



Special Communication

Selecting and Implementing Patient-Reported Outcome and Experience Measures to Assess Health System Performance

Claudia Bull, BNutr (Hons); Helena Teede, PhD, MBBS; Diane Watson, PhD, MBA; Emily J. Callander, PhD

Abstract

IMPORTANCE Psychometrically robust patient-reported outcome measures (PROMs) and patient-reported experience measures (PREMs) are critical to evaluating quality and performance across health services and systems. However, the adoption and implementation of PROMs and PREMs remain a challenge in many countries. The aim of this guide is to support instrument selection and implementation to measure health system performance.

OBSERVATIONS The guide is split into 3 step-by-step sections. Step 1: Knowing What to Measure discusses what PROMs and PREMs capture and how they differ from related instruments. Step 2: Choosing the Right Instrument describes the critical psychometric properties of validity, reliability, and responsiveness, and provides resources to support instrument selection and evaluation. Step 3: Mitigating Potential PROM and PREM Implementation Barriers outlines key barriers and supports for instrument implementation at system, service, and individual levels.

CONCLUSIONS AND RELEVANCE This guide aims to provide practical resources for the identification of psychometrically robust PROMs and PREMs, as well as support for their implementation to drive improvements across health systems globally.

Author affiliations and article information are listed at the end of this article.

JAMA Health Forum. 2022;3(4):e220326. doi:10.1001/jamahealthforum.2022.0326

Introduction

The systematic and routine use of rigorously developed patient-reported outcome measures (PROMs) and patient-reported experience measures (PREMs) is heralded as critical to evaluating quality and performance across health services and systems.¹⁻⁵ Use of PROMs and PREMs captures a multitude of health, well-being, and experiential dimensions,^{6,7} supporting the provision of person-centered, value-based health care.⁸⁻¹² However, the adoption and implementation of PROMs and PREMs remain a challenge in many countries. Furthermore, while there have been substantial advancements in the development and psychometric evaluation of PROMs and PREMs in recent decades,^{13,14} using patient data to drive quality improvement and system changes is still variable and ambiguous.⁵

The aims of this guide are 2-fold: first, to provide practical support for the identification of appropriate and psychometrically sound PROMs and PREMs, and second, to support PROM and PREM implementation efforts by highlighting potential barriers and how they may be mitigated. This guide follows a step-by-step approach (**Figure**).

Step 1: Knowing What to Measure

PROMs

Patient-reported outcome measures are standardized, validated instruments that are completed by patients to measure their health and well-being.^{7,15} They can capture a wide range of outcomes,

Open Access. This is an open access article distributed under the terms of the CC-BY License.

including physical functioning, social functioning, psychological well-being, symptom severity, disability, and impairment.^{6,7,16} Use of PROMs is common to assess changes in outcomes after the implementation of a new intervention (eg, therapy, service, policy), allowing inferences to be made about effectiveness and safety.^{6,17} They are also used to monitor health conditions of individuals and cohorts over time, as well as the effects of treatments longitudinally.⁴ They can be used in isolation or linked with other sources of information about individuals and cohorts.

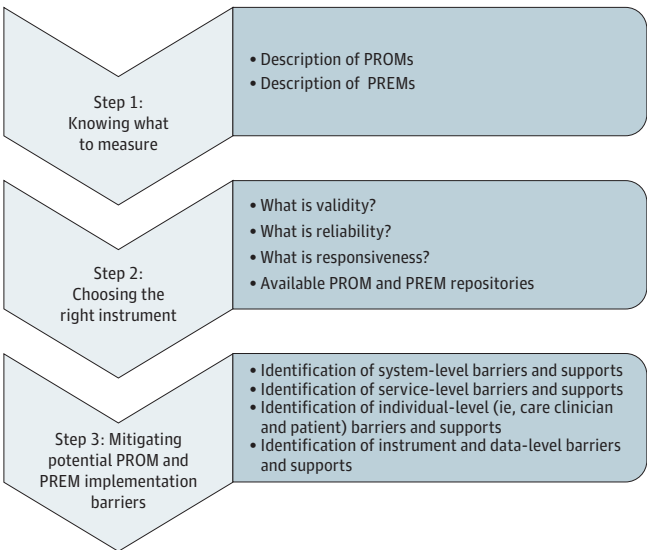
Patient-reported outcome measures can be either generic or condition specific. Generic PROMs are designed to apply to a broad range of patients, conditions, and treatments, capturing aspects of health and well-being that are applicable to everyone. Examples include the 12-item Short Form Health Survey (SF-12), 36-item Short Form Health Survey (SF-36), and the PROMIS (Patient-Reported Outcome Measurement Information System) adult global health measure. Alternatively, condition-specific PROMs apply to specific groups of patients relative to their health condition or the procedure they are undergoing (eg, Oxford Knee Score, Oxford Hip Score).

Despite being referred to and used synonymously in literature and research,¹⁸⁻²¹ PROMs differ from quality of life (QoL) and utility measures. Quality of life and utility measures are preference-based measures primarily used in the context of cost-utility analysis (a form of economic evaluation).²² They comprise 2 components: (1) the measure itself, and (2) a valuation system.²² While QoL and utility measures may capture similar health concepts to those captured by PROMs (eg, physical functioning), they are operationalized (put to use) differently; PROMs are typically scored on an item-by-item or domain (subscale) basis, based on the scale structure determined through psychometric analyses, whereas an individuals' QoL is scored as one of a finite number of health states.²² The QoL valuation system then enables a number between -1 and +1 to be assigned to the individuals' health state, subsequently representing the value of that health state (where 1 indicates full health, 0 indicates death, and <0 indicates health states worse than death).²² Examples of QoL and utility measures include the EQ-5D (EuroQol 5 dimensions), the SF-6D (Short Form 6 dimensions), and the HUI (Health Utility Index).

