NATIONAL QUALITY FORUM
Driving measurable health improvements together

Building a Roadmap From Patient-Reported Outcome Measures to Patient-Reported Outcome Performance Measures: Environmental Scan

Final Report (Updated)

June 28, 2022

This report is funded by the Centers for Medicare & Medicaid Services under contract HHSM-500-2017-00060I – 75FCMC20F0003.

Contents

Executive Summary

The National Quality Forum (NQF) and the Centers for Medicare & Medicaid Services (CMS) have a long-standing partnership to advance patient-reported outcomes (PROs) as a tool to empower patients and elevate their voices in healthcare quality measurement. This shared commitment has contributed to progress towards the development and use of patient-reported outcome measures (PROMs) and patient-reported outcome performance measures (PRO-PMs). However, PRO-PMs make up less than 7 percent of all NQF-endorsed quality measures, and only one NQF-endorsed PRO-PM is included on the CMS list of Merit-based Incentive Payment System (MIPS) quality measures for 2022.[1,2] These gaps demonstrate the opportunity to increase the number of PRO-PMs that measure what matters to patients in ways that are suitable for use in CMS value-based purchasing (VBP) programs or alternative payment models (APMs).

CMS has heard from measure developers and measurement experts about the lack of detailed technical guidance for developing high impact outcome measures that are based on patient-reported data. Providing this guidance is one strategy to increase the prevalence of PRO-PMs. Additionally, as part of its goal to reduce measurement burden and transform measures to fully digital by 2025, CMS is particularly interested in advancing PRO-PMs that are used in clinical care and stored in the electronic health record (EHR).[3] This CMS-funded initiative, Building a Roadmap From Patient-Reported Outcome Measures to Patient-Reported Outcome Performance Measures (henceforth referred to as *Building a Roadmap*), seeks to lower barriers to the development and use of "digital" PRO-PMs by providing guidance on developing PRO-PMs that utilize data from high quality PROMs, are suitable for use in CMS' VBP programs and APMs, and can be calculated and transmitted electronically.

The Building a Roadmap initiative includes four reports:

- The Environmental Scan Report identifies and summarizes existing information relevant to the use of high quality PROMs as the basis for PRO-PMs in accountability programs. It does not directly address challenges or offer specific guidance to developers, but rather provides background for the other documents in the Building a Roadmap project.
- The Interim Report identifies and describes attributes of PROMs that are suitable data collection instruments for high quality digital PRO-PMs and provides guidance to measure developers on selecting PROMs for use in a PRO-PM.
- The Technical Guidance Report describes a series of stages and tasks that measure developers can follow when developing and testing digital PRO-PMs that are suitable for CMS VBP programs and APMs.
- The Developer Feedback Report (to be released in summer 2022) identifies recommendations from measure developers and other key audience members on improving the Technical Guidance Report.

This report is an updated version of the Environmental Scan Report that NQF first published in spring 2021. Information that NQF retained from the 2021 report includes:

- a discussion of the role of PROMs and PRO-PMs in quality-based models, including CMS' strategic aim to elevate patients' voices through the increased use of interoperable digital measures[4,5];
- a description of existing resources that measure developers can reference when identifying PROMs for use in high quality PRO-PMs;
- a review of guidance to assist with the development of PRO-PMs and digital PRO-PMs; and
- an identification of gaps and challenges that measure developers commonly face when developing PRO-PMs.

This updated version of the Environmental Scan Report, published in spring 2022, contains several additions recommended by CMS and the technical expert panel (TEP), including:

- an expanded discussion of modes of PROM administration and methods of data collection;
- an explanation of gaps between PRO-PMs and other quality measures and domains;
- a more comprehensive view of the implications of interoperability on PRO-PMs; and
- throughout the report, additional links to web sites that are constantly updated with new information on digital quality measurement.

Although the Environmental Scan Report focuses on digital PRO-PMs, its findings are relevant to the development of all PRO-PMs. At the recommendation of NQF, CMS, and the TEP, the Building a Roadmap initiative focuses on patient-reported outcomes (specifically health related quality of life [HRQoL], functional status, and symptoms and symptom burden) as distinct from experience of care measures.

Additional materials related to the Building a Roadmap initiative, including reports and summaries of TEP meetings, are available on the Building a Roadmap project page.

Introduction

The patient voice is essential to healthcare quality measurement, and PRO-PMs are a powerful way to elevate patients' voices. NQF and CMS have a shared commitment to measuring outcomes that are meaningful to patients using data that patients provide. Many barriers hinder the development of PRO-PMs, however, including a lack of robust guidance to assist measure developers. Additionally, an opportunity exists for measure developers to create digital PRO-PMs—in which health IT systems not only collect and share data but also calculate and submit aggregate scores for regulatory and reimbursement purposes—that are based on high quality PROMs. To facilitate measure developers' efforts to develop and test PRO-PMs, other reports in the Building a Roadmap initiative present a list of attributes of high quality PROMs for use in performance measures (the Interim Report) and a roadmap to follow when creating digital PRO-PMs for accountability purposes (the Technical Guidance Report).

The Building a Roadmap initiative supports the development of digital PRO-PMs that are based on high quality PROMs (i.e., PROMs that are suitable to be the foundation of a digital PRO-PM that can be used to evaluate the performance of healthcare entities) and that may be appropriate for CMS VBP programs or APMs. The purpose of this Environmental Scan Report is to document the current state of developing digital PRO-PMs, including available guidance on best practices for development. Digital PRO-PMs

collect outcome data from patients with minimal burden, maximize response rates to PROMs to increase representativeness, and leverage health IT systems for data collection, storage, and measure calculations. Although the Environmental Scan Report focuses on digital PRO-PMs, its findings are relevant to the development of all PRO-PMs.

NQF first published the Environmental Scan Report in 2021, and published this updated edition in spring 2022 to replace the previous report. The updated edition includes several improvements that were identified by the TEP, including an expanded discussion of PROM modes and methods, an explanation of PRO domains and which are within the scope of the Building a Roadmap work, and a more comprehensive view of the current state of digital quality measurement as it pertains to PRO-PMs. The updated Environmental Scan Report includes new information on digital measurement and PRO-PMs that was not available when the original report was published, as well as links to webpages that are constantly updated with new information on digital quality measurement.

NQF's Work on Patient-Reported Outcomes

This section of the Environmental Scan Report briefly describes previously published NQF reports that can inform PROM selection and/or PRO-PM development, the current Building a Roadmap initiative, and terminology that is relevant to performance measures. This information helps the reader to understand how the Building a Roadmap project fits within NQF's body of work on PROs.

Previous Work With Patient-Reported Outcomes in Performance Measurement

Over the past decade, NQF has actively participated in the development of numerous reports to further the use of PROs and PROMs in clinical settings as well as the use of PRO-PMs to assess the performance of healthcare organizations.

In 2012, with funding from CMS, NQF launched the PROs in Performance Measurement project. The project included two commissioned background reports. The first report, Methodological Issues in the Selection, Administration and Use of Patient-Reported Outcomes in Performance Measurement in Health Care Settings (henceforth referred to as *Methodological Issues*), focused on selecting PROMs for use in performance measurement and was updated in 2015 by David Cella and his colleagues. The second, PRO-Based Performance Measures for Healthcare Accountable Entities (henceforth referred to as *Accountable Entities*), focused on the reliability and validity of PRO-PMs. As part of the project, NQF also convened two meetings of an Expert Panel that contributed to the development of the 2013 report titled Patient-Reported Outcomes in Performance Measurement.[7] This work brought together diverse experts to lay the groundwork for developing, testing, endorsing, and implementing PRO-PMs.

In 2017, NQF partnered with PatientsLikeMe, and with funding from the Robert Wood Johnson Foundation, developed and published Measuring What Matters to Patients: Innovations in Integrating the Patient Experience into Development of Meaningful Performance Measures.[8] This report reiterated the importance of patient-centered quality measurement and demonstrated the value that online patient communities could offer to measure developers and other stakeholders involved with PROs.

Many challenges have remained unaddressed since these reports were published, such as burdensome workflows related to data collection and lack of funding for the use of PROs. In 2019, CMS funded the first of two new projects with NQF, Patient-Reported Outcomes: Best Practices on Selection and Data Collection (henceforth referred to as *PRO Best Practices*). In September 2020, NQF published the PRO Best Practices Final Report from that initiative, in which a TEP identified best and promising practices to help clinicians and administrators select and implement PROs and PROMs in care settings.[9] This report presented solutions to common challenges, such as:

- increasing clinician support by securing physician buy-in before launching a PRO program;
- improving patient buy-in by including patients, family members, and caregivers in the process of selecting PROs and PROMs;
- engaging staff in developing feasible and effective clinical workflows; and
- collaborating with leadership to identify funding sources to offset the costs of collecting and using PRO data.

The TEP for the 2020 initiative designed a PROM Attribute Grid that guides the selection of high quality PROMs for use within a clinic or health system (See page 12 of the PRO Best Practices Final Report). While this grid is intended for clinicians who are selecting PROs and implementing PROMs, it can provide insight into the attributes of high quality PROMs for use in performance measures.

Current Work to Build a Roadmap From PROMs to PRO-PMs

In the first year of the Building a Roadmap initiative, the TEP revisited NQF's past work to understand what has—and has not—worked well in identifying high quality PROMs as the basis for PRO-PMs that can be used for accountability. Additionally, the TEP considered new knowledge about PRO-PMs that has emerged since the previous reports, particularly regarding digital quality measures (dQMs). NQF is publishing four reports related to the Building a Roadmap initiative that provide information and guidance to advance the development of digital PRO-PMs based on high quality PROMs.

- The Environmental Scan Report identifies and summarizes existing information relevant to the use of high quality PROMs as the basis for PRO-PMs in accountability programs, and provides background for the other documents in the Building a Roadmap project; this updated edition replaces the original edition that was published in spring 2021
- The Interim Report identifies and describes attributes of PROMs that are suitable data collection instruments for high quality digital PRO-PMs and provides guidance to measure developers on selecting PROMs for use in a PRO-PM.
- The Technical Guidance Report describes a series of stages and tasks that measure developers can follow when developing and testing digital PRO-PMs that are suitable for CMS VBP programs and APMs; an updated edition will be published in autumn 2022.
- The Developer Feedback Report (to be released in summer 2022) identifies recommendations from measure developers and other key audience members on improving the Technical Guidance Report.

The TEP for the Building a Roadmap initiative comprises multistakeholder experts who represent measure developers, health IT professionals, payers, researchers, clinicians, and other healthcare

perspectives. Importantly, the TEP also includes patients and patient advocates, who are critical voices for this effort. Because of the technical nature of this topic and the focus on PROMs and PRO-PMs that are used by federal agencies, NQF intentionally included individuals and organizations that are involved in the development of PROMs and/or the development and stewardship of PRO-PMs, including:

- Patient-Reported Outcomes Measurement Information System (PROMIS)
- Kansas City Cardiomyopathy Questionnaire – 12 item (KCCQ-12)
- Hip disability and Osteoarthritis Outcome Score, Joint Replacement (HOOS, JR)
- Knee injury and Osteoarthritis Outcome Score, Joint Replacement (KOOS, JR)
- PRO-PMs related to depression and orthopedic outcomes

While PROMs are not the explicit focus of this work, NQF recognizes that these tools are important building blocks in the development of high impact performance measures. NQF does not endorse instruments or scales (including PROMs),[7] nor does it advocate for the need to develop new PROMs.

Terminology Used in Describing Performance Measures

In this report, NQF uses terminology from the CMS Measures Management System (MMS) Blueprint (henceforth referred to as *the CMS Blueprint*) to distinguish between PROs, PROMs, and PRO-PMs (Table 1). While the literature contains several different definitions for PROs, PROMs, and PRO-PMs, measure developers should be aware of the CMS Blueprint's definitions, particularly when developing digital PRO-PMs that may be used in CMS VBP programs or APMs. Other terms, including dQMs and electronic clinical quality measures (eCQMs) are included in the glossary (Appendix A):

Performance measures are standards that can be used to measure and quantify healthcare processes, outcomes, patient perceptions, organizational structures, and/or systems that are associated with the ability to provide high quality care.[10] NQF has endorsed several hundred healthcare performance measures. PRO-PMs are a unique type of outcome measure that can be used to evaluate and compare healthcare entities (e.g., clinicians or health plans) using patient health outcome data that are provided directly by patients. Performance measurement is not unique to healthcare. The Baldrige Performance Excellence Program, the nation's only Presidential awards program that recognizes exceptional organizational achievement, prioritizes performance measures in its points-based scoring model, which is an example of the importance of performance measurement as it applies to a broad range of industries and organizations.[11]

Table 1. Distinctions Among PROs, PROMs, and PRO-PMs

Concept	Definition	Example
Patient-Reported Outcome (PRO)	Any report of the status of a patient's health condition or health behavior that comes directly from the patient, without interpretation of the patient's response by a clinician or anyone else[12]	Symptoms of depression

Concept	Definition	Example
Patient-Reported Outcome Measure (PROM)	Tools used to collect patient-reported outcomes[12]	Patient Health Questionnaire 9 (PHQ-9)©, a standardized tool to assess depression
Patient-Reported Outcome Performance Measure (PRO-PM)	A way to aggregate the information from patients into a reliable, valid measure of performance at the measured entity, level, e.g., clinician[12]	NQF #0711: Adult patients age 18 and older with major depression or dysthymia and an initial PHQ-9 score > 9 who demonstrate remission at six months (defined as a PHQ-9 score less than 5)[1]

Environmental Scan Methodology

To consolidate resources for PRO-PM developers, users, and policy makers in this report, NQF conducted an Environmental Scan using the following methodology. First, NQF conducted a literature review to assess the body of literature related to PROMs and PRO-PMs and identify those articles most relevant to this initiative (Part One). Second, NQF conducted a related scan of existing PROMs and PRO-PMs, as well as the organizational bodies that assess the quality of these measures (Part Two). Third, NQF convened a multistakeholder TEP comprised of experts in fields related to PROs and held a series of web meetings in which key topics were discussed (Part Three). Fourth, NQF conducted nine key informant interviews (KIIs) with measure developers, IT experts, and patients who are not employed by federal agencies, as well as three interviews with federal employees. Each interviewee possesses unique experience with or perspective on measure development and digital quality measurement (Part Four). Each of these approaches is outlined in more detail below.

