

2025

ePPOC Adult Benchmarks

INTERIM REPORT – OCTOBER 2025

ACKNOWLEDGEMENTS

We acknowledge the Traditional Custodians of the lands on which the University of Wollongong is situated. We pay our respects to Aboriginal Elders past and present, who are the knowledge holders and teachers. We acknowledge their continued spiritual and cultural connection to Country. As we share knowledge, teaching, learning and research within this University we also pay respect to the knowledge embedded forever within the Aboriginal Custodianship of Country.

We also want to acknowledge the valuable contribution made by:

- The many staff from pain management services who have spent considerable time collecting, collating and correcting the data and without whose effort this report would not be possible
- Members of the ePPOC Clinical and Management Advisory Committee

DISCLAIMER

ePPOC has made every effort to ensure that the data used in this report are accurate. Data submitted to ePPOC are checked for anomalies and services asked to resubmit information where relevant.

COPYRIGHT

This work is copyright. It may be produced in whole or in part for study or training purposes subject to the inclusion of an acknowledgement of the source. It is not for commercial use or sale. Reproduction for purposes other than those above requires the written permission of ePPOC.

SUGGESTED CITATION

Phillips, B (2025) Electronic Persistent Pain Outcomes Collaboration Adult Benchmarks: 2025 Review. Interim Report. Faculty of Science, Medicine and Health, University of Wollongong

CONTENTS

| | |
|-------------------------------------------------------------------|----|
| 1. EXECUTIVE SUMMARY | 3 |
| 2. INTRODUCTION | 3 |
| 3. DATA AND METHODS | 4 |
| 4. AVERAGE PAIN | 5 |
| 4.1 Overview and Definitions | 5 |
| 4.2 Benchmark Performance | 5 |
| 4.3 Discussion and Future Directions | 6 |
| 5. PAIN INTERFERENCE | 6 |
| 5.1 Overview and definitions | 6 |
| 5.2 Benchmark Performance | 7 |
| 5.3 Discussion and Future Directions | 7 |
| 6. PAIN CATASTROPHISING | 7 |
| 6.1 Overview and definitions | 7 |
| 6.2 Benchmark Performance | 8 |
| 6.3 Discussion and Future Directions | 8 |
| 7. PAIN SELF-EFFICACY | 9 |
| 7.1 Overview and definitions | 9 |
| 7.2 Benchmark Performance | 9 |
| 7.3 Discussion and Future Directions | 10 |
| 8. DEPRESSION, ANXIETY AND STRESS | 10 |
| 8.1 Overview and definitions | 10 |
| 9. CONCLUSIONS AND NEXT STEPS | 13 |
| 10. REFERENCES | 14 |
| APPENDIX A: ASSESSMENT TOOLS | 15 |
| Brief Pain Inventory | 15 |
| Depression Anxiety Stress Scales | 16 |
| Pain Catastrophising Scale | 16 |
| Pain Self-Efficacy Questionnaire | 17 |
| APPENDIX B: CRITERIA FOR DEVELOPING EPPOC BENCHMARKS | 18 |
| Criteria | 18 |
| Domains | 18 |
| APPENDIX C: ASSESSMENT TOOL VALIDITY AND COMPLETION | 19 |
| APPENDIX D: MEAN DAS-21 SUB-SCORES | 19 |

1. EXECUTIVE SUMMARY

Benchmarking drives continuous quality improvement by comparing performance against established standards. ePPOC’s benchmarks were developed through a collaborative consensus to identify core domains and set aspirational targets for pain services. This review of seven of ePPOC benchmarks analyses performance at the service- and patient-level and aims to ensure the benchmarks and their underlying assessment tools remain relevant, evidence-based, and aligned with evolving care models and consumer and clinician priorities.

2. INTRODUCTION

Benchmarking is a widely used method for driving continuous quality improvement across various sectors, including healthcare¹. At its core, benchmarking involves identifying a reference point or threshold, known as a benchmark, against which performance can be compared. The impact of benchmarking can be amplified by combining it with complementary strategies, such as audit and feedback and peer learning², which help uncover the causes of variation and highlight opportunities for improvement.

The electronic Persistent Pain Outcomes Collaboration (ePPOC) has nine established benchmarks for adult services that provide a structured foundation for measuring and comparing performance across participating pain services.³ To enhance its effectiveness, ePPOC’s benchmarking approach integrates data analytics with dedicated clinical support. Clinically trained Improvement Facilitators (IFs) work directly with services to embed ePPOC processes into routine care and help translate results into actionable quality improvement initiatives. Additionally, collaborative learning opportunities, including the annual Australasian Benchmarking Workshop, create a feedback loop for services to reflect on outcomes, share models of care, and refine practice based on the collective experience.

The nine adult ePPOC benchmarks and their aspirational targets were developed through a collaborative consensus process involving member services and the ePPOC governance committee. Established in 2014, the benchmarks were designed to reflect meaningful outcomes while balancing the feasibility of consistent measurement across diverse pain services.

This review focuses on seven ePPOC benchmarks: average pain, pain catastrophising, pain self-efficacy, pain interference, depression, anxiety, and stress. We explore performance metrics at both the service- and patient-level, and discuss the whether the selected domains and their underlying assessment tools continue to capture outcomes that matter most. This discussion is particularly timely, given the concurrent major revision of the overarching ePPOC dataset (‘Version 3’). The two remaining benchmarks, which both relate to opioid use, have been excluded from this review. This is due to a concurrent review of the definition of ‘high-dose opioid use’, which may influence how these benchmarks are interpreted and applied.

This review also revisits the criteria for benchmark development and the steps involved in defining benchmark targets. The ePPOC benchmark criteria were originally endorsed in May 2016 by the ePPOC governance committee, which endorsed the following six principles for benchmark development (see Appendix 2):

- The domain is clinically important
- There is evidence in the literature demonstrating that the domain is amenable to intervention
- The benchmark threshold reflects good practice rather than average practice
- The benchmark is the same in all services regardless of public, private, tier or role delineation
- Clinical benchmarks should include patients whose baseline score reflects at least moderate severity
- There is sufficient data available to effectively benchmark

At the 2025 ePPOC Australasian Benchmarking Workshop, member services will provide input on the ongoing relevance of the benchmark criteria and the seven domains reviewed here, and may propose additional

outcome domains for future benchmarks, helping to ensure that ePPOC measures remain clinically meaningful and responsive to the changing landscape of participating services. The benchmarking workshop will additionally provide opportunity to discuss the underlying definitions of clinically significant improvement that the benchmark calculations rely on. While some of the current definitions are grounded in established IMPACT guidelines⁴, others were developed through consensus by the ePPOC governance committee. The committee has recognised the need to revisit these definitions and to consider alternative or additional metrics, such as minimal clinically important differences (MCIDs)—the smallest change in a score that patients perceive as beneficial.

The current ePPOC benchmark target levels also warrant review. According to the established criteria, these targets have been designed to reflect good practice rather than average performance; however, the concept of ‘good practice’ has never been formally defined and historically the proportion of services meeting each benchmark has varied widely. Informal feedback from services, as well as discussions among the ePPOC governance committee, suggest that some ePPOC benchmarks may be so difficult to achieve that they discourage active engagement. Additionally, evidence from reviews of clinical quality registries and audit programs also indicates that benchmarks that are perceived as unattainable or that lack clear definitions can reduce engagement and useability.⁵ To address this inconsistency, the 2025 benchmarking workshop will explore alternative approaches to target setting, including defining a benchmark so that a fixed proportion of services meet it or setting service-level targets a specified margin or percentile above the average result.