PREMs

Patient-reported experience measures are instruments that capture a patient's experience of receiving care,^{7,23} specifically the patient's perception of what happened during their care encounter and how it happened.¹³ Several researchers describe patient experiences broadly in terms of

Figure. Stepwise Approach to Supporting Patient-Reported Outcome Measure (PROM) and Patient-Reported Experience Measure (PREM) Selection and Implementation



relational and functional aspects of care.²⁴⁻²⁶ Relational aspects can include the provision of emotional and psychological support, being treated with respect and dignity, being involved in decisions about care and treatment, receiving support for family and carer involvement, the provision of clear and comprehensive information, and transparent and honest communication.²⁴ Functional aspects may include effective and timely treatment, expert management of physical symptoms, attention to physical support needs (eg, wheelchairs and blankets), attention to environmental needs (eg, clean, safe, and comfortable care environments), and the coordination and continuity of care.²⁴

Patient-reported experience measures are primarily used to describe one-off patient experiences across a range of health care contexts and to demonstrate trends in experience scores within health services and systems over time. Notably, PREMs differ from patient satisfaction measures, despite being referred to interchangeably throughout the literature. Where PREMs ask patients to provide a report of their care experience (objective), patient satisfaction measures ask patients to evaluate their care experience (subjective).^{23,27,28} Evaluating involves the concepts of duty and culpability.²⁷ That is, when evaluating a service, patient responses are influenced by their expectations of what a service should and should not do for them (duty) and whether the service is to blame when things that should not happen do happen, or when things that should happen do not happen (culpability).²⁷ Thus, responses to satisfaction questionnaires are more likely to reflect patient expectations, attitudes of appreciation, and social acceptability^{27,29,30} as opposed to an objective report of what they experienced during their care encounter. There are also differences in the response categories used in PREMs and satisfaction measures. Where PREMs typically employ frequency-based response scales (eg, never, sometimes, often, always),³¹⁻³³ patient satisfaction measures use agreement-based response scales (eg, strongly disagree, disagree, neither agree nor disagree, agree, strongly agree).³⁴⁻³⁶ Agreement-based scales are criticized for being biased by acquiescence (tendency to agree with an item irrespective of what is being asked³⁷) and straightlining (tendency to give identical or near identical responses to consecutive questions³⁸). For these reasons, patient experience data are argued to be more actionable for evaluating health care quality.

Thus, a critical component of implementing PROMs and PREMs is understanding whether users want to capture patient outcomes of care (PROMs), experiences of care (PREMs), QoL (QoL measures), or satisfaction with care (patient satisfaction measures). The decision of what to measure should be based on the goal or aim of the study or quality improvement project. Additionally, the selection of an instrument(s) should ideally be driven by key stakeholders, such as patients and health care professionals, relative to their priorities for optimizing person-centered, value-based health care.

Step 2: Choosing the Right Instrument

Psychometrics is a branch of psychology that seeks to measure any number of behavioral and social phenomena using rigorous statistical methods.³⁹ Validity, reliability, and responsiveness are key psychometric properties that provide an assessment of instrument quality, thereby supporting the quality of the resultant information collected.

Validity

Validity is defined as the extent to which an instrument measures what it was designed to measure.⁴⁰ The American Psychological Association recognizes 3 overarching types of validity.⁴¹

Criterion validity is the extent to which an instrument correlates with a "gold standard."⁴² There are 2 types of criterion validity: concurrent criterion validity (ie, the extent of correlation between an instrument and gold standard comparator administered at the same time) and predictive criterion validity (ie, the extent of correlation between a current instrument and gold standard comparator administered in the future). Criterion validity is criticized, however, for the general lack of gold

standard instruments to compare against. The Consensus-Based Standards for the Selection of Health Measurement Instruments (COSMIN) group notes that an exception to this general lack of a gold standard is when a short version of an instrument is compared with a longer or existing version⁴³ (eg, comparing the SF-12 to the SF-36, where the SF-36 is considered the gold standard).⁴⁴ Thus, it can be difficult for instruments to demonstrate criterion validity.

Construct validity is the extent to which an instrument behaves as expected relative to the construct being measured.⁴⁵ For example, it is expected that a patient's overall experience of emergency department care is going to be positively correlated with their outcomes of care (ie, better outcomes correlate with better experiences) but negatively correlated with the duration of their wait time (ie, a shorter wait correlates with a better experience). Thus, a PREM with good construct validity would demonstrate this type of relationship because it aligns with the hypotheses of how these variables correlate.

Content validity is the extent to which the content of an instrument adequately reflects the theoretical construct(s) being measured.⁴³ This is considered to be the most important property of an instrument because it reflects whether items are relevant, comprehensive, and comprehensible to the target population. Given that the target population of PROMs and PREMs is patients (irrespective of the health care context), the involvement of patients in the development of PROM and PREM items is essential. Construct and face validity are sometimes referred to interchangeably in the literature.⁴⁶ However, face validity is an informal assessment of whether items look (on face value) to capture the intended theoretical construct(s)^{43,47} and should not be used in isolation or as a surrogate for content validity. Thus, it is critical that PROMs and PREMs demonstrate sound levels of validity because this ensures that they are accurately measuring the construct(s) under investigation.

Reliability

Reliability refers to the extent to which an instrument performs in consistent and predictable ways.⁴⁸ There are 3 main types of reliability.

Test-retest reliability is the extent that participant responses remain stable over consecutive data collections in relation to the same event (construct stability).^{40,49} Undertaking this type of repeated measurement is likely to produce different results unrelated to the instruments' level of reliability (eg, random error, subject inconsistency in scoring).⁴² As such, a shorter time frame between instrument administrations is likely to be more suitable for producing a robust test-retest reliability score. It is well documented that patient experiences deteriorate over time because of an individual's changing expectations, exposure to similar stimuli (eg, other health care experiences), and misremembering the events that occurred.⁵⁰⁻⁵⁴ Indeed, some researchers purposefully refrain from assessing PREM test-retest reliability for this reason,⁵⁵ and it is one of the least commonly assessed PREM psychometric properties.¹³ Thus, neither PROMs nor PREMs should be discounted if they do not demonstrate sound test-retest reliability.

Internal consistency reliability is the degree of interrelatedness among items of the same scale (unidimensional instrument) or subscale (multidimensional instrument).^{43,46} Cronbach α is most frequently used to indicate an instrument's internal consistency.¹³ Additionally, factor analysis (commonly used to demonstrate construct validity) is another means of supporting internal consistency assessment because items within a PROM or PREM subscale (assuming the instrument is multidimensional) should correlate strongly as they are measuring the same latent (unobservable) trait.^{46,48}

Measurement error represents the extent of systemic and random error in respondent scores that is not attributable to true changes in the construct being measured.⁴⁶ Measurement error may also be described as noise (variance in scores that is owing to randomness), where we are more interested in the signal (variance in true scores).^{48,56} Adequate sample sizes in the development and psychometric evaluation of PROMs and PREMs are critical for reducing measurement error because a greater number of responses reduces randomness in the results.⁵⁶

Responsiveness

Responsiveness is considered an aspect of construct validity and refers to the ability of an instrument to detect change over time where change has occurred.⁵⁷ This is commonly reported in terms of clinically important change(s) to the construct being measured.⁴⁶ However, PROMs and PREMs do not necessarily need to demonstrate responsiveness if used in a discriminatory capacity (eg, to detect differences in experiences between different health care professionals) but should evidence responsiveness if used in an evaluative capacity (eg, evaluating the effect of a new intervention).^{46,58} Notably, clinical importance is different from statistical significance. Where statistical significance ($P < .05$) indicates the probability of results being due to chance, clinical significance relates to the magnitude of a treatment effect and if the effect is meaningful.⁵⁹ Results that are clinically important are those most likely to change current practice. While there are several articles providing guidance on responsiveness and clinically important differences for PROMs and QoL or utility measures,^{60,61} the same cannot be said for PREMs. This is an area of PREM psychometric evaluation that warrants greater investigation; determining a clinically important difference in patient experience scores may present an alternative means to supporting value-based performance (VBP) programs, where incentives are based on meaningful changes in scores, not just improvements on previous scores.