Part One: Literature Review

To support the goals and objectives, NQF conducted a literature review to identify current measure gaps and challenges in developing and implementing PRO-PMs. The literature review included a search for sources that detail attributes of high quality PROMs.

Methods

Databases for the literature review included PubMed/Medline and Google Scholar. NQF conducted a targeted search within these databases using various combinations of keywords that were derived terms related to guidance on developing PRO-PMs (including digital PRO-PMs) as well as general terms to capture broader work that may include relevant information (Appendix B). In order to maintain focus on current recommendations and practices, NQF confined the search to English-language work published between 2016 and present day, unless an older source remains an important part of the body of literature (i.e., it is noted as important by the TEP, it is widely recognized or cited by experts in the field, and/or its conclusions or recommendations remain relevant and have not been significantly revised or disproven). In order to identify new literature, searches were constrained to information published on or after January 1, 2021.

NQF also included grey literature in the literature review and considered papers and websites from government, not-for-profit, and corporate organizations for the Environmental Scan Report. The project team conducted additional searches using Google, with the intent of identifying grey literature that did not appear in the database searches. NQF reviewed various sections of the CMS website in order to accurately represent the current state of PROMs and PRO-PMs in VBP programs and APMs.[13]

NQF reviewed and listed the websites related to each PROM discussed in detail in the Environmental Scan Report (Appendix C). The project team located these websites via Google searches focused on the copyright, licensing, and/or developer information for each PROM.

Part Two: Measure Scan of Existing PRO-PMs

NQF conducted a measure scan for NQF-endorsed PRO-PMs (Appendix D, Table 2). Because this initiative includes a roadmap that guides measure developers to the beginning of the NQF endorsement process, the Environmental Scan Report centered on PRO-PMs that are currently endorsed by NQF. NQF's scan for PRO-PMs included NQF's Quality Positioning System (QPS) and the CMS Measure Inventory Tool (CMIT) (Appendix D, Table 3).

Part Three: Discussions With Experts

NQF selected 25 experts to serve on the TEP, 22 of whom continued to serve for the second year of the project. These experts bring diverse perspectives on developing PRO-PMs for use in VBP programs and APMs, including viewpoints of measure developers and patients. This report synthesizes information gained in the discussions that occurred during meetings of the TEP. Key topics addressed during these meetings included:

- Unique roles of PROMs and PRO-PMs and the intersection between them
- Implications and requirements of digital PROMs and PRO-PMs
- Modes of administration and methods of data collection for PROMs
- Domains of PROs and rationale to focus on outcomes instead of experience
- Meaning of "high quality" as it pertains to PROMs and PRO-PMs
- Attributes of a PROM that is suitable for use in a high quality digital PRO-PM
- Evaluation of public comments and related revisions to reports

These discussions were moderated by NQF staff and facilitated by the co-chairs of the TEP. Information elicited during the discussions included anecdotal experiences that were common to multiple TEP members, as well as professional activities related to performance measurement that are not represented in the literature. Members of the TEP also provided resources to NQF staff, including peer reviewed articles.

Part Four: Key Informant Interviews (KIIs)

To gather additional information, NQF conducted nine, one-hour KIIs with measure developers, IT experts, and patients who are not employed by federal agencies, as well as three interviews with federal employees. Each interviewee offered a unique perspective on measure development, digital measurement, and/or PRO-PMs. NQF led each interview using an Interview Guide to promote consistency across the interviews. (Federal employees were interviewed using a separate interview

guide designed to elicit information about digital quality measurement at federal healthcare agencies.) Although the interviews followed a standard guide, each interview accommodated the interviewee's individual expertise and background. Interviewees were identified and selected based on recommendations from the TEP and CMS.

The interview guide included questions to elicit information on specific content areas, including:

- introductory questions that inquired about the interviewee's background and experience with measure development;
- exploratory questions on the structure and content of the Interim Report and the Technical Guidance Report;
- guidance questions related to digital measurement and development of PRO-PMs; and
- general questions on how the Technical Guidance Report could be more useful to its target audience of PRO-PM developers.

Information in the Environmental Scan Report that is not explicitly attributed to a specific source has been synthesized by NQF using parts 1-4 of the methodology described above.

Environmental Scan Findings

NQF and the TEP identified key findings about the current state of selecting high quality PROMs and developing digital PRO-PMs. The majority of the Environmental Scan Report focuses on these findings:

- The role of PROMs and PRO-PMs in quality-based models in both the public and private sectors, and the importance of digital measurement in advancing this role.
- The modes of administering PROMs, the methods of data collection, and the implications of each on digital measurement.
- Gaps between PRO-PMs and other types of quality measures (e.g., process and outcome measures), as well as the gap between patient-reported outcome and experience measures.
- Existing resources that measure developers can use to identify candidate PROMs as data collection tools for PRO-PMs.
- Currently available guidance to assist measure developers with the creation of PRO-PMs.
- An overview of the implications of interoperability on PRO-PMs.
- Common challenges to the development of PRO-PMs.

The Environmental Scan Report does not attempt to provide recommendations or novel solutions to these findings. Instead, it documents current information about each area. In cases where multiple approaches exist (such as with the challenge of determining whether a new PRO-PM should utilize data from a single questionnaire or multiple different questionnaires), the scan presents the tradeoffs of the respective approaches. Other reports in the Building a Roadmap initiative (i.e., the Interim Report and the Technical Guidance Report) do provide recommendations that are pertinent to these findings.

Role of PROMs and PRO-PMs in Quality-Based Models

The Environmental Scan Report confirms the importance placed on PROMs and PRO-PMs by a broad range of healthcare stakeholders, including federal agencies, payers, health systems, professional societies, patient advocacy organizations, and quality improvement organizations.

CMS and Industry-Wide Perspective

An industry-wide shift away from fee-for-service reimbursement is occurring, and it includes discussions about the role of PROMs in value-based payment. The use of PRO-PMs in accountability and value-based initiatives has the potential to promote patient-centeredness, improve care, and lower cost.[14,15]

CMS has supported the shift toward value-based care through a variety of programs and initiatives that encourage the use of PROMs and PRO-PMs in quality measurement and improvement. In 2017, CMS launched the Meaningful Measures Initiative, which identifies and prioritizes areas for quality measurement and improvement.[16] This initiative also helps to identify and close important measurement gaps, align measures across both the continuum of care and payers, and spur innovation in new types of measures, such as patient-reported measures and electronic measures.[17] CMS identified PRO-PMs in this initiative as a way of amplifying the patient voice and driving measures toward patient-centeredness.[17] In addition to the Meaningful Measures Initiative, CMS also sets priorities based on input from the National Impact Assessment of CMS Quality Measure Reports, further emphasizing the importance of prioritizing measures that use patient-generated data.[17]

Although attention to PROM and PRO-PM adoption is increasing, the scan identified few examples of payer PRO-PM implementation. One payer example of a large-scale implementation of PROMs is the effort by Blue Cross Blue Shield of Massachusetts (BCBSMA) to incorporate PROMs into clinical care.[15] The implementation started with a phased adoption of PHQ-9 (utilized for depression screening) and HOOS, JR / KOOS, JR (utilized for orthopedic hip and knee pain, respectively), and ultimately, to the adoption, use, and data sharing on PROMs in six clinical areas. The effort paid provider systems for participation rather than performance, though the goal was to eventually develop PRO-PMs that could be used for performance-based payment and accountability.[15] While the BCBSMA effort did not ultimately produce PRO-PMs, the case demonstrates potential ways in which financial incentives that reward PROM adoption can improve diagnosis and treatment, such as improved diagnosis and longitudinal tracking of depression as well as accurate prediction of outcomes from baseline functioning scores for hip and knee replacement patients.[15]

CMS Goals Related to PRO-PMs and Digital PRO-PMs

One aim of the CMS Meaningful Measures 2.0 initiative is to promote better collection and integration of patient perspectives through the use of PROMs.[3] CMS has identified several strategies within this aim:

- Simplifying the use of PROMs
- Better integrating PROMs into EHR systems and their related workflows
- Developing new PROMs that are embedded into workflows, accessible to patients through digital means, and helpful in reducing reporting burden
- Expanding the use of the PROMIS tools
- Identifying “self-reported health” as a key result across CMS for patient-reported information[3]

Another aim of the Meaningful Measures 2.0 initiative relates to transforming 100 percent of quality measures to be fully digital by 2025. To accomplish this, dQMs need to be fully interoperable, meaning they allow data entry, storage, integration, calculation, and reporting to be conducted by health IT systems, and enable data to be used in multiple ways across systems. The ambitious goal of modernizing and digitizing quality measures and programs includes key steps, such as finalizing a digital measure strategy and advancing the electronic data infrastructure.[4]

The move to dQMs promotes important patient-centered goals, such as increasing support for value-based programs across payers and improving care coordination. One of the ways in which CMS plans to integrate PROMs into the EHR workflow is by aligning the EHR certification process with other CMS reporting requirements.[4]

Modes of PROM Administration and Methods of Data Collection

Stakeholders agree that modes of PROM administration and methods of data collection are important to consider when developing PRO-PMs. NQF's PRO Best Practices Report focused on the importance of using multiple modes to maximize patient participation and response, because directly capturing and reflecting the patient voice is a critical goal of PRO-PMs. This focus remains relevant throughout the Building a Roadmap project.

Digital input directly from the patient is widely viewed as the ideal. However, it is commonplace for clinicians to collect PROM results via a paper survey or telephone interview, then transpose those data into an EHR. Some TEP members noted anecdotal examples in which administration of a PROM by a clinician appeared to result in higher scores, and the literature reflects statistically significant differences in PROM scores obtained via telephone compared to those from self-administered instruments.[18] Stakeholders can digitally capture PROM data with technology such as iPads, email, or patient portals to maximize the chances of collecting accurate information directly from the patient, but they should provide alternative methods for those patients who are unable to navigate these options.[9]

The capture of PROM data from patients presents a variety of challenges. Nine patient-level factors may impact PROM completion: platform design, print literacy, health literacy, technology literacy, language proficiency, physical functioning, vision, cognitive functioning, and time.[19] These factors should be carefully considered when implementing PROMs as part of a workflow, as they can contribute to poor response rates and inaccurate completion of PROMs.[19] These patient-level factors may disproportionately affect minority populations, thus contributing to healthcare inequities.[19]

Gaps Between PRO-PMs and Other Types of Quality Measures

CMS' commitment to elevating the patient voice is constrained by the small number of NQF-endorsed PRO-PMs. Though there are hundreds of PROMs, at the time of this publication, NQF has endorsed 28 PRO-PMs that span different domains (i.e., HRQoL, functional status, symptoms and symptom burden, health behaviors, and experience with care), conditions and diseases (e.g., joint replacement, depression, and pain), and settings (e.g., ambulatory, inpatient, long-term care, and hospice) (Appendix D, Table 2).[1,20] While the existence of 28 NQF-endorsed measures is a positive step in the relatively short lifespan of PRO-PMs, these measures compose less than 7 percent of all NQF-endorsed measures.[1] In prioritizing the Building a Roadmap initiative, CMS noted the gap between the number of NQF-endorsed

PRO-PMs and the limited number that are currently used for accountability programs; for example, only one NQF-endorsed PRO-PM is included in the MIPS Performance Year 2022 Quality Measures List.[2]

NQF, CMS, and the TEP agreed that the Building a Roadmap project should focus on digital PRO-PMs that measure three domains: HRQoL, functional status, and symptoms and symptom burden. While all five domains (including experience with care and health behaviors) are important, three primary drivers influence this recommendation.

