new or additional services. Results of the screen would then trigger one or more assessments for some clients. However, in the field trial, participating sites screened both new and existing clients, even when there was no request for additional services. They were also asked to assess clients, irrespective of the results of the screen, even though this did not reflect routine practice. This would enable the linking of data from screens and assessments across the full spread of need—not just the complex or high-need end—in order to establish whether the screen selected the appropriate clients for assessment.

Agencies were given the choice between the two self-care measures selected for field-testing. As all agencies selected the Barthel Index, the FIM was not tested.

Research questions

The data were analysed to answer three broad questions. First, were the screen scores able to provide useful information on this population, and, if so, did they provide evidence of a functional hierarchy? Second, was the use of the instruments acceptable to staff and clients? Third, did the screen do what it was designed to do?

To answer the first of these questions, the tools were investigated for floor and ceiling effects. In other words, when used on HACC clients, the distribution of scores should be across the whole spectrum of functional ability measured by an instrument; the majority of scores should not be clumped at either the top (ceiling effect) or the bottom (floor effect) of the scale.

Since the rationale behind the two-tiered assessment approach is that people lose abilities in a predictable order, it was important to establish whether or not a functional hierarchy was evident or not for this particular client group. To do this, clients were grouped according to the number of dependencies they reported. The most commonly occurring combinations were then investigated for evidence of a hierarchy of loss.

The data collection forms completed after each screen and assessment asked staff to record their qualifications and their years of experience. The time taken to complete the screen or assessment was recorded, as was an indication of their confidence in the ratings and any problems experienced. These data were used to assess both the burden placed on staff by using these tools and the acceptability of the tools themselves in terms of the clarity of the instructions, their relevance to clients, and the usefulness of the information provided by the items.

Three different analyses were used to gauge whether the screen was doing what it was designed to do, namely to recommend assessments for the clients who needed it. First, staff were asked to independently recommend further assessment for clients. The agreement between these staff recommendations and recommendations based on the functional screen was investigated.

Second, a detailed statistical analysis of the screening tool was undertaken to determine the sensitivity and specificity of the measures that triggered each of the four assessments. The sensitivity of the instrument is measured as the proportion of those with higher dependency levels (based on the assessment) for whom an assessment was recommended. The specificity is measured as the proportion of those with little or no dependency (based on the assessment) for whom an assessment was not recommended.

To do this, it was necessary to identify the client's "need" for an assessment. Because staff had assessed many clients—whether or not it was recommended by the screen—a broad range of assessment scores was available. High need and moderate to low need groups were defined using the scores on the assessment tools. High need for the respective domain was defined as scoring less than 19 on the self-care tool, less than 26 on the domestic tool, less than 20 on the behaviour tool, and less than 25 on the cognition tool.

In the third analysis to determine the appropriateness of the screen recommendations, a set of alternative rules that could be applied to the screen items to trigger further assessment were devised. These rules were compared using sensitivity and specificity analysis. A sensitivity of 90%, say, means that 10% of clients with higher dependency are missed. A specificity of 60%, say, means that 40% of clients with low levels of dependency are assessed unnecessarily. There are costs associated with these errors in terms of too many clients missed versus too many unnecessarily recommended. The decision regarding appropriate thresholds for triggering further assessments has to consider the relative trade-offs between these costs.

Data on the completed screen and assessment forms were entered into a Microsoft Access database and analysed using Microsoft Excel and SAS.

Results

In total, 746 screens and 462 assessments were returned. There were both a screen and one or more assessment forms completed for 456 clients.

Question 1: Were the screen scores able to provide useful information on this population and, if so, did they provide evidence of a functional hierarchy?

Scores on all items in the screen were fairly evenly distributed across the range of possible values, thereby providing no evidence of floor or ceiling effects. When clients were grouped according to the number of reported dependencies, the first five items— which all relate to IADL functions—were indeed found to be hierarchical, the order being housework (Item 1), shopping (Item 3), transport (Item 2), finances (Item 5), and medication (Item 4). The coefficient of reproducibility was 0.9674. Any value over 0.9 is considered to be good.