Additional Considerations

Floor and ceiling effects occur when a substantial proportion (15%-20%) of participant scores for an item cluster at the bottom end (floor) or top end (ceiling) of a response scale.⁶² The presence of floor and ceiling effects makes it difficult to distinguish differences between respondents at respective ends of the response scale, thereby affecting the sensitivity of the instrument. If floor and ceiling effects occur for a number of items within a PROM or PREM, this may indicate that the instrument lacks variability, suggestive of inadequate content validity.⁶³

Cross-cultural adaptation refers to the process whereby a measure developed in one language or culture is adapted to be equivalently meaningful and applicable in another language or culture.⁶⁴ The adaptation process is supported by *cross-cultural validation* whereby the performance of translated items are assessed to determine whether they are an accurate reflection of the items in the original instrument.⁶⁴ Best practice guidance by Guillemin et al⁶⁵ for cross-cultural adaptation of instruments suggests 5 sequential steps be followed when translating instruments, including:

1. Forward translation of items into new language by qualified translators
2. Back translation of items into existing language to determine consistency of content
3. Convene a stakeholder committee to compare the original and new instrument versions
4. Undertake pretesting with both the original and new instrument to check for content equivalence
5. Consider weighting items depending on the cultural context (ie, there may be items that are more important in certain contexts than the original)

An adapted measure is considered to demonstrate cross-cultural validity when respondents from different groups (eg, English speaking and Spanish speaking) respond similarly to items of respective instruments as evidenced through regression analyses, confirmatory factor analysis, or non-differential item functioning.⁶⁶

Supporting Robust PROM and PREM Selection

There are several resources that can support the selection of psychometrically robust PROMs and PREMs. **Table 1**^{8,13,14,67-74} presents a comprehensive list of repositories that include PROMs and PREMs for a variety of health care contexts. Many of these also provide information on the extent of psychometric evaluation that instruments have undergone.

Another important resource to support robust PROM and PREM selection is COSMIN, an initiative that seeks to improve both the selection and development of PROMs and other health outcome measurement instruments.⁷⁵ Given the general lack of PREM-specific guidance, COSMIN is

also used to support the development and psychometric evaluation of PREMs.^{13,76,77} This also highlights the impetus for PREM-specific guidance. The COSMIN guidance can support users to assess the psychometric quality of an instrument where no formal assessment has been documented or to support the development and psychometric evaluation of a new instrument. Despite numerous useful resources to support the identification and selection of psychometrically sound PROMs and PREMs, uniform adoption and implementation is challenged by a range of barriers.

Table 1. Repositories of PROMs and PREMs With Established Validity and Reliability

| Source | Description | How it can support PROM/PREM selection |
|--|--|--|
| Churrua et al, ¹⁴ 2021 | Authors identified 315 generic and condition-specific PROMs with established validity and reliability ^a | Provides information on the PROMs identified, including the clinical settings in which they are used, where they were developed, and the extent to which they have been validated |
| ACSQHC 2019 List of generic PROMs ⁶⁷ | ACSQHC Compiled a list of valid generic PROMs whose validation analyses were conducted on an English-language version of the instrument (either the original validation paper or subsequently) | ACSQHC provides information about each PROM, including: <ul style="list-style-type: none"> • Short summary of the instrument • Domains/dimensions captured • Self-reported or other • Validation • Reference to original validation paper(s) • Licensing or restrictions on PROM use • Fees applicable for reuse • Contact information regarding licensing • Published papers and reports using the instrument • Other relevant information (eg, translation into other languages) |
| ACSQHC 2019 List of condition-specific PROMs ⁶⁸ | ACSQHC Compiled a list of valid condition-specific PROMs (based on the AIHW list of chronic diseases and burden of diseases data) whose validation analyses were conducted on an English-language version of the instrument (either the original validation paper or subsequently) | |
| Bull et al, ¹³ 2019 | Authors identified 88 PREMs and assessed instrument validity and reliability using a revised version of the COSMIN checklist | Provides information about the PREMs identified (eg, where they were developed, number of items, patient experience domains captured) and scores each PREM against the COSMIN criteria, identifying which criteria were met, not met, or not assessed |
| BHI Patient Surveys ⁶⁹ | BHI Undertakes regular patient experience surveying, including among admitted adults and children/youth; those treated in the emergency department or outpatient care; rural adults admitted and/or treated in the emergency department; patients receiving maternity, cancer, or virtual care; Aboriginal patients; and patients with long-standing health conditions ⁶⁹ | The BHI website includes all current patient experience surveys used to monitor and report on health system performance, development reports to describe the rigor of each survey, and technical supplements that describe sampling, weighting, standardization, suppression protocols, and other methods |
| de Silva, ⁸ 2014 | The author identified 160 PREMs and person-centered care measures targeting a range of patients (including children) and health care professionals | The downloadable spreadsheet reports information such as whether instruments were developed or well tested, or tested in samples >300 |
| Grid-Enabled Measures Database ⁷⁰ | The Grid-Enabled Measures Database contains more than 1400 behavioral, social science, and other instruments relevant to a range of health conditions and services ^b | Each instrument is briefly described, the construct it captures is identified, and it is either available for download or instructions on where to access it are provided; references to the instruments' development/psychometric evaluation are provided |
| HealthMeasures ⁷¹ | HealthMeasures hosts all of the PROMIS, Neuro-QoL, NIH Toolbox, and ASCQ-Me instruments (>600 measures) | All instruments are freely downloadable and in a ready-to-use format; information is also provided about scoring options for instruments, administration platforms they can be used with, and the difference between measures |
| Person Centred Coordinated Care ⁷² | The Person Centred Coordinated Care website is a compendium of PROMs and PREMs that can be used to support the delivery or evaluation of person-centered and coordinated care programs | The full database provides information for >340 PROMs and PREMs, including a description of what they measure, year and country of development, links to original publications, the domains captured by the instrument, and the psychometrics underpinning instruments |
| PROQOLID ⁷³ | PROQOLID is a comprehensive online database of clinical outcome assessments for the purpose of supporting successful patient-centered outcomes research | The full database contains >2300 instruments. Free access members can obtain basic information about instruments such as the name of authors/developers, copyright, the instruments' objective(s), target therapeutic area and indication, type of instrument, original language, and associated references for describing the instrument. By paying access, members can obtain greater detail about instruments, including documentation related to the instruments' development and psychometric properties, scoring, available translations, and e-versions of instruments |
| Standard Sets of ICHOM ⁷⁴ | ICHOM Develops Standard Sets (ie, standardized outcome sets with associated PROMs to be administered at specific time points to support person-centered outcome measurement) | Each standard set comprises validated PROMs with supporting licensing information |