1. **Representation of each domain in currently endorsed NQF PRO-PMs:** Nearly half of currently endorsed PRO-PMs measure the patient experience domain. PRO-PMs related to quality of life and symptoms are especially underrepresented.
2. **Assessment of healthcare entity performance:** While health behaviors and experience with care are important domains, they are farther removed from direct clinical interventions than HRQoL, functional status, and symptoms.
3. **Differences in Data Collection Methodology:** Entities that collect PROM data on HRQoL, functional status, and symptoms typically have a high degree of autonomy around data collection, storage, and analysis. Experience data, however, typically rely on stringent methodologies that require external partners to collect and analyze data, in part to ensure patient responses are confidential and will not affect the delivery of care.[21]

NQF endorses performance measures but does not endorse instruments or scales (including PROMs).[7] If a PROM is explicitly identified in the specification of a PRO-PM, the NQF Scientific Methods Panel (SMP) reviews that PROM for reliability and validity as part of the endorsement process (Appendix E).[22] However, NQF remains agnostic to the specific instrument and reviews it only to evaluate whether it meets an acceptable scientific standard as an element of the PRO-PM. Measure developers who are interested in learning more about the PROMs that are used in NQF-endorsed PRO-PMs can review measure specifications on NQF's Quality Positioning System.

Potential Resources for Identifying Data Collection Instruments for PRO-PMs

As a rule, the measure developers who work on PRO-PMs are not the same people who develop and test PROMs. While exceptions to this rule do exist, the Building a Roadmap initiative treats PROM developers and PRO-PM developers as separate roles. The measure developers who work on PRO-PMs assess the extensive library of existing PROMs, consider the attributes of each instrument (e.g., psychometric soundness, usability and feasibility, and interpretability), and determine which should be used to collect data for performance measurement. This section of the Environmental Scan Report presents resources that measure developers commonly use when identifying candidate PROMs that might be suitable for data collection with a high quality PRO-PM. (The Interim Report provides extensive detail on assessing each candidate PROM against a set of defined attributes and determining its suitability for use with a performance measure.)

PROMs in Use With CMS VBP Programs or APMs

With the passing of the Affordable Care Act of 2010, the U.S. healthcare system has shifted towards improving and rewarding value.[23] CMS designed the VBP program to increase the quality of care and

experience for patients.[24] There are several VBP programs that can apply to various provider settings, such as hospitals and outpatient centers.

The APM is one track of CMS' Quality Payment Program (QPP), and it provides incentives to eligible participants to ensure high quality and cost-efficient care is provided.[25] Because the patient perspective is critical to quality measurement, payers need to leverage instruments that measure and account for the patient voice. VBPs and APMs are likely to interact during the shift towards improving value, given that incentives linked to APMs and VBP programs may be similar for some providers.[23]

PROMs are critical tools for capturing the patient voice. The following list highlights seven candidate PROMs that may be suitable for use with PRO-PMs. While this list identifies specific PROMs that CMS has selected for use in accountability programs, this is *not* an endorsement of any individual PROM.

- **National Cancer Institute's (NCI) Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE):** This tool was designed to measure and evaluate symptoms and adverse events for participants in cancer clinical trials.[26] The PRO-CTCAE Measurement System should be utilized with the Common Terminology Criteria for Adverse Events (CTCAE) due to the supplemental information that the PRO-CTCAE can provide to clinicians.[26] The PRO-CTCAE functions to enhance the precision of adverse-event reporting in cancer clinical trials, to provide useful data for clinicians, and to ensure that the patient perspective related to experiencing an adverse event is collected. Given the favorable test-retest reliability (median ICC 0.77) in a sample of 975 patients, as well as being linguistically validated, PRO-CTCAE demonstrates strong validity, reliability, and responsiveness.[27]

- **National Institutes of Health (NIH)-funded Patient-Reported Outcomes Measurement Information System® (PROMIS):** This initiative was established in 2004 with the goal of standardizing measures to allow for different PRO domains to be assessed.[28] The set of standards for the development and validation of item banks and instruments within PROMIS provides a useful tool for developers. PROMIS offers short forms, computerized adaptive testing (CAT), and profiles (i.e., fixed collection of short forms from multiple domains), as well as appropriate use across a range of patient populations. In using PROMIS measures with CAT, measures usually only require four to six items for precise measurement of health-related constructs, thereby reducing respondent burden.[28]

- **Patient Health Questionnaire 9 (PHQ-9):** This instrument has been in use since 1999 after it was developed through a grant from Pfizer. As a nine-question instrument, the PHQ-9 is shorter than its 16-question predecessor (the PHQ) and assesses the presence and intensity of depression and depression symptoms. The PHQ-9 is defined in the specification for four NQF-endorsed PRO-PMs that are stewarded by MN Community Measurement and are related to depression remission and depression response at six and 12 months.[1] Given its use in various medical specialty areas, the PHQ-9 and the related performance measures offer examples of how a widely adopted PROM can be used as the basis for NQF-endorsed PRO-PMs.

- **Kansas City Cardiomyopathy Questionnaire – 12 item (KCCQ-12) and Minnesota Living with Heart Failure Questionnaire:** The KCCQ-12 PROM is a sensitive and specific HRQoL measure for patients with heart failure (HF).[29] Similar to PHQ-9, the KCCQ-12 is a truncated version of the 23-

question KCCQ. Although the KCCQ has been shown to be valid, reliable, and sensitive, its length (23 questions) has been a barrier to gaining insight on the patient experience. An additional instrument that has been utilized to assess quality of life among HF patients is the Minnesota Living with Heart Failure Questionnaire (MLHFQ). The MLHFQ has two domains—physical and emotional—and is a self-administered instrument.[30] Both the KCCQ-12 and the MLHFQ are commonly used in clinical research and have the potential to predict outcomes important to clinicians of HF patients.[31]

- **Hip disability and Osteoarthritis Outcome Score, Joint Replacement (HOOS, JR) and Knee injury and Osteoarthritis Outcome Score, Joint Replacement (KOOS, JR):** Within orthopedics, PROMs are utilized for several conditions, including ligament injuries and joint replacements.[32] Two examples of validated and commonly used PROMs for knee injury and joint replacement include the KOOS, JR and the HOOS, JR. These PROMs assess outcomes after total knee arthroplasty (TKA) and total hip arthroplasty (THA), respectively, and are short-form measures that help to reduce patient survey fatigue. In separate validation studies,[34,35] high internal consistency and high responsiveness were seen with a Pearson Separation Index of 0.84 for KOOS, JR and 0.86 for HOOS, JR. An additional attribute of both measures is the output of a single score that relays knee and hip "health" to the clinician. The KOOS, JR and HOOS, JR have crosswalks that allow scores from these PROMs to be converted to Oxford Knee Scores and Oxford Hip Scores (other PROMs that are widely used after TKA and THA), and vice versa.[34] The KOOS, JR and HOOS, JR are also both included as acceptable PROMs to meet the reporting requirements for CMS' Comprehensive Care for Joint Replacement Model.[37]

NQF Resources

NQF offers numerous resources that can assist measure developers in identifying candidate PROMs that might be appropriate for use in PRO-PMs. The reports from 2012 and 2013, particularly the Methodological Issues white paper, describe eight characteristics of PROMs that measure developers should assess when considering PROMs as data collection instruments in PRO-PMs.[5] The Interim Report from the current Building a Roadmap initiative expands on this guidance and presents 12 attributes that measure developers should assess when selecting PROMs for use with high quality PRO-PMs.[38] Although the core audience for the PRO Best Practices Report is clinicians and administrators who are choosing PROMs for use in a clinical setting, it features guidance that developers may consider when determining if a PROM is suitable for use with a PRO-PM.[9] Measure developers can also benefit from the specifications of PRO-PMs that are listed within the NQF Quality Positioning System database.[1] These specifications typically list any PROMs that are identified as data collection tools for PRO-PMs, including those that are no longer or were never endorsed.

International Consortium for Health Outcomes Measurement (ICHOM) Resources

ICHOM was founded in 2012 with the intent of creating "critical foundations for value-based healthcare."[39] Part of the organization's work has focused on convening clinical experts and patients to develop standard Sets of Patient-Centered Outcome Measures. As of 2022, ICHOM has published 41 Sets, each of which is a pragmatic measurement recommendation based on a working group's comprehensive review of relevant PROMs and other measures and data sources. As an example, the Hip and Knee Osteoarthritis Set identifies a minimum data set of case-mix variables, treatment variables,

and outcomes, then recommends three potential HRQoL PROMs, a pain scale, and a hip- or knee-specific physical function PROM.[40] Because of ICHOM's focus on outcomes that matter most to patients and a vetting process that involves clinical experts and consumers, the Sets are a potential source of high quality PROMs. The Sets include common chronic diseases, such as diabetes; population-specific sets, such as older person primary/preventive care and hypertension in low- and middle-income countries; and behavioral health conditions, including dementia, depression, and anxiety.

PROMs Identified by Professional Societies

Professional societies can be a valuable resource for identifying PROMs that may provide strong foundations for future PRO-PMs. Many professional societies have convened working groups to evaluate PROMs and recommend those that meet certain criteria, such as patient-centeredness, cost, modality of administration, methods of data collection, completion time, clinical meaningfulness, and widespread clinical adoption. Approaches and recommendations from three societies are listed below, but there are numerous associations and societies that have published comparable recommendations on their websites and in white papers, journals, and other media.

- **American Academy of Orthopaedic Surgeons (AAOS):** This academy established the Quality Outcomes Data (QOD) Work Group in 2015. The workgroup evaluated instruments for PROs in orthopedics against the criteria of free use, inclusion of only patient-reported data, multiple modalities, number of questions, responsiveness, one generic quality of the PROM, no more than three joint or disease-specific PROMs, and availability of CAT.[13] As a result of this work, AAOS developed a set of recommended PROMs for upper extremities (e.g., shoulder and shoulder instability, along with elbow, wrist, and hand), lower extremities (e.g., foot and ankle, knee, and hip), spine, and disease-agnostic quality of life.[41–44]

- **Society of Gynecologic Oncology (SGO):** This society convened a daylong meeting of its Policy, Quality, and Outcomes Taskforce in 2018 that resulted in disease-specific recommendations for PRO data collection using the Functional Assessment of Cancer Therapy (FACT)-G7 as a general HRQoL questionnaire; disease-specific PROMs for ovarian, uterine, cervical, vulvar, and vaginal cancers; and instruments that specifically address sexual health in women with cancer.[45,46]

- **American Academy of Neurology (AAN):** Some societies opt to list PROMs that are common in their field. While these lists may not be as rigorously vetted as those from societies that assign dedicated working groups to recommend PROMs, they can still be useful in identifying PROMs that may provide meaningful data collection for PRO-PMs. AAN provides a brief list of PROMs used in neurology, including cross-cutting instruments, such as PROMIS and PHQ-9, as well as condition-specific scales and tools for dementia, headache, epilepsy, and multiple sclerosis.[47]

Regardless of how societies assemble a list of preferred PROMs, their research and recommendations can be useful for identifying PROMs that may be suitable for high quality digital PRO-PMs.

Currently Available Guidance for Developing PRO-PMs

The literature review identified limited guidance on the development of PRO-PMs, a finding that aligns with CMS' decision to fund the Building a Roadmap initiative based on a lack of detailed technical

guidance for developing high impact digital PRO-PMs. While the NQF and CMS documents discussed earlier in this Environmental Scan Report are valuable resources for PRO-PM developers, only one other substantive guidance resource emerged from the literature review.

The two NQF-commissioned white papers from 2012, as well as the 2013 report that is based on the work of the Expert Panel, are widely recognized as foundational resources for the development and use of PRO-PMs, and all three remain relevant. The 2012 Methodologic Issues white paper, along with its 2015 update, focuses on the selection, implementation, and use of PROs in clinical settings, as well as best practices in using PROs in performance measurement.[5,6] The 2012 Accountable Entities white paper focuses on the validity and reliability of PRO-PMs, as well as considerations of measure construction, interpretation, score calculation, and risk adjustment.[48] The 2013 report, *Patient-Reported Outcomes (PROs) in Performance Measurement*, reflects the Expert Panel's work to identify key characteristics of PROMs that are appropriate for use in PRO-PMs, define considerations for evaluating PRO-PMs, and establish pathways to move from PROs to PRO-PMs. While overlap exists between these reports and the Building a Roadmap initiative, the latter builds upon these foundations and expands their value and usefulness by providing step-by-step guidance designed for a diverse audience of measure developers.

In addition to the above reports that are specific to PRO-PMs, NQF has published extensive information and guidance on the Consensus Development Process (CDP) it uses to evaluate and endorse measures. Information about the CDP is available on the NQF website, and relevant resources include:

- The CDP homepage
- Overview of the CDP Process (PDF, 12/2020)
- Measure Developer Guidebook for Submitting Measures to NQF (PDF, 8/2021)
- Measure Evaluation Criteria and Guidance for Evaluating Measures for Endorsement (PDF, 9/2021)

The CMS Blueprint and its related documents are indispensable resources for measure developers. At the time of this Environmental Scan Report, the CMS Blueprint was on its 17th version, which is available through the CMS Measure Management System Blueprint section of the CMS website. Previous versions of the CMS Blueprint offered in-depth guidance for measure developers, but beginning with version 16, CMS began optimizing the report for a broad variety of audiences.[17] In conjunction with this change, CMS started publishing a new document, *The Blueprint for the CMS Measures Management System: Contractual Edition*, that offers detailed information for measure developers; the Contractual Edition is only available to Measure and Instrument Development and Support (MIDS) contractors, their government leads, and CMS.[17] As with the foundational NQF reports mentioned in the previous paragraph, the CMS Blueprint and its related materials provide essential guidance to measure developers. However, these materials are not specifically focused on PRO-PMs, and the Contractual Edition is not available to an expansive audience of measure developers. As such, the Building a Roadmap initiative fills an important gap in PRO-PM guidance: The Interim Report assists measure developers with the identification of high quality PROMs, while the Technical Guidance Report outlines a series of stages and tasks that guides the development of PRO-PMs.