In summary, while ePPOC benchmarking is designed to provide a simple and rapid view of service performance, that simplicity depends on a robust foundation of clear definitions, reliable data, and shared standards. Only with these in place can services trust the results and use them to drive improvement. This review strengthens that foundation, ensuring the benchmarking system remains practical, credible, and meaningful. The 2025 ePPOC benchmarking workshop will give services a central role in completing this review, providing an opportunity to both reflect on performance and contribute critical insights that will help refine the foundations of the ePPOC benchmarks.

3. DATA AND METHODS

This review is based on an analysis of the ePPOC ‘all-services’ benchmarking results, which are generated every six months. The benchmarking reports include aggregated outcome data submitted by adult specialist pain services participating in ePPOC.

The current review includes data from six consecutive reporting periods, covering submissions between July 2022 and January 2025. Each submission reflects outcomes captured over a 12-month period. For example, the 2022 Mid-Year (MY) report includes data from 1 July 2021 to 30 June 2022, while the 2022CY report covers 1 January 2022 to 31 December 2022.

Two key types of data are presented in this review:

- **Service-level benchmark attainment:** The percentage of services that met each benchmark target in a given reporting period. To be eligible for assessment against a benchmark, a service must have captured outcomes for at least 10 patients relevant to that benchmark. The reported n values represent the number of services eligible for assessment against each benchmark.
- **Patient-level outcome attainment:** The percentage of individual patients who achieved a clinically significant improvement in that area. The n values associated with these data represent the number of patients with moderate to severe symptoms at referral who also had end of treatment data that enabled calculation of an outcome result. The definitions of clinically significant improvement for each measure are included in the relevant sections of this report.

Appendix A provides further details about the outcome measures and assessment tools included.

4. AVERAGE PAIN

SUMMARY: The ePPOC average pain benchmark is challenging for services to achieve. Due to the low valid completion rate of the item, and the difficulty patients have estimating ‘average’ pain, alternative measures—such as worst pain, pain frequency, or composite scores—may provide a more reliable reflection of treatment impact. Additionally, due to accumulating evidence that pain severity outcomes are influenced by baseline characteristics, benchmarking in the pain severity domain may require a more nuanced or case-mix adjusted approach.

4.1 OVERVIEW AND DEFINITIONS

The Brief Pain Inventory (BPI) is a core component of the ePPOC dataset and is used to capture patients’ experiences of pain intensity and its impact on daily life. Average pain is one of four pain intensity items included in the BPI, alongside worst, least, and current pain. The item asks patients to rate their average level of pain over the past week on a scale from 0 (no pain) to 10 (worst imaginable pain).

In the context of ePPOC benchmarking, average pain outcomes are calculated for the subset of patients that report moderate to severe average pain at episode start, and a clinically significant improvement at episode end is defined as a reduction in score of at least 30% from baseline, which is consistent with IMMPACT recommendations.⁴

The target for the ePPOC average pain benchmark is 40%. That is, for the benchmark to be achieved, at least 40% of applicable patients must report a clinically significant improvement in their average pain score between episode start and end.

4.2 BENCHMARK PERFORMANCE

This average pain benchmark has consistently been one of the most challenging targets to achieve, with only 11.2–19.6% of services meeting the benchmark between the 2022MY and 2024CY reporting periods (Figure 1A). While the patient-level results show less noticeable variation across the reporting periods, hovering between 30.4% and 27.9% (Figure 1B), both service- and patient-level data show a decline in performance after the 2023MY reporting period.

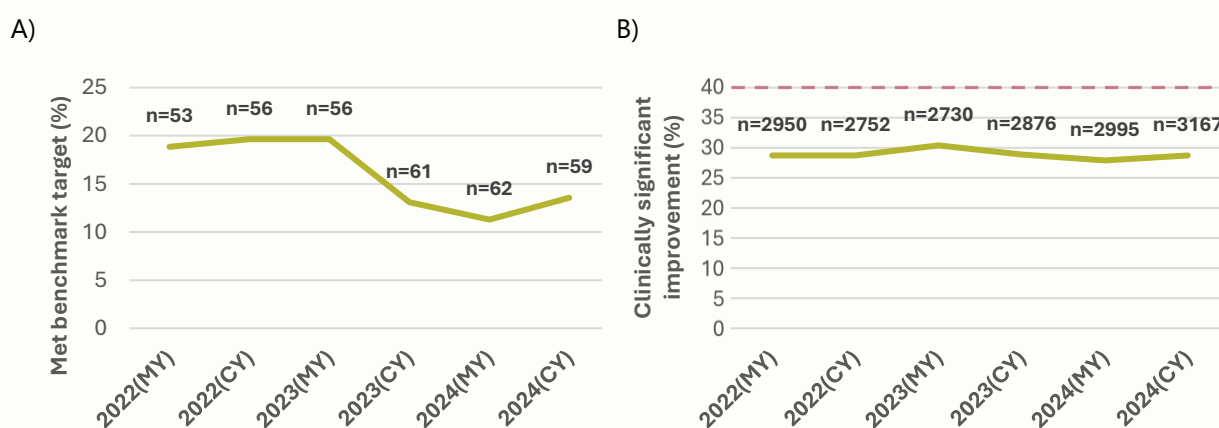


Figure 1 Average pain benchmark performance and patient outcomes between the 2022MY and 2024CY reporting periods. (A) Percentage of services meeting the average pain benchmark target of 40% of eligible patients achieving clinically significant improvement. Services must record outcomes for at least 10 eligible patients for a benchmark result to be calculated. n represents the number of services for whom a benchmark result was calculated. (B) Percentage of patients who achieved clinically significant improvement. n represents the number of patients included in each reporting period. Clinically significant improvement is defined as a $\geq 30\%$ reduction in average pain score⁴. Results are calculated for patients with moderate or severe average pain at referral.

4.3 DISCUSSION AND FUTURE DIRECTIONS

The ePPOC governance committee have acknowledged the complexity of achieving a clinically significant improvement in average pain and the subsequent challenge for services to meet the benchmark target. Some committee members suggested that the current benchmark target may feel unattainable, which can discourage services from focusing on improving outcomes in this area. Others suggested that success in improving average pain is largely influenced by baseline clinical and/or demographic factors, which may limit the extent to which services can meaningfully shift outcomes through program delivery alone. From this perspective, lowering the benchmark target was considered unlikely to result in meaningful improvements in service performance. It was subsequently suggested that if average pain is retained as an ePPOC benchmark, a spectrum of aspirational targets, adjusted for patient profile and complexity, may be more appropriate than a single target. This case-mix-adjusted approach is supported by research using ePPOC data, which has identified multiple patient and clinical factors associated with treatment outcomes. Tardif et al. (2023)⁶ found that sociodemographic and clinical factors, including unemployment, language barriers, widespread pain, long pain duration, and pain attributed to work injury, were associated with less favourable pain severity and interference outcomes.