Abbreviations: ACSQHC, Australian Commission for Safety and Quality in Health Care; AIHW, Australian Institute of Health and Welfare; ASCQ-Me, Adult Sickle Cell Quality of Life Measurement Information System; BHI, Bureau of Health Information; COSMIN, Consensus-Based Standards for the Selection of Health Measurement Instruments; ICHOM, International Consortium for Health Outcomes Measurement; Neuro-QoL, Quality of Life in Neurological Disorders; NIH, National Institutes of Health; PREM, patient-reported experience measure; PROM, patient-reported outcome measure; PROMIS, Patient-Reported Outcome Measurement Information System.

^a Authors only included PROMs if instruments were "validated, that is, there was published statistical analyses establishing reliability (eg Cronbach's alpha) and validity of the scale(s), including construct validity (eg factor analyses, item-response modelling, convergent and discriminant validity), criterion-related validity (eg concurrent and predictive validity), or analyses of known group differences."^{14(p1017)}

^b Requires the user to create a free account to access instruments.

Step 3: Mitigating Potential PROM and PREM Implementation Barriers

The implementation of PROMs and PREMs to measure health system performance is challenged by factors within and across health systems. The Organisation for Economic Co-operation and Development notes that “if each country continues to do its own thing on patient-reported performance, opportunities to identify excellence, support poor performers and drive improvements across the board will be missed.”⁷⁸ Among the key recommendations for international improvement in health system performance measurement, the Organisation for Economic Co-operation and Development suggests that where valid PROMs and PREMs do not yet exist for priority diseases, sectors, or services, new measures should be developed.⁷⁸ They also highlight the importance of formally assessing and piloting PROMs and PREMs in different languages and settings to ensure that they are rigorously developed and tested for the purposes of international comparison.⁷⁸ However, given that implementation challenges are rife at the microlevel (eg, at the service level), it may be more appropriate to consider the barriers and supports for PROM and PREM implementation within the context of a single system.

Table 2^{15,79-89} summarizes a range of barriers and supports for PROM and PREM implementation at the system, service, and individual levels. There are also challenges associated with some instruments themselves, such as their burdensome length,⁸² lack of sensitivity (eg, poor ability to discriminate results between different services and patient groups), and susceptibility to floor and ceiling effects.⁷⁹ However, these are well documented elsewhere in the literature and will not be addressed here.

It is important to note that changes occurring at any one level are likely to have a ripple effect on other levels. Critically, this is important for shifting system cultures and priorities toward understanding the importance of PROMs and PREMs for value-based health care. However, there is also an important balance to be struck between producing results that are tailored to individual services yet supportive of initiatives at the system level.

System Level

A key barrier at the system level includes cultural resistance toward substantial advancements in person-centered, value-based health care. System-level priorities that align with person-centered and value-based health care necessitate a focus on PROMs and PREMs because the principles underpinning these frameworks champion consumer outcomes, values, preferences, and perspectives.^{90,91} Thus, embedding these priorities in national health care policy, funding, safety and

Table 2. Barriers and Supports for PROM and PREM Implementation at the System, Service, and Individual Levels

| Barriers | Supports |
|---|--|
| System level | |
| Resistance to change from pay for service to pay for performance ^{79,80} | Mandating value-based funding and pay-for-performance models of care |
| Lack of PROMs/PREMs appropriate for system-level performance measurement | Development and implementation of nationally mandated PROM/PREM programs |
| Lack of performance benchmarking frameworks inclusive of patient-reported indicators (eg, Hospital Value-Based Purchasing Program in the US) | Development and implementation of value-based funding programs for public health care services |
| Service level | |
| PROM/PREM implementation superseded by other organizational priorities ⁷⁹⁻⁸¹ | Integration of PROM/PREM implementation into service strategic priorities ⁸⁵ |
| Lack of or poorly designed information technology infrastructure that fails to support PROM/PREM implementation or real-time feedback ^{79,82-84} | Integration of PROM/PREM data into existing information technology infrastructure (ie, solutions tailored to individual services) ^{82,84,86,87} |
| High costs associated with PROM/PREM implementation ^{79-81,84,85} | Funding to support robust PROM/PREM implementation and quality improvement ^{84,86} |
| Individual level | |
| Staff lack the skill and power to implement PROMs/PREMs, undertake data analysis, develop reports, disseminate findings, and engage in quality improvement ^{79,82-85,88} | Outsourcing PROM/PREM implementation, data analysis, and interpretation to minimize staff burden ⁷⁹ or supporting staff training relative to data analysis, results interpretation, and quality improvement methodologies ^{79,83,89} |
| Staff skepticism (eg, questioning the reliability and validity of PROM/PREM data, questioning the value and purpose of implementation) ^{79,82-85} | Relaying PROM/PREM scores to staff in an accessible and meaningful manner ^{15,79,83,85,89} and improving transparency of PROM/PREM psychometric quality through staff education initiatives ⁸⁵ |
| Concerns that PROM/PREM data are not representative of specific instances of care but conflated by patients’ opinion of staff and the service as a whole ⁸⁵ | Presenting PROM/PREM scores in both aggregate (national and state level) and individualized (service and staff level) forms ^{79,85,86,89} |

Abbreviations: PREM, patient-reported experience measure; PROM, patient-reported outcome measure.

quality standards, accreditation expectations, and service agreements in meaningful ways will have flow-on effects for establishing system-level PROM and PREM implementation programs and benchmarking.