One PRO-specific resource that is related to the CMS Blueprint is the CMS Blueprint Supplement on Patient Reported Outcome Measures, which offers high level guidance to measure developers on the development and evaluation of PRO-PMs. However, the *CMS Blueprint Supplement on Patient Reported Outcome Measures* does not offer detailed, step-by-step guidance on the development of digital PRO-PMs. For example, the section on choosing and defining a PRO addresses only three key tasks: the importance of identifying quality issues that are meaningful to a targeted population, determining relevance, and ensuring usability by the measured entities.[12] While these tasks are important, the document does not provide additional detail to help measure developers, particularly novices, understand how to accomplish these tasks. The document does link to key resources (e.g., the CMS Blueprint Supplement on Risk Adjustment in Quality Measurement and NQF's *Patient-Reported Outcomes in Performance Measurement*), but it lacks the level of detail to be useful to measure developers who are navigating the early steps of PRO-PM development. It does not provide information on how to choose and define a PRO that is meaningful to the population being served, what stakeholders (including patients) should be involved in choosing the PRO, or how to ensure the measure will be usable by the necessary entities.

The literature review identified one additional guidance document: a 2015 peer-reviewed article published in *Value in Health*.[14] A TEP assembled by the American Medical Association (AMA) prepared a document that provides detailed recommendations on PRO-PM development. The article, written by Basch et al., identified nine best practices for developing PRO-PMs (Appendix D, Table 4).[14] Several of the best practices correspond to NQF's measure evaluation criteria on reliability, validity, usability, and feasibility.[14] The best practices were developed with the goal of supporting future development of robust approaches to better understand the impact of care on the patient experience.[14] While these best practices remain valuable and relevant, the guidance documents from which they were drawn were all published between 2000 and 2011 and generally addressed the development and use of PROMs rather than PRO-PMs.[14] The Basch et al. paper does not specifically address the development of digital PRO-PMs, which further emphasizes the need for the Building a Roadmap guidance.

Because limited guidance is available regarding PRO-PM development, there is a need to create more detailed instructions for guiding measure developers, regardless of experience, through the PRO-PM development process.

Implications of Pending Advances in Interoperability for PRO-PMs

Developers and users should understand pending advances in interoperability as they plan an approach to PRO-PM design and implementation. These changes, which are being phased in over the next few years, should facilitate PRO-PM adoption. The Environmental Scan Report summarizes likely pending advances in interoperability that are relevant to digital PRO-PMs.

While the glossary (Appendix A) includes an extensive list of definitions, readers of this section will benefit from understanding current definitions for dQMs and eCQMs:

- **Digital quality measures (dQMs):** The definition of dQMs is evolving. In May 2021, CMS shared the approach of defining dQMs in a published Request for Information that was open for feedback, and in August 2021, CMS published a Final Rule that included feedback on the

proposed approach.[49,50] The definition describes dQMs as, "Software that processes digital data to produce a measure score or measure scores. Data sources for dQMs may include administrative systems, electronically submitted clinical assessment data, case management systems, EHRs, instruments (for example, medical devices and wearable devices), patient portals or applications (for example, for collection of patient-generated health data), health information exchanges (HIEs) or registries, and other sources."[50]

- **Electronic clinical quality measures (eCQMs):** eCQMs are expressed and formatted to use data from EHRs and/or health IT systems to measure healthcare quality, ideally data captured in structured form during the process of patient care.[17] They are the most common type of digital quality measures and are specified for use in the Medicare and Medicaid EHR Incentive Programs. Eligible professionals, eligible hospitals, and critical access hospitals are required to submit eCQM data from certified EHR technology to help measure and track the quality of healthcare services provided within the healthcare system.[51]

As noted previously, CMS has stated the goal of fully transitioning to only using dQMs (including eCQMs, which typically rely on data stored within a single EHR) in its quality and VBP programs by 2025.[3] CMS' overarching vision for digital measurement aims to "improve the patient experience including quality of care, improve the health of populations, and/or reduce costs."[50] Using interoperable digital data for measurement reduces burden on providers by eliminating time-consuming tasks such as chart abstraction and manual data entry. Interoperable data can also be aggregated for the purposes of calculating quality measure scores across different EHR systems at different organizations, as well as across other health IT systems such as registries or claims databases.[4,52]

Near-term advances in interoperability could remove data collection and measure score calculation barriers that currently impede PRO-PM use. For example, calculating risk-adjusted PRO-PMs requires centralizing PROM data collected by providers with disparate EHRs to build, test, and run risk-adjusted measure score calculations. Advances in interoperability will ease data sharing and create the infrastructure and specificity of shared data formats needed to aggregate data across measured providers and use shared data in centralized measure score calculations.

The U.S. is advancing data interoperability through several mechanisms that will lower the burden associated with collecting and sharing data used for quality measurement. Many significant advances in interoperability have occurred in recent years, including the development and widespread support for using Fast Healthcare Interoperability Resources (FHIR) as the standard for interoperable data. FHIR is a Health Level Seven International (HL7) standard that defines how healthcare information can be exchanged between different computer systems regardless of how it is stored in those systems.[53] FHIR can be leveraged to define common specific data formats that allow for data sharing across diverse systems without sacrificing information integrity.[54]

CMS and the Office of the National Coordinator for Health Information Technology (ONC) finalized regulations in 2020 that will accelerate providers' use and low burden sharing of FHIR data.[55] The ONC 21st Century Cures Act Final Rule requires providers using certified electronic health record technology (CEHRT) systems to map a specified scope of EHR data to FHIR resources and make those data accessible in the FHIR standard through FHIR application programming interfaces (APIs).[55–57] This requirement will

expand the scope of EHR data, including data used for quality measurement, that can be accessed without burden.

Not all the data needed for PROMs and PRO-PMs, however, will be interoperable when the ONC requirement takes effect December 31, 2022. Rather, the interoperability of relevant data will likely expand over time. The initial scope of data required to be interoperable is defined in the United States Core Data for Interoperability (USCDI) Version 1, and the detailed FHIR formats required in the US Core Implementation Guide.[56] USCDI Version 1 does not require that PROM data be interoperable. The ONC's data requirements grow as health information technology (IT) matures, meaning the USCDI is an evolving standard: The second version of the standard was published in July 2021 and added data elements related to social determinants of health (SDOH), sexual orientation, and gender identity (SOGI), while the draft third version proposes the addition and/or reclassification of data elements related to health insurance, health status, demographics, and other areas.[57,58] While Version 2 and Version 3 of the USCDI are not yet required for CEHRT, the inclusion of health status as assessed by standardized tools as a data class in the draft Version 3 lends momentum to the widespread uptake of the specific FHIR standards needed for PROMs and PRO-PMs.

While ONC has defined FHIR standards requirements at the data class and data element level for CEHRT through the USCDI and US Core Implementation Guide, ONC also recently launched an initiative, USCDI+, that defines and advances interoperable datasets for specific federal use cases (e.g., unique programmatic requirements for CDC surveillance programs).[52,57,59] The initiative could potentially be used to advance and accelerate a core set of interoperable data for quality measures, including PRO-PMs. USCDI+ allows harmonization to occur across federal programs, so a single data element can be consistently used across multiple use cases, even those with different coding structures.[52] While USCDI+ is focused on federal agencies, ONC understands that other organizations with unique datasets and use cases (e.g., specialty societies) can benefit from the development of use-case specific federally-defined datasets.

In summary, FHIR, USCDI, USCDI+, and other standards and technologies create the infrastructure for interoperability and, by extension, digital quality measures, including PRO-PMs. These standards are evolving, informed by stakeholder input through a structured process. These advances in interoperability can be leveraged by a range of stakeholders in the coming years to reduce PRO-PM data collection and sharing costs. The TEP encourages these stakeholders, including federal agencies and EHR vendors, to support and adopt standards for the storage and reporting of data elements to enhance interoperability and ultimately elevate the patient voice.

Challenges and Barriers of Developing PRO-PMs

The environmental scan identified several challenges related to developing PRO-PMs:

- Challenges with developing and testing PRO-PMs
- Difficulties navigating the NQF endorsement process
- Tradeoffs with determining whether a PRO-PM should utilize data from one PROM or multiple PROMs
- Technical challenges related to the development of digital PRO-PMs

- Issues related to patient burden, including low response rates and under-detection of poor performance when sicker patients might not be able to self-report[14,60–62]

Development and Testing Challenges

PRO-PMs are, by definition, complex measures. As discussed in the overview of the NQF endorsement process (Appendix E), the word "complex" has a specific meaning with regard to the CDP: it requires an SMP review of the PRO-PM and its underlying PROM(s). As a general word, "complex" also applies to the PRO-PM measure development process. Developers must identify candidate PROMs, review the literature for each, analyze how well each aligns with the attributes of a high quality PROM (as identified in the Interim Report), and select a PROM that will be used to collect data for the PRO-PM. If the measure developers consider using multiple PROMs as data collection tools, they must identify crosswalks to harmonize data across different instruments and create specifications that identify how those data are used in the PRO-PM. Testing a PRO-PM also adds complexities that do not affect most quality measures, such as identifying test sites that either are already using the required PROM as part of their standard protocols or are willing to add new PROMs not only to their clinical workflows but also to their EHR. These are only a few of the challenges that can emerge during the development and testing of a PRO-PM, and measure developers should be prepared to address these barriers.

Challenges With NQF Endorsement

Unique challenges arise for PRO-PMs undergoing the NQF endorsement process (Appendix E), specific to three of the CDP criteria. Additionally, developing an endorsement-ready, PROM-based performance measure requires a PROM that has been used extensively for data collection, which can eliminate the consideration of recently developed PROMs or those that have not been widely adopted within the clinical or research communities.

Criterion 2: Reliability and Validity – Scientific Acceptability of Measure Properties

According to NQF's Measure Evaluation Criteria and Guidance for Evaluating Measure for Endorsement, reliability *and* validity must be demonstrated at the data element level (i.e., the PROM) as well as the computed performance score (i.e., the PRO-PM) for complex, instrument-based measures (including PRO-PMs).[22] Any related threats to validity must also be addressed (i.e., exclusions, risk adjustment, performance comparability if multiple PROMs are used). Generally, for other types of measures, empirical testing at the data-element or measure score level is sufficient. This requirement for PRO-PMs adds a unique challenge to the endorsement process.

Criterion 3: Feasibility

An important aspect of feasibility is to ensure there is an achievable and implementable plan for collecting data or information without undue burden. Measures tend to be more feasible if they occur during the normal process of care, such as an intake survey to check functional status on an iPad in a waiting room. Electronic measures are generally preferred, as they utilize a data collection methodology that is more efficient and less burdensome than traditional measurement approaches. Examples of considerations for undue burden include instrument length and timing when requesting feedback from a patient.

Criterion 4: Usability and Use

NQF-endorsed measures should be included in accountability programs and publicly reported to ensure they remain in use. Otherwise, measures may not be ideal for endorsement since they may not be maintained or updated over time.

Importance of Relationships Between PROMs and Performance Measures

Some PRO-PMs rely upon data from a single PROM (i.e., a 1-to-1 relationship between a PROM and a PRO-PM) while others are designed to accept data from different PROMs that address the same domain (i.e., a many-to-1 relationship). There are advantages and disadvantages to both approaches, and not understanding their respective benefits and drawbacks can present a challenge to measure developers who are considering candidate instruments as data collection tools for performance measures.

The lack of standardized structured data fields across both PROMs and EHRs underlies one argument in favor of the 1-to-1 relationship between PROMs and PRO-PMs: The developer can tailor the measure specification to the unique data structure of a single PROM. Because different instruments use different scores to measure change, a performance measure based on a single PROM only needs to consider one score; this eliminates the need to identify crosswalks that harmonize PROM scores and build different instruments into the measure specification. Measure developers and stewards can more easily maintain a PROM-based performance measure that depends on a single instrument. A 1-to-1 relationship also allows the measure developer to select the PROM that is most aligned with the attributes identified in the Interim Report.