Question 2: Was the use of the instruments acceptable to staff and clients?

The burden on staff in terms of time taken to complete the tools was acceptable. More than 75% of screens were completed within 10 minutes, while more than 60% of consultations where all four assessment tools were used were completed within 30 minutes. Staff with professional health qualifications took significantly less time to complete the screen than those with welfare or other relevant qualifications and those with no relevant qualifications ($p<0.001$). This effect was not found among staff who completed the assessment tools.

Staff from agencies providing predominantly post-acute care took significantly less time to complete screens and assessments than staff from other agencies. A related finding was that screens and assessments conducted at the client's home (or place of usual residence) or in the assessor's office took significantly longer than those completed from a patient's hospital record or by telephone interview.

In total, responses for 94 clients addressed the issue of relevance and appropriateness. Some questioned a rule employed in the study in which items relating to management of finances and medication triggered a cognitive and a behavioural assessment. Staff questioned whether this was necessary when such difficulties were due to physical or medical problems. Further, if the screen flagged cognition for further assessment, behaviour was also automatically flagged, but often only one of these was a problem. The screen did not pick up motivation of clients and it was hard to apply to those with varying levels of function due to having episodic conditions. Some questions could not be answered by very young clients or by those with a very high level of disability. Client sensitivity to discussing functional abilities was also raised as an issue.

Staff were asked about the difficulty and accuracy of rating. Ninety-one per cent of the staff rated the assessment tools as very or moderately easy to score and 89% were very or fairly confident of the accuracy. Some reported specific difficulties, such as rating functional ability without assessing broader contextual factors. Fluctuating functional ability was also mentioned as a reason for finding a rating difficult.

Recommendations for further assessment by the screen are cross-tabulated with staff recommendations in Table 1. In most cases, the screening tool was overly sensitive and prone to select more clients for further assessment than the staff felt necessary. However, in some cases, the interviewer felt that the client should be assessed further, even though the screening tool had not identified a problem. There was a higher level of agreement between staff and screen recommendations when the assessor had conducted more then 10 screens.

In some cases, the interviewer recorded why they disagreed with the screen's recommendation. Comments included: other information being known about the client; being unsure that the client had understood the question; and feeling that the screen had missed an important attribute of the client. Numerous staff raised the issue of the frequency of administration and whether all questions would be needed for all clients or at all agencies. Staff already knew many of the clients and usually this meant that adequate support services were in place, so they considered further assessment unnecessary. Many clients needed help in a specific area and staff indicated a preference for identifying specific service needs or interventions, such as falls prevention, home maintenance and home nursing assistance.

Question 3: Did the screen do what it was designed to do?

Clients were categorised on their need for assessment into high and low dependency groups, based on their assessment scores. Using these categories of need, the screen was found to have performed better than the interviewers on sensitivity and, for self-care and domestic assessments, on specificity (as shown in Table 2).

In the field trial, the screening tool triggered a domestic assessment if a dependency was recorded in any of the first three items. These covered housework, shopping and transport. This was considered to be too sensitive. Only 15% of clients could independently manage housework. Some of these clients were not totally independent on the shopping and transport items. In fact, the screen triggered a domestic assessment for 89% of clients. A number of alternatives were tested to assess whether there are better rules for triggering a domestic assessment than the one used in the trial. Sensitivity and specificity of each of these alternatives were compared, resulting in a change to the rule for triggering a domestic assessment to be based on a count of the number of 2s scored in the first five screen items, rather than being based on their total score.

