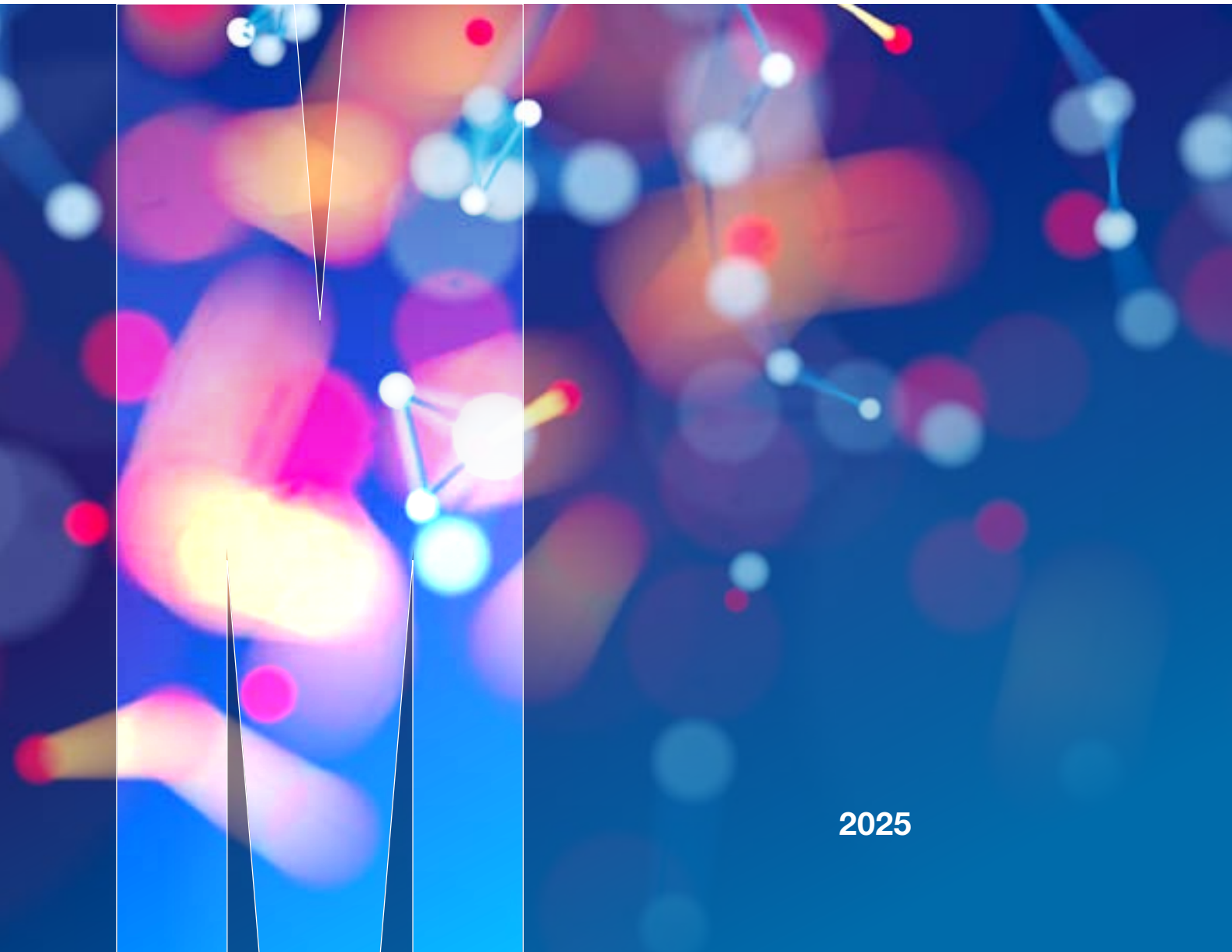
MONASH University

MONASH
PUBLIC HEALTH AND
PREVENTIVE MEDICINE

Using Patient Reported Measures in Clinical Quality Registries for Healthcare Improvement: A Guide



2025

Publication Details

Publication Title: Using Patient Reported Measures in Clinical Quality Registries for Healthcare Improvement: A Guide
Published: 2025
Publisher: Monash University
Doi: 10.26180/28801313
Suggested citation: Using Patient Reported Measures in Clinical Quality Registries for Healthcare Improvement: A Guide. Monash University. 2025, Version 1.

Copyright

Attribution

Use of Images

Contact Us

For further information or to provide feedback about the Guide, please contact:
E: med-clinicalregistries@monash.edu

Contents

Acronyms

ABDR	Australian Breast Device Registry
ACSQHC	Australian Commission of Safety and Quality in Health Care
ANOVA	Analysis of Variance
AOANJRR	Australian Orthopaedic Association National Joint Replacement Registry
APFQ	Australian Pelvic Floor Questionnaire
APFPR	Australasian Pelvic Floor Procedure Registry
CALD	Culturally and Linguistically Diverse
CQR	Clinical Quality Registry
EORTC	European Organization for Research and Treatment of Cancer
EQ-5D	EuroQol Group Five Dimension Measure of Quality of Life
EQ-5D VAS	EuroQol Group Visual Analogue Scale
HRQoL	Health-Related Quality of Life
ICHOM	International Consortium for Health Outcomes Measurement
ISOQOL	International Society for Quality of Life
ITS	Interrupted Time Series
MAR	Data Missing at Random
MCAR	Data Missing Completely at Random
MCID	Minimum Clinically Important Difference
MDC	Minimum Detectable Change
MNAR	Data Missing Not at Random
ODI	Oswestry Disability Index
POP	Pelvic Organ Prolapse
PREMs	Patient Reported Experience Measures
PRMs	Patient Reported Measures
PRO	Patient Reported Outcomes
PROMs	Patient Reported Outcome Measures
PROMIS	Patient-Reported Outcomes Measurement Information System
QALYs	Quality-Adjusted Life Years
QLQ-30	Quality of Life Questionnaire
QLQ-LC13	Quality of Life Questionnaire Lung Cancer Module
SF-36	Short Form Questionnaire
SPCC	Statistical Process Control Charts
SUI	Stress Urinary Incontinence
WHOQOL	World Health Organization Quality of Life

Foreword

Australian Government Department of Health, Disability and Ageing Preface

The National Clinical Quality Registry Program (the Program) aims to improve the quality of health care and ensure better health outcomes for Australian patients.

As part of this Program, the Department is leading a range of activities under the National Strategy for Clinical Quality Registries and Virtual Registries 2020-2030 (the Strategy).

Increasing use of patient-reported outcome and experience measures in national Clinical Quality Registries is a key priority of the Strategy. These measures tell us about people's health outcomes and quality of life post treatment, and whether our health care system is responding to the preferences, needs and values of patients.

The 'Using Patient Reported Measures in Clinical Quality Registries for Healthcare Improvement: A Guide' provides – for the first time – a set of national principles for how Clinical Quality Registries can best report and use patient reported data for health care quality improvement.

This guide complements the 'Reporting Patient Reported Measures in Clinical Quality Registries to Consumers: A Guide' and forms part of a suite of best practice materials being developed under the Program.

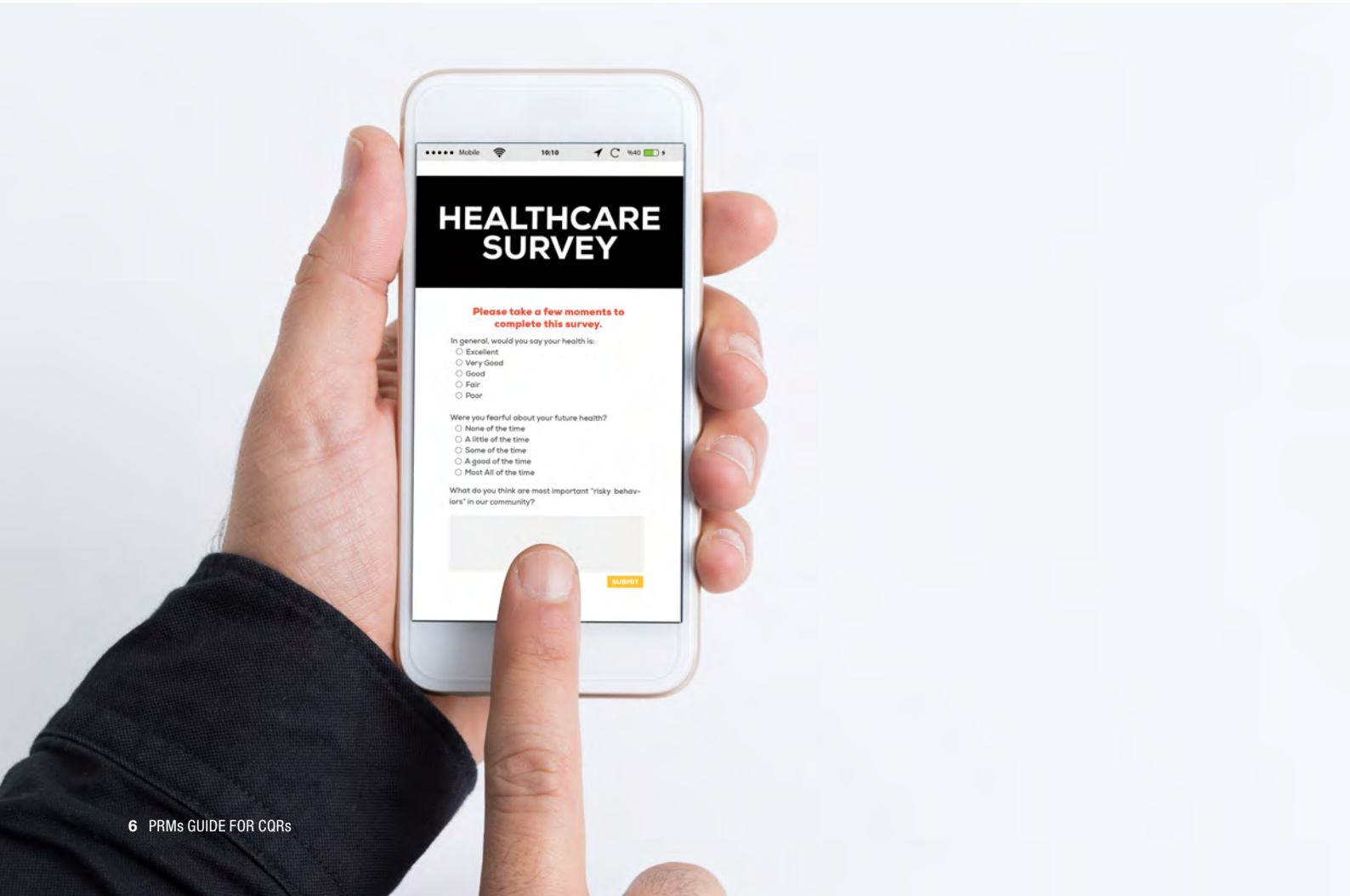
We thank Monash University for partnering with us on this important initiative.

Andrew Lalor
Assistant Secretary
Health Modelling, Partnerships and Evaluation Branch
Date: 24/07/2025

Introduction

Clinical registries are defined as systematically collected databases of health-specific information that are managed by operational and research teams and designed to monitor patient care and outcomes[1]. They can operate in different jurisdictions and have local, state, national or international level reach, and population coverage[2]. The term **"clinical quality registry" (CQR)** is growing in popularity as the focus shifts towards using registry data to improve the quality of care[3, 4]. CQRs monitor the appropriateness and effectiveness of healthcare by collecting clinical care and outcome information in relation to individuals who undergo a particular procedure, are diagnosed with a particular disease, or use a particular healthcare service.

Traditionally, registries report clinician-collected data back to health services to inform quality improvement (QI). However, patient reported measures (PRMs) are increasingly being introduced into CQRs in Australia, where they can be used to monitor outcomes from both acute and chronic conditions and clinical interventions. Yet, limited published evidence indicates that reporting of CQR-collected PRMs to health services for QI is highly variable in its extent and nature. Patient reported outcome and experience measures are key components of patient-centred care that can underpin CQR monitoring and QI efforts. These data provide important information about how healthcare systems are performing from the perspective of the people accessing care.



Purpose and Target Audience of the Guide

This Guide aims to facilitate standardised approaches to PRM data reporting across registries.

This Guide has been designed to provide a set of guiding principles and recommendations primarily for CQR operators, but may also be relevant for clinicians and health service managers for the planning and collection of PRM data for QI purposes, including analysis, reporting and interpretation of hospital or clinician-level PRMs from CQRs.

A separate resource will be available for reporting PRM data to consumers and patients.

Target audience for the Guide:

- Clinicians and health service managers who are involved in PRMs data collection for QI purposes;
- Registry staff who are involved in PRMs data analysis and reporting;
- Healthcare quality and safety roles and professionals, hospital managers, researchers;
- Staff working in clinical registries that are not CQRs (e.g., research only focused clinical registries) and are interested in PRMs data collection and reporting.