One example of an NQF-endorsed PRO-PM that is based on a single PROM is NQF #0711 *Depression Remission at Six Months*. This measure, stewarded by MN Community Measurement, assesses improvement in depression scores on the PHQ-9 by measuring the number of adults with a diagnosis of major depression and an initial PHQ-9 score greater than nine who have achieved a six-month PHQ-9 score of less than five.[1] According to its measure specification, MN Community Measurement selected the PHQ-9 as the PROM for this measure because it is "(1) validated with a sensitivity of 0.080 and a specificity of 0.92 with substantial heterogeneity I2 = 82%, (2) widely accepted and utilized in Minnesota, (3) available for clinical use, (4) translated into many languages, and (5) easy for the patient to complete and the provider to score."[1] This PRO-PM was initially endorsed in January 2011, making it a relevant and time-tested example of a single-PROM performance measure.[1]

There is also rationale for a many-to-1 relationship between PROMs and PRO-PMs, in part because it provides clinicians with the flexibility to use instruments based on appropriateness for their setting (e.g., language translations, licensing costs, and brevity of instrument) rather than requiring them to use a specific questionnaire. However, different instruments use different scoring systems and cut points (i.e., markers in PROMs that indicate the need to screen for a diagnosis or provide treatment), so measure developers must be careful when combining these different approaches into a single measure.[48] Crosswalks are necessary to harmonize different PROMs and ensure different data elements and scoring structures map to common elements in the PRO-PM. While crosswalks exist for some PROM pairs (e.g., the KOOS JR and the Oxford Knee Score, and the 12-Item Short Form Survey [SF-12] and the Veterans Rand 12-Item Health Survey [VR-12]), the number of possible PROM pairings far exceeds the number of available crosswalks.[35,40] An additional challenge with a many-to-1 relationship is the fact that the

technical specification of the measure becomes significantly more complex with every additional PROM that is added, which results in additional time for measure development, implementation, and maintenance.

Regardless of whether the PRO-PM utilizes data from one PROM or many different PROMs, the PROM quality shapes the effectiveness of the performance measure. If a PROM suffers from poor design or inaccurate data collection, the PRO-PM will suffer as well.

Technical Challenges in Digital PRO-PM Development

In 2011, CMS established the Medicare and Medicaid EHR Incentive Programs to encourage providers (including skilled nursing facilities, dialysis facilities, and hospitals) to adopt, implement, upgrade, and demonstrate meaningful use of CEHRT. These programs are now known as the Promoting Interoperability Programs, moving beyond the requirements of meaningful use to a new phase of EHR measurement with an increased focus on interoperability and improving patient access to health information.[63] One such opportunity is implementing universal standards for data.

One aspect of standardized data is the use of Logical Observation Identifiers Names and Codes (LOINC) to provide a separate code for each discrete unit of measure, which creates the ability to account for differences in these measures.[64] As an example, LOINC can enable the interpretation of clinical laboratory test result data by accurately and reliably coding the units of measure.[65] LOINC is critical for collecting, storing, and analyzing PROM data in EHRs and other health IT systems, and using these data in the calculation and reporting of PRO-PMs. Results of completed PROMs cannot be easily shared with EHRs and other health IT systems unless there are accepted vocabulary standards.[66] LOINC supports the structure and content of questionnaires by creating a model to capture the essential aspects of assessments. This model represents the hierarchical panel structure, global item attributes, panel-specific item attributes, and structured answer lists.[66] LOINC has embraced adapting standardized scales (e.g., the Glasgow Coma Score and the Apgar Score) and PROMs (including the PHQ-2 and PHQ-9, Confusion Assessment Method [CAM], PROMIS, and Outcome and Assessment Information Set [OASIS] assessments).[66] Today, LOINC supports more than 500 survey instruments.[67] Despite progress within standardized code sets, such as LOINC, a gap remains between coding and the storage of individualized codes required for each PROM.

The successful implementation and use of PROMs for PRO-PM development depends on integration within EHRs, whether from the vendor or built locally.[9] These technical challenges point to the importance of both embedding the instruments in the EHR and displaying the results in an actionable way. While interoperability is an industry-wide initiative, the burden falls on the implementor of the PROMs and PRO-PMs rather than the measure developers. Widespread use of PROMs and PRO-PMs requires improved integration with EHRs and other health IT systems, which can be achieved through a combination of interoperability standards (e.g., FHIR and USCDI) and coding schemes (e.g., LOINC).[68]

Challenges Due to Patient Burden

PROM developers, clinicians, patients, and caregivers have expressed concerns about the inherent burden on patients to complete PROMs.

Commonly noted barriers to PROM completion include lack of clarity for patients regarding the importance of PROM completion, excessive time to complete a questionnaire, questions that may be perceived as intrusive or irrelevant, selection bias, and low participation from vulnerable populations.[9] While these insights have led to changes, such as shortened versions of existing questionnaires and expanded use of CAT, obstacles remain to capturing patient-reported data on health outcomes.

Burden on patients to complete PROMs causes downstream issues for performance measurement. Factors such as pain or functional limitations, recall difficulties, and negative survey perception can add to response burden and affect the amount and quality of patient data collected. Additionally, issues regarding social determinants of health and health disparities, such as patients' access to digital tools and language barriers, can lead to less patient engagement with PROMs. PRO-PMs cannot exist if patients do not complete questionnaires.

Physical and cognitive impairments can also have an impact on the completion of PROMs. Patients with severe physical or cognitive impairments may require proxies (i.e., caregivers, family members, or other people who complete PROMs on a patient's behalf). While it is important to ensure all patients can complete PROMs and that caregiver voices are also captured and measured, mixing patient-reported data with proxy-reported data can create data fidelity issues that affect PRO-PMs.

Attempts to decrease patient burden and increase data collection can have unintended consequences, and alleviating one burden can create a new burden. As an example, one goal of digital data collection (e.g., sending PROMs to patients via email, text message, or a patient portal) is so patients can easily complete questionnaires. However, a single patient might need to access multiple websites or applications (apps) to complete PROMs for different providers at different health systems.

While PROM completion is a shared responsibility between patients and clinicians, NQF and the TEP emphasize that the healthcare delivery system and measure developers are ultimately accountable for minimizing patient burden.

Other Challenges and Considerations

The implementation of PRO-PMs will have an impact on clinical workflows, data flow, patient experience and satisfaction, clinician engagement, and much more. Patients are particularly affected by workflows that require the completion of lengthy PROMs at redundant intervals, and active engagement by patients is critical to the success of PRO-PMs. All stakeholders must be engaged as active partners in addressing data collection and workflows.

In addition to workflow challenges, there can be gaps between what existing PROMs measure and what is valued by patients. A theme in elevating the patient voice through PRO-PMs, or ensuring PRO-PMs are patient-directed, is to focus on what is important to patients.[9] Not all PRO-PMs align with patients' priorities, though, and not all patients share the same priorities. Performance measurement must be guided by what matters most to patients, and the incorporation of diverse patient perspectives (including viewpoints that represent disenfranchised populations) must inform this guidance. Digital PRO-PMs will continue to grow as innovative payment models and data standards evolve, and stakeholders must ensure these changes serve patients as well as healthcare professionals. Future work

on digital PRO-PMs should utilize patient partnerships to address and overcome complex issues (e.g., health equity, the digital divide, and data collection burden) that hamper patients' voices.

Measure development is a time-consuming and costly endeavor, and PRO-PMs are classified by NQF as complex measures. Future work should consider whether PRO-PM development would benefit from incentives, such as increased funding for measure developers or streamlined development requirements, which could lead to a larger range of organizations and stakeholders involved in development.

Limitations of the Environmental Scan Report

To note, there are three main limitations of the Environmental Scan Report.

1. The Environmental Scan Report presents resources that measure developers can use to identify candidate PROMs as data collection tools for PRO-PMs, but it does not offer a list of vetted or approved PROMs. The PROMs discussed in this report were included as potential examples of high quality PROMs due to their inclusion in one or more federal programs. However, the quality of a PROM cannot be assessed by its inclusion in this report.
2. The literature search returned only a short listing of guidance for developing PRO-PMs. To supplement this gap, NQF engaged with TEP members and measure development experts to inform the reports in the Building a Roadmap initiative. This limitation does, however, speak to the need for the Building a Roadmap initiative.
3. Guidance on digital quality measurement continues to rapidly evolve. The Environmental Scan Report and other reports within the Building a Roadmap initiative provide links to external resources with up-to-date information on digital measurement and interoperability. The Building a Roadmap initiative focuses on development of new digital PRO-PMs and does not fully address the steps of digitizing an existing PRO-PM.

Conclusion

The road from high quality PROMs to NQF-endorsed PRO-PMs is fraught with barriers and delays. Although there are many PROMs in today's healthcare environment, there is not a clear way to identify high quality PROMs that will provide a foundation for a digital PRO-PM. Development and endorsement processes can be challenging for PRO-PMs, and there are currently only a small number of NQF-endorsed PRO-PMs. While healthcare's technical infrastructure is at an unprecedented level of sophistication, developing and implementing digital PRO-PMs remains difficult. Understanding the current state of these opportunities and barriers, as described in this Environmental Scan Report, is the first step in navigating the road from high quality PROMs to PRO-PMs. The Interim Report lists the attributes of high quality PROMs for use in performance measures and offers guidance to measure developers on selecting a data collection instrument. Finally, the Technical Guidance Report will identify stages and tasks that measure developers—from entry-level employees to veteran developers with decades of experience—can follow when creating digital PRO-PMs for CMS accountability programs.

References

1. National Quality Forum (NQF). NQF: Quality Positioning System ™. Accessed March 1, 2022. https://www.qualityforum.org/QPS

2. Centers for Medicare & Medicaid Services (CMS). MIPS Explore Measures - QPP. Quality Payment Program. Accessed March 28, 2022. https://qpp.cms.gov/mips/explore-measures?tab=qualityMeasures&py=2022

3. Centers for Medicare & Medicaid Services (CMS). Meaningful Measures 2.0: Moving from measure reduction to modernization. CMS. Published March 31, 2021. Accessed March 1, 2022. https://www.cms.gov/meaningful-measures-20-moving-measure-reduction-modernization

4. Centers for Medicare & Medicaid Services (CMS). *Understanding Clinical Quality Measures: How CMS Is Modernizing Its Approach to Digital Measurement*.; 2020. Accessed December 27, 2020. https://www.youtube.com/watch?v=fsChOWzM2GI&feature=youtu.be

5. Cella D, Hahan E, Jensen S, Butt Z, Nowinski C, Rothrock N. Methodological issues in the selection, administration and use of patient-reported outcomes in performance measurement in health care settings. Published online September 28, 2012. http://www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=72156

6. Cella DF, Hahn EA, Jensen SE, et al. *Patient-Reported Outcomes in Performance Measurement*. Vol 97. RTI Press Research Triangle Park, NC; 2015.

7. National Quality Forum (NQF). Patient reported outcomes (PROs) in performance measurement. Published online January 10, 2013. https://www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=72537

8. National Quality Forum (NQF). Measuring what matters to patients: Innovations in integrating the patient experience into development of meaningful performance measures. Published online August 28, 2017. http://www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=85770

9. National Quality Forum (NQF). *Patient-Reported Outcomes: Best Practices on Selection and Data Collection - Final Technical Report*.; 2020. Accessed December 21, 2020. http://www.qualityforum.org/Publications/2020/09/Patient-Reported_Outcomes__Best_Practices_on_Selection_and_Data_Collection_-_Final_Technical_Report.aspx

10. National Quality Forum (NQF). Measuring performance. Accessed March 3, 2022. https://www.qualityforum.org/Measuring_Performance/Measuring_Performance.aspx

11. National Institute of Standards and Technology (NIST). Baldrige Performance Excellence Program. NIST. Published October 7, 2021. Accessed March 28, 2022. https://www.nist.gov/baldrige/how-baldrige-works/about-baldrige

12. Centers for Medicare & Medicaid Services (CMS). *Patient Reported Outcome Measures*.; 2021:6. Accessed March 1, 2022. https://www.cms.gov/files/document/blueprint-patient-reported-outcome-measures.pdf

13. Centers for Medicare & Medicaid Services (CMS). Innovation Models | CMS Innovation Center. Accessed January 5, 2021. https://innovation.cms.gov/innovation-models#views=models

14. Basch E, Spertus J, Dudley RA, et al. Methods for developing patient-reported outcome-based performance measures (PRO-PMs). *Value in Health*. 2015;18(4):493-504. doi:10.1016/j.jval.2015.02.018

15. Gelb Safran D, Higgins A. Getting to the next generation of performance measures for value-based payment. Health Affairs Blog. Accessed December 27, 2020. https://www.healthaffairs.org/do/10.1377/hblog20190128.477681/full/

16. Centers for Medicare & Medicaid Services (CMS). Meaningful Measures Framework | CMS. Centers for Medicare and Medicaid Services. Accessed January 4, 2021. https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/QualityInitiativesGenInfo/CMS-Quality-Strategy

17. Centers for Medicare & Medicaid Services (CMS). *CMS Measures Management System Blueprint.*; 2021:111. Accessed March 1, 2022. https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/MMS/Downloads/Blueprint.pdf

18. Acosta J, Tang P, Regal S, et al. Investigating the Bias in Orthopaedic Patient-reported Outcome Measures by Mode of Administration: A Meta-analysis. *JAAOS Global Research & Reviews*. 2020;4(12):e20.00194. doi:10.5435/JAAOSGlobal-D-20-00194

19. Long C, Beres LK, Wu AW, Giladi AM. Patient-level barriers and facilitators to completion of patient-reported outcomes measures. *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation*. Published online 2021. doi:10.1007/s11136-021-02999-8