Service Level

Commonly cited barriers for PROM and PREM implementation at the service level include a lack of leadership^{79,83,85} and information technology infrastructure to support PROM and PREM implementation.^{79,82,84} Additionally, implementation is costly,^{79-81,84,85} further compounding other recognized service-level barriers. Strong and visible support for PROM and PREM implementation at executive and senior levels, as well as the integration of implementation in service-level strategic priorities, is critical for fostering cultures engaged in PROM and PREM implementation.^{79,80,82,84,85} Integrating PROM and PREM data collection and feedback mechanisms within services' existing information technology infrastructure will also support implementation, as well as reduce costs and the burden on health care professionals.^{82,84,86,87}

Individual Level

Implementation can be hugely burdensome on health care professionals in terms of the volume of data they are asked to collect from patients,^{80,83} as well as the amount of time they are required to invest in the analysis and interpretation of PROM and PREM feedback.^{79,82-85,88} Skepticism arises when health care professionals question the reliability and validity of PROM and PREM data and the value and purpose of instrument implementation.^{79,82-85} There are also patient-related barriers, including burden,^{82,84} and the view that PROM and PREM data are only useful for health care professionals (not themselves).^{82,88} Thus, potential supports include outsourcing PROM and PREM implementation, data analysis, and interpretation to third parties to minimize the burden on health care professionals. This presently occurs in the US and England^{92,93} but needs to be considered relative to a services' desire to promote local ownership of PROM and PREM implementation. Additionally, the interpretation and presentation of PROM and PREM feedback to health care professionals is viewed as more useful when provided alongside other important metrics (eg, service-related data such as wait times, patient importance scores, qualitative case examples) and adjusted for confounding patient variables.^{15,79,85,89}

Examples of System-Level PROM and PREM Programs

Two examples of successful system-level PROM and PREM programs include (1) the English National Health Service (NHS) PROMs program and (2) the US Consumer Assessment of Healthcare Providers and Services (CAHPS) program. The NHS PROMs program was established in 2009 and saw PROMs administered preoperatively and postoperatively to all patients undergoing hip and knee replacement across NHS-funded services.⁹³ The purpose of the program was to gauge the extent to which patients recovered or deteriorated in the 6 months after surgery.⁹³ The PROM data are summarized and publicly reported on NHS Digital annually,⁹⁴ supporting service benchmarking and the National Tariff Payment System,⁹⁵ a value-based performance scheme that offers bonus payments to health care professionals who provide quality care in line with risk-adjusted national average improvements in patient health status.^{96,97}

The US CAHPS program is arguably the largest PREM program internationally. Since launching in 1995, close to 20 CAHPS measures have been implemented across Medicare and Medicaid services in the US.⁹² These PREMs have been increasingly used to support health plan accreditation, consumer choice of health care professionals (through public reporting websites such as Care Compare⁹⁸ and Plan Finder⁹⁹), services and health plans, and VBP programs. Arguably the best-known VBP program is the Hospital VBP (HVBP) Program that uses PREM data from the Hospital CAHPS. This program adjusts payments to hospitals under the Inpatient Prospective Payment System based on the quality of care delivered to patients.¹² The HVBP Program is funded by a 2% reduction in participating hospitals' Medicare severity diagnosis-related group payments for the

fiscal year. This pool of funding is then redistributed to hospitals relative to their Total Performance Score, which is based on 4 equally weighted quality domains: (1) clinical outcomes, (2) person and community engagement (informed by Hospital CAHPS), (3) safety, and (4) efficiency and cost reduction.¹² The HVBP Program made roughly \$1.9 billion in incentive payments during the 2020 financial year.¹⁰⁰

Thus, while the implementation of PROMs and PREMs at the system level can be challenging, it is also clear that these types of programs can be successful. Learnings from the English and US programs should be duly considered as we look toward the future of PROMs and PREMs in value-based health care.

Conclusions

Despite important progress in the development and psychometric evaluation of PROMs and PREMs over recent years, there remains challenges to their implementation. Moreover, learnings from countries that have successfully integrated PROMs and PREMs into health system performance measurement are yet to be effectively transferred to countries that are further behind. While it is clear that changes at the system, service, and individual levels need to occur to support PROM and PREM implementation, these efforts will be critical to driving person-centered, value-based health care globally.

ARTICLE INFORMATION

Accepted for Publication: February 2, 2022.

Published: April 1, 2022. doi:10.1001/jamahealthforum.2022.0326

Open Access: This is an open access article distributed under the terms of the [CC-BY License](#). © 2022 Bull C et al. *JAMA Health Forum*.

Corresponding Author: Claudia Bull, BNutr (Hons), School of Public Health and Preventive Medicine, Monash University, 553 St Kilda Rd, Melbourne, VIC, 3141, Australia (claudia.bull@monash.edu).

Author Affiliations: School of Public Health and Preventive Medicine, Monash University, Melbourne, Victoria, Australia (Bull, Callander); Monash Centre for Health Research and Implementation, Monash University, Clayton, Victoria, Australia (Bull, Teede, Callander); Monash Partners Academic Health Science Centre, Clayton, Victoria, Australia (Teede); Bureau of Health Information, New South Wales Health, St Leonards, New South Wales, Australia (Watson).

Author Contributions: Ms Bull had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Concept and design: Bull, Callander.

Acquisition, analysis, or interpretation of data: Bull, Teede, Watson.

Drafting of the manuscript: Bull, Callander.

Critical revision of the manuscript for important intellectual content: All authors.

Statistical analysis: Bull.

Administrative, technical, or material support: Bull, Callander.

Supervision: Watson, Callander.

Other—Expertise and guidance in the field: Teede.

Conflict of Interest Disclosures: Profs Teede and Callander reported grants from the Australian National Health and Medical Research Council during the conduct of the study. No other disclosures were reported.

REFERENCES

1. High quality care for all: NHS next stage review final report. Department of Health. June 2008. Accessed February 27, 2022. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/228836/7432.pdf