Various options were also tested as alternative rules for triggering a self-care assessment. These included using the first five (domestic) items on the screen only, selecting two different items from

Table 1: Number of clients classified by assessment recommendation

Domestic assess recommended by screen?	Domestic assess recommended by interviewer?	
	Yes	No
Yes	444 (65%)	168 (24%)
No	11 (2%)	62 (9%)
Self-care assess recommended by screen?	Self-care assess recommended by interviewer?	
	Yes	No
Yes	256 (39%)	105 (16%)
No	63 (9%)	240 (36%)
Behaviour assess recommended by screen?	Behaviour assess recommended by interviewer?	
	Yes	No
Yes	130 (22%)	214 (37%)
No	13 (2%)	225 (39%)
Cognitive assess recommended by screen?	Cognitive assess recommended by interviewer?	
	Yes	No
Yes	200 (32%)	183 (29%)
No	22 (4%)	217 (35%)
Agreement between screen and interviewer?	Yes	No
Domestic	506 (74%)	179 (26%)
Self care	496 (75%)	168 (25%)
Behaviour	355 (61%)	227 (39%)
Cognitive	417 (67%)	205 (33%)

the Barthel Index to be included in the screen rather than those initially chosen and/or counting the number of 2s scored on the self-care items in the screen rather than adding them. Based on this analysis, the screening tool was redesigned to replace the item on dressing with an item on mobility. However, the rule for triggering a self-care assessment remained unchanged.

The existing rules for triggering both a behavioural and a cognitive assessment displayed high sensitivity but low specificity. Although a number of different rules were tested, none improved on this result. The alternative approach adopted was to change the way that the final two screen items were scored. In the field trial there was no capacity for the interviewer to rate based on his or her own observations of the client. This was changed to allow these items to be rated based on all available information, including the observations of the interviewer. This change to the rules is expected to make a difference, especially in relation to specificity. However, the actual impact can only be determined at a later date through a formal test. The final version of the functional screening instrument is provided in Attachment 1.

Discussion and Conclusions

The use of the screening instrument and associated assessment instruments (Barthel Index, Lawton's IADL measure, Modified RCS behaviour measure, MMSE) for home and community care services in Australia was supported by the results. The tools were designed as generic assessments of function to complement (not replace) specialist assessments undertaken by individual professions. These selected tools from the literature satisfactorily answered the three research questions in terms of providing useful information about the target population, acceptability for staff and clients, and screening ability.

However, the field trial had limitations. Ideally, the screen should have been formally tested for inter-rater reliability but the agencies did not have the resources for this. Thus, the trial was limited to an alternative, less powerful, approach in which staff rated their confidence in their application of the tools and the results.

The sample size in the field test was insufficient to test the appropriateness of the screening instrument for special needs groups such as ATSI and NESB consumers. This can only be assessed when it is introduced into routine practice.

Table 2: Sensitivity and specificity of recommendations for each of the assessments

	Self-care		Domestic		Behaviour		Cognition	
Recommended by	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity
Screen	84%	83%	99%	35%	92%	29%	95%	28%
Staff	71%	71%	80%	31%	72%	81%	84%	54%

While every attempt was made to ensure the sample in the field trial was representative of the total HACC population, there is no way of verifying this. It may be that results such as sensitivity and specificity will differ with other samples, or when the data are routinely collected on a national basis. Reviewing the results after sufficient data have been collected on a routine basis would be valuable.

Measures of behaviour and cognition were originally included on the assumption that, while the volume was likely to be low, the impact on carers and on the need for community support services would be high. There is nothing to suggest this is not the case. Accordingly, the Modified RCS behaviour measure and the MMSE appear suitable, at least until such time as more appropriate instruments for use in non-clinical community care services are available. A study to design such a cognitive instrument has recently been reported (Storey, Rowland, Conforti, & Dickson, 2004).

Further work is required to develop a data collection protocol that addresses who should receive functional screening and assessment and who should undertake them as well as when and how often they should happen, including rules for re-screening clients already receiving HACC services. The reliability of the data will depend on the number of people in each agency who carry out assessment; this needs to be considered in the development of the protocol. The protocol will also need to address data management and consumer participation and confidentiality.