Introduction to PRMs

Patient Reported Outcome Measures

Patient Reported Outcome Measures (PROMs) are designed to assess various dimensions of a person's health and well-being from the perspective of the individuals themselves. Beyond assessing treatment effectiveness and the monitoring of well-being in the context of clinical trials and other research activities, PROMs have been used in clinical practice, supporting patient-centred care and shared clinical decision making[5], and more recently in CQRs.

PROMs are most often self-completed questionnaires. PROMs may include instruments that measure functional status, health-related quality of life (HRQoL), symptoms and symptom burden, and health-related behavioural conditions such as anxiety and depression[6].

PROMs can be either **generic or condition-specific**[7].

Generic PROMs (e.g., EQ-5D, SF-36, PROMIS) are designed to assess general aspects of health that are not specific to a particular disease. They usually focus on general person's well-being, mental health and/or HRQoL. Some generic PROMs, such as the EQ-5D, enable cost-effectiveness analysis to examine the cost of a health-related intervention and the benefit it produces in terms of the number of years lived in full health. As a short instrument, the EQ-5D (www.euroqol.org) is very popular amongst clinical registries and it is often used alongside a condition-specific PROM for each clinical area[8]. Generic PROMs also allow comparison of impact on individual well-being influenced by different diseases and conditions.

Condition-specific PROMs are designed to identify specific symptoms arising from specific conditions or diseases, and their impact on functional status. Condition-specific PROMs often have greater clinical utility than generic PROMs, but are not intended for comparisons across a variety of conditions. Condition-specific PROMs capture elements of health relevant to a particular patient group or condition. For example, the European Organisation for Research and Treatment of Cancer (EORTC)[9] developed widely used PROMs that assess the HRQoL of patients with cancer (EORTC QLQ-30). There is a core generic questionnaire for all patients with cancer, as well as modules targeting symptoms and outcomes of different cancer diagnoses (e.g., the Lung Cancer Module EORTC QLQ-LC13[10]).

Condition-specific and generic PROMs are both important for understanding and improving patient care at multiple levels of the health-care systems. Often, generic and condition-specific PROMs are used in combination as they can provide complementary information[11].

PROMs are also categorised as **profile or preference-based**.

Profile measures (e.g., SF-36, WHOQOL, PROMIS) are used to determine the position of some characteristic on specific domains, which are measured by multiple items and reported by domain[12]. Profile measures are useful where the aim is to assess health or an aspect of it from the patient's perspective in the evaluation of treatments or services[11].

Preference-based measures (e.g., EQ-5D, Health Utilities Index) are increasingly used in health economic evaluations to calculate quality-adjusted life years (QALYs). Such measures usually comprise a number of domains (or descriptive set) that patients can use to describe various aspects of their health (e.g., limitations in daily activities and mobility, pain and discomfort). These patient reported values (profile scores) are then converted to an index score using an algorithm or scoring weights, which can be specific to a particular country. The index scores (also referred to as 'utilities') usually range between 0 and 1, where 1 is usually taken to reflect a valuation of "perfect health" and 0 refers to valuation of "death". In some of these measures values below zero may be possible, representing health states perceived to be worse than death.

Proxy Measures

A major challenge for the assessment of patient outcomes is how to collect this information for people who are unable to self-report their outcomes due to cognitive or physical limitations. Examples include cognitive or linguistic impairment that inhibits comprehension of items, self-awareness or self-expression, or symptom burden and clinical deterioration in terminal illness[13, 14]. Some aspects of patient health may be assessable via clinician observation or performance-based measurement, but others require self-report or proxy report. **Proxy reports** are often used when patients are unable to self-report or where instruments are not available in the patient's native language. These measures are obtained from a patient's caregiver, guardian, or other individual who can speak to the patient's behaviours and symptoms. In some cases, proxy measures will ask the proxy-informant about their own feelings regarding their responsibilities as a caregiver, guardian and etc.

Patient Reported Experience Measures

Patient reported experience measures (PREMs) are instruments often used to capture the overall patient experience of healthcare[7].

PREMs do not look at the outcomes of care but rather the impact of the process of the care on the patient's experience, for example, communication and the timeliness of assistance. They differ from satisfaction surveys by reporting objective patient experiences, removing the ability to report subjective views.

PREMs are also classified as **functional** or **relational**.

Functional PREMs examine practical issues, such as the facilities available.

Relational PREMs identify the patients' experience of their relationships during treatment, e.g., did they feel listened to.

Traditional Uses of PRMs

PRMs can be used to compare and benchmark processes and health outcomes **against selected criteria, such as industry standards or the performance of other healthcare providers**. Such comparisons can be used to highlight best practice and to identify areas of potential improvement.

Individual Patient Level (Micro)

At the individual level, PROMs and PREMs data can be used to enhance patient-provider communication, informing the care pathway for patients and encouraging shared decision making[5].

Originally, PROMs were used in pharmacological research to assess treatment effects in conditions such as cancer, particularly where cure was not possible, and HRQoL was the main clinical goal. In the last 20 years, the use of PROMs has increased considerably, and, currently, PROMs outcomes are used to assess the effects of treatment and quality of care, and to evaluate policies and inform health economic analysis[15]. In the UK PROMs have been regularly collected since 2009 for certain surgeries to support health services evaluation and to inform patient treatment choices[16].

PRMs are also used within the NHS's Outcomes Framework to enhance the HRQoL for people and maximise healthcare experiences for patients with chronic conditions[17].

Health Services Level (Meso)

At this level, aggregate PRM data can be used to guide QI efforts for benchmarking as well as patient safety[18]. When linked with traditional, clinical-based outcomes, these data provide a more comprehensive understanding of outcomes and effectiveness and can be used to identify gaps in care, evaluate health programs, assess and monitor outcomes of a group of patients over time, and evaluate the impact of healthcare services.

Health Systems Level (Macro)

At the health systems level, PRM data can be used to help decision-makers establish and evaluate policies intended to benefit a given population[5]. This includes comparing outcomes over time, locally and regionally; informing QI activities at a system level, such as adherence to clinical guidelines; and comparing performance measurement across organisations. Aggregate PRM data, therefore, provide important information to support evaluations of the impact of health system transformation changes.

PRMs in Clinical Quality Registries for Quality Improvement

Australia currently has over one hundred clinical registries, as listed on the Australian Commission of Safety and Quality in Health Care (ACSQHC)'s Register of Clinical Registries, https://www.safetyandquality.gov.au/publications-and-resources/australian-register-clinical-registries. However, not all of these meet the Commission's definition of CQRs, which are activities that provide data back to participating health services and/or clinicians for QI purposes as part of a Learning Health System (LHS).

PRMs are being increasingly introduced in CQRs, providing a personal perspective on the expectations and impacts of treatment[19]. Including PRMs in CQRs offers many advantages. First, incorporating a form of patient voice ensures that measurement of healthcare outcomes is patient-centric. Second, symptom burden, HRQoL and satisfaction with care are essentially lost if not captured in "real time", which may not coincide with clinical care episodes. Third, capturing of comprehensive PRM data in a registry can inform health service planning, research and evaluation, and facilitate benchmarking of participating health services[20].

When reported, CQR-collected PRMs may drive improvements in patient care and/or outcomes at the **clinician** or **health services** level via;

Clinician-level improvement activities:

- Compare outcomes of different patient cohorts and sub-groups
- Compare outcomes from different interventions, services or models of care
- Facilitate shared decision making between providers and consumers (e.g., incorporate registry data into clinician-patient/carer discussion, or use of predictive modelling)

Health service-level improvement activities:

- Monitor access to/satisfaction with health services or journeys of care
- Monitor consumer/patient experience of care
- Identify gaps in service delivery/poor outcomes that may benefit from new or changes to existing clinical services and models of care
- Evaluate outcomes following changes to clinical services and models of care
- Evaluate outcomes from other QI initiatives
- Benchmark quality of care and patient outcomes against peers
- Track patient clinical presentations and/or outcomes over time at the health services

Use of PRMs data for QI at the health service-level is often **limited by the volume and quality of health service-level PRM data**. This is particularly the case when collected PRMs are related to **low incidence outcomes (e.g., complications or poor experiences)**. Therefore, caution should be applied when utilising PRMs at a health services level in these situations. Conversely, reporting PRMs at a health services level is particularly appropriate for moderate to high incidence outcomes, e.g., satisfaction with service or overall improvement post intervention. Such outcomes are less affected by low volumes and the resulting increased statistical uncertainty.

Guide for Clinical Quality Registries Reporting of PRMs to Healthcare Providers

SECTION 1

PLANNING PRM ACTIVITIES

This section of the Guide provides brief recommendations for CQRs on the initial steps required for setting up PRMs program and data collection.

PRMs Program Governing Group

The CQR should have a governing group that oversees the PRMs program. A designated person should be appointed to oversee the PRM program to maximise chances for successful PRM implementation.

Consumers (patients, carers and their families) are important partners in PRM programs and should be involved in the PRMs governing group. The PRM governing group should include representatives from First Nations and culturally and linguistically diverse (CALD) communities.

Rationale for PRM Program Activities

Designing a PRM program in CQRs requires careful **planning**. For CQRs, **PRMs are complementary to clinical data**. At the time of planning, decisions should be made regarding how the PRMs and clinical data may be used together for QI.

There should be a clear **rationale** for the CQR to report PRMs for QI purposes that is informed by consultation with participating clinicians and health services and other stakeholders (e.g., jurisdiction health departments). The rationale should be documented prior to commencement of data collection.

As part of the **explanatory information**, patients/consumers should be informed of whether their identified patient-level PRMs results will be made available to their treating clinician for **clinical care purposes**, and/or whether de-identified data or aggregate reports only will be provided to their clinician and other stakeholders (e.g., health service managers, jurisdictions, industry, insurers, the public) for **QI purposes**.

A **sufficient budget should be allocated** to support PRM data collection, analysis and reporting activities.

Considerations for PRM Data Collection Methods and Population

PRMs can be collected pre- and post-intervention, longitudinally, and cross-sectionally. CQRs are encouraged to collect the data in a **longitudinal, repeated measures method, including baseline data**.

Depending on the **purpose(s)** of data collection, PRMs may be collected from the whole CQR population or a particular **subpopulation** (Example 1).

In general, use of PRMs for clinical care purposes requires the PRMs to be administered at time points prior to **expected clinician consultation and review**.

EXAMPLE 1: THE AUSTRALIAN BREAST DEVICE REGISTRY PROMs PROGRAM

The Australian Breast Device Registry (ABDR) is an Australian Government health initiative that records information on surgeries involving breast devices, such as breast implants[21].

A PROMs program was established in 2017 to assess if PROMs could predict breast device failure. Review of the program in 2021 revealed that during a 5-year period, response rates for PROMs had progressively declined. Response rates (proportion of contacted patients who completed PROMs) were higher for reconstructive patients compared with cosmetic patients. Reconstructive patient follow-up decreased for each calendar year; for Year 1 follow-up from 75% (2017) to 45% (2020); and for Year 2 follow-up from 74% (2017) to 42% (2020). For cosmetic patients, PROMs completeness was overall lower; from Year 1 follow-up of 46% (2017) to 40% (2020); and for Year 2 follow-up from 40% (2017) to 38% (2020).

Continuation of the PROMs program required redefining its objective given the risk of selection bias from the low response rates. Given that the cosmetic cohort were low PROMs responders and had a generally uncomplicated course, the ABDR decided that there was little value in continuing to collect PROMs from this group. The reconstructive cohort had higher rates of complications and revisions, and had a higher response rate, therefore it was considered there would be value in collecting PROMs in this cohort.

The existing PROMs program was paused with the plan to be relaunched for women undergoing breast reconstruction only, aiming to monitor and compare post-operative satisfaction and wellbeing between surgical techniques and peers.

SECTION 2

INSTRUMENT SELECTION

This section of the Guide provides guidance regarding instrument selection for CQRs wanting to collect PRMs for QI purposes. QI opportunities will differ depending on the condition and the information collected in the CQR (e.g., measuring access to care vs monitoring adherence to clinical care guidelines). The choice of PRM tool(s) should be driven by outcomes capable of contributing to improved patient care for the CQR condition/disease, that can be feasibly monitored.

Using Previously Existing Resources for Instrument Selection

A **literature review** is encouraged to inform the selection of appropriate PRMs for CQRs. Registries should also be aware of locally implemented PRMs (e.g., by jurisdictions) that may complement or align with the needs of the registry.

PRMs can be chosen from the existing measurement systems (e.g., EORTC or FACIT) and **core outcome sets** (e.g., ICHOM) containing individually validated and standardised instruments and scales that allow greater opportunity to compare outcomes across settings and populations.