Additionally, it is worth considering whether the average pain item provides the most accurate indicator of improvement in pain severity. Across the 6 reporting periods included in this review, average pain consistently had the lowest valid completion rate among the four BPI pain severity items (appendix C), which may reflect how challenging it can be for patients to estimate an “average” level of pain. This observation aligns with broader research that has shown that patients struggle to interpret the concept of average pain. For instance, when asked to report average pain over different time periods, patients consistently rate their pain higher when asked to consider the average over a 2-week period than over a 24 h period.⁷ Taken together, the low valid completion rate for average pain and the known challenges with interpreting the concept lend weight to the suggestion that alternative measures may provide a more reliable reflection of treatment impact.

Reduced pain is often a treatment priority for patients. Therefore, any change to the ePPOC average pain benchmark must ensure that the benchmarked domain remains meaningful to both consumers and clinicians. If the benchmark is to be modified or replaced, consumer consultation will be essential. Potential alternatives—such as worst pain, pain frequency, or composite indices like the BPI pain severity score, which combines the four pain severity items into a single measure—will be explored further with ePPOC member services at the 2025 Annual Benchmarking Workshop.

5. PAIN INTERFERENCE

SUMMARY: Given its stable patient-level outcomes, high prevalence at episode start and strong valid completion rate, the pain interference measure remains robust, and no changes are proposed to its use as a core ePPOC benchmark.

5.1 OVERVIEW AND DEFINITIONS

Pain interference is the second of the two core domains assessed by the BPI, following pain intensity. The pain interference assessment asks patients to rate how much pain has interfered with seven aspects of daily life over the past week, such as general activity, mood, and sleep, on a scale from 0 (does not interfere) to 10 (completely interferes). In the context of ePPOC benchmarking, pain interference outcomes are calculated only for the subset of patients that report moderate to severe pain interference at episode start, and a clinically significant improvement at episode end is defined as a change of 1 point over the average of the 7 items on the BPI interference scale, which is consistent with IMMPACT recommendations.⁴

ePPOC services are considered to have met the pain interference benchmark target if more than 70% of applicable patients make a clinically significant improvement between episode start and end.

5.2 BENCHMARK PERFORMANCE

The proportion of services achieving the pain interference benchmark rose steeply from 12.5% in the 2022MY report to 22% in the 2022CY report. However, this upward trend did not continue, with performance initially plateauing and then gradually declining to 17.7% in the 2024CY report (Figure 2A). Alternatively, at the patient level, pain interference outcomes have been more stable. Across the reporting periods reviewed, between 58.2% and 60.5% of patients with moderate to severe pain interference at referral reported a clinically significant improvement at episode end (Figure 2B).

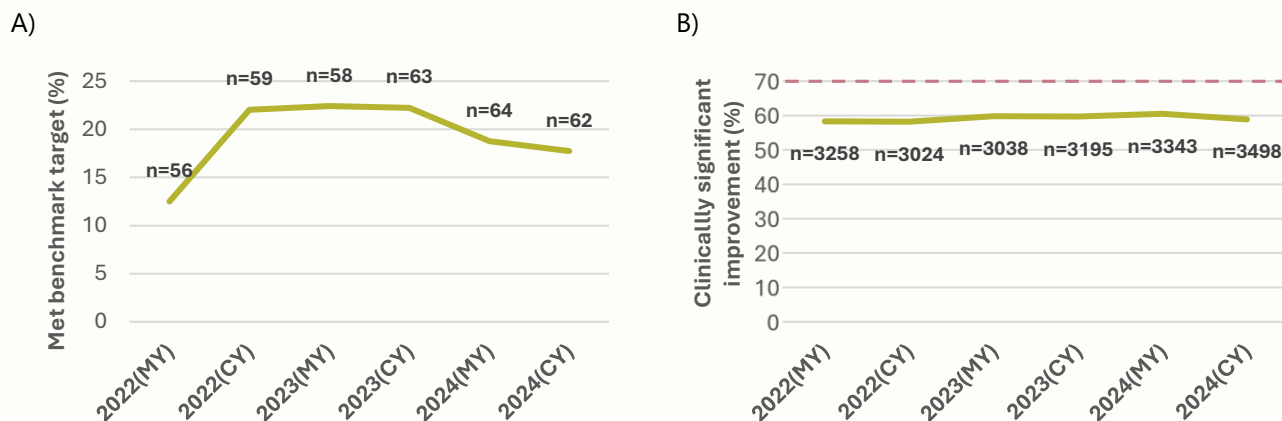


Figure 2 Pain interference benchmark performance and patient outcomes between the 2022MY and 2024CY reporting periods. (A) Percentage of services meeting the pain interference benchmark target of 70% of eligible patients achieving clinically significant improvement. Services must record outcomes for at least 10 eligible patients for a benchmark result to be calculated. n represents the number of services for whom a benchmark result was calculated. (B) Percentage of patients who achieved clinically significant improvement. n represents the number of patients included in each reporting period. Clinically significant improvement is defined as a ≥ 1 -point reduction in the average of the seven items on the BPI interference scale⁴. Results are calculated for patients with moderate or severe pain interference at referral.

5.3 DISCUSSION AND FUTURE DIRECTIONS

The importance of improved pain interference as a treatment goal is underscored by its high prevalence at episode start, with 87% of patients reporting moderate to severe pain interference at referral in the 2024CY period. Additionally, the high pain interference questionnaire completion rate (99% in the 2024CY period, Appendix C), also suggests that patients find the questions clear, easy to interpret, and relevant to their experience. Performance trends over recent reporting periods show stable outcomes, providing no indication of concern or need for change. Together, these data highlight the value of the pain interference measure and support its continued inclusion as a core ePPOC benchmark.

6. PAIN CATASTROPHISING

SUMMARY: There has been a recent decline in pain catastrophising outcomes, which is evident at both the service- and patient-level. ePPOC will continue to monitor this trend. This review also recommends removing the required change in severity category from the current definition of clinically significant improvement for this benchmark. Simplifying the definition of improvement to a reduction of 6 or more points on the PCS scale, regardless of start and end category, may result in more consistent assessment.

6.1 OVERVIEW AND DEFINITIONS

Pain catastrophising is measured using the Pain Catastrophising Scale (PCS), which assesses the extent to which patients have negative, exaggerated thoughts and feelings about their pain. The PCS asks patients to rate

13 statements related to rumination, magnification, and helplessness on a scale from 0 (not at all) to 4 (all the time).

Clinically significant improvement in pain catastrophising is currently defined as a reduction of 6 or more points on the PCS, accompanied by a shift to a lower severity category⁸. Pain catastrophising outcomes are calculated only for the subset of patients that report high or severe pain catastrophising at episode start.

The current ePPOC benchmark target for pain catastrophising is 70%.

6.2 BENCHMARK PERFORMANCE

The ePPOC pain catastrophising benchmark target (70%) is among the highest ePPOC benchmark targets, alongside pain interference and depression. This high aspirational target reflects the importance of reducing maladaptive pain-related thinking as a key treatment goal. However, the relatively low proportion of services meeting this benchmark, peaking at 27.5% in the 2023 MY report and declining to 14.8% in the 2024CY report, indicates that it is challenging for services to meet this BM (Figure 3A). At the patient level, a similar yet more subtle trend in results was observed across the reporting periods. The proportion of patients achieving clinically significant improvement peaked at 57.6% in the 2023MY report, but subsequently fell to 54.9% in the 2024CY report (Figure 3B). This performance trend will continue to be monitored by ePPOC.