Staff training and guidelines are needed on the various functional screening and assessment tools and on the data collection protocol. Training is required to address issues such as how to rate a person with fluctuating functional ability and how to rate the cognition and behaviour items. For existing clients, staff need to be sensitive to decline in their functional status over time. Specific protocols regarding the frequency of repeat screening will help detect such changes, and training will also improve staff skills in this area.

Functional measurement appears warranted on a routine basis across the whole spectrum of HACC services in Australia and functional data need to be included in the HACC MDS. The level of detail incorporated in the national MDS needs to be determined by the purposes for which the data will be used, taking into account the administrative burden of collection. This project was the first necessary step toward the standardisation of the assessment of function for Home and Community Care Services in Australia. How the tools are used and who uses them will critically influence the long-term benefits of routinely collecting this information.

Acknowledgements

This study was funded by the Commonwealth Department of Health and Ageing. The study was approved by the University of Wollongong Ethics Committee.

References

Collin, C., Wade, D.T., Davies, S., & Horne, V. (1988). The Barthel ADL Index: A reliability study. *International Disability Studies, 10*(2), 61-63.

Eagar, K., Owen, A., Green, J., Cromwell, D., Poulos, R., Gordon, R., Quinsey, K., & Adamson, L. (2001). *Towards a national measure of functional dependency for Home and Community Care services in Australia: Stage 2 Report of the HACC dependency data items project.* University of Wollongong: Centre for Health Service Development.

Eagar, K., Owen, A., Marosszeky, N., & Poulos, R. (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, 12*(1), 73-91.

Folstein, M.F., Folstein, S.E., & McHugh, P.R. (1975). "Mini-Mental State": a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*, 189-98.

Granger, C.V., Hamilton, B.B., Linacre, J.R. et al. (1993). Performance profiles of the Functional Independence Measure. *Archives of Physical Medicine and Rehabilitation, 72, 84-9.*

Katz, S., Ford, A.B., Moskowitz, R.W., Jackson, B.A., & Jaffe, M.W. (1963). Studies ofillness in the aged. The index of ADL: a standardized measure of biological and psychosocial function. *Journal of the American Medical Association, 185*, 914-9.

Lincoln Gerontology Centre. (1998). *National Framework for Comprehensive Assessment.* Aged and Community Care Service Development and Evaluation Report No 34. Canberra: Australian Government Publishing Service.

Mahoney, F., & Barthel, D. (1965). Functional evaluation: the Barthel Index. *Maryland State Medical Journal, 14*, 41-3.

Storey, J., Rowland, J., Conforti, A., & Dickson, H. (2004). The Rowland Universal Dementia Assessment Scale (RUDAS): a multicultural cognitive assessment scale. *International Psychogeriatrics, 16*(1), 13-31.

Thomas, V.S., Rockwood, K., & McDowell, I. (1998). Multidimensionality in Instrumental and Basic Activities of Daily Living. *Journal of Clinical Epidemiology, 51*(4), 315-321.

Janette Green
Centre for Health Service Development
University of Wollongong
Wollongong NSW 2522
AUSTRALIA

Kathy Eagar
Centre for Health Service Development
University of Wollongong
Wollongong NSW 2522
AUSTRALIA
E-mail: keagar@uow.edu.au

Alan Owen
Centre for Health Service Development
University of Wollongong
Wollongong NSW 2522
AUSTRALIA

Rob Gordon
Centre for Health Service Development
University of Wollongong
Wollongong NSW 2522
AUSTRALIA

Karen Quinsey
Centre for Health Service Development
University of Wollongong
Wollongong NSW 2522
AUSTRALIA

Correspondence to Kathy Eagar

Attachment 1

National HACC Functional Screening Instrument

Part One: Questions to ask the client (or the person who represents the client)[1]:

Unique Client ID_______________________________ **Date screened**_______________________

I would like to ask you about some of the activities of daily living, things that we all need to do as part of our daily lives. I would like to know if you can do these activities without any help at all, or if you need some help to do then or if you can t do them at all. The questions refer to how you are managing at the moment.