2. Partnering with consumers standard. Australian Commission on Safety and Quality in Health Care. Accessed February 27, 2022. <https://www.safetyandquality.gov.au/standards/nsqhs-standards/partnering-consumers-standard>
3. About the NHS Outcomes Framework (NHS OF). National Health Service. Accessed February 27, 2022. <https://digital.nhs.uk/data-and-information/publications/ci-hub/nhs-outcomes-framework>
4. Patient-centred measurement and reporting in Canada. Canadian Institute for Health Information. 2017. Accessed February 27, 2022. <https://www.cihi.ca/sites/default/files/document/visioning-day-paper-en-web.pdf>
5. Menear M, Blanchette MA, Demers-Payette O, Roy D. A framework for value-creating learning health systems. *Health Res Policy Syst*. 2019;17(1):79. doi:10.1186/s12961-019-0477-3
6. Williams K, Sansoni J, Morris D, Grootemaat P, Thompson C. Patient-reported outcome measures: literature review. Centre for Health Service Development. 2016. Accessed February 27, 2022. <https://www.safetyandquality.gov.au/sites/default/files/migrated/PROMs-Literature-Review-December-2016.pdf>
7. Kingsley C, Patel S. Patient-reported outcome measures and patient-reported experience measures. *BJA Educ*. 2017;17(4):137-144. doi:10.1093/bjaed/mkw060
8. de Silva D. Helping measure person-centred care: a review of evidence about commonly used approaches and tools to help measure person-centred care. The Health Foundation. March 2014. Accessed February 27, 2022. <https://www.health.org.uk/publications/helping-measure-person-centred-care>
9. Patient-centred care: improving quality and safety through partnerships with patients and consumers. Australian Commission on Safety and Quality in Health Care. 2011. Accessed February 27, 2022. <https://www.safetyandquality.gov.au/publications-and-resources/resource-library/patient-centred-care-improving-quality-and-safety-through-partnerships-patients-and-consumers>
10. Withers K, Palmer R, Lewis S, Carolan-Rees G. First steps in PROMs and PREMs collection in Wales as part of the prudent and value-based healthcare agenda. *Qual Life Res*. 2021;30(11):3157-3170. doi:10.1007/s11136-020-02711-2
11. What are value-based programs? Centers for Medicare & Medicaid Services. Updated December 1, 2021. Accessed February 27, 2022. <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/Value-Based-Programs>
12. The Hospital Value-Based Purchasing Program. Centers for Medicare & Medicaid Services. Updated December 1, 2021. Accessed February 27, 2022. <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/HVBP/Hospital-Value-Based-Purchasing>
13. Bull C, Byrnes J, Hettiarachchi R, Downes M. A systematic review of the validity and reliability of patient-reported experience measures. *Health Serv Res*. 2019;54(5):1023-1035. doi:10.1111/1475-6773.13187
14. Churrua K, Pomare C, Ellis LA, 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):1015-1024. doi:10.1111/hex.13254
15. Black N, Burke L, Forrest CB, et al. Patient-reported outcomes: pathways to better health, better services, and better societies. *Qual Life Res*. 2016;25(5):1103-1112. doi:10.1007/s11136-015-1168-3
16. What are PROMs? Centre for Health Services and Policy Research at University of British Columbia. Accessed February 27, 2022. <https://patientreportedoutcomes.ca/what-are-pros/proms/>
17. Black N, Varaganum M, Hutchings A. Relationship between patient reported experience (PREMs) and patient reported outcomes (PROMs) in elective surgery. *BMJ Qual Saf*. 2014;23(7):534-542. doi:10.1136/bmjqs-2013-002707
18. Jonsson Å, Orwelius L, Dahlstrom U, Kristenson M. Evaluation of the usefulness of EQ-5D as a patient-reported outcome measure using the Paretian classification of health change among patients with chronic heart failure. *J Patient Rep Outcomes*. 2020;4(1):50. doi:10.1186/s41687-020-00216-7
19. Benson T, Williams DH, Potts HW. Performance of EQ-5D, howRu and Oxford hip & knee scores in assessing the outcome of hip and knee replacements. *BMC Health Serv Res*. 2016;16(1):512. doi:10.1186/s12913-016-1759-x
20. Patient reported outcome measures (PROMs) in England—2011-2012, special topic, EQ-5D™ Index analysis. NHS Digital. May 8, 2014. Accessed February 27, 2022. <https://digital.nhs.uk/data-and-information/publications/statistical/patient-reported-outcome-measures-proms/patient-reported-outcome-measures-proms-in-england-2011-2012-special-topic-eq-5d-index-analysis>
21. Weldring T, Smith SM. Patient-reported outcomes (PROs) and patient-reported outcome measures (PROMs). *Health Serv Insights*. 2013;6:61-68.
22. Whitehurst DG, Noonan VK, Dvorak MF, Bryan S. A review of preference-based health-related quality of life questionnaires in spinal cord injury research. *Spinal Cord*. 2012;50(9):646-654. doi:10.1038/sc.2012.46

23. Ahmed F, Burt J, Roland M. Measuring patient experience: concepts and methods. *Patient*. 2014;7(3):235-241. doi:10.1007/s40271-014-0060-5
24. Doyle C, Lennox L, Bell D. A systematic review of evidence on the links between patient experience and clinical safety and effectiveness. *BMJ Open*. 2013;3(1):e001570. doi:10.1136/bmjopen-2012-001570
25. Bull C, Latimer S, Crilly J, Gillespie BM. A systematic mixed studies review of patient experiences in the ED. *Emerg Med J*. 2021;38(8):643-649. doi:10.1136/emermed-2020-210634
26. Wood L, Alsawy S. Patient experiences of psychiatric inpatient care: a systematic review of qualitative evidence. *J Psychiatr Intensive Care*. 2016;12(1):35-43. doi:10.20299/jpi.2016.001
27. Williams B, Coyle J, Healy D. The meaning of patient satisfaction: an explanation of high reported levels. *Soc Sci Med*. 1998;47(9):1351-1359. doi:10.1016/S0277-9536(98)00213-5
28. Bull C. Patient satisfaction and patient experience are not interchangeable concepts. *Int J Qual Health Care*. 2021;33(1):mzab023. doi:10.1093/intqhc/mzab023
29. Noyes RW, Levy MI, Chase CL, Udry JR. Expectation fulfillment as a measure of patient satisfaction. *Am J Obstet Gynecol*. 1974;118(6):809-814. doi:10.1016/0002-9378(74)90493-1
30. Williams B. Patient satisfaction: a valid concept? *Soc Sci Med*. 1994;38(4):509-516. doi:10.1016/0277-9536(94)90247-X
31. Hospital Consumer Assessment of Healthcare Providers and Systems hospital survey. Centers for Medicare & Medicaid Services. Updated February 23, 2022. Accessed February 27, 2022. <https://hcahpsonline.org/en/survey-instruments/>
32. Emergency department CAHPS (ED CAHPS). Centers for Medicare & Medicaid Services. Updated December 1, 2021. Accessed February 27, 2022. <https://www.cms.gov/Research-Statistics-Data-and-Systems/Research/CAHPS/ED>
33. Adult Inpatient Survey 2021. Care Quality Commission. Accessed February 27, 2022. <https://nhssurveys.org/surveys/survey/02-adults-inpatients/>
34. Grogan S, Conner M, Willits D, Norman P. Development of a questionnaire to measure patients' satisfaction with general practitioners' services. *Br J Gen Pract*. 1995;45(399):525-529.
35. Kinnerley P, Stott N, Peters T, Harvey I, Hackett P. A comparison of methods for measuring patient satisfaction with consultations in primary care. *Fam Pract*. 1996;13(1):41-51. doi:10.1093/fampra/13.1.41
36. Hojat M, Louis DZ, Maxwell K, Markham FW, Wender RC, Gonnella JS. A brief instrument to measure patients' overall satisfaction with primary care physicians. *Fam Med*. 2011;43(6):412-417.
37. Acquiescence response bias. In: Lavrakas PJ, ed. *Encyclopedia of Survey Research Methods*. Sage Publications; 2008.
38. Dykema J, Schaeffer NC, Garbarski D, Assad N, Blixt S. Towards a reconsideration of the use of agree-disagree questions in measuring subjective evaluations. *Res Social Adm Pharm*. 2022;18(2):2335-2344. doi:10.1016/j.sapharm.2021.06.014
39. Kelderman H, Molenaar P, Borsboom D, Thissen D. What is psychometrics? Psychometric Society. Accessed February 27, 2022. <https://www.psychometricsociety.org/what-psychometrics>
40. Babbie ER. *The Practice of Social Research*. Cengage Learning; 2015.
41. APA dictionary of psychology: validity. American Psychological Association. Accessed February 27, 2022. <https://dictionary.apa.org/validity>
42. Bellamy N. Principles of clinical outcome assessment. In: Hochberg MC, Silman AJ, Smolen JS, Weinblatt ME, Weisman MH, eds. *Rheumatology*. 6th ed. Mosby; 2015:9-19. doi:10.1016/B978-0-323-09138-1.00002-4
43. Mokkink LB, de Vet HCW, Prinsen CAC, et al. COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. *Qual Life Res*. 2018;27(5):1171-1179. doi:10.1007/s1136-017-1765-4
44. Lin Y, Yu Y, Zeng J, Zhao X, Wan C. Comparing the reliability and validity of the SF-36 and SF-12 in measuring quality of life among adolescents in China: a large sample cross-sectional study. *Health Qual Life Outcomes*. 2020;18(1):360. doi:10.1186/s12955-020-01605-8
45. DeVillis RF. Validity. In: Bickman L, Rog DJ, eds. *Scale Development: Theory and Applications*. 4th ed. SAGE Publications; 2017.
46. Mokkink LB, Terwee CB, Patrick 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):737-745. doi:10.1016/j.jclinepi.2010.02.006
47. Chiang I-CA, Jhangiani RS, Price PC. *Research Methods in Psychology—2nd Canadian Edition*. BCCampus; 2015.