Psychometric Properties

It is recommended that **PRMs** with robust psychometric evidence (e.g., validated tools) are selected for CQRs.

In addition to ensuring a selected PRM serves the identified purpose, the instrument must demonstrate adequate measurement properties, **such as validity, reliability, and responsiveness**, in the population in which they intend to be used ("fit-for-purpose" PRMs)[22, 23].

Multiple Instruments and Scales

Should we use more than one tool?

More than one instrument (e.g., generic vs condition-specific, or profile vs preference) is encouraged as this helps to meet the objective(s) of the PRMs program.

For example, **condition-specific** PROMs are more sensitive to change as they may provide more detail regarding specific outcomes for clinicians, whereas health service managers may be particularly interested in **generic** PROMs which can provide comparative information between conditions.

Deciding between scales

In deciding between **single- or multi-item scales** (tools containing more than one question), registries must consider that single-item instruments are often less reliable for measuring changes over time. Single-item instruments provide a very limited picture of the patient's perspective/experience.

Multi-item scales have greater reliability, sensitivity and content validity, but they may take longer to complete compared to single-item ones, and may generate domain/subscale scores rather than an overall summary score.

Consumer Involvement

When selecting PRMs, **consumers should be involved** to:

- Ensure **cultural appropriateness, response burden, appropriate literacy level** (e.g., in languages other than English, cultural relevance), and the real-world context in which people with lived experience and their families live, work, and play;

- Confirm that **selected instruments address health outcomes or experiences** that are relevant to patients and capture these in a comprehensive and understandable manner (Example 2).

Other Practical Considerations

When selecting PRM instruments, **practical considerations should include the length of the instrument, scoring requirements, the cost of licensing, and translation in languages that will serve the target population**.

The time taken to complete PRMs should be considered for PRMs instrument selection in CQRs to maximise response rates and minimise responder burden.

EXAMPLE 2. CONSUMER INVOLVEMENT IN THE AUSTRALIAN PELVIC FLOOR PROCEDURE REGISTRY

The Australasian Pelvic Floor Procedure Registry (APFPR) is an Australian Government Department of Health, Disability and Ageing initiative established in 2019 to record information about the safety and effectiveness of procedures for Stress Urinary Incontinence (SUI) and Pelvic Organ Prolapse (POP) that involve the use of devices or prostheses including mesh[24].

Careful consideration is required to choose the appropriate PROMs to include in a registry. In the context of a registry, limitations of using PROMs may be the combined length of multiple tools, and therefore the potential time burden for patients in their ease of completion.

The APFPR conducted acceptability studies with patients to assess the feasibility of incorporating PROMs into the APFPR[25, 26]. This included the evaluation of the preferred mode of administration and determining which instruments were the most suitable by assessing their relevance, clarity of wording and ease of use.

Most participants believed the PROMs would be useful in the APFPR and have potential for use in individual care.

SECTION 3

PRMs ADMINISTRATION

This section provides recommendations regarding PRMs administration in CQRs, timing and frequency for data collection, ethical requirements and response rates.

Timing and Frequency

The timing and frequency of data collection will depend on the purpose for data collection, the clinical context, and the resources available.

When considering the **frequency** for data collection, successful PRM initiatives should take into consideration the burden of data collection on patients, healthcare professionals, CQR staff, and it should be guided by consumers.

In general, PRMs administration for CQRs involves the use of the **same tool** at **multiple timepoints**.

PRMs in CQRs may be collected at various **time-points** including:

- Baseline (at time of diagnosis, or pre-intervention)
- Early post-operative/post-acute medical care (e.g., 3-4 weeks)
- At time of clinical stability (e.g., 6-12 months)
- Longer-term follow up (e.g., annual)

PRMs Data Collection Relating to Various Conditions

PRMs for CQRs relating to chronic medical conditions are generally collected at baseline and periodically thereafter (e.g., annually). PRMs may be administered after acute admissions e.g., to hospital.

For cancer patients PRMs may likely start from the time of diagnosis, and continue throughout cancer journey e.g., following different aspects of care.

PRMs for surgical condition CQRs and acute medical interventions are generally collected at baseline (if possible), early post-intervention, at the point of clinical stability, and for medium term follow up (e.g., 12 – 24 months).

Baseline PRMs are required for a **pre-post intervention evaluation.**

The usefulness of PRMs collected beyond 2 years may be limited by lower response rates, however if possible may provide very valuable insights into longer term outcomes.

PRMs **analysis** should reflect the nature of CQR follow up and be **longitudinal** i.e., the analysis compares outcomes for the same patient over time.

Ethical Considerations

When PRM data are collected and used in the course of routine care, for QI or healthcare delivery, ethical approval and patient consent may not be required. Patients should be provided information regarding the PRMs captured.

Providing patients with information about why and how their personal health information is being used is still an important privacy consideration and can be an opportunity to talk to patients about the value of PRM data to guide patient care and improve the quality of care provided.

Response Rates

PRMs may be administered via multiple methods to increase response rate. **Mode of PRMs administration should take into consideration patient factors, such as the age, gender and digital literacy**. Digital data collection methods (e.g., text, QR code, email) should be encouraged. Paper forms should be used only when necessary (e.g., for older people or people without digital access).

Electronic assessment systems should be considered (where appropriate) to minimise patient's time and data entry burden.

SECTION 4

STAKEHOLDER ENGAGEMENT AND EDUCATION

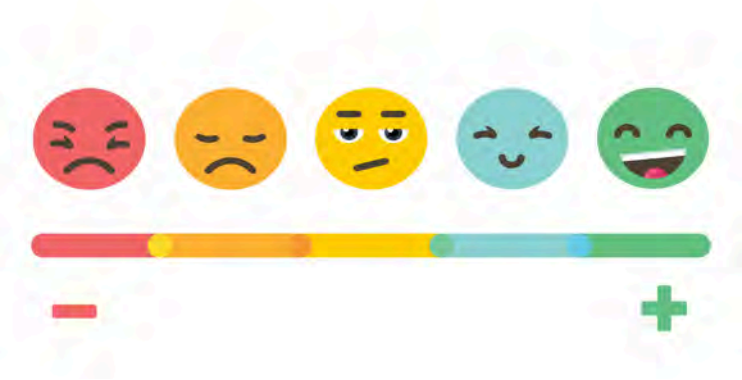
This section provides recommendations regarding stakeholder engagement, data presentation and education of clinicians and health services.

PRMs' use is relatively new in CQRs, so there is a variable level of knowledge regarding PRMs among clinicians and health service managers[27]. This should be considered when reporting PRMs, i.e., results should be presented clearly and informatively.

Education of Stakeholders

Education of clinicians/health service providers and consumers regarding how CQR PRMs can contribute to safety and QI and shared decision making, as well as **how to interpret PRMs** is important to maximise buy-in to the PRMs program[19].

Education **regarding PRM data validity, statistical methods, and instructions on how to interpret** the results (e.g., scoring range, scoring direction, any relevant thresholds/cut-off scores) should be considered to enhance acceptance and use of PRM data by clinicians[28].



SECTION 5

PRM DATA PREPARATION AND ANALYSIS

This section is intended to provide high-level guidance on the appropriate statistical techniques to consider, but we strongly recommend approaching a biostatistician in the planning stages to develop an analysis plan for any project reporting or interpretation of PRMs outcomes, in order to address data and design issues relevant to the specific CQR and/or project.

Biostatisticians and/or epidemiologists should be involved in processing and reporting CQR PRMs results.

Data analysis methods used for PRMs data analysis should be planned prior to the commencement of data collection and clearly described and documented.

More details on statistical analyses is provided in the **Statistical considerations section, Appendix A.**

Data Missingness

The most significant issue regarding PRMs data quality is **data missingness**[29].

Two main types of missing data include:

- Patients omitting certain questions, and
- A large quantity of missing data from multiple variables overall.

Missing data can be classified into three categories[30]: data missing completely at random (MCAR), data missing at random (MAR) and data missing not at random (MNAR).

Data missing completely at random (MCAR) such that no systematic differences exist between participants with missing data and those with complete data. This can happen due to an administrative error. Missing data are randomly distributed across the variables and are unrelated to other variables.

Data missing at random (MAR) have a systematic relationship between observed data and the nature of the missing outcomes. Missing data are not randomly distributed but they are accounted for by other observed variables. When data are Missing at Random, the absence of data is systematically related to the observed data but not to the unobserved data. For instance, in a study about depression, MAR data could occur if male participants are less likely than female participants to complete a survey on depression severity. This means that the likelihood of survey completion is associated with their gender (which is observed) but not with the severity of their depression (which is unobserved). Analyses using only complete cases—those records where no data are missing—can potentially be biased or unbiased with MAR data. If a bias exists in the complete case analysis, correctly adjusting for the known factors (like gender in this example) can help achieve unbiased results in the analysis.

Data missing not at random (MNAR) has a strong association between the missing values and the cohort, where the cause of the missing values is dependent on the patient and their environment. The missing data systematically differs from the observed values. Continuing the earlier example, a depression registry might have MNAR data if participants with severe depression are more likely to decline participation in a survey about depression severity. Like MAR, analysing a dataset with MNAR data using only complete cases can result in either biased or unbiased results. However, if there is bias in the complete case analysis due to MNAR, it is generally challenging to correct because the factors causing the missing data are unmeasured, often leading to a biased estimation of effects. It is crucial to identify this type of data to inform analytical strategies to address related biases that may result[29, 30].

Multiple Imputation

With missing outcomes, the assessment of provider performance requires careful consideration. A common approach for dealing with missing data is to discard patients for whom any outcome or covariate is missing. **Multiple imputation methods can be used to generate missing PRM data**[31]. With multiple imputation, each missing value is replaced by a set of plausible values, which are drawn from the posterior distribution of the missing outcomes given the observed data.

Reference Values

There are multiple **reference options** to inform the interpretation of PRMs results. These include **the baseline comparison to reference groups, including a) reference values obtained from other patients with the same indication or b) general population norm data for generic PROMs, and minimal important difference (MID) thresholds to interpret the comparison across time**[32, 33].

The values used for data interpretation need to be determined early with a clear justification provided why the values were used.

Minimum Clinically Important Difference

CQRs should use, where appropriate and previously defined, the **Minimum Clinically Important Difference (MCID) when a measure is responsive to change**. The MCID can distinguish between clinically relevant and statistically significant changes.

The MCID is the smallest change in a treatment outcome that an individual patient would identify as important and which would indicate a change in the patient's management[34].

The MCID is a threshold used to measure the effect of clinical treatments, and the Minimum Detectable Change (MDC) which is considered the most appropriate MCID value are two relevant metrics. MDC refers to the smallest change in score that can be detected statistically with some degree of certainty.

Ceiling and Floor Effect

The **ceiling effect** describes a situation where at least 15% of patients report the highest possible score, thereby losing discrimination at the upper end of the scale[22].

Similarly, the **floor effect** describes where at least 15% of patients report the lowest possible scores. These domains may not accurately reflect the real-life situation in specific diseases.

Appropriate modelling techniques should be considered to account for floor and ceiling effects in order to enable detection of change and reduce unwanted bias. However, properly validated tools should not exhibit ceiling nor floor effects.

PRM Scoring

Depending on PRM scoring guidelines, PRM data presented can include **the individual item, the subscale/domain score**, and/or **the overall instrument score** (Example 3).