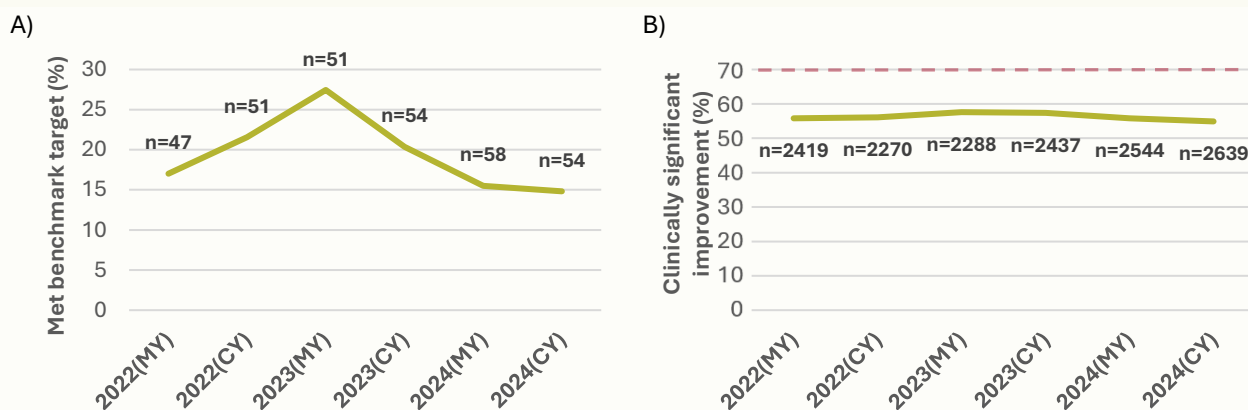


Figure 3 Pain catastrophising benchmark performance and patient outcomes between the 2022MY and 2024CY reporting periods. (A) Percentage of services meeting the pain catastrophising benchmark target of 70% of eligible patients achieving clinically significant improvement. Services must record outcomes for at least 10 eligible patients for a benchmark result to be calculated. n represents the number of services for whom a benchmark result was calculated. (B) Percentage of patients who achieved clinically significant improvement. n represents the number of patients included in each reporting period. Clinically significant improvement is defined as a ≥ 6 -point reduction in pain catastrophising score, with movement to a different severity category⁸. Results are calculated for patients with high or severe pain catastrophising at referral.

6.3 DISCUSSION AND FUTURE DIRECTIONS

The challenge in reaching the pain catastrophising benchmark may be compounded by the definition of clinically significant improvement that is applied to the domain. The current definition of a reduction of 6 or more points on the PCS accompanied by a shift to a lower severity category was established by consensus of the ePPOC governance committee in 2014⁸. It has since been suggested that the required change in severity category may limit the recognition of meaningful improvement because the categories themselves are based on arbitrary thresholds. For example, an improvement of 6 points likely represents a comparable degree of improvement, regardless of whether it corresponds to a change in severity category or not. This suggests some patients' improvement may not be captured under the current definition. These considerations raise important questions about how clinically significant improvement is defined and suggest that the current approach may warrant further discussion with services and revision.

7. PAIN SELF-EFFICACY

SUMMARY: Service- and patient-level outcomes show no consistent trends across the reporting periods assessed. No changes to the domain or benchmark target are proposed, however the definition of clinically significant improvement warrants revision. Specifically, the required change in severity category may limit recognition of meaningful improvement.

7.1 OVERVIEW AND DEFINITIONS

Pain self-efficacy is assessed using the Pain Self-Efficacy Questionnaire (PSEQ), which rates a patient's confidence in their ability to carry out daily activities, despite experiencing pain. The PSEQ includes 10 statements covering activities such as work, social engagement, and coping with pain without medication. Patients rate each item on a scale from 0 (not at all confident) to 6 (completely confident), providing an overall indication of how capable they feel in managing and functioning with their pain.

Clinically significant improvement in pain self-efficacy is defined as a change in score of ≥ 7 points, with movement to a different severity category⁹, with results calculated for the sub-group that report moderate or severely impaired self-efficacy at referral.

The ePPOC benchmark target for pain self-efficacy is 60%.

7.2 BENCHMARK PERFORMANCE

Across the six reporting periods included in this review (2022MY to 2024CY), pain self-efficacy results have remained relatively stable at both the service- and patient-level, with the proportion of services meeting the benchmark ranging from 21.6–28.6% (Figure 4A), and the proportion of patients achieving clinically significant improvement ranging from 48.1–50.7% (Figure 4B). In the 2024CY reporting period specifically, 57 services captured sufficient data to calculate a benchmark result; of these, 24.6% met the benchmark. At the patient level in the same reporting period, 48.7% of the 2,985 patients with moderate or severe impairment at referral recorded clinically significant improvement at episode end.

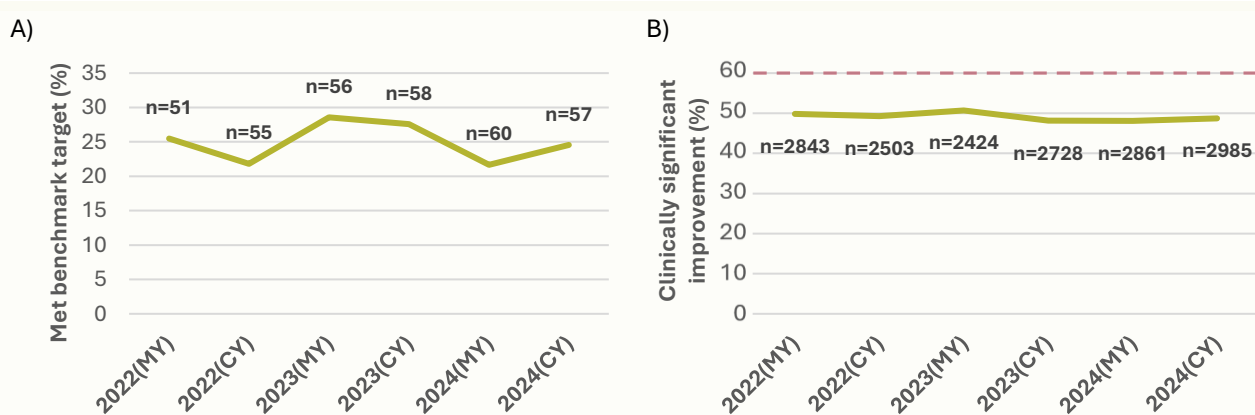


Figure 4 Pain self-efficacy benchmark performance and patient outcomes between the 2022MY and 2024CY reporting periods.

(A) Percentage of services meeting the pain self-efficacy benchmark target of 60% of eligible patients achieving clinically significant improvement. Services must record outcomes for at least 10 eligible patients for a benchmark result to be calculated. n represents the number of services for whom a benchmark result was calculated. (B) Percentage of patients who achieved clinically significant improvement. n represents the number of patients included in each reporting period.

Clinically significant improvement is defined as a ≥ 7 -point increase in pain self-efficacy score, with movement to a different severity category⁹. Results are calculated for patients with moderate or severe impairment in pain self-efficacy at referral.