48. DeVellis RF. Reliability. In: Bickman L, Rog DJ, eds. *Scale Development: Theory and Applications*. 4th ed. SAGE Publications; 2017.
49. APA dictionary of psychology: reliability. American Psychological Association. Accessed February 27, 2022. <https://dictionary.apa.org/retest-reliability>
50. Bendall-Lyon D, Powers TL, Swan JE. Time does not heal all wounds: patients report lower satisfaction levels as time goes by. *Mark Health Serv*. 2001;21(3):10-14.
51. Bjertnaes OA. The association between survey timing and patient-reported experiences with hospitals: results of a national postal survey. *BMC Med Res Methodol*. 2012;12:13. doi:10.1186/1471-2288-12-13
52. Festinger L. *A Theory of Cognitive Dissonance*. Stanford University Press; 1957.
53. Jensen HI, Ammentorp J, Kofoed PE. User satisfaction is influenced by the interval between a health care service and the assessment of the service. *Soc Sci Med*. 2010;70(12):1882-1887. doi:10.1016/j.socscimed.2010.02.035
54. O'Neill M, Palmer A. Survey timing and consumer perceptions of service quality: an overview of empirical evidence. *Manage Serv Quality Int J*. 2001;11(3):182-190. doi:10.1108/09604520110391351
55. Christalle E, Zeh S, Hahlweg P, Kriston L, Härter M, Scholl I. Assessment of patient centredness through patient-reported experience measures (ASPIRED): protocol of a mixed-methods study. *BMJ Open*. 2018;8(10):e025896. doi:10.1136/bmjopen-2018-025896
56. Staggs VS, Cramer E. Reliability of pressure ulcer rates: how precisely can we differentiate among hospital units, and does the standard signal-noise reliability measure reflect this precision? *Res Nurs Health*. 2016;39(4):298-305. doi:10.1002/nur.21727
57. Cohen RJ. *Psychological Testing and Assessment: An Introduction to Tests and Measurement*. 8th ed. McGraw Hill; 2013.
58. Holmes C, Briffa N. Patient-reported outcome measures (PROMS) in patients undergoing heart valve surgery: why should we measure them and which instruments should we use? *Open Heart*. 2016;3(1):e000315. doi:10.1136/openhrt-2015-000315
59. Ranganathan P, Pramesh CS, Buyse M. Common pitfalls in statistical analysis: clinical versus statistical significance. *Perspect Clin Res*. 2015;6(3):169-170. doi:10.4103/2229-3485.159943
60. Revicki DA, Cella D, Hays RD, Sloan JA, Lenderking WR, Aaronson NK. Responsiveness and minimal important differences for patient reported outcomes. *Health Qual Life Outcomes*. 2006;4:70. doi:10.1186/1477-7525-4-70
61. Su F, Allahabadi S, Bongbong DN, Feeley BT, Lansdown DA. Minimal clinically important difference, substantial clinical benefit, and patient acceptable symptom state of outcome measures relating to shoulder pathology and surgery: a systematic review. *Curr Rev Musculoskelet Med*. 2021;14(1):27-46. doi:10.1007/s12178-020-09684-2
62. Garin O. Ceiling effect. In: Michalos AC, ed. *Encyclopedia of Quality of Life and Well-Being Research*. Springer Netherlands; 2014:631-633. doi:10.1007/978-94-007-0753-5_296
63. Terwee CB, Prinsen CAC, Chiarotto A, et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. *Qual Life Res*. 2018;27(5):1159-1170. doi:10.1007/s11136-018-1829-0
64. Huang WY, Wong SH. Cross-cultural validation. In: Michalos AC, ed. *Encyclopedia of Quality of Life and Well-Being Research*. Springer; 2014:1369-1371. doi:10.1007/978-94-007-0753-5_630
65. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J Clin Epidemiol*. 1993;46(12):1417-1432. doi:10.1016/0895-4356(93)90142-N
66. Mokkink LB, Prinsen CAC, Patrick DL, et al. COSMIN methodology for systematic reviews of patient-reported outcome measures (PROMs) user manual. COSMIN. February 2018. Accessed February 27, 2022. https://www.cosmin.nl/wp-content/uploads/COSMIN-syst-review-for-PROMs-manual_version-1_feb-2018-1.pdf
67. List of generic PROMs. Australian Commission on Safety and Quality in Health Care. Accessed February 27, 2022. <https://www.safetyandquality.gov.au/list-generic-proms>
68. Condition specific PROMs. Australian Commission on Safety and Quality in Health Care. Accessed February 27, 2022. <https://www.safetyandquality.gov.au/condition-specific-proms>
69. BHI patient surveys. Bureau of Health Information. Accessed February 27, 2022. https://www.bhi.nsw.gov.au/nsw_patient_survey_program
70. Grid-Enabled Measures Database. GEM. Accessed February 27, 2022. <https://www.gem-measures.org/Login.aspx>