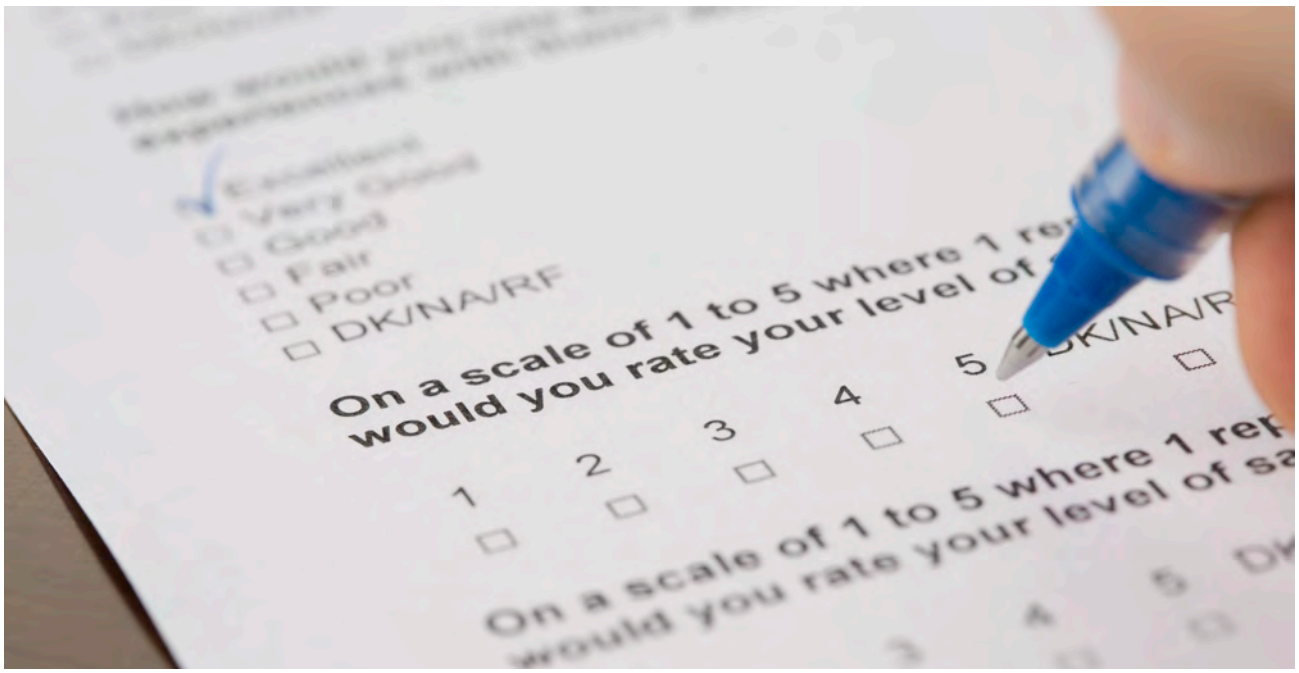
A variety of **summary statistics** can be **displayed** to provide additional context to the PRM results. These can include descriptive statistics (mean, median, frequency, range), quartiles, and confidence intervals[35, 36].

PRM Patient Numbers for Reporting

For the purposes of **QI reporting to clinicians or sites, a minimum volume of PRMs data per reporting period** should be considered to (a) ensure patient confidentiality (where appropriate) and (b) minimise statistical uncertainty due to low volumes.

It is recommended that **minimum patient numbers** for regular site/clinician reporting was between 5 – 20 patients depending on the outcome(s) measured, where low incidence outcomes should have a higher minimum patient number.

It is generally recommended to have at **least 30 responses** to approximate a normal distribution, in line with the central limit theorem, which is a fundamental concept in statistical analysis. Consequently, interpretations of results based on smaller sample sizes should be approached with caution, as they may not provide a reliable basis for statistical conclusions.



EXAMPLE 3. APFPR APFQ MULTISCALE SCORING

The APFPR commenced the collection of PROMs in 2022, using the Australian Pelvic Floor Questionnaire (APFQ). The APFQ is composed of 42 questions and consists of four independent domains: bladder function (15 items), bowel function (12 items), prolapse symptoms (5 items) and sexual function (10 items)[37].

Thirty-eight of the items assess pelvic floor symptoms, and four questions assess the bother caused by the symptoms in all areas. The last question in each domain of the questionnaire is the evaluation of bother, *"How much do your bladder/bowel/prolapse/sexual symptoms bother you?"* If the participant chose any of *"a little"*, *"quite a lot"* and *"very much"*, then it is regarded as bother and the variable is dichotomised accordingly.

Most of the items in the questionnaire use a 4-point scoring system, apart from a few. The symptom scores and bothersome scores are added to create a global score. The score of relevant questions is divided by the total score for each domain and multiplied by 10, with a value of 0–10 for each domain and a global maximum pelvic floor dysfunction score of 40.

The higher the score, the worse the symptoms/function i.e., a low score is 'good'. If a woman is not sexually active, she will have a score of 8.5 (18/21) in the sexual domain. In this case, the maximum score is 30.

A pilot study was conducted in 2022 to determine response rates, mode and frequency of administration of PROMs. Baseline and post-operative responses were collected from 156 and 185 women respectively.

Response rates were 80% at baseline and 76.2% at 6 months. There were clinically significant improvements in mean scores associated with bladder and prolapse symptoms pre- and post-surgery.

These results were reported to consumers and clinicians at Clinical Advisory and Steering Committee Meetings. Following further consultations with consumers and clinicians it was agreed to collect follow-up PROMs at 6, 12 and 24 months. In addition, consumers recommended to add the EQ-5D general wellbeing survey and a global improvement questionnaire to the PROMs program.

SECTION 6

USE CASES FOR QUALITY IMPROVEMENT PURPOSES

This section is intended to provide **some examples of QI use cases using PROMs and PREMs.**

- These include measuring and **comparing mean/median PRM** scores **between groups** or over time.
- Reviewing **variation** in PRMs scores via comparison between multiple individuals/groups using **visual mechanisms** such as **box and whisker** and **funnel plots**
- Reviewing a **single cohort's** PRM scores through the use of e.g., **histograms** to visualise PRM variation within a single scale.
- Visualising the distribution of the **change in PRMs** (pre- vs post- clinical intervention) to evaluate the effectiveness of the intervention.

Comparison of PRM Scores with an External Threshold

PRMs scores may be unfamiliar to healthcare providers, making interpretation of PRMs challenging. One way of assisting interpretation and use of PRM scores is to set specific **score thresholds**.

There are two commonly used methods for establishing score thresholds (cut-scores) for PRMs in CQRs:

1. **Normative values** - The first approach compares the PRM score to reference values (e.g., externally-derived values) derived from normative data, or data that is considered "usual" for a specific population. PRM scores are reported as the same, worse, or better than the reference population.

2. **Criterion validity** - The second approach involves using a gold standard or clear external criterion to determine an optimal threshold based on diagnostic accuracy. This threshold is then used for categorising patients into clinically relevant groups[38].

This type of analysis can be used to compare one site's PRMs scores against the various normative values to assess variation in clinical practice or quality of care.

Comparison of PRM Scores with an External Peer Group Data

In this type of analysis one site's PRMs scores are compared to a **pre-determined comparator** and may include all participating clinicians/providers, or a 'like' subset e.g., the comparator group may be similar hospital settings such as public or private, metro or regional, or hospitals that undertake certain procedural volumes.

Comparison of PRM Scores between two or more Provider Groups (e.g., patient cohorts/ interventions)

This type of analysis will display PRM scores between two provider groups or interventions.

For example, scores from electronic PRMs (Group 1) and paper PRMs (Group 2) can be compared using mean and SD. A Pearson's correlation or the non-parametric Spearman coefficient will be used to calculate the association between the measures and a paired t-test to compare means between the electronic PRMs and paper collected PRMs. Reliability analyses can be conducted using an intraclass correlation coefficient calculation.

Comparison of PRMs for the same cohort at different time points: longitudinal analysis (e.g., early post-operative vs at 12 months; or annually over time)

Longitudinal PRM scores can **describe trends/ changes over time** e.g., early and later outcomes following an intervention, and may also be used to describe health status pre/post specific interventions.

Longitudinal data analysis is often **complicated by complex correlation structures**, irregularly spaced visits, time-varying and static covariate effects, and, especially, by **missing data**, particularly patients who are lost to follow up[39].

A good example of longitudinal PROMs data analysis from the Swedish Spine Registry is described below (Example 4).

EXAMPLE 4. THE SWEDISH SPINE REGISTRY: LONGITUDINAL PROMs DATA ANALYSIS

Trends of PROMs at 1-, 2-, and 5-years following surgery in patients with traumatic cervical spine injuries were assessed[40]. For between-group statistical comparison of categorical variables the Chi-squared test was calculated while a one-way ANOVA was used for continuous variables. A mixed ANOVA was used to evaluate the effect resulting from the interaction of both follow-up time and surgical approach on the PROMs. Linear regression was used to assess the predictive ability of 1-year PROMs in predicting outcome at 2 and 5 years.

This study demonstrated the utility of 1-year PROMs in predicting longer-term outcomes after surgery for sub-axial traumatic injuries; however, a big limitation was loss to follow-up post-operatively, which may induce a risk of attrition bias.

Comparison of PRM scores specifically pre- and post- a clinical or QI intervention

Such analysis may be conducted to evaluate changes in PRMs pre- and post- a clinical or QI intervention at site level. Analysis can be undertaken via **Interrupted Time Series (ITS) analysis, ARIMA or mixed models**.

ITS is a quasi-experimental design used to **evaluate the impact of interventions or exposures**. Multiple statistical methods are available to analyse data from ITS studies[41, 42].

A natural experimental study using ITS was performed using National Joint Registry (NJR) data for patients with **primary hip replacement surgery** in England[43]. The intervention comprised a new policy from the Clinical Commissioning Group (CCG) in relation to waiting time to access to hip replacement for overweight patients (i.e., to delay surgery rather than speed up access). PROMs comprising pre- and post-operative Oxford Hip Score (OHS) questionnaire data were linked to the NJR dataset. The main outcome measures were rate of surgery.

Segmented linear regression models were used to estimate the trend **before policy introduction, and how this trend changed after policy introduction**, also allowing for an immediate step change at the date the policy was introduced.

A controlled ITS analysis was conducted using segmented linear regression of the differences between the groups, to compare the differences in trends and estimate an overall national effect of intervention compared to control CCGs.

The main outcome of this study showed that rates of surgery fell after policy introduction, whereas rates rose in localities with no policy. Policies enforcing extra waiting time before surgery were associated with worsening mean pre-operative OHS symptom scores and rising obesity.

Comparison of PRM scores for newer cohorts over time (e.g., comparison of 1-year PRMs for patients from 2020, 2021 and 2022) to assess for long term improvement

Such analysis may be conducted to evaluate changes in PRMs over time following CQR feedback reports to providers.

For example, a site will compare their post-operative PRM scores for patients who they operated on in 2020, to those they operated on from 2021, and then 2022 to see if specific functionality and/or overall wellbeing have reduced over time. This type of analysis can also be undertaken via ITS (please see above).

Integration of PRMs with clinical data to support predictive modelling and/or patient decision-making

The use of PRMs alongside clinical data is increasingly common in CQRs for predictive modelling. Predictive models can now leverage both clinical and PRMs data to enhance forecasts of individual outcomes, such as the likelihood of success or complications following medical interventions.

Risk prediction models can use aggregate or site data to estimate the risk of a particular outcome at the population or site level. Traditionally these models have been based on clinician-derived data. However, if PRMs data is available then it may be included as part of an overall **composite (clinical and PRMs) measure**.

Development of a composite score is a complex process and undertaken via a five-step approach: (i) theoretical framework and metric selection, (ii) initial data analysis, (iii) rescaling, (iv) weighting and aggregation, and (v) sensitivity and uncertainty analysis[44].

Composite measures that combine different types of indicators are widely used in medical research; to evaluate health systems, as outcomes in clinical trials and PRMs. The potential advantages of such indices are clear; however, composite variables can obscure details about the individual variables measured. Also, if individual indicators have problems, the resulting index can also be affected. At a minimum, all variables employed in a composite measure must be valid, unidimensional measures of quality or outcome[45]. **Composite measures should not be used if they fail to apply measurement theory, as they would then produce invalid and misleading scores**.

In **predictive modelling**, PRMs may be used as predictors (risk factors) or as outcomes. Predictive modelling uses risk factors and statistical techniques to predict which patients and populations are at high risk of increased utilisation, inappropriate service utilisation, increased cost, or poor outcomes. Factors such as comorbidities, severity of underlying disease, frailty scores and socioeconomic status are established risk adjustment variables, and may need to be adjusted to address these differences.

Risk adjustment is the process of statistically accounting for observed or hypothesised differences in patient case mix that influence healthcare outcomes. In a multivariable regression model, patient-related risk factors can be added to control for their contribution to the outcome of interest[46].

A recently conducted study of the Australian Breast Device Registry (ABDR) aimed to develop a risk adjustment model for PROMs to enable equitable benchmarking of satisfaction outcomes of breast implants in the ABDR. Clinical factors associated with PROMs outcomes following breast implant procedures within reconstructive and cosmetic procedure cohorts were identified using regression analysis[47]. It was found that the main operation indication and site type were important factors associated with PROMs satisfaction from reconstructive procedures. Age was the most common factor of importance identified for risk adjustment for benchmarking cosmetic breast implant PROMs.

SECTION 7

VARIATION AND OUTLIER MANAGEMENT

The previous use cases have generally involved reporting and comparing a central PRMs measure e.g., mean/median. An alternative use of PRMs data for QI reporting is to measure PRMs variation. This section provides recommendations and examples for PRMs outlier management.

Funnel Plots

This analysis provides a **visual representation of variation**, with funnel plot curves providing for lower certainty at sites that have lower patient numbers. Funnel plots can also identify **outliers**, typically at 3 standard deviations.

Depending on the significance of the PRMs results, CQRs may choose to develop a process for review and notification of sites of sub-optimal PRMs scores e.g., outliers.