Item	Question	Score	Record score
1	**Can you do housework...**		
	Without help (can clean floors etc)?	2	
	With some help (can do light housework but need help with heavy housework)?	1	
	Or are you completely unable to do housework?	0	
2	**Can you get to places out of walking distance...**		
	Without help (can drive your own car, or travel alone on buses or taxis)?	2	
	With some help (need someone to help you or go with you when travelling)?	1	
	Or are you completely unable to travel unless emergency arrangements are made for a specialised vehicle like an ambulance?	0	
3	**Can you go out for shopping for groceries or clothes (assuming you have transportation)...**		
	Without help (taking care of all shopping needs yourself)?	2	
	With some help (need someone to go with you on all shopping trips)?	1	
	Or are you completely unable to do any shopping?	0	
4	**Can you take your own medicine...**		
	Without help (in the right doses at the right time)?	2	
	With some help (able to take medication if someone prepares it for you and/or reminds you to take it)?	1	
	Or are you completely unable to take your own medicines?	0	
5	**Can you handle your own money...**		
	Without help (write cheques, pay bills etc)?	2	
	With some help (manage day-to-day buying but need help with managing your chequebook and paying your bills)?	1	
	Or are you completely unable to handle money?	0	
Do not ask the following 2 questions if the client scored 2 on all of the above 5 items (ie, can do all 5 activities without help). Instead, for clients who scored 2 on all of the above items, record a 9 on each of the following 2 items to indicate that you did not ask the question.			
6	**Can you walk...**		
	Without help (except for a cane or similar)?	2	
	With some help from a person or with the use of a walker, or crutches etc	1	
	Or are you completely unable to walk?	0	
7	**Can you take a bath or shower...**		
	Without help?	2	
	With some help (eg, need help getting into or out of the bath)?	1	
	Or are you completely unable to bathe yourself?	0	

NOTES:

- *If unanswered, score X.*
- *Rate what the person is* **currently capable** *of doing rather than what they actually do. In assessing capability, take into account not only physical function but also cognition (such as problems caused by dementia or an intellectual disability) and behaviour (such as unpredictable challenging behaviour). Consumers able to complete a task with verbal prompting should not be rated as independent (anc therefore should be rated as a 1). In rating an item that is irrelevant (for example, the person has no shops in the vicinity or does not use any medications), rate based on what the person would be capable of doing if the item was actually relevant to their situation.*
- *Item 6 (walking). Clients who are in a wheelchair should be rated as (1) if they are independent including corners etc or (0) if they are not wheelchair independent.*

1 Reproduced from the OARS/MFAQ. Copyright: the Center for the Study of Aging and Human Development, Duke University Medical Cen Durham, North Carolina. Used with permission. Questions 1, 6 and 7 have been modified.

Part Two: Questions for you to complete

Complete the following based on all information available to you – your judgement based on interviewing or observing the client, information contained in a referral letter, client notes or information provided by a proxy respondent, such as a friend, relative, carer or referring agency.

Note that the client should not be asked to answer these questions.

Item	Question	Record score
8	Does the person have any memory problems or get confused? No – score 2 Yes – score 0	
9	Does the person have behavioural problems for example, aggression, wandering or agitation? No – score 2 Yes – score 0	

Recommended functional assessments based on this functional screen

Domestic

Look solely at items 1 to 5. Count the number of these items that scored 2 (ie, count the number of activities that the person can do without help).
Refer for a domestic functional assessment if the person can do less than 3 activities without assistance – ie, the count is 2 or less (a count of 0, 1 or 2).

Self-care

Refer for a self-care functional assessment if the client SCORED LESS THAN 2 on either Item 6 (mobility) or Item 7 (bathing).