20. Churruca 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

21. Health Services Advisory Group. CAHPS Hospital Survey (HCAHPS) quality assurance guidelines, v. 17.0. Published online March 2022. Accessed March 1, 2022. https://hcahpsonline.org/globalassets/hcahps/quality-assurance/2022_qag_v17.0.pdf

22. National Quality Forum (NQF). Measure evaluation criteria and guidance for evaluating measures for endorsement. National Quality Forum. Accessed March 1, 2022. https://www.qualityforum.org/show_content.aspx?id=322

23. Chee TT, Ryan AM, Wasfy JH, Borden WB. Current state of value-based purchasing programs. *Circulation*. 2016;133(22):2197-2205. doi:10.1161/CIRCULATIONAHA.115.010268

24. Centers for Medicare & Medicaid Services (CMS). Hospital value-based purchasing | CMS. Accessed December 29, 2020. https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/HVBP/Hospital-Value-Based-Purchasing

25. Quality Payment Program (QPP). Advanced APMs - QPP. Accessed December 29, 2020. https://qpp.cms.gov/apms/advanced-apms

26. Background of the PRO-CTCAE. Accessed December 30, 2020. https://healthcaredelivery.cancer.gov/pro-ctcae/background.html

27. National Institutes of Health (NIH). Development, testing, and implementation of the PRO-CTCAE Measurement System. Development, testing, and implementation. Published April 12, 2021. https://healthcaredelivery.cancer.gov/pro-ctcae/testing.html

28. National Institutes of Health (NIH). Patient-Reported Outcomes Measurement Information System (PROMIS). National Institute on Aging. Accessed December 29, 2020. http://www.nia.nih.gov/research/resource/patient-reported-outcomes-measurement-information-system-promis

29. Spertus JA, Jones PG. Development and validation of a short version of the Kansas City Cardiomyopathy Questionnaire. *Circulation: Cardiovascular Quality and Outcomes*. 2015;8(5):469-476. doi:10.1161/CIRCOUTCOMES.115.001958

30. Bilbao A, Escobar A, García-Perez L, Navarro G, Quirós R. The Minnesota Living With Heart Failure questionnaire: comparison of different factor structures. *Health Qual Life Outcomes*. 2016;14(1):23. doi:10.1186/s12955-016-0425-7

31. Yee D, Novak E, Platts A, Nassif ME, LaRue SJ, Vader JM. Comparison of the Kansas City Cardiomyopathy Questionnaire and Minnesota Living With Heart Failure Questionnaire in predicting heart failure outcomes. *The American Journal of Cardiology*. 2019;123(5):807-812. doi:10.1016/j.amjcard.2018.11.037

32. Gagnier JJ. Patient reported outcomes in orthopaedics. *Journal of Orthopaedic Research*. 2017;35(10):2098-2108. doi:https://doi.org/10.1002/jor.23604

33. Hospital for Special Surgery (HSS). Osteoarthritis & joint replacement outcome surveys. Hospital for Special Surgery. Accessed December 30, 2020. https://www.hss.edu/hoos-jr-koos-jr-outcomes-surveys.asp

34. Polascik BA, Hidaka C, Thompson MC, et al. Crosswalks Between Knee and Hip Arthroplasty Short Forms: HOOS/KOOS JR and Oxford. *J Bone Joint Surg Am*. 2020;102(11):983-990. doi:10.2106/JBJS.19.00916

35. Lyman S, Lee YY, Franklin PD, Li W, Cross MB, Padgett DE. Validation of the KOOS, JR: A short-form knee arthroplasty outcomes survey. *Clin Orthop Relat Res*. 2016;474(6):1461-1471. doi:10.1007/s11999-016-4719-1

36. Lyman S, Lee YY, Franklin PD, Li W, Mayman DJ, Padgett DE. Validation of the HOOS, JR: A short-form hip replacement survey. *Clin Orthop Relat Res*. 2016;474(6):1472-1482. doi:10.1007/s11999-016-4718-2

37. Centers for Medicare & Medicaid Services (CMS). Comprehensive Care for Joint Replacement Model | CMS Innovation Center. Accessed January 5, 2021. https://innovation.cms.gov/innovation-models/cjr

38. National Quality Forum (NQF). *Building a Roadmap From Patient-Reported Outcome Measures to Patient-Reported Outcome Performance Measures: Interim Report.*; 2021:53. Accessed April

4, 2022.
https://www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=95979

39. ICHOM | About Us | Our Mission | History | Michael Porter. ICHOM – International Consortium for Health Outcomes Measurement. Accessed December 30, 2020. https://www.ichom.org/mission/

40. International Consortium for Health Outcomes Measurement (ICHOM). Hip & knee osteoarthritis data collection reference guide. Published online 2022. Accessed March 1, 2022. https://connect.ichom.org/wp-content/uploads/2022/02/16-Hip_Knee_Osteoarthritis_Reference_Guide_V4.0_24.01.2022.docx

41. Patient Reported Outcome Measures. Accessed December 30, 2020. https://www.aaos.org/quality/research-resources/patient-reported-outcome-measures/upper-extremity-proms/

42. Patient Reported Outcome Measures. Accessed December 30, 2020. https://www.aaos.org/quality/research-resources/patient-reported-outcome-measures/lower-extremity-performance-measures/

43. Patient Reported Outcome Measures. Accessed December 30, 2020. https://www.aaos.org/quality/research-resources/patient-reported-outcome-measures/spine/

44. Patient Reported Outcome Measures. Accessed December 30, 2020. https://www.aaos.org/quality/research-resources/patient-reported-outcome-measures/treatment-outcome-quality-of-life/

45. SGO Calls on Clinicians to Measure Patient-Reported Outcomes Routinely. Mass General Advances in Motion. Accessed December 30, 2020. https://advances.massgeneral.org/obgyn/journal.aspx?id=1695

46. Sisodia RC, Dewdney SB, Fader AN, et al. Patient reported outcomes measures in gynecologic oncology: A primer for clinical use, part I. *Gynecologic Oncology*. 2020;158(1):194-200. doi:10.1016/j.ygyno.2020.04.696

47. Patient Reported Outcomes. Accessed December 30, 2020. https://www.aan.com/policy-and-guidelines/quality/quality-measures2/patient-reported-outcome-pro/

48. Deutsch A, Smith L, Gage B, Kelleher C, Garfinkel D. Patient-reported outcomes in performance measurement commissioned paper on PRO-based performance measures for healthcare accountable entities. Published online October 22, 2012. http://www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=72157

49. Centers for Medicare & Medicaid Services (CMS). *Medicare Program; Hospital Inpatient Prospective Payment Systems for Acute Care Hospitals and the Long-Term Care Hospital Prospective Payment System and Proposed Policy Changes and Fiscal Year 2022 Rates; Quality Programs and Medicare Promoting Interoperability Program Requirements for Eligible Hospitals and Critical Access Hospitals; Proposed Changes to Medicaid Provider Enrollment; and Proposed Changes to the Medicare Shared Savings Program*.; 2021:25550. Accessed April 2,

2022. https://www.federalregister.gov/documents/2021/05/10/2021-08888/medicare-program-hospital-inpatient-prospective-payment-systems-for-acute-care-hospitals-and-the

50. Centers for Medicare & Medicaid Services (CMS). *Medicare Program; Hospital Inpatient Prospective Payment Systems for Acute Care Hospitals and the Long-Term Care Hospital Prospective Payment System and Policy Changes and Fiscal Year 2022 Rates; Quality Programs and Medicare Promoting Interoperability Program Requirements for Eligible Hospitals and Critical Access Hospitals; Changes to Medicaid Provider Enrollment; and Changes to the Medicare Shared Savings Program*.; 2021:45344. Accessed March 28, 2022. https://www.govinfo.gov/content/pkg/FR-2021-08-13/pdf/2021-16519.pdf

51. National Quality Forum (NQF). NQF health information technology glossary: A guide to HIT jargon. https://www.qualityforum.org/Measuring_Performance/HIT_Glossary.aspx..

52. Andress J, Cobb K. Overview of Digital Quality Measurement. Presented at: NQF Leveraging Electronic Health Record (EHR) Sourced Measures to Improve Care Communication and Coordination: Option Year, Web Meeting 4; March 21, 2022. https://www.qualityforum.org/ProjectMaterials.aspx?projectID=94203

53. Office of the National Coordinator for Health Information Technology (ONC). *What Is FHIR?*; 2019. Accessed March 4, 2022. https://www.healthit.gov/sites/default/files/2019-08/ONCFHIRFSWhatIsFHIR.pdf

54. Health Level 7 International. Overview - FHIR v4.0.1. Accessed February 3, 2021. https://www.hl7.org/fhir/overview.html

55. Office of the National Coordinator for Health Information Technology (ONC). *21st Century Cures Act: Interoperability, Information Blocking, and the ONC Health IT Certification Program*.; 2020. Accessed April 2, 2022. https://www.federalregister.gov/documents/2020/05/01/2020-07419/21st-century-cures-act-interoperability-information-blocking-and-the-onc-health-it-certification

56. Office of the National Coordinator for Health Information Technology (ONC). *Cures Act Final Rule: United States Core Data for Interoperability*.; 2020. Accessed March 22, 2022. https://www.healthit.gov/cures/sites/default/files/cures/2020-03/USCDI.pdf

57. Argentieri R, Myers E, Posnack S, Tripathi M. Thinking Outside the Box: The USCDI+ Initiative. Health IT Buzz. Published October 8, 2021. Accessed March 22, 2022. https://www.healthit.gov/buzz-blog/health-it/thinking-outside-the-box-the-uscdi-initiative

58. Office of the National Coordinator for Health Information Technology (ONC). United States Core Data for Interoperability (USCDI). HealthIT.gov. Accessed March 22, 2022. https://www.healthit.gov/isa/united-states-core-data-interoperability-uscdi

59. Office of the National Coordinator for Health Information Technology (ONC). USCDI+. HealthIT.gov. Published October 28, 2021. Accessed March 22, 2022. https://www.healthit.gov/topic/interoperability/uscdi-plus

60. Fairclough DL, Peterson HF, Cella D, Bonomi P. Comparison of several model-based methods for analysing incomplete quality of life data in cancer clinical trials. *Statistics in Medicine*. 1998;17(5-7):781-796.

61. Brundage M, Bass B, Davidson J, et al. Patterns of reporting health-related quality of life outcomes in randomized clinical trials: implications for clinicians and quality of life researchers. *Quality of Life Research*. 2011;20(5):653-664.

62. National Quality Forum (NQF). *Patient-Reported Outcomes: Approaches and Challenges in Selection and Data Collection - An Environmental Scan*.; 2019. Accessed December 28, 2020. http://www.qualityforum.org/Publications/2019/12/Patient-Reported_Outcomes_Environmental_Scan.aspx

63. Centers for Medicare & Medicaid Services (CMS). Promoting Interoperability Programs | CMS. Accessed February 3, 2021. https://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms

64. Forrey AW, McDonald CJ, DeMoor G, et al. Logical Observation Identifier Names and Codes (LOINC) database: a public use set of codes and names for electronic reporting of clinical laboratory test results. *Clin Chem*. 1996;42(1):81-90.

65. Hauser RG, Quine DB, Ryder A, Campbell S. Unit conversions between LOINC codes. *Journal of the American Medical Informatics Association*. 2018;25(2):192-196. doi:10.1093/jamia/ocx056

66. Vreeman DJ, McDonald CJ, Huff SM. Representing patient assessments in LOINC®. *AMIA Annu Symp Proc*. 2010;2010:832-836.

67. LOINC. Survey instruments archives. LOINC. Accessed February 3, 2021. https://loinc.org/panels/category/survey-instruments/

68. 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/bmjoq-2019-000789

69. Centers for Medicare & Medicaid Services (CMS). Alternative payment models (APMs) overview - QPP. CMS.gov. Accessed April 20, 2021. https://qpp.cms.gov/apms/overview

70. McGlothlin AE, Lewis RJ. Minimal clinically important difference: Defining what really matters to patients. *JAMA*. 2014;312(13):1342-1343. doi:10.1001/jama.2014.13128

71. National Quality Forum (NQF). *NQF: Attribution - Principles and Approaches*.; 2016. https://www.qualityforum.org/Publications/2016/12/Attribution_-_Principles_and_Approaches.aspx.