7.3 DISCUSSION AND FUTURE DIRECTIONS

Overall service- and patient-level results remain strong, and the current level of about 25% of services meeting the target could be viewed as representing ‘good practice’, setting an aspirational yet achievable goal for services. The benchmark domain and its 60% target are therefore proposed to continue unchanged. However, the definition of clinically significant improvement warrants revision. As with pain catastrophising, requiring both a PSEQ score increase of seven or more, as well as a shift to a higher severity category, may miss meaningful raw-score gains when patients do not cross a category boundary. Updating this definition would better capture clinically important change while preserving the benchmark’s relevance. These proposals, including the concept of defining ‘good practice’, will be explored further with ePPOC member services at the 2025 Annual Benchmarking Workshop.

8. DEPRESSION, ANXIETY AND STRESS

SUMMARY: Psychological distress is assessed in ePPOC using the DASS-21, which includes subscales for depression, anxiety, and stress. Clinically significant improvement is defined as a ≥ 5 -point reduction in subscale score, accompanied by a shift to a less severe category. As with pain self-efficacy and catastrophising, this categorical requirement may warrant revision.

Among the sub-domains, anxiety and stress warrant close, ongoing monitoring and more detailed statistical assessment. Anxiety results show a consistent downward trend at both service and patient levels, while stress results have declined more recently. In contrast, while depression outcomes show no clear trend, the proportion of services meeting this benchmark is low. Feedback from services will be sought at the 2025 Benchmarking Workshop to inform whether the depression benchmark target should be reduced to a more achievable level.

8.1 OVERVIEW AND DEFINITIONS

Psychological distress is assessed by ePPOC using the DASS-21, which comprises three seven-item subscales measuring depression, anxiety, and stress. Patients rate how much each statement applied to them over the past week on a 4-point scale from 0 (did not apply to me at all) to 3 (applied to me very much or most of the time).

For ePPOC benchmarking, clinically significant improvement in each of the sub-domains is calculated for the subset of patients that report moderate to extremely severe symptoms at referral and defined as an improvement in sub-score of five or more points, accompanied by movement to a less severe category. This definition was established through consensus of the ePPOC governance committee in 2014. As with other ePPOC benchmark domains that require a category shift to qualify as improvement, these definitions of clinically significant change will be discussed at the upcoming benchmarking workshop to ensure they do not underestimate meaningful gains.

The ePPOC benchmark targets for depression, anxiety and stress are 70%, 50% and 60%, respectively.

8.2 DEPRESSION

The depression benchmark has consistently been one of the most difficult ePPOC targets to achieve. Between 2022 and 2024, the proportion of services meeting the benchmark target of 70% ranged from 10.7% (2024MY) to 20.0% (2023CY) (Figure 6A). At the individual patient level, the proportion of individuals that recorded a clinically significant improvement by episode end ranged from 53.6% (2022CY) to 56.1% (2024CY) (Figure 6B). While this is a substantial proportion of patients, it falls well below the 70% target.

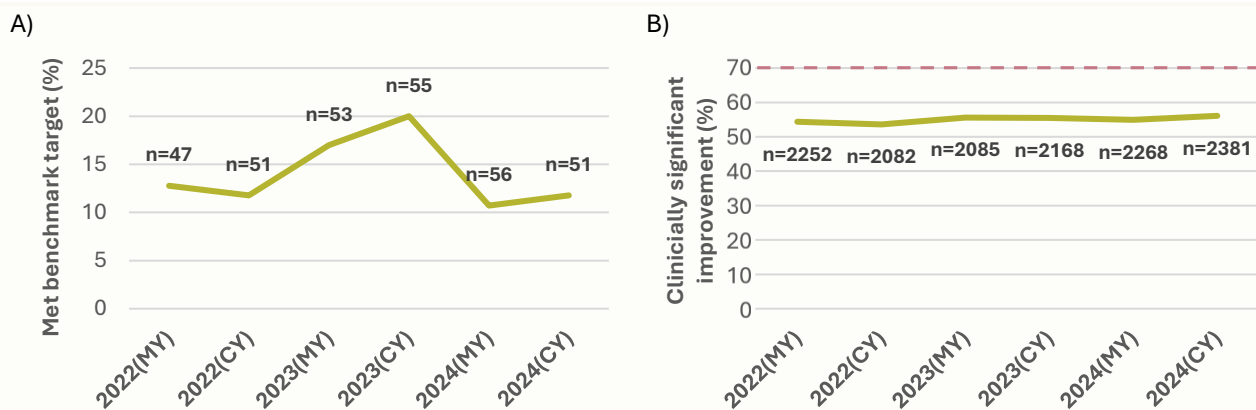


Figure 6 Depression benchmark performance and patient outcomes between the 2022MY and 2024CY reporting periods. (A) Percentage of services meeting the depression benchmark target of 70% of eligible patients achieving clinically significant improvement. Services must record outcomes for at least 10 eligible patients for a benchmark result to be calculated. *n* represents the number of services for whom a benchmark result was calculated. (B) Percentage of patients who achieved clinically significant improvement. *n* represents the number of patients included in each reporting period. Clinically significant improvement is defined as a ≥ 5 -point reduction in depression score on the DASS depression subscale, with movement to a different severity category. Results are calculated for patients with moderate, severe, or extremely severe depression at referral.

Given the consistently low proportion of services meeting the current target, and the limitations of the existing definition of clinically significant improvement, there is a strong case for revision in this area. Removing the required categorical shift from the definition of clinically significant improvement may help capture additional meaningful reductions in depression scores. Additionally, feedback will be sought from ePPOC member services at the 2025 Benchmarking Workshop regarding lowering the benchmark target to a more attainable level.

8.3 ANXIETY

This review identifies a downward trend in performance, both in the proportion of services meeting the 50% anxiety benchmark target and in the percentage of patients achieving clinically significant improvement in this domain. The proportion of services achieving the benchmark declined from 35.7% in the 2022MY report to 22.0% in the 2024CY report, reaching a low of 17.0% in the 2024MY report (Figure 7A). Similarly, the percentage of patients who reported clinically significant improvement in the domain fell from 45.1% to 40.4% over the same period (Figure 7B).