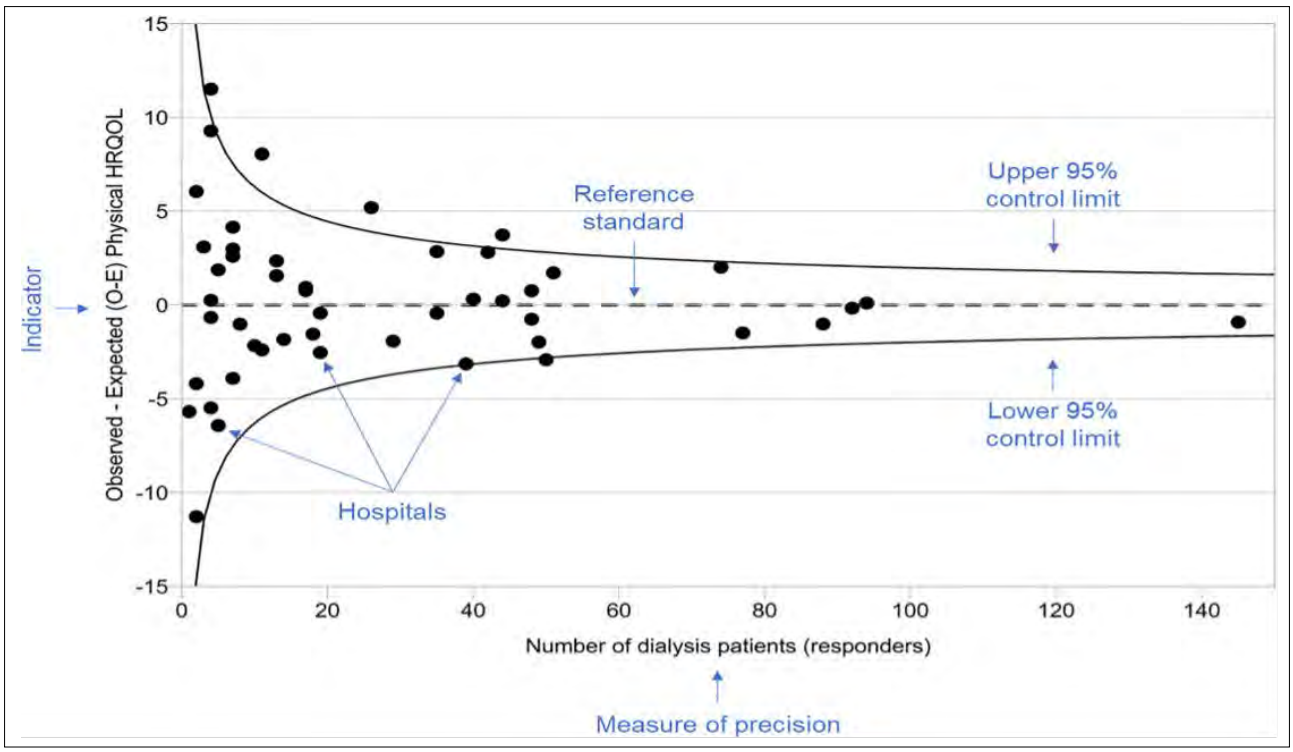


Figure 1A: A funnel plot on physical HRQoL in 48 Dutch dialysis centres that participated in the Dutch registry of PROMs in 2019. Adapted from van der Willik et al[48]

Figure 1A shows an example of **a funnel plot** of physical HRQoL in 48 Dutch dialysis centres that participated in the Dutch registry of PROMs in 2019[48]. The indicator shows the comparisons between the centres' observed and expected scores on physical HRQoL. The total study population of Dutch dialysis patients is used as **a reference standard**. The 95% control limits are provided around the reference standard.

Figure 1B shows funnel plot of post-operative EQ-5D VAS Health in Primary Total Knee Replacement in patients from the Australian Orthopaedic Association National Joint Replacement Registry (AOANJRR).

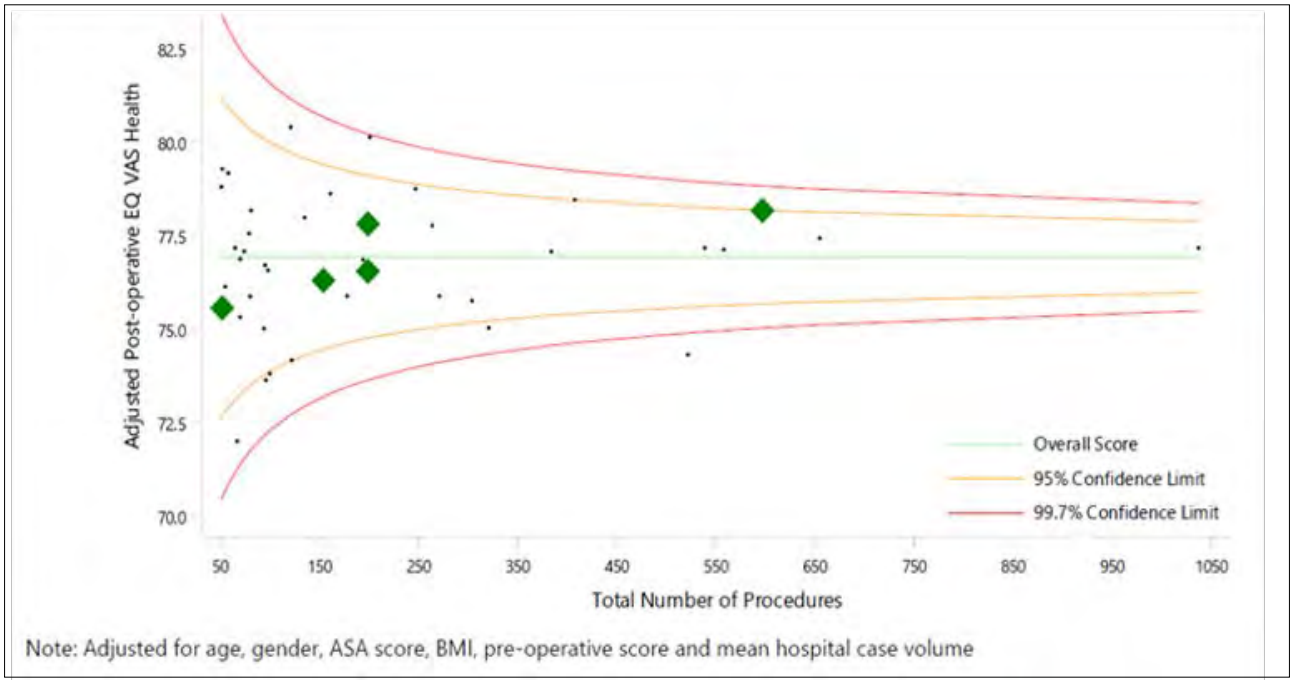


Figure 1B: A funnel plot of post-operative EQ-5D VAS Health in Primary Total Knee Replacement in patients from the AOANJRR report to a state health department

Histograms

PRMs outcomes may also be **visualised** for a **single PRMs tool/scale** for **one cohort** at the provider/site level at a single or multiple point in time.

This is undertaken via a descriptive analysis of the PRMs score results for one cohort, and usually presenting these as a histogram with score options (e.g., 1-10) along the x-axis and frequency (e.g., 0% - 100%) along the y-axis.

These histograms can be presented at different time-points to show variation in PRMs spread over time. Results may be used by providers/health services to show e.g., the proportion of patients with persistent poor outcomes at longer term follow up - e.g., at 12 months.

The example below is adapted from the Australian Spine Registry Annual Report[49] and shows the Oswestry Disability Index (ODI) scores for the discectomy cohort pre-operatively and at 6-months post-operatively. Figure 2 shows a shift to the left (lower scores) in the overall ODI for the discectomy cohort at the 6-month follow up time point indicating improvement over the 6-month period.

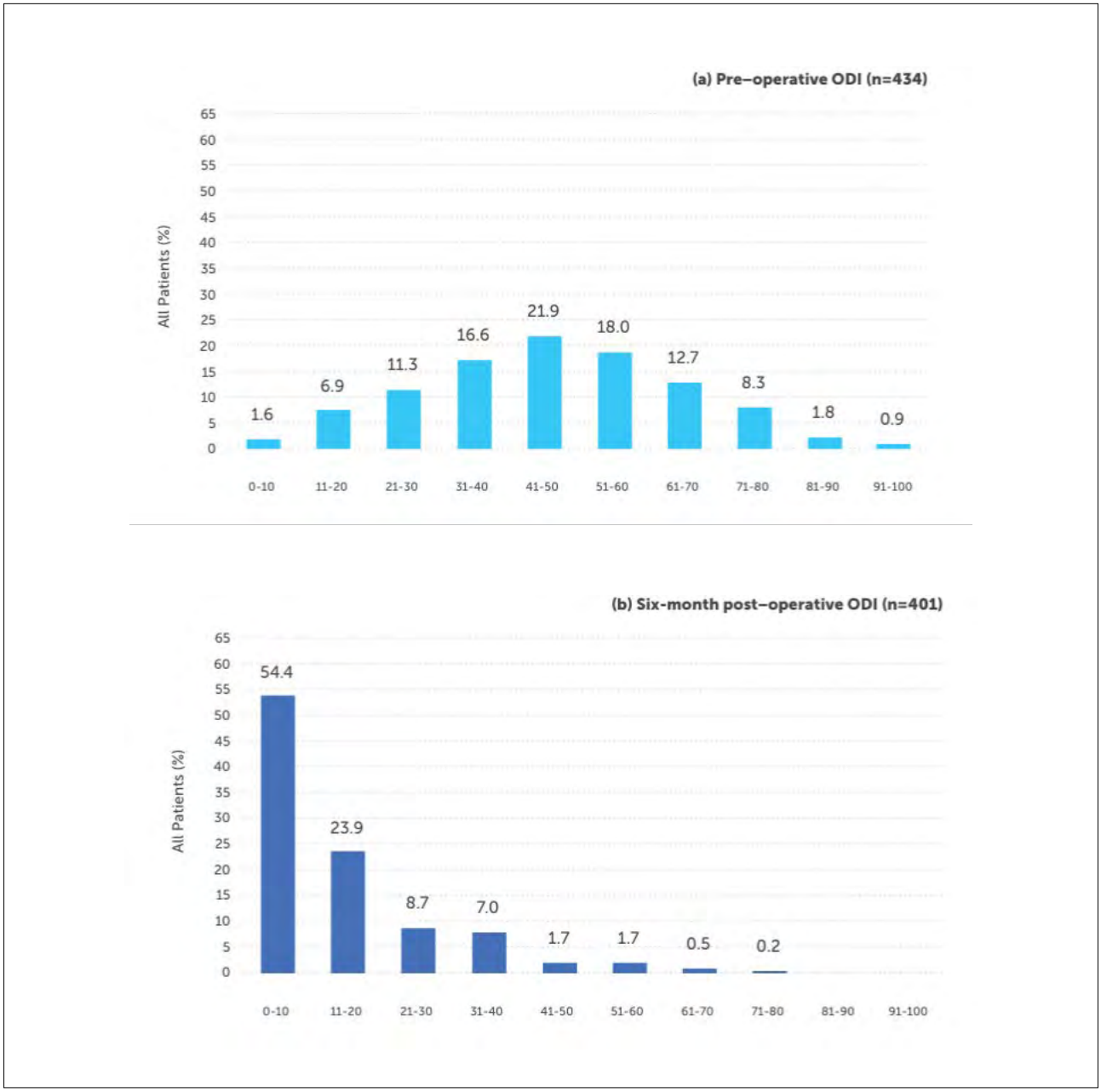


Figure 2: The Oswestry Disability Index scores for the discectomy cohort pre-operatively and at 6-months post-operatively. The histogram adapted from the Australian Spine Registry Annual Report 2023[49]

Control Charts

Another useful method for analysing PRM data is the use of control charts. **Control charts, also known as Shewhart Charts or Statistical Process Control Charts (SPCC) are tools used to determine if a process is in a state of statistical control, or how much variation exists in a process**.

Control charts are time-series charts. Depending on the type of data, a particular type of chart is used, with a centre line and upper and lower control limit marked on the chart. These lines are used to distinguish between common and special cause variation, which determines the most appropriate improvement approach.

They are particularly useful in tracking PRMs data over time to identify trends, shifts, or any variations that are outside of normal process limits.

Application in PRMs analysis:

- Trend Analysis: Control charts can track changes in PRMs scores over time across different providers or treatment modalities. This helps in identifying any emerging trends that may indicate improvements or deteriorations in patient outcomes

- Process Stability: They help in assessing the stability of healthcare processes by monitoring performance and highlighting any variations that may suggest a deviation from standard practice

- Quality Control: By setting upper and lower control limits, typically at three standard deviations from the process mean (Figure 3), control charts help in pinpointing times when the process may be out of control, necessitating further investigation or corrective actions.

Wide process limits indicate a process or system that is volatile and unreliable. It is important to understand what is causing the instability and look to reduce or eliminate this. If performance is predictable you will find it easier to plan and manage services, and provide your patients with more certainty.