71. PROMIS. HealthMeasures. Accessed February 27, 2022. <https://www.healthmeasures.net/explore-measurement-systems/promis>
72. Measures for Person Centred Coordinated Care. Plymouth University. Accessed February 27, 2022. <http://p3c.org.uk/>
73. PROQOLID. Mapi Research Trust. Accessed February 27, 2022. <https://eprovide.mapi-trust.org/about/about-proqolid>
74. Patient-centered outcome measures. International Consortium for Health Outcomes Measurement. Accessed February 27, 2022. <https://www.ichom.org/standard-sets/>
75. Mokkink LB, Prinsen CAC, Terwee CB. About the COSMIN initiative. COSMIN. Accessed February 27, 2022. <https://www.cosmin.nl/about/>
76. Beecher C, Greene R, O'Dwyer L, et al. Measuring women's experiences of maternity care: a systematic review of self-report survey instruments. *Women Birth*. 2021;34(3):231-241. doi:10.1016/j.wombi.2020.05.002
77. Gärtner FR, Bomhof-Roordink H, Smith IP, Scholl I, Stiggelbout AM, Pieterse AH. The quality of instruments to assess the process of shared decision making: a systematic review. *PLoS One*. 2018;13(2):e0191747. doi:10.1371/journal.pone.0191747
78. Recommendations to OECD Ministers of Health from the High Level Reflection Group on the future of health statistics: strengthening the international comparison of health system performance through patient-reported indicators. Organisation for Economic Co-operation and Development. January 2017. Accessed February 27, 2022. <https://www.oecd.org/health/Recommendations-from-high-level-reflection-group-on-the-future-of-health-statistics.pdf>
79. Davies E, Cleary PD. Hearing the patient's voice? factors affecting the use of patient survey data in quality improvement. *Qual Saf Health Care*. 2005;14(6):428-432. doi:10.1136/qshc.2004.012955
80. Wolf JA. The state of patient experience 2021: transforming the human experience. The Beryl Institute. Accessed February 27, 2022. <https://www.theberylinstitute.org/store/viewproduct.aspx?id=18489732>
81. Morton RL, Lioufas N, Dansie K, et al. Use of patient-reported outcome measures and patient-reported experience measures in renal units in Australia and New Zealand: a cross-sectional survey study. *Nephrology (Carlton)*. 2020;25(1):14-21. doi:10.1111/nep.13577
82. Wheat H, Horrell J, Valderas JM, Close J, Fosh B, Lloyd H. Can practitioners use patient reported measures to enhance person centred coordinated care in practice? a qualitative study. *Health Qual Life Outcomes*. 2018;16(1):223. doi:10.1186/s12955-018-1045-1
83. Sheard L, Peacock R, Marsh C, Lawton R. What's the problem with patient experience feedback? a macro and micro understanding, based on findings from a three-site UK qualitative study. *Health Expect*. 2019;22(1):46-53. doi:10.1111/hex.12829
84. Stover AM, Haverman L, van Oers HA, Greenhalgh J, Potter CM; ISOQOL PROMs/PREMs in Clinical Practice Implementation Science Work Group. Using an implementation science approach to implement and evaluate patient-reported outcome measures (PROM) initiatives in routine care settings. *Qual Life Res*. 2021;30(11):3015-3033. doi:10.1007/s11136-020-02564-9
85. Baldie DJ, Guthrie B, Entwistle V, Kroll T. Exploring the impact and use of patients' feedback about their care experiences in general practice settings-a realist synthesis. *Fam Pract*. 2018;35(1):13-21. doi:10.1093/fampra/cmx067
86. Benson T. Measure what we want: a taxonomy of short generic person-reported outcome and experience measures (PROMs and PREMs). *BMJ Open Qual*. 2020;9(1):e000789. doi:10.1136/bmjopen-2019-000789
87. Benson T, Potts HW. A short generic patient experience questionnaire: howRwe development and validation. *BMC Health Serv Res*. 2014;14:499. doi:10.1186/s12913-014-0499-z
88. Nelson EC, Eftimovska E, Lind C, Hager A, Wasson JH, Lindblad S. Patient reported outcome measures in practice. *BMJ*. 2015;350:g7818. doi:10.1136/bmj.g7818
89. Porter I, Gonçalves-Bradley D, Ricci-Cabello I, et al. Framework and guidance for implementing patient-reported outcomes in clinical practice: evidence, challenges and opportunities. *J Comp Eff Res*. 2016;5(5):507-519. doi:10.2217/ce-2015-0014
90. Harding E, Wait S, Scrutton J. The state of play in person-centred care: a pragmatic review of how person-centred care is defined, applied and measured. The Health Policy Partnership. December 2015. Accessed February 27, 2022. <https://www.healthpolicypartnership.com/app/uploads/The-state-of-play-in-person-centred-care-summary.pdf>
91. Porter ME. What is value in health care? *N Engl J Med*. 2010;363(26):2477-2481. doi:10.1056/NEJMp1011024

92. Consumer Assessment of Healthcare Providers & Systems (CAHPS). Centers for Medicare & Medicaid Services. Updated January 4, 2022. Accessed February 27, 2022. <https://www.cms.gov/research-statistics-data-and-systems/research/cahps>
93. Provisional monthly patient reported outcome measures (PROMs) in England: a guide to PROMs methodology. National Healthcare Service. April 2010. Accessed March 2, 2022. <https://files.digital.nhs.uk/publicationimport/pub03xxx/pub03765/prom-data-arch-guid-v1-0804-2010-rep2.pdf>
94. Patient reported outcome measures (PROMs). NHS Digital. Updated February 21, 2022. Accessed February 27, 2022. <https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/patient-reported-outcome-measures-proms>
95. National tariff payment system. National Health Service. Accessed February 27, 2022. <https://www.england.nhs.uk/pay-syst/national-tariff/national-tariff-payment-system/>
96. Gutacker N, Street A, Gomes M, Bojke C. Should English healthcare providers be penalised for failing to collect patient-reported outcome measures? a retrospective analysis. *J R Soc Med*. 2015;108(8):304-316. doi:10.1177/0141076815576700
97. 2020/21 national tariff payment system. National Health Service. November 2020. Accessed February 27, 2022. https://www.england.nhs.uk/wp-content/uploads/2021/02/20-21_National-Tariff-Payment-System.pdf
98. Care Compare. Centers for Medicare & Medicaid Services. Accessed February 27, 2022. <https://www.medicare.gov/care-compare/>
99. How to compare plans using the Medicare Star Rating System. Medicare Interactive. Accessed February 27, 2022. <https://www.medicareinteractive.org/get-answers/medicare-health-coverage-options/changing-medicare-coverage/how-to-compare-plans-using-the-medicare-star-rating-system>
100. CMS Hospital Value-Based Purchasing Program results for fiscal year 2020. Centers for Medicare & Medicaid Services. October 29, 2019. Accessed February 27, 2022. <https://www.cms.gov/newsroom/fact-sheets/cms-hospital-value-based-purchasing-program-results-fiscal-year-2020>