72. Polascik BA, Hidaka C, Thompson MC, et al. Crosswalks between knee and hip arthroplasty short forms: HOOS/KOOS JR and Oxford. *JBJS*. 2020;102(11):983-990. doi:10.2106/JBJS.19.00916

73. Centers for Medicare & Medicaid Services (CMS). *Supplemental Materials to MMS Blueprint: Patient Reported Outcome Measures*.; 2021:6. Accessed March 3, 2022. https://www.cms.gov/files/document/blueprint-patient-reported-outcome-measures.pdf

74. Borsboom D, Molenaar D. Psychometrics. In: Wright JD, ed. *International Encyclopedia of the Social & Behavioral Sciences (Second Edition)*. Elsevier; 2015:418-422. doi:10.1016/B978-0-08-097086-8.43079-5

75. Centers for Medicare & Medicaid Services (CMS). Value-based programs. CMS.gov. Published December 1, 2021. Accessed March 3, 2022. https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/Value-Based-Programs

76. National Quality Forum (NQF). 2017 Consensus Development Process redesign. National Quality Forum. https://www.qualityforum.org/docs/2017_Kaizen__CDP_Redesign_Report.aspx

77. NQF: Scientific Methods Panel. Accessed December 30, 2020. https://www.qualityforum.org/Measuring_Performance/Scientific_Methods_Panel.aspx

78. National Quality Forum (NQF). Intent to submit. Accessed January 5, 2021. http://www.qualityforum.org/Measuring_Performance/Consensus_Development_Process_s_Principle/Intent_to_Submit.aspx

79. National Quality Forum (NQF). Call for nominations. Accessed January 5, 2021. http://www.qualityforum.org/Measuring_Performance/Consensus_Development_Process_s_Principle/Call_for_Nominations.aspx

80. Office of Management and Budget, Executive Office of the Presiden. CIRCULAR NO. A-119 Revised: Federal participation in the development and use of voluntary consensus standards and in conformity assessment activities. Published February 10, 1998. https://www.whitehouse.gov/wp-content/uploads/2017/11/Circular-119-1.pdf

81. National Quality Forum (NQF). Measure review. Accessed January 5, 2021. http://www.qualityforum.org/Measuring_Performance/Consensus_Development_Process_s_Principle/Measure_Review.aspx

82. National Quality Forum (NQF). Public commenting with member support. Accessed January 5, 2021. http://www.qualityforum.org/Measuring_Performance/Consensus_Development_Process_s_Principle/Public_Commenting_with_Member_Support.aspx

83. National Quality Forum (NQF). Measure endorsement. Accessed January 5, 2021. http://www.qualityforum.org/Measuring_Performance/Consensus_Development_Process/CSAC_Decision.aspx

84. NQF: Consensus Standards Approval Committee. Accessed December 30, 2020. https://www.qualityforum.org/About_NQF/CSAC/Consensus_Standards_Approval_Committee.aspx

85. National Quality Forum (NQF). Measure appeals. Accessed January 5, 2021. http://www.qualityforum.org/Measuring_Performance/Consensus_Development_Process_s_Principle/Measure_Appeals.aspx

Appendices

Appendix A: Glossary of Terms

Alternative Payment Models (APMs)

A payment approach that gives added incentive payments to provide high quality and cost-efficient care. APMs can apply to a specific clinical condition, care episode, or population.[69]

Anchors

Anchor-based methods are one of three types of methods used to determine minimal clinically important difference (MCID). A numerical scale for an outcome is "anchored" to a subjective and independent assessment of improvement. For example, a response of "a little better" to a question about how the patient feels post-treatment can be anchored to a numeric outcome.[70]

Attribute

A characteristic or trait of a PROM. Past National Quality Forum (NQF) reports have used *attribute* and *characteristic* synonymously.[7,9] Throughout the Building a Roadmap initiative, *attributes* primarily refer to the characteristics that make a PROM suitable for use in a PRO-PM.

Attribute Grid

A table designed to provide a systematic method to perform a side-by-side comparison of PROMs on the basis of meaningful PROM attributes.[9]

Attribution

A process used in quality measurement that aims to assign accountability for a patient's outcomes to a clinician, groups of clinicians, or a facility.[71]

Burden

Burden refers to the time, effort, or other demands placed on respondents or those administering the PROM. This can include the number and complexity of items and the literacy level needed to understand and complete the measure.[48]

Crosswalk

A concordance table to convert scores from one scale to the other and vice versa.[72] Crosswalks can allow harmonization of PROMs that measure similar outcomes (e.g., HRQoL after a knee replacement surgery), which may facilitate multicenter collaboration or allow sites to switch PROMs without loss of historic comparison data.[72]

Cut Points

Clinically meaningful thresholds of a score change within a PROM that is often associated with either improvement in patient outcome or indication of need for treatment.[70]

Digital Quality Measures (dQMs)

Software that processes digital data to produce a measure score or measure scores. Data sources for dQMs may include administrative systems, electronically submitted clinical assessment data, case management systems, EHRs, instruments (for example, medical devices and wearable devices), patient portals or applications (for example, for collection of patient-generated health data), HIEs or registries, and other sources. We also note that dQMs are intended to improve the patient experience including quality of care, improve the health of populations, and/or reduce costs.[50]

Electronic Clinical Quality Measures (eCQMs)

eCQMs are expressed and formatted to use data from EHRs and/or health IT systems to measure healthcare quality, ideally data captured in structured form during the process of patient care.[17] They are the most common type of digital quality measures and are specified for use in the Medicare and Medicaid EHR Incentive Programs. Eligible professionals, eligible hospitals, and critical access hospitals are required to submit eCQM data from certified EHR technology to help measure and track the quality of healthcare services provided within the healthcare system.[51]

Fast Healthcare Interoperability Resources (FHIR)

A Health Level Seven International (HL7) standard that defines how healthcare information can be exchanged between different computer systems regardless of how it is stored in those systems.[53]

Interoperability

The degree to which the meaning of the scores can be easily understood by any group requiring use of the scores. A PRO measure should have documentation to support interpretation of scores, including the meaning of low and high scores and guidance on the minimally important difference in scores between groups and/or over time.[7]

Logical Observation Identifiers, Names, and Codes (LOINC)

LOINC is a database and universal standard for identifying medical observations. It was developed in 1994 and is maintained by the Regenstrief Institute, a U.S. nonprofit medical research organization. LOINC was created in response to the demand for an electronic database for clinical care and management and is publicly available at no cost.[51]

Method

How PROM data are collected, such as via a paper form or a patient portal.

Minimal Clinically Important Difference (MCID)

MCID is the smallest improvement needed after treatment that would be considered worthwhile from the patient's perspective.[70] MCID can be calculated using three different methods: consensus or delphi method, which depends on consensus of an expert panel; anchors (described above); and a distribution-based method, which relies on the statistical analysis of the distribution of outcome scores.[70]

Mode

How a PROM is administered, such as self-administration or verbal administration by a clinician.

Patient-Reported Outcome (PRO)

Any report of the status of a patient's health condition or health behavior that comes directly from the patient, without interpretation of the patient's response by a clinician or anyone else.[73]

Patient-Reported Outcome Measure (PROM)

Tools used to collect patient-reported outcomes.[73]

Patient-Reported Outcome Performance Measure (PRO-PM)

A way to aggregate the information from patients into a reliable, valid measure of performance at the measured entity, level, e.g., clinician.[73]

Performance Measures (PMs)

These are standards that can be used to measure and quantify healthcare processes, outcomes, patient perceptions, organizational structure, and/or systems that are associated with the ability to provide high quality care.[10]

Psychometric Soundness

How consistently and accurately an assessment measures what it purports to measure.[48] Validity and reliability are key aspects to attaining psychometric soundness. Psychometrics is a scientific discipline concerned with the construction of measurement models for psychological data.[74]

United States Core Data for Interoperability (USCDI)

A standardized set of health data classes and constituent data elements for nationwide, interoperable health information exchange.[58]

USCDI+

An ONC initiative that supports the identification and establishment of domain or program-specific datasets that will operate as extensions to the existing USCDI. It is a service for federal agencies who have a need to establish, harmonize, and advance the use of interoperable datasets that extend beyond the core data in the USCDI in order to meet agency-specific programmatic requirements. The three pillars of USCDI+ are collaboration, harmonization, and specification.[59]

Value-Based Purchasing (VBP) Program

Value-based programs reward healthcare providers with incentive payments or penalties for the quality of care they give to people with Medicare. These programs are part of CMS' larger quality strategy to reform how healthcare is delivered and paid for.[75]

Appendix B: Search Terms

To gain a broad understanding of literature related to PRO-PMs and existing guidance for developing PRO-PMs, various search terms were included within PubMed and Google Scholar queries. An initial PubMed search that incorporated potentially relevant Medical Subject Heading (MeSH) terms yielded zero results. Given that MeSH terms are not in existence for PROMs or PRO-PMs, general search terms and phrases were utilized. Such phrases included the following:

- "PROM"
- "PRO-PM"
- "Patient-reported outcome-performance measure" and "Patient-reported outcome-performance measures"
- "Patient-reported outcome measure" and "patient-reported outcome measures"
- "Attributes of patient-reported outcome measures"
- "Development of patient-reported outcome measures"
- "PRO-PM guidance"
- "Logical Observation Identifiers, Names, and Codes (LOINC)"
- "Patient assessments" and "LOINC"
- "HL7 FHIR"
- "Interoperability"
- "Digital quality measures"
- "Digital quality measurement"
- "dQMs"

Terms also included specific searches for those PROMs referenced in Appendix C.

Appendix C: PROMs Discussed in This Report

The following PROMs were used as potential examples of high quality PROMs based on recommendations by CMS, documentation for CMS VBP programs and/or APMs, the TEP, and information identified during literature reviews. PROMs are linked to a homepage or developer site where possible.

- Hip disability and Osteoarthritis Outcome Score, Joint Replacement (HOOS, JR)
- Kansas City Cardiomyopathy Questionnaire – 12 item (KCCQ-12)
- Knee injury and Osteoarthritis Outcome Score, Joint Replacement (KOOS, JR)
- Minnesota Living with Heart Failure Questionnaire (MLHFQ)
- Patient Health Questionnaire (PHQ-9)
- Patient-Reported Outcomes Measurement Information System (PROMIS)
- Patient Reported Outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE)

Appendix D: Reference Tables

Table 2: Current NQF-Endorsed PRO-PMs

Measure Title	NQF#	Measure Steward	Updated Date
Adolescent Assessment of Preparation for Transition (ADAPT) to Adult-Focused Health Care	2789	Center of Excellence for Pediatric Quality Measurement	May 04, 2016
Bereaved Family Survey	1623	Department of Veterans Affairs / Hospice and Palliative Care	October 11, 2017
CAHPS Clinician & Group Surveys (CG-CAHPS) Version 3.0 -Adult, Child	0005	Agency for Healthcare Research and Quality	October 25, 2019
CAHPS® Home- and Community-Based Services Measures	2967	Centers for Medicare & Medicaid Services	November 09, 2020
CAHPS® Home Health Care Survey (experience with care) NQF#: 0517	0517	Centers for Medicare & Medicaid Services	November 09, 2020
CAHPS® Hospice Survey (experience with care)	2651	Centers for Medicare & Medicaid Services	November 20, 2020
CollaboRATE Shared Decision Making Score	3227	The Dartmouth Institute for Health Policy & Clinical Practice	October 25, 2019
Consumer Assessment of Healthcare Providers and Systems (CAHPS) In-Center Hemodialysis Survey (ICH CAHPS)	0258	Centers for Medicare & Medicaid Services	October 25, 2019
Consumer Assessment of Healthcare Providers and Systems (CAHPS) Health Plan Survey, Version 5.0 (Medicaid and Commercial)	0006	Agency for Healthcare Research and Quality	October 25, 2019
Consumer Assessment of Healthcare Providers and Systems (CAHPS)® Surgical Care Survey Version 2.0	1741	American College of Surgeons	June 05, 2018
CoreQ: AL Family Satisfaction Measure	3422	American Health Care Association	October 26, 2018
CoreQ: AL Resident Satisfaction Measure	3420	American Health Care Association	October 26, 2018
CoreQ: Long-Stay Family Measure	2616	American Health Care Association/National Center for Assisted Living	November 20, 2020
CoreQ: Long-Stay Resident Measure	2615	American Health Care Association	November 20, 2020
CoreQ: Short Stay Discharge Measure	2614	American Health Care Association/National Center for Assisted Living	November 20, 2020

Measure Title	NQF#	Measure Steward	Updated Date
Depression Remission at Six Months	0711	MN Community Measurement	March 06, 2015
Depression Remission at Twelve Months	0710e	MN Community Measurement	March 06, 2015
Depression Response at Six Months- Progress Towards Remission	1884	MN Community Measurement	February 08, 2016
Depression Response at Twelve Months- Progress Towards Remission	1885	MN Community Measurement	October 26, 2016
Functional Status Change for Patients with Low Back Impairments	0425	Focus on Therapeutic Outcomes	July 31, 2020
Functional Status Change for Patients with Neck Impairments	3461	Focus on Therapeutic Outcomes	October 25, 2019
Gains in Patient Activation (PAM) Scores at 12 Months	2483	Insignia Health	April 07, 2016
Hospital-Level, Risk-Standardized Patient-Reported Outcomes Following Elective Primary Total Hip and/or Total Knee Arthroplasty (THA/TKA)	3559	Centers for Medicare & Medicaid Services	November 20, 2020
Informed, Patient Centered (IPC) Hip and Knee Replacement Surgery	2958	Massachusetts General Hospital	October 26, 2016
National Core Indicators for Intellectual and Developmental Disabilities (ID/DD) Home- and Community-Based Services (HCBS) Measures	3622	Human Services Research Institute	November 30, 2021
Person-Centered Contraceptive Counseling (PCCC) measure	3543	University of California, San Francisco	February 08, 2021
Person-Centered Primary Care Measure PRO-PM	3568	American Board of Family Medicine	July 02, 2021
Shared Decision Making Process	2962	Massachusetts General Hospital	September 06, 2017

Table 3: PRO-PMs Used in CMS Programs

The following PRO-PMs are located within CMS' Measure Inventory Tool as active in federal programs. Please note, those measures listed below are not necessarily NQF endorsed. Those that are NQF-endorsed are indicated with an asterisk (*).