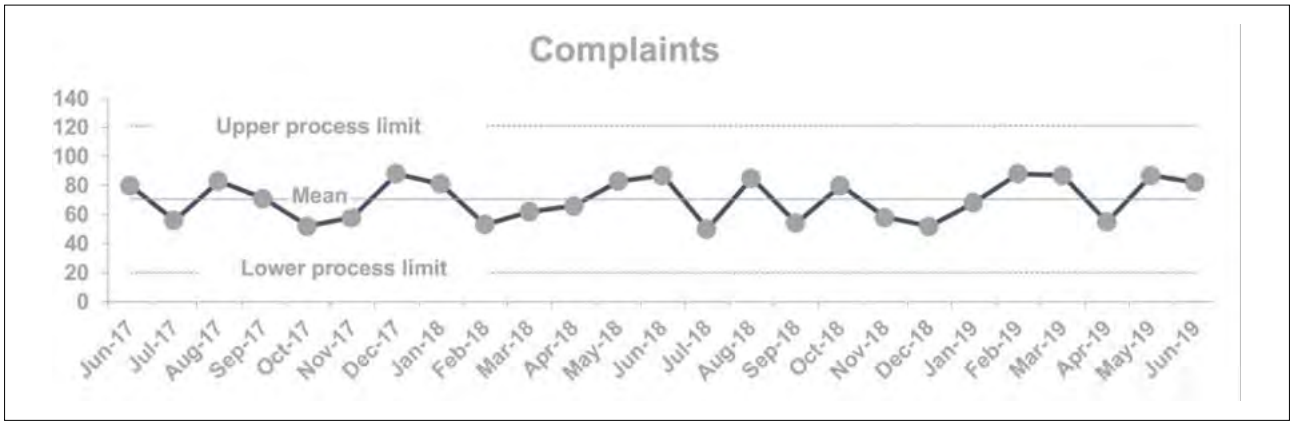


Figure 3: Control chart example. Figure adapted from "Making data count. Strengthening your decisions", https://www.england.nhs.uk/wp-content/uploads/2019/12/making-data-count-strengthening-your-decisions.pdf.

SECTION 8

PRM DATA VISUALISATION

This section is intended to provide a brief overview of PRM data visualisation to aid the interpretation of PRM results. For more detailed information please see the following sources:

- Making a Picture Worth a Thousand Numbers: Recommendations for Graphically Displaying Patient-Reported Outcomes Data[50]
- ISOQOL Users' Guide for Implementing PRO Assessment in Clinical Practice[51]
- The PROTEUS Guide to Implementing Patient-reported Outcomes in Clinical Practice. A synthesis of resources (https://theproteusconsortium.org/proteus-practice/proteus-practice-guide/visualizations-to-aid-interpretation/)
- Additional suggestions and tips regarding data visualisation are provided in **Appendix B.**

General Recommendations

In general, when **displaying PRM data visually**, ensure the following:

- Clear figure title
- Both axes are labelled, with units shown (if appropriate)
- There is a key/description of the PRMs tool that clearly indicates what a high vs a low score means clinically (e.g., good vs poor outcome)
- Keep figures simple where possible

Visual/graphical presentation includes charts, graphs, and pictographs, and may be more intuitive to ease interpretation of the data[52]:

- May be preferred for displaying information about changes in health

Can more effectively communicate health to low literacy groups with specific visualisation methods, such as pictographs.

Examples of PRM Data Presentation

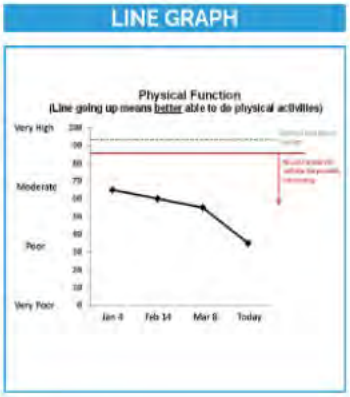
Tables, graphs and figures

PRM data may be presented in **tables** or visually as **graphs/figures or pictographs**.

Figure 4 shows different types of graphs that can be used to visualise individual-level PRO data with examples.

TABLE

	Scores on Each Visit Date			
Date	3/12	3/26	4/9	4/23
Physical Function	70	75	68	56
Emotional Function	80	80	85	80
Overall Quality of Life	75	70	70	65
Nausea or Vomitting	10	15	15	25
Pain	5	5	5	10
Fatigue	20	25	25	35



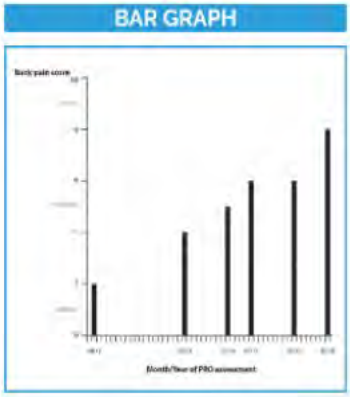


Figure 4: Tables, line and bar graph examples to visualise PRM data. Adapted from the PROTEUS Guide, ePROS in Clinical Care website and from Recommendations for PRO Data Display

Pictographs

Figure 5 shows a pictograph of PROMs data comparing one group of patients to another group of patients.

An example in Figure 6 from the Australian Dementia Network highlights the display of PREMs data through bar charts. Here, the patient reported experience regarding their treatment is displayed as an example. The bar graph below displays key components of optimal visualisation of PREMs data by incorporating a figure title; labelled axes; and a simple explanation of the PREMs tool.

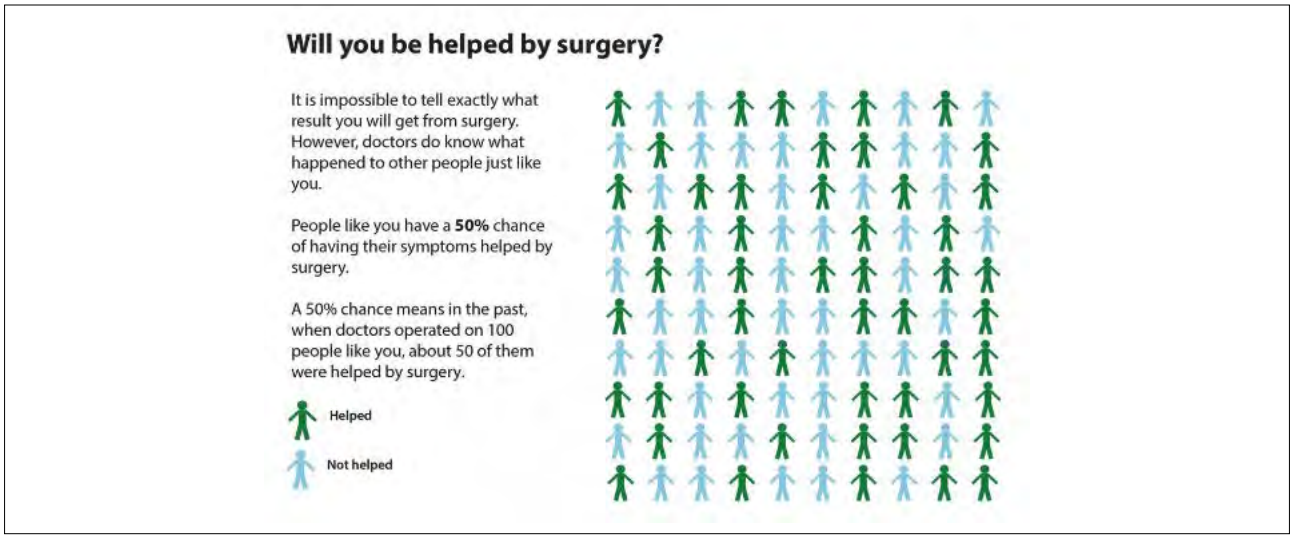


Figure 5: Pictograph comparing one group of patients to another group of patients. Adapted from https://becertain.org/epros

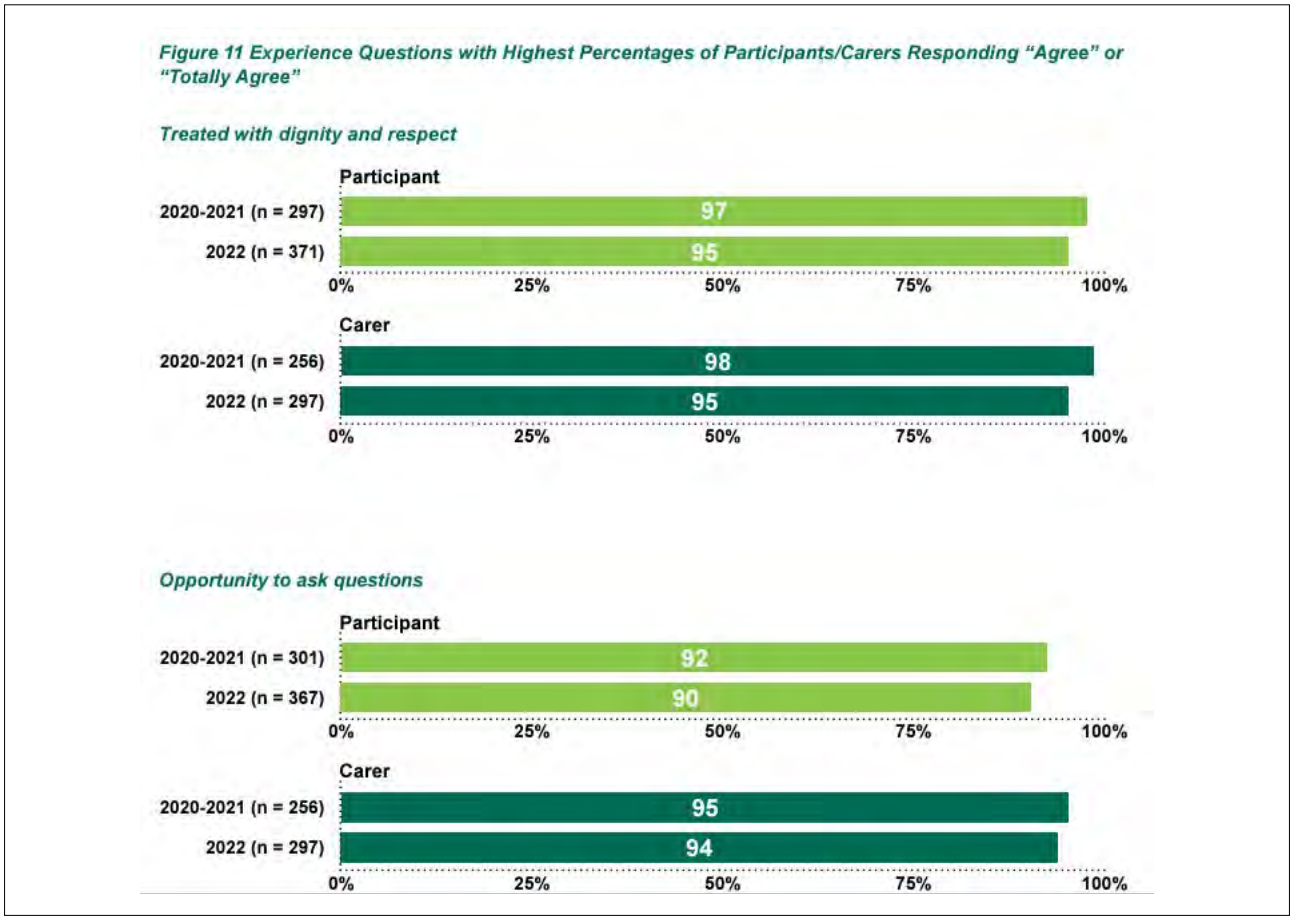


Figure 6: Experience Questions with Highest Percentages of Participants/Carers responding "Agree" or "Totally Agree". Adapted from Australian Dementia Network Registry 2022 Annual Report[53]

Lines, Colour, Bolding and Symbols

Lines, colour, bolding and symbols can be used to draw attention to aspects of the PRM data display[54]:

- Lines on the graph can be used to visualise discrimination of scores (e.g., lines can indicate the threshold for mild, moderate, severe scores) and to reflect reference values such as those from the general population/another comparator group;

- Traffic light colours (green, yellow, red) can be used to designate severity. Pairings of colour and shading should consider the needs of people with visual impairments such as colour blindness/colour confusion;

- Cultural associations of patient populations should be considered when relevant as colours may have different meanings in different cultures.

The example below (Figure 7) displays the use of traffic light colours to highlight physical, emotional functioning, pain and fatigue QLQ-C30 scores in cancer patients by visualizing clinically unimportant scores in green and clinically important scores in red. The thresholds for clinical importance were predefined. Thresholds for clinical importance were distribution-based: scores in the range of the lowest decile of the reference population (brain tumour patients monitored at the department) were considered clinically important[55].