Cognition

Refer for a cognitive assessment if:

- the client scored LESS THAN 2 on either Item 4 (medicine) or Item 5 (financial management) AND you havı determined that the client has no physical disabilities or problems with English literacy that may account for tl client not being independent on these items OR
- the client scored 0 on Item 8.

Behaviour

Refer for a behavioural assessment if:

- the client scored LESS THAN 2 on either Item 4 (medicine) or Item 5 (financial management) AND you havε determined that the client has no physical disabilities or problems with English literacy that may account for tł client not being independent on these items OR
- the client scored 0 on Item 9.

Economic Evaluation of the Good Life Club Intervention for Diabetes Self-management

Duncan Mortimer[1] and Jill Kelly[2]

Centre for Health Economics, Faculty of Business and Economics, Monash University[1]
Good Life Club Project, Whitehorse Division of General Practice[2]

This paper reports results from an economic evaluation of the Good Life Club (GLC), a self-management intervention for type II diabetes. Based on intention to treat data at 18-month follow-up, 39% of GLC clients were well-managed compared to just 30% at baseline. This gives a risk difference of 0.09 (95%CI: 0.01, 0.16) and a relative risk of being well-managed equal to 1.28 (95%CI: 1.04, 1.58) for the GLC intervention as compared to usual care. Based on detailed description of the GLC intervention, client-level data as to service utilisation and published unit costs, the incremental cost per client of the GLC intervention as compared to usual care was calculated at $1,457 in the base case analysis. Combining estimates of incremental cost and treatment effect implies an incremental cost-effectiveness ratio of just over $16,000 per additional responder (client with well-managed type II diabetes) in the base-case analysis. Additional cost savings are, however, expected to accrue in future years because well-managed diabetes is less likely than usual care to result in diabetes-related complications.

Key words: *Economic evaluation, Diabetes, Chronic disease, Self-management*

The Good Life Club (GLC) was a demonstration project funded by the Commonwealth Department of Health and Ageing under the Sharing Health Care Initiative. The GLC adopted a "club" approach to supporting/educating patients and aimed to build on existing health service delivery structures to increase the capacity of clients to self-manage their condition. The GLC employed a number of strategies in pursuit of its objectives, which included telephone coaching, group activities and the GLC website. A detailed description of the rationale, design and implementation of the GLC intervention has been provided in a previous issue of this journal (see Kelly, Menzies, & Taylor, 2003; Lindner, Menzies, & Kelly, 2003), but a brief description is given here to assist in interpreting findings.

GLC clients completed a self-management plan in consultation with the GLC Project Manager upon recruitment to the study and were referred either directly to a GLC coach or to the Coach Co-ordinator for the local Community Health Service, which then allocated each client to a GLC coach. Each client was also advised to make an appointment with their usual GP; firstly, to request that the GLC be kept informed of that client's medical management plan, and secondly, to discuss self-management objectives. During the first 12 months of membership, GLC clients were to receive a telephone coaching session once every month averaging 20 minutes' duration. In practice, the average number of coaching sessions was just six per client, such that the treatment effect observed in the data is associated with a slightly less intensive intervention than planned. Coaches completed a consultation map during each coaching session and three-monthly reports to each client's usual GP.

In addition to coaching sessions, the GLC intervention included a number of group activities designed to address risk factors for disease progression. GLC activities included social get-togethers, physical activity sessions (including group walks and aqua aerobics), educational talks, supermarket tours, focus groups and internet/email training. A total of 224 GLC activity sessions were delivered over the 18-month intervention period with educational talks and internet/email training accounting for 190 of the 224 sessions. Client attendance at GLC activity sessions was 1189 client sessions.

The GLC intervention also included a number of strategies to educate/support clients and to encourage collaboration between clients, carers and health care providers. For example, a community resource kit was provided at baseline, carers were encouraged to attend GLC activity sessions and GLC coaches were instructed to