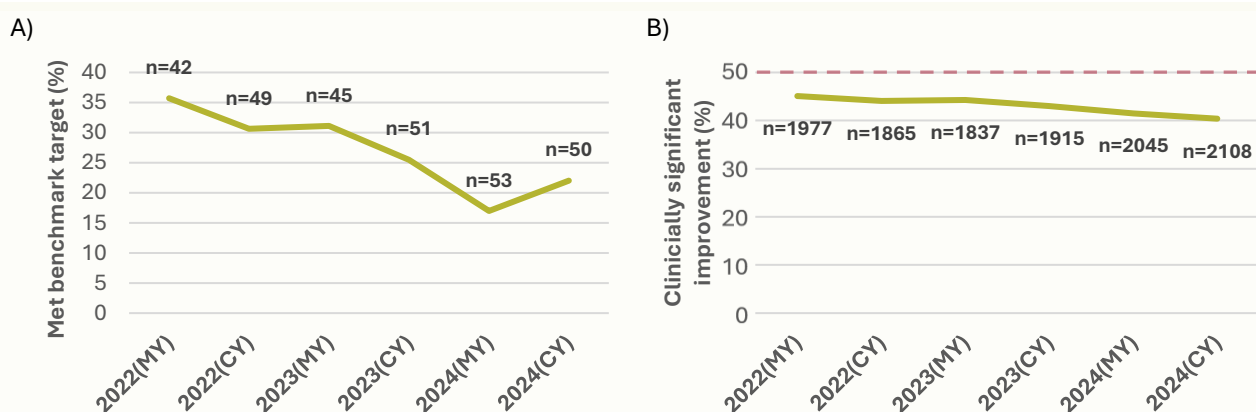


Figure 7 Anxiety benchmark performance and patient outcomes across the 2022MY and 2024CY reporting periods. (A) Percentage of services meeting the anxiety benchmark target of 50% of eligible patients achieving clinically significant improvement. Services must record outcomes for at least 10 eligible patients for a benchmark result to be calculated. *n* represents the number of services for whom a benchmark result was calculated. (B) Percentage of patients who achieved clinically significant improvement. *n* represents the number of patients included in each reporting period. Clinically significant improvement is defined as a ≥ 5 -point reduction in anxiety score on the DASS anxiety subscale, with movement to a different severity category. Results are calculated for patients with moderate, severe, or extremely severe anxiety at referral.

To explore possible reasons for the decline in outcomes, we examined both baseline anxiety scores and the proportion of patients presenting with moderate to extremely severe anxiety at referral, as higher initial severity can influence both the likelihood and the magnitude of improvement. Additionally, population data show that anxiety rates have risen in recent years and were affected by global events such as the COVID-19 pandemic¹⁰, so we considered whether similar shifts might be evident in our cohort. We found that average anxiety referral scores remained relatively stable across the reporting period (Appendix D), and the proportion of patients in the moderate-to-extremely-severe category varied only slightly, from 52 % to 54 %. This suggests that changes in baseline severity are unlikely to account for the reduced treatment gains.

A more detailed statistical analysis of the data is required to explore the impact of other potentially relevant baseline patient characteristics and to assess whether the decline reflects an overall reduction in service performance or is driven by a specific subset of services. Identifying such variations will enable targeted support and training to be directed where it is most needed. Additionally, ePPOC will prioritise anxiety management as a focus area for additional education, aiming to equip all ePPOC member services with enhanced strategies to improve outcomes in this domain. At this stage, no change to the anxiety benchmark is recommended. The results will continue to be monitored closely in upcoming reporting periods to determine whether the decline persists.

8.4 STRESS

The stress benchmark has historically been one of the most attainable of the ePPOC benchmarks. In the 2023CY reporting period, 64% of services met the benchmark target of 60% of patients achieving clinically significant improvement. At the time, this prompted discussion about whether the benchmark threshold should be raised to maintain its aspirational intent. However, subsequent data show a decline in the proportion of services reaching the target, reducing to 34.1% in the 2024CY report (Figure 5A). This downturn suggests that the 2023CY result may have been an outlier rather than evidence of a sustained improvement.

A similar trend is evident in the patient-level outcomes. The proportion of patients reporting clinically significant improvement peaked at 60.2% in the 2023CY report, but subsequently reduced to 56.8% in the 2024CY report (Figure 5B). While these strong results highlight the meaningful impact of pain management interventions on reducing stress symptoms, given this recent decline at both the service and patient level, raising the benchmark target at this stage is not recommended. The current 60% threshold is proposed to remain in place to ensure it remains both aspirational and achievable, however ongoing monitoring of this domain will help determine whether the recent decline persists.

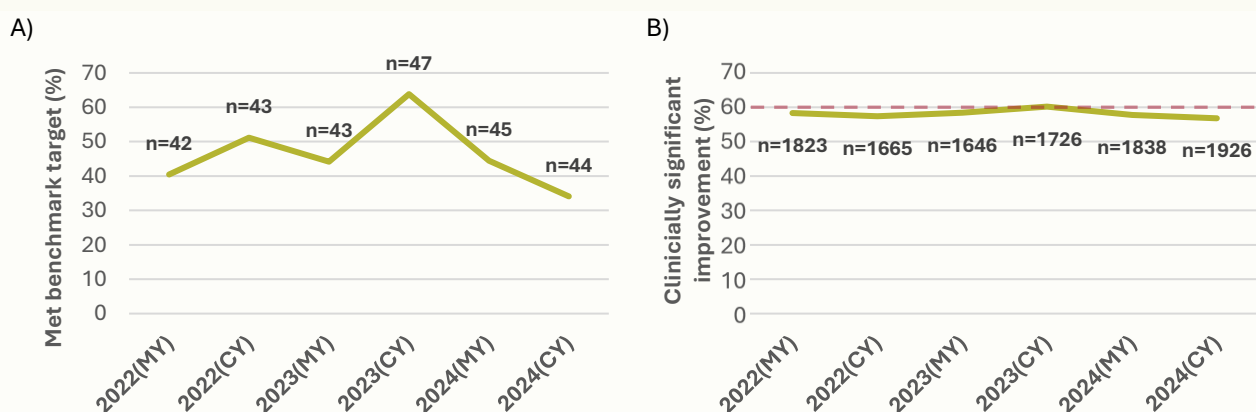


Figure 5 Stress benchmark performance and patient outcomes across the 2022MY and 2024CY reporting periods. (A) Percentage of services meeting the stress benchmark target of 60% of eligible patients achieving clinically significant improvement. Services must record outcomes for at least 10 eligible patients for a benchmark result to be calculated. n represents the number of services for whom a benchmark result was calculated. (B) Percentage of patients who achieved clinically significant improvement. n represents the number of patients included in each reporting period. Clinically significant improvement is defined as a ≥ 5 -point reduction in stress score, with movement to a different severity category. Results are calculated for patients with moderate, severe, or extremely severe stress at referral.

9. CONCLUSIONS AND NEXT STEPS

This analysis provides a high-level overview of ePPOC member service performance against seven adult benchmarks, alongside patient-level rates of clinically significant improvement. Its purpose is to guide discussion about areas of benchmarking and aspects of the ePPOC dataset that may benefit from refinement or further investigation.

Further statistical analysis of benchmarking data may provide additional insights into the attainability and fairness of current targets. Examining systemic or patient-factors, such as service type or baseline severity, may clarify why some benchmarks are more difficult to achieve and if they should be uniformly applied across services. The ePPOC governance committee have also emphasised the importance of interpreting benchmark results in context. Variations in patient populations, such as differences in pain duration or compensable status, can significantly influence outcomes. Research, including Tardif et al. (2023)⁶, supports these factors predictors of treatment response. Incorporating case-mix adjustment or stratified reporting could enhance fairness in service comparisons and improve accuracy in estimating patient outcomes.

The 2025 ePPOC Australasian Benchmarking Workshop, scheduled for 12 November, will provide a forum to discuss these findings with member services. Structured discussions will focus on:

1. Domain-specific discussion

- a. Average pain
- b. Anxiety

2. Methodological discussion

- a. Review of the original benchmark-setting criteria and the definition of 'good practice'
- b. Review of the definitions of clinically significant improvement

Ongoing engagement from member services is essential to ensuring the benchmarks remain relevant and reflective of the diversity of pain services across the network. Insights gained from over a decade of data collection have laid a strong foundation for refining the benchmarking framework. Discussion at the 2025 Benchmarking Workshop will guide the finalisation of this review, ensuring it reflects the collective insights and priorities of the network.