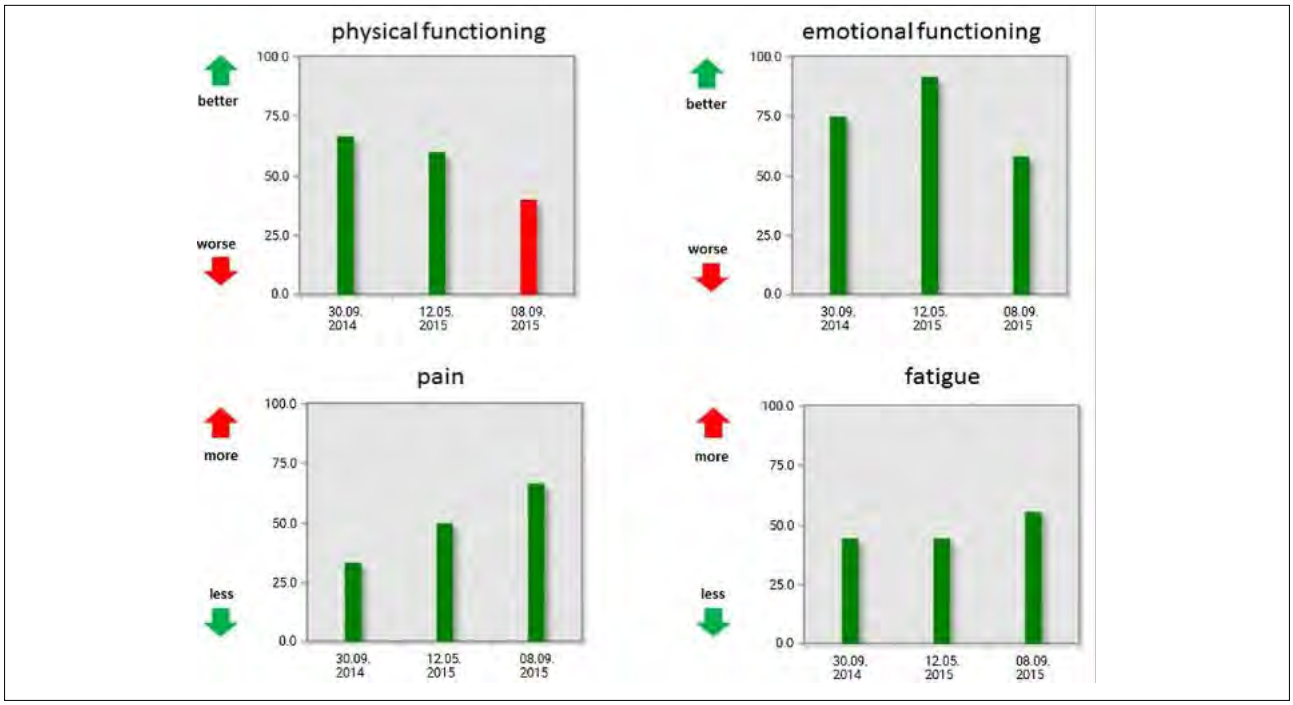


Figure 7: Graphical presentation of PROMs using traffic light system, adapted from Loth et al.[55]

As displayed in Figure 8, the traffic light system is used to visualise the severity of PRM results. Red is used to display severe results; yellow depicts moderate results; and green displays mild problems.

Figure 8: Presentation of PROMs data as traffic light. Adapted from Santana et al.[58]

SECTION 9

PRM DATA REPORTING AND DISSEMINATION

The success of a PRM program lies not only in the successful collection and analysis of data but also in how this information is reported and translated to advance patient care and health system goals.

This includes determining the target audience of the information gathered from the PRM program. The reporting and use of PRM data require that decisions relating to how the data will be used be made early on in the planning phase. This section provides a brief guidance regarding PRM Data Reporting and dissemination.

Timing for PRM Data Reporting

CQRs should generate routine standard reports that include PRM data and send these to identified institutions, clinicians and consumers.

Time intervals for PRM data reporting should, in general, align with the frequency of reporting of clinical data to sites, as well as taking into consideration the volume of PRMs data available for each time period.

Frequency for PRM Data Reporting

It is recommended that PRM data at single/multiple timepoints be provided to sites cumulatively as well as periodically (e.g., an annual time series).

Ethical Considerations

All ethical considerations that were applied in collecting and analysing PRM data should be reported clearly stating the details regarding patient's confidentiality and privacy, and any relevant ethics/governance approvals.

SECTION 10

PRMs PROGRAM EVALUATION

Systematic evaluation of PRMs usage within a registry should be performed regularly and should consider facilitators/ barriers that were identified during preparation and implementation phase. This section provides guidance regarding PRM program evaluation.

Evaluation Components

Evaluation of PRMs program includes:

- **PRM instruments:** Evaluating the effectiveness and relevance of the tools used for data collection.
- **Frequency of PRM data collection**: Evaluating the relevance of follow up data points, burden to patients, clinicians and registry staff.
- **PRM response rates:** Assessing the completion rates to ensure high engagement and representativeness of the data.
- **PRM use cases:** Reviewing how the data is applied in clinical and administrative settings to ensure it meets intended goals.

Additional Considerations

Additional considerations should include:

- **Quality assurance:** Implementing regular checks for data accuracy and integrity, and the volume of missing data.
- **Stakeholder feedback:** Integrating feedback from clinicians and consumers of the PRMs to refine tools and processes.
- **Impact assessment**: Analysing the actual impact of PRMs on patient outcomes and healthcare practices.

There may be opportunities to regularly refine the PRM program e.g., by reducing PRM time points, or the frequency of reporting PRM to sites, once the PRM program has matured.

Conclusion

The aim of this project was to develop a set of guiding principles and recommendations for CQRs, clinicians and health service managers for the planning and collection of PRM data, including analysis, reporting and interpretation of hospital or clinician-level PRMs from CQRs.

Developed in collaboration with partners across CQRs, this document presents a framework and recommendations to guide the collection, analysis, reporting, and use of PROMs and PREMs. It is the first step in a coordinated approach, where PRM data are collected from patients and made readily available to guide continuous QI, support clinician-patient decision making, and inform health system policy decisions as part of a learning health system.

Acknowledgements

We acknowledge the funding provided by the Australian Government Department of Health, Disability and Ageing for this project. We acknowledge the participants of focus groups for sharing their perspectives and views. We also thank the workshop participants for their guidance and assistance in reviewing the draft Best Practice Guide. We thank Bec Harvey and Dr Esther Apos for sharing cases from their clinical quality registries. Finally, we would like to thank Professor Arul Earnest and Dr Ahmad Reza Pourghaderi who provided input on the statistical analysis section of the document.

References and Further Readings

1 Garrubba A, Melder M. Clinical Registries: Informing a framework for hospital participation. Centre for Clinical Effectiveness, Monash Health, Melbourne, Australia 2016.

2 Evans SM, Scott IA, Johnson NP, Cameron PA, McNeil JJ. Development of clinical-quality registries in Australia: the way forward. Med J Aust. 2011;194(7):360-3.

3 Wynne R, Jackson D, Prince J, O'Regan A, Kirk A, Ferguson C. Clinical quality registries: An approach to support research capacity building in clinical academic partnerships. J Clin Nurs. 2021;30(7-8):e29-e31.

4 Parker KJ, Hickman LD, Ferguson C. The science of clinical quality registries. Eur J Cardiovasc Nurs. 2023;22(2):220-5.

5 Williams K, Sansoni J, Morris D, Grootemaat P CT. Patient-reported outcome measures: Literature review. Sydney: ACSQHC https://wwwsafetyandqualitygovau/sites/default/files/migrated/PROMs-stakeholder-interviews-report-2018pdf. 2016.

6 MacLean CH, Antao VC, Fontana MA, Sandhu HS, McLawhorn S. PROMs: Opportunities, challenges, and unfinished business. NEJM Catal Innov Care Deliv 2021 2021;2(11) DOI: 10.1056/CAT.21.0280 VOL. 2 NO. 11.

7 Weldring T, Smith SM. Patient-Reported Outcomes (PROs) and Patient-Reported Outcome Measures (PROMs). Health Serv Insights. 2013;6:61-8.

8 Ernstsson O, Janssen MF, Heintz E. Collection and use of EQ-5D for follow-up, decision-making, and quality improvement in health care - the case of the Swedish National Quality Registries. J Patient Rep Outcomes. 2020;4(1):78.

9 Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. J Natl Cancer Inst. 1993;85(5):365-76.

10 Bergman B, Aaronson NK, Ahmedzai S, Kaasa S, Sullivan M. The EORTC QLQ-LC13: a modular supplement to the EORTC Core Quality of Life Questionnaire (QLQ-C30) for use in lung cancer clinical trials. EORTC Study Group on Quality of Life. Eur J Cancer. 1994;30a(5):635-42.

11 Churruca K, Pomare C, Ellis LA, Long JC, Henderson SB, Murphy LED, et al. Patient-reported outcome measures (PROMs): A review of generic and condition-specific measures and a discussion of trends and issues. Health Expect. 2021;24(4):1015-24.

12 Elsevier. PFMKAP. The measurement of health and health status: concepts, methods and applications from a multidisciplinary perspective: 1st edn. Academic Press, Berlin; 2017.

13 Jones JM, McPherson CJ, Zimmermann C, Rodin G, Le LW, Cohen SR. Assessing agreement between terminally ill cancer patients' reports of their quality of life and family caregiver and palliative care physician proxy ratings. J Pain Symptom Manage. 2011;42(3):354-65.

14 Roydhouse JK, Cohen ML, Eshoj HR, Corsini N, Yucel E, Rutherford C, et al. The use of proxies and proxy-reported measures: a report of the international society for quality of life research (ISOQOL) proxy task force. Qual Life Res. 2022;31(2):317-27.

15 Dawson J, Doll H, Fitzpatrick R, Jenkinson C, Carr AJ. The routine use of patient reported outcome measures in healthcare settings. Bmj. 2010;340:c186.

16 Devlin A. Getting the Most Out of PROMs: Putting Health Outcomes at the Heart of NHS Decision-Making. 2010.

17 Canadian Institute for Health Information (CIHI). PROMs Background Document. Available at: https://www.cihi.ca/en/patient-reported-outcome-measures-proms. Accessed 2024.

18 Franklin P, Chenok K, Lavalee D, Love R, Paxton L, Segal C, et al. Framework to guide the collection and use of patient-reported outcome measures in the Learning Healthcare System. EGEMS (Wash DC). 2017;5(1):17.

19 Australian Commission on Safety and Quality in Health Care, Framework for Australian clinical quality registries - a national standard for clinical safety and quality data collections and reporting Second Edition. Sydney. ACSQHC, December 2022.

20 Ruseckaite R, Maharaj AD, Dean J, Krysinska K, Ackerman IN, Brennan AL, et al. Preliminary development of recommendations for the inclusion of patient-reported outcome measures in clinical quality registries. BMC Health Serv Res. 2022;22(1):276.

21 Hopper I, Ahern S, Best RL, McNeil J, Cooter RD. Australian Breast Device Registry: breast device safety transformed. ANZ J Surg. 2017;87(1-2):9-10.

22 Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol. 2010;63(7):737-45.

23 Edwards MC, Slagle A, Rubright JD, Wirth RJ. Fit for purpose and modern validity theory in clinical outcomes assessment. Qual Life Res. 2018;27(7):1711-20.

24 Jayasinghe RT, Ruseckaite R, Dean J, Kartik A, Wickremasinghe AC, Daly O, et al. Establishment and initial implementation of the Australasian Pelvic Floor Procedure Registry. Int Urogynecol J. 2023;34(8):1697-704.

25 Ruseckaite R, Bavor C, Marsh L, Dean J, Daly O, Vasiliadis D, et al. Evaluation of the acceptability of patient-reported outcome measures in women following pelvic floor procedures. Qual Life Res. 2022;31(7):2213-21.

26 Ruseckaite R, Jayasinghe R, Bavor C, Dean J, Daly O, Ahern S. Evaluation and acceptability of patient-reported outcome measures in women following pelvic organ prolapse procedures. BMC Health Serv Res. 2023;23(1):624.

27 Santana MJ, Haverman L, Absolom K, Takeuchi E, Feeny D, Grootenhuis M, et al. Training clinicians in how to use patient-reported outcome measures in routine clinical practice. Qual Life Res. 2015;24(7):1707-18.

28 Foster A, Croot L, Brazier J, Harris J, O'Cathain A. The facilitators and barriers to implementing patient reported outcome measures in organisations delivering health related services: a systematic review of reviews. J Patient Rep Outcomes. 2018;2:46.