CMIT Ref. No	Measure Title	Program
02802-C-MQRS	Access to Care	Marketplace Quality Rating System
02803-C-MQRS	Access to Information	Marketplace Quality Rating System
05597-C-MIPS	Back Pain After Lumbar Discectomy/Laminectomy	Merit-Based Incentive Payment System Program
05598-C-MIPS	Back Pain After Lumbar Fusion	Merit-Based Incentive Payment System Program
02517-C-PC, 02517-C-MIPS	CAHPS for MIPs Clinician/Group Survey	Doctors & Clinicians Compare, Merit-Based Incentive Payment System Program
02575-C-DFC	CAHPS In-Center Hemodialysis Consumer Assessment of Healthcare Providers and Systems (ICH CAHPS) Survey Administration	Dialysis Facility Compare
05142-X-MSSP	CAHPS: Care Coordination	Medicare Shared Savings Program
05141-X-MSSP	CAHPS: Courteous and Helpful Office Staff	Medicare Shared Savings Program
02830-C-MQRS, 04007-C-PARTC	Care Coordination	Marketplace Quality Rating System, Medicare Part C Star Rating
01049-C-ASCQR, 01049-C-HC, 01049-C-MIPS	Cataracts: Improvement in Patient's Visual Function within 90 Days Following Cataract Surgery	Ambulatory Surgical Center Quality Reporting, Hospital Compare, Merit-Based Incentive Payment System Program
01052-C-MIPS	Cataracts: Patient Satisfaction within 90 Days Following Cataract Surgery	Merit-Based Incentive Payment System Program
06186-C-HEDIS	Children With Chronic Conditions (CCC)	HEDIS Quality Measure Rating System
04015-C-PARTC, 04015-C-PARTD	Complaints about the Drug Plan / Complaints about the Health Plan	Medicare Part C Star Rating, Medicare Part D Star Rating

CMIT Ref. No	Measure Title	Program
02840-C-HEDIS	Consumer Assessment of Healthcare Providers and Systems (CAHPS) Health Plan Survey 5.0H Child Version Including Medicaid and Children with Chronic Conditions Supplemental Items	HEDIS Quality Measure Rating System
02841-C-HEDIS, 02841-C-MACS	Consumer Assessment of Healthcare Providers and Systems (CAHPS) Health Plan Survey 5.0H, Adult Version (Medicaid)	HEDIS Quality Measure Rating System, Medicaid: Adult Core Set
02840-X-MCCS	Consumer Assessment of Healthcare Providers and Systems (CAHPS) Health Plan Survey 5.1H Child Version Including Medicaid and Children with Chronic Conditions Supplemental Items (CPC-CH)	Medicaid: Child Core Set
02155-C-HPEC, 02155-C-HQR	Consumer Assessment of Healthcare Providers and Systems (CAHPS) Hospice Survey*	Hospice Care Compare, Hospice Quality Reporting
04018-C-PARTC	Customer Service	Medicare Part C Star Rating
02853-C-HEDIS, 02853-C-MQRS, 02853-C-MACS	Flu Vaccinations for Adults Ages 18 to 64	HEDIS Quality Measure Rating System, Marketplace Quality Rating System, Medicaid: Adult Core Set
06194-C-HEDIS	Flu Vaccinations for Adults Ages 65 and Older (FVO)	HEDIS Quality Measure Rating System
05878-C-MIPS	Functional Status After Lumbar Discectomy/Laminectomy	Merit-Based Incentive Payment System Program
05877-C-MIPS	Functional Status After Lumbar Fusion	Merit-Based Incentive Payment System Program
05876-C-MIPS	Functional Status After Primary Total Knee Replacement	Merit-Based Incentive Payment System Program
05828-E-PC, 05828-E-MIPS	Functional Status Assessment for Total Hip Replacement	Doctors & Clinicians Compare, Merit-Based Incentive Payment System Program

CMIT Ref. No	Measure Title	Program
05833-E-PC, 05833-E-MIPS	Functional Status Assessment for Total Knee Replacement	Doctors & Clinicians Compare, Merit-Based Incentive Payment System Program
05840-E-MIPS	Functional Status Assessments for Congestive Heart Failure	Merit-Based Incentive Payment System Program
01263-C-PC, 01263-C-MIPS	Functional Status Change for Patients with Elbow, Wrist or Hand Impairments	Doctors & Clinicians Compare, Merit-Based Incentive Payment System Program
01251-C-PC, 01251-C-MIPS	Functional Status Change for Patients with Hip Impairments	Doctors & Clinicians Compare, Merit-Based Incentive Payment System Program
01248-C-PC, 01248-C-MIPS	Functional Status Change for Patients with Knee Impairments	Doctors & Clinicians Compare, Merit-Based Incentive Payment System Program
01257-C-PC, 01257-C-MIPS	Functional Status Change for Patients with Low Back Impairments	Doctors & Clinicians Compare, Merit-Based Incentive Payment System Program
01254-C-MIPS	Functional Status Change for Patients with Lower Leg, Foot or Ankle Impairments	Merit-Based Incentive Payment System Program
06045-C-MIPS	Functional Status Change for Patients with Neck Impairments	Merit-Based Incentive Payment System Program
01260-C-MIPS	Functional Status Change for Patients with Shoulder Impairments	Merit-Based Incentive Payment System Program
04025-C-PARTC	Getting Appointments and Care Quickly	Medicare Part C Star Rating
04028-C-PARTC	Getting Needed Care	Medicare Part C Star Rating
04029-C-PARTD	Getting Needed Prescription Drugs	Medicare Part D Star Rating
00113-C-HC, 00113-C-HIQR	Hospital Consumer Assessment of Healthcare Providers and Systems Survey (HCAHPS)	Hospital Compare, Hospital Inpatient Quality Reporting
05599-C-MIPS	Leg Pain After Lumbar Discectomy/Laminectomy	Merit-Based Incentive Payment System Program
05875-C-MIPS	Leg Pain After Lumbar Fusion	Merit-Based Incentive Payment System Program

CMIT Ref. No	Measure Title	Program
06201-C-HEDIS	Medicare Health Outcomes Survey (HOS)	HEDIS Quality Measure Rating System
06151-C-MACS	National Core Indicators Survey (NCIDDS-AD)	Medicaid: Adult Core Set
00549-C-PC, 00549-C-MIPS	Nuclear Medicine: Correlation with Existing Imaging Studies for All Patients Undergoing Bone Scintigraphy	Doctors & Clinicians Compare, Merit-Based Incentive Payment System Program
02885-C-MQRS	Plan Administration	Marketplace Quality Rating System
02569-C-PC	Quality of Life Assessment For Patients With Primary Headache Disorders	Doctors & Clinicians Compare
02898-C-MQRS	Rating of All Health Care	Marketplace Quality Rating System
04089-C-PARTD	Rating of Drug Plan	Medicare Part D Star Rating
04090-C-PARTC	Rating of Health Care Quality	Medicare Part C Star Rating
02899-C-MQRS, 04091-C-PARTC	Rating of Health Plan	Marketplace Quality Rating System, Medicare Part C Star Rating
02900-C-MQRS	Rating of Personal Doctor	Marketplace Quality Rating System
02901-C-MQRS	Rating of Specialist	Marketplace Quality Rating System
05874-E-MIPS	Urinary Symptoms Score Change 6-12 Months After Diagnosis of Benign Prostatic Hyperplasia	Merit-Based Incentive Payment System Program
02554-C-MIPS	Varicose Vein Treatment with Saphenous Ablation: Outcome Survey	Merit-Based Incentive Payment System Program

Table 4: Best Practices and Considerations

The table below consists of methodological best practices and associated considerations for developing and evaluating proposed PRO-PMs, as identified by Basch et al.[14]

Best Practice	Considerations
A rationale for measuring the outcome should be described.	Is a knowledge gap described and justified? Is there evidence that the outcome is meaningful and important to patients, caregivers, and/or other stakeholders? How does patient self-reporting, in particular, address the gap? Are patients the most appropriate source of information?
The intended context of use should be described and justified.	Is the intended context of use clearly described and justified? How is information from the measure expected to inform change in practice to improve performance in the intended context of use? How will the nominated measure complement other measures to improve understanding of performance in the intended context of use? Is there variability in the outcome at the practice or practitioner level?
The measure should be adequately developed for the intended context of use (or a similar context of use), including demonstration of meaningfulness and importance to patients as well as adequate psychometric properties.	Is the underlying concept to be measured clearly identified (e.g., post-chemotherapy nausea)? Is there prior or planned qualitative work in a patient population similar to the intended context of use that demonstrates understanding of terminology and mapping of the terminology to the underlying concept(s) of interest? Is there evidence of adequate psychometric properties of the measure, including construct validity and reliability, meaningfulness of score changes in a comparable population, and reasonableness of the recall period?

Best Practice	Considerations
There should be prior or planned work using the measure in the intended context of use (or a similar context of use), demonstrating that it is sensitive to change and clinically actionable.	Has the measure been shown to detect changes over time or differences between known patient groups, practices, and/or procedures? Does the measure detect change in clinical action(s)? Is there evidence that there is not a floor or ceiling effect of the measure in the intended context of use?
There should be a recommended implementation strategy for the measure in the intended context of use.	Is there a rationale for an administration mode (e.g., paper, electronic) and schedule (e.g., timing of follow-up evaluations)? Is there a plan to maximize recruitment and response rates (e.g., backup data collection plan for nonrespondents)? Is proxy or surrogate reporting considered allowable? Is there a plan to accurately identify patients in the target population and calculate the denominator (i.e., number of people who were asked to complete the measure)?
There should be a recommended analysis plan, including a risk adjustment strategy, missing data approach, and power calculation.	Is there a well-justified, *a priori* risk adjustment or stratification strategy based on evidence? Is there a plan to adjust analyses for case mix, recruitment bias, and response bias? Is there a plan for imputing missing data with sensitivity analyses? What sample sizes are necessary for planned analyses?

Best Practice	Considerations
There should be a recommended framework for interpreting results, including unit(s) of analysis and meaningful score thresholds.	What unit of analysis is recommended (e.g., hospital system, hospital, individual practice, individual practitioner, and patient-level)? What metrics should be used to reflect performance (e.g., proportion of patients achieving a specific score change, proportion of providers who are outliers)? How are the results of different PRO measures that may not agree with each other considered?
There should be a recommended approach for reporting and disseminating results.	Is there a suggested approach for packaging and presenting reports to practices, providers, and/or patients?

Appendix E: Review of NQF PRO-PM Endorsement Process

Themes Related to NQF Analysis

NQF's endorsed PRO-PMs vary in scope but have similar goals of gathering and quantifying PROs. They reflect various conditions and topic areas, including experience with care (e.g., CAHPS or CoreQ [i.e., patient, resident, and family satisfaction for skilled nursing care centers and assisted living communities]), depression response or remission, shared decision making/patient activation, transitions of care, pain management, and contraception.

PRO-PMs are considered complex measures and are evaluated against more stringent requirements than most other measures. These include an evaluation of the psychometric properties (e.g., reliability and validity) of the PROM(s) that collect data for the PRO-PM. All complex measures, which include but are not limited to PRO-PMs, require an evaluation by the SMP.[76]

Scientific Methods Panel

The SMP is composed of approximately 25 individuals with methodological expertise who provide NQF Standing Committees with evaluations of measures' scientific acceptability. Panel members use NQF's standard measure evaluation criteria to assess new and maintenance measures.[77] The SMP's feedback is critical for endorsement recommendations by the Standing Committees and for endorsement decisions by the Consensus Standards Approval Committee (CSAC). Although the number of PRO-PMs that come through the SMP process is relatively low compared with other types of measures, the SMP recognizes the inherent complexity of PRO-PMs.

Consensus Development Process and Standing Committee Reviews

The CDP is NQF's formal, cyclical process to evaluate and endorse measures. It is designed to allow input and discussion from stakeholder groups across the industry. The CDP involves six principal steps, which are described in more detail on the Consensus Development Process webpage:

1. **Intent to Submit:** The measure developer notifies NQF at least three months prior to the designated cycle's submission deadline.[78]
2. **Call for Nominations:** NQF seats a Standing Committee to offer expert advice, ensure input is obtained from relevant stakeholders, and make recommendations to NQF membership about standards that are proposed for endorsement.[79]
3. **Measure Review:** The Standing Committee evaluates the measure, and is expected to reach consensus on whether the measure continues toward possible NQF endorsement or is returned to the developer and/or steward for refinement.[80,81]
4. **Public Commenting with Measure Support:** All Standing Committee recommendations are included in a draft report, which is opened to the public for commenting. The Standing Committee may revise a recommendation based on public comment.[82]
5. **Measure Endorsement:** The CSAC, whose members are appointed by the NQF Board of Directors, makes an endorsement decision that either upholds the Standing Committee's decision or sends the measure back for further consideration.[83,84]
6. **Measure Appeals:** After the CSAC