10. REFERENCES

1. Ettorchi - Tardy A, Levif M, Michel P. Benchmarking: A method for continuous quality improvement in health. *Healthc Policy*. 2012;7(4):101–119.
2. Willmington C, Belardi P, Murante AM, Vainieri M. The contribution of benchmarking to quality improvement in healthcare. A systematic literature review. *BMC Health Serv Res*. 2022 Feb 2;22(1):139. doi: 10.1186/s12913-022-07467-8. PMID: 35109824; PMCID: PMC8812166.
3. Holloway D, Allingham S, Bryce M, Cameron K, Cook M, Shebeshi D. A decade of outcomes: The evolution of an australasian outcomes collaboration for chronic pain services. *Front Pain Res (Lausanne)*. 2023 Apr 12;4:1153001. doi: 10.3389/fpain.2023.1153001. PMID: 37139341; PMCID: PMC10150651.
4. Dworkin, RH, et al 2008, 'Interpreting the Clinical Importance of Treatment Outcomes in Chronic Pain Clinical Trials: IMMPACT Recommendations.' *The Journal of Pain*, vol. 9, no. 2, pp 105-121.
5. Heenan MA, Randall GE, Evans JM. Selecting Performance Indicators and Targets in Health Care: An International Scoping Review and Standardized Process Framework. *Risk Manag Healthc Policy*. 2022;15:747-764 <https://doi.org/10.2147/RMHP.S357561>
6. Tardif H, Allingham SF, Rahman M, Daly A. Patterns of patient outcomes following specialist pain management in Australasia: a latent class analysis using the electronic Persistent Pain Outcomes Collaboration database. *Pain*. 2023 May 1;164(5):967-976. doi: 10.1097/j.pain.0000000000002799. Epub 2022 Oct 18. PMID: 36448970.
7. Dorfman, David, George, Mary Catherine, Robinson-Papp, Jessica, Rahman, Tanni, Tamler, Ronald and Simpson, David M.. "Patient reported outcome measures of pain intensity: Do they tell us what we need to know?" *Scandinavian Journal of Pain*, vol. 11, no. 1, 2016, pp. 73-76. <https://doi.org/10.1016/j.sjpain.2015.12.004>
8. Sullivan, MJL, personal communication with Nicholas, MK July 2014.
9. Nicholas, MK, personal communication, July 2014.
10. Amini-Rarani, M., Azami-Aghdash, S., Mousavi Isfahani, H. *et al*. Estimation of the prevalence of anxiety during the COVID-19 pandemic: A meta-analysis of meta-analyses. *BMC Public Health* **24**, 2831 (2024). <https://doi.org/10.1186/s12889-024-19729-7>

APPENDIX A: ASSESSMENT TOOLS

The assessment tools used in ePPOC are:

- Brief Pain Inventory (BPI)
- Depression, Anxiety, Stress Scale (DASS)
- Pain Catastrophising Scale (PCS)
- Pain Self-Efficacy Questionnaire (PSEQ)
- ePPOC Patient Impression of Change (formerly referred to as the Global Rating of Change (GRC))
- Work Productivity and Impairment (WPAI)
- CARRA Body Chart.

Each of these assessment tools are briefly described below.

BRIEF PAIN INVENTORY¹

The BPI items used in the ePPOC dataset measure the severity of pain and the degree to which the pain interferes with common activities of daily living. Pain severity questions are rated on a scale of 0 to 10, where 0 = 'No pain' and 10 = 'Pain as bad as you can imagine', with patients asked to rate their pain in four items including, average, worst and least pain over the last week, and their pain right now. Pain severity is then calculated as an average of these four items. Whereas the benchmark for average pain is based on the single average pain item only.

The IMMPACT group's recommendations for assessing clinical significance for 0-10 numeric pain scales are that a change of:

≥ **10%** represents minimally important change

≥ **30%** represents moderate clinically important change

≥ **50%** represents substantial clinically important change.

The interference questions are rated on a scale of 0 to 10, where 0 = 'Does not interfere' and 10 = 'Completely interferes'. The interference subscale is an average of the seven interference questions. At least 4 of 7 questions must be completed for this subscale to be valid. The IMMPACT recommendation for assessment of clinically significant change on the BPI interference scale is a change of 1 point over the average of the 7 items².

Updated BPI severity bands

From the 2022 mid-year report onwards, ePPOC had implemented a modified version of the previously reported severity bands: 0 = none | 1-4 = mild | 5-6 = moderate | 7-10 = severe. A score of 0 – no pain (or 0 – no interference) was previously included in the mild category, but is now reported separately³.

The above severity bands can be applied directly to the individual items reported on the integer scales (e.g. the 'average pain' item). When applied to the sub-scale scores for pain severity and interference, the mean values are rounded to the closest integer before the severity ranges are applied. The 'none' category should be interpreted as 'no or negligible pain' or 'no or negligible interference' for the pain severity and pain interference subscales respectively.

¹ Modified Brief Pain Inventory, reproduced with acknowledgement of the Pain Research Group, University of Texas, MD Anderson Cancer Centre, USA

² Dworkin, RH, et al 2008, 'Interpreting the Clinical Importance of Treatment Outcomes in Chronic Pain Clinical Trials: IMMPACT Recommendations.' *The Journal of Pain*, vol. 9, no. 2, pp 105-121.

³ Palos, GR, et al 2006, 'Asking the community about cutpoints used to describe mild, moderate, and severe pain.' *Journal of Pain*, vol. 7, no. 1, pp 49-56

DEPRESSION ANXIETY STRESS SCALES⁴

The DASS measures the negative emotional states of depression, anxiety and stress. Due to the large number of questions in the full DASS (42 questions), the DASS21 (21 questions) is administered. The questions are rated on a scale of 0 to 3, where 0 = 'did not apply to me at all', 1 = 'applied to me to some degree, or some of the time', 2 = 'applied to me to a considerable degree, or a good part of the time', or 3 = 'applied to me very much, or most of the time'. Scores are multiplied by 2 to enable comparison with the full-scale DASS42 for which norms exist.

For each subscale (Depression, Anxiety and Stress), the 7 items are summed and then multiplied by 2. The test developers suggest that at least 6 of 7 items should be complete for each subscale to be considered valid. Table 1 shows the range of scores associated with severity categories for each subscale.

| Table 1 – DASS severity ratings | Depression | Anxiety | Stress |
|---------------------------------|------------|---------|--------|
| Normal | 0-9 | 0-7 | 0-14 |
| Mild | 10-13 | 8-9 | 15-18 |
| Moderate | 14-20 | 10-14 | 19-25 |
| Severe | 21-27 | 15-19 | 26-33 |
| Extremely Severe | 28+ | 20+ | 34+ |

Clinical significance on each of the DASS subscales requires a change of 5 or more points coupled with a move to a different severity category.

PAIN CATASTROPHISING SCALE⁵

The PCS measures a patient's thoughts and feelings related to their pain. This includes three subscales measuring the dimensions of Rumination, Magnification and Helplessness. The PCS comprises 13 questions (Rumination – 4 items, Magnification – 3 items, Helplessness – 6 items) which are rated on a scale of 0 to 4, where 0 = 'not at all', 1 = 'to a slight degree', 2 = 'to a moderate degree', 3 = 'to a great degree' and 4 = 'all the time'. For each subscale, all items must be completed to be valid. For the total to be valid, at least 12 of 13 items must be completed.

Severity bands for the PCS are:

- <20 = mild
- 20 to 30 = high
- >30 = severe.

Clinically significant change requires a change in score of 6 or more points, combined with movement to a different severity category⁶.

⁴ Lovibond, SH and Lovibond, PF 1995, *Manual for the Depression Anxiety Stress Scales*, Psychology Foundation Monograph, Sydney, Australia.

⁵ Sullivan, MJL, et al 1995, 'The Pain Catastrophizing Scale: Development and Validation', *Psychological Assessment*, vol. 7, num. 4, pp 524-532.

⁶ Sullivan, MJL, personal communication with Nicholas, MK July 2014.

PAIN SELF-EFFICACY QUESTIONNAIRE

The PSEQ measures how confident a patient is that he or she can do a range of activities despite their pain. The PSEQ Total is a sum of scores from 10 questions which are rated on a scale from 0 = 'Not confident at all' to 6 = 'Completely confident'. At least 9 of 10 items must be complete for the PSEQ Total to be valid. Increases in score represent an improvement in self-efficacy.

Severity bands for the PSEQ are:

- <20 = severe
- 20 to 30 = moderate
- 31 to 40 = mild
- >40 = minimal impairment.

Clinically significant change requires a change in score of 7 or more points, combined with movement to a different severity category⁷.

⁷ Nicholas, MK, personal communication, July 2014.

APPENDIX B: CRITERIA FOR DEVELOPING EPPOC BENCHMARKS

CRITERIA

The ePPOC SCAC agreed to accept 6 of the 7 suggested criteria

Table.

Table B1. Criteria for developing ePPOC benchmarks

| No. | Criteria | Accept/Reject |
|-----|------------------------------------------------------------------------------------------------------|-------------------|
| 1. | The domain is clinically important | Accepted |
| 2. | <i>There is sufficient variation between services and patients</i> | <i>Rejected *</i> |
| 3. | There is evidence in the literature demonstrating that the domain is amenable to intervention | Accepted |
| 4. | The benchmark threshold reflects good practice rather than average practice | Accepted |
| 5. | The benchmark is the same in all services regardless of public, private, tier or role delineation | Accepted |
| 6. | Clinical benchmarks should include patients whose baseline score reflects at least moderate severity | Accepted |
| 7. | There is sufficient data to effectively benchmark | Accepted |

* Outcomes for an important domain may be poor across all services

DOMAINS

The SCAC agreed on the domains to benchmark, shown below.

Table B2. List of domains for benchmarks

| No. | Domain | Comments |
|-----|----------------------|--------------------------------------------------------------|
| 1. | Pain Severity | Agreed |
| 2. | Pain interference | Agreed |
| 3. | Depression | Agreed |
| 4. | Anxiety | Agreed |
| 5. | Stress | Deleted following discussion of individual benchmarks |
| 6. | Pain catastrophising | Agreed |
| 7. | Pain self-efficacy | Agreed |
| 8. | Opioid use | Agreed |
| 9. | Waiting time | Agreed |

APPENDIX C: ASSESSMENT TOOL VALIDITY AND COMPLETION

Table C1. Assessment tool validity and completion results for the 2024 calendar year report.

| | Validity – number of items that must be completed | All Adult Services | |
|------------------------------------------|---------------------------------------------------|-----------------------------------|---------------------------------------|
| | | Average number of completed items | % of validly completed questionnaires |
| Brief Pain Inventory | | | |
| - Worst pain | 1/1 | - | 98.6 |
| - Least pain | 1/1 | - | 97.5 |
| - Average pain | 1/1 | - | 96.8 |
| - Pain now | 1/1 | - | 97.1 |
| Pain severity | 4/4 | - | 94.4 |
| Pain interference | 4/7 | 6.9 | 99.0 |
| Depression, Anxiety, Stress Scale | | | |
| - Depression | 6/7 | 6.8 | 97.5 |
| - Anxiety | 6/7 | 6.8 | 97.5 |
| - Stress | 6/7 | 6.8 | 97.4 |
| Pain Catastrophising Scale | | | |
| - Rumination | 4/4 | 3.9 | 97.1 |
| - Magnification | 3/3 | 2.9 | 96.7 |
| - Helplessness | 6/6 | 5.9 | 95.6 |
| - Total | 12/13 | 12.7 | 97.2 |
| Pain Self-Efficacy Questionnaire | 9/10 | 9.8 | 97.7 |
| Global Rating of Change | | | |
| - Overall | 1/1 | - | 95.7 |
| - Physical | 1/1 | - | 93.6 |

APPENDIX D: MEAN DAS-21 SUB-SCORES

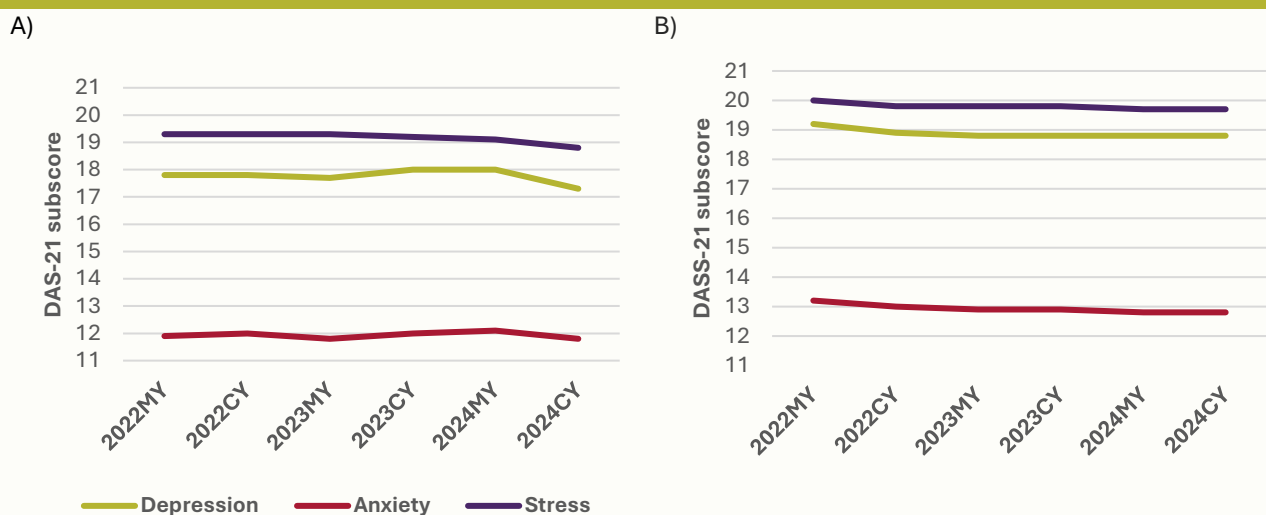


Figure D1 Mean DASS-21 subscale scores for depression, anxiety, and stress at referral between the 2022MY and 2024CY reporting periods.

Scores represent the average severity reported by patients at the start of their episode of care. A) Results are based on patients who completed both referral and end-of-episode questionnaires. B) Results based on patients who completed a referral questionnaire.