29 Rubin DB, Schenker N. Multiple imputation in health-are databases: An overview and some applications. Stat Med. 1991; https://doi.org/10.1002/sim.4780100410

30 Mack C, Su Z, Westreich D. AHRQ Methods for Effective Health Care. Managing Missing Data in Patient Registries: Addendum to Registries for Evaluating Patient Outcomes: A User's Guide, Third Edition. Rockville (MD): Agency for Healthcare Research and Quality (US); 2018.

31 Gomes M, Gutacker N, Bojke C, Street A. Addressing missing data in Patient-Reported Outcome Measures (PROMS): Implications for the use of PROMS for comparing provider performance. Health Econ. 2016;25(5):515-28.

32 Scott NW, Fayers PM, Aaronson NK, Bottomley A, de Graeff A, Groenvold M, Gundy C, Koller M, Petersen MA, Sprangers MA & on behalf of the EORTC Quality of Life Group. EORTC QLQ-C30 Reference Values. 2008.

33 Nolte S, Liegl G, Petersen MA, Aaronson NK, Costantini A, Fayers PM, et al. General population normative data for the EORTC QLQ-C30 health-related quality of life questionnaire based on 15,386 persons across 13 European countries, Canada and the Unites States. Eur J Cancer. 2019;107:153-63.

34 Copay AG, Subach BR, Glassman SD, Polly DW, Jr., Schuler TC. Understanding the minimum clinically important difference: a review of concepts and methods. Spine J. 2007;7(5):541-6.

35 Earnest A. Essentials of a Successful Biostatistical Collaboration (1st ed.) Chapman and Hall/CRC. https://doi.org/10.1201/9781315373447

36 NSW Agency for Clinical Innovation. Patient-reported outcome measures: Methods for analysis and reporting. Sydney: ACI; 2023.

37 Baessler K, O'Neill SM, Maher CF, Battistutta D. A validated self-administered female pelvic floor questionnaire. Int Urogynecol J. 2010;21(2):163-72.

38 Cohen RJ, Swerdlik ME. Psychological testing and assessment: An introduction to tests and measurement (6th ed.): New York: McGraw-Hill; 2005.

39 Zeger SL, Liang KY. An overview of methods for the analysis of longitudinal data. Stat Med. 1992;11(14-15):1825-39.

40 El-Hajj VG, Singh A, Blixt S, Edström E, Elmi-Terander A, Gerdhem P. Evolution of patient-reported outcome measures, 1, 2, and 5 years after surgery for subaxial cervical spine fractures, a nation-wide registry study. Spine J. 2023;23(8):1182-8.

41 Wagner AK, Soumerai SB, Zhang F, Ross-Degnan D. Segmented regression analysis of interrupted time series studies in medication use research. J Clin Pharm Ther. 2002;27(4):299-309.

42 Schaffer AL, Dobbins TA, Pearson SA. Interrupted time series analysis using autoregressive integrated moving average (ARIMA) models: a guide for evaluating large-scale health interventions. BMC Med Res Methodol. 2021;21(1):58.

43 McLaughlin J, Kipping R, Owen-Smith A, McLeod H, Hawley S, Wilkinson JM, et al. What effect have commissioners' policies for body mass index had on hip replacement surgery?: an interrupted time series analysis from the National Joint Registry for England. BMC Med. 2023;21(1):202.

44 National Quality Forum. Composite Measure Evaluation Framework and National Voluntary Consensus Standards for Mortality and Safety: Composite measures: a consensus report 2009.

45 McKenna SP, Heaney A. Composite outcome measurement in clinical research: the triumph of illusion over reality? J Med Econ. 2020;23(10):1196-204.

46 PROMs risk adjustment methodology guide for general surgery and orthopaedic procedures. https://www.england.nhs.uk/statistics/wp-content/uploads/sites/2/2013/07/proms-ris-adj-meth-sur-orth.pdf

47 Hansen J, Ahern S, Gartoulla P, Khu Y, Elder E, Moore C, et al. Identification of predictive factors for patient-reported outcomes in the prospective Australian Breast Device Registry. Aesthet Surg J. 2022;42(5):470-80.

48 van der Willik EM, Hemmelder MH, Bart HAJ, van Ittersum FJ, Hoogendijk-van den Akker JM, Bos WJW, et al. Routinely measuring symptom burden and health-related quality of life in dialysis patients: first results from the Dutch registry of patient-reported outcome measures. Clin Kidney J. 2021;14(6):1535-44.

49 Apos E, Bulmer S, Truong T, Garduce P, Pourgadheri A, Ahern S, et al. The Australian Spine Registry Annual Report, 2023: Spine Society of Australia and Public Health and Preventive Medicine, Monash University, April 2024, Report No.06.

50 Snyder C, Smith K, Holzner B, Rivera YM, Bantug E, Brundage M. Making a picture worth a thousand numbers: recommendations for graphically displaying patient-reported outcomes data. Qual Life Res. 2019;28(2):345-56.

51 Chan EKH, Edwards TC, Haywood K, Mikles SP, Newton L. Implementing patient-reported outcome measures in clinical practice: a companion guide to the ISOQOL user's guide. Qual Life Res. 2019;28(3):621-7.

52 Bantug ET, Coles T, Smith KC, Snyder CF, Rouette J, Brundage MD. Graphical displays of patient-reported outcomes (PRO) for use in clinical practice: What makes a pro picture worth a thousand words? Patient Educ Couns. 2016;99(4):483-90.

53 Ward SA, Lin X, Wallis K, Honardoost MA, Tsui A, Lassetter C, Rowe C, Anstey K, Brodtmann A, Chong T, Darling G, Inacio M, Jeon Y-H, Kain B, Loi S, McCabe M, Naismith S, Natarajan K, Nelson M, Newton L, Pietsch A, Quirke L, Rand E, Yates M, Arsenova V, Earnest A, McAloney K, Pourghaderi AR, Rahja M, Richardson J, Tan S and Ahern S. The Australian Dementia Network (ADNeT) Registry 2022 Annual Report. Monash University, School of Public Health and Preventive Medicine, May 2023.

54 Ryan OF, Hancock SL, Marion V, Kelly P, Kilkenny MF, Clissold B, et al. Feedback of aggregate patient-reported outcomes (PROs) data to clinicians and hospital end users: findings from an Australian codesign workshop process. BMJ Open. 2022;12(7):e055999.

55 Loth FL, Holzner B, Sztankay M, Bliem HR, Raoufi S, Rumpold G, et al. Cancer patients' understanding of longitudinal EORTC QLQ-C30 scores presented as bar charts. Patient Educ Couns. 2016;99(12):2012-7.

56 van der Willik EM, Terwee CB, Bos WJW, Hemmelder MH, Jager KJ, Zoccali C, et al. Patient-reported outcome measures (PROMs): making sense of individual PROM scores and changes in PROM scores over time. Nephrology (Carlton). 2021;26(5):391-9.

57 Painter E, Zwar J, Carino S, Kermonde Z. Best practice data visualisation: guidelines and case study. Monash University Monash Climate Change Communication Research Hub 2021.

58 Santana MJ, Haverman L, Absolom K, Takeuchi E, Feeny D, Grootenhuis M, Velikova G. Training clinicians in how to use patient-reported outcome measures in routine clinical practice. Qual Life Res. 2015 Jul;24(7):1707-18.

APPENDIX A: Statistical Considerations for PRM Data Analysis

This document is intended to provide high-level guidance on the appropriate statistical techniques to consider, but we strongly recommend approaching a biostatistician in the planning stages to develop an analysis plan for any project reporting or interpretation of PRMs outcomes, in order to address data and design issues relevant to the specific CQR and/or project.

PRMs generally comprise predominantly **Likert or visual analogue scales**. Both methods generally produce scores where higher scores correlate with a better HRQoL.

For example, a positively phrased statement such as "I enjoy my hobbies" will be scored more highly. It is important to consider negatively phrased items such as "I am in pain." A higher score in this case reflects a more negative outcome. In these scenarios, reverse scoring should be performed where the reported value is subtracted from the maximum score in that statement (for example a 4-point Likert scale means the maximum score is 4). The interpretation of scores can be challenging, for instance due to the complexity of the PRO (e.g., HRQoL, which includes various physical, mental and social domains) or the instrument (e.g., a complex scoring method). Some methods that may facilitate the interpretation of scores have been described before[56].

When performing statistical analysis that **compares two or more groups**, it is important to determine whether the data distribution is **parametric** or **non-parametric**.

Parametric data follow a **"normal distribution"**, where the data are spread symmetrically around the midpoint.

Appropriate statistical tests for parametric data include:

- t-tests for continuous data
- chi-squared tests for categorical data, and
- analysis of variance (ANOVA) for repeated measures from the same population.

For data that does not follow a "normal distribution" and are skewed towards one end of the bell curve, **non-parametric tests** are required, such as:

- Mann-Whitney U test
- Wilcoxon signed-rank test
- Kruskal-Wallis test
- Friedman's test for repeated measures. The Friedman test is an extension of the Wilcoxon signed-rank test and the non-parametric analog of one-way repeated-measures.

To check for normality, visual inspection methods like histograms or Q-Q plots can be employed alongside statistical tests such as the Kolmogorov-Smirnov or Shapiro Wilk tests.

For comparing paired data which follow a normal distribution, for example pre- and post-treatment scores, **paired t-tests** can be utilised.

If data are found to be non-normally distributed, non-parametric alternatives such as the Mann-Whitney U test for two independent samples, or the Wilcoxon signed-rank test for paired data, should be used. **Repeated measures** in a longitudinal model can use regression analysis (linear or multiple), while studies comparing different groups can use ANOVA tests.

Another critical consideration is the assumption of homoscedasticity, which implies that the variances among the groups are equal. This is particularly relevant for ANOVA tests.

Tools like Levene's test can be employed to verify this assumption. If variances are unequal, using adjustments such as Welch's ANOVA can provide a more accurate analysis.

Independence of observations is a fundamental requirement for most statistical tests. In PRMs studies, where data might be collected at multiple time points from the same patients, this assumption could be violated, leading to clustered or correlated data. In such cases, techniques such as **mixed-effects models** or **generalised estimating equations** (population averaged) can be applied to accommodate the correlation within subjects over time.

When conducting statistical analysis on PRMs data, it is essential to understand the assumptions underlying each statistical test.

Incorporating advanced statistical methods to address assumption violations:

To ensure robust and reliable results in PRMs data analysis, researchers must be prepared to use advanced statistical methods when standard assumptions are not met. For example, transformation of data (e.g., log transformation) might be necessary to achieve normality or homogeneity of variances. Alternatively, robust statistical methods, which are less sensitive to assumptions, can be used.

Additionally, in the analysis of repeated measures data or when dealing with multiple testing issues, corrections for multiple comparisons (such as Bonferroni or Benjamini-Hochberg procedures) should be considered to control the family-wise error rate or the false discovery rate, respectively. It is highly recommended that statistician input be sought for these advanced analysis methods.

Other factors to consider when analysing the data:

- Comparative sizes of provider groups for cohort analysis

- Potential confounders which may lead to selection bias. For example, if you were concerned about age as a confounder, you would "control for" the effect of age in your statistical modelling. In these instances, the PRMs data may need to be adjusted for confounders if possible (this would require the collection of other relevant variables).

APPENDIX B:
Best Practices for Data Visualisation

To maximise the effectiveness of presenting PRM data visually, it is important to adhere to established best practices in data visualisation. These practices ensure that the information is accessible, interpretable, and actionable for all stakeholders, including clinicians, patients, and policymakers[57].

Consistency in Design

Use a consistent design language across all visualisations (e.g., colour schemes, typography, and layout) to aid in recognition and understanding. Consistency helps in building a coherent narrative across different sets of data.

Clarity and Simplicity

Strive for simplicity by avoiding cluttered visuals. Ensure that each element of the visualisation serves a clear purpose. Use white space effectively to enhance readability.

Accessibility

Design the visualisations with accessibility in mind to ensure that everyone, including individuals with disabilities, can understand and interact with the data. This includes using colour-blind friendly palettes and providing text equivalents for all visual information.

Interactivity

Where possible, incorporate interactive elements that allow users to engage with the data actively. Features such as hover details, drill-down capabilities, and filters can enable users to explore the data in ways that are most relevant to their needs.

Narrative Integration

Integrate visual elements with narrative text to guide the viewer through the data, providing context and highlighting key insights. A well-crafted narrative can significantly enhance the impact of the data presented.

Feedback and Evaluation

Provide mechanisms for feedback on visualisations to allow for continuous improvement. Regular evaluation of how effectively the visual data communicates information to different stakeholders will help refine and optimise the presentation of these data.

MONASH University

MONASH
PUBLIC HEALTH AND
PREVENTIVE MEDICINE

CRICOS PROVIDER: MONASH UNIVERSITY 00008C
DESIGNED BY MONASH PRINT AND DESIGN SERVICES: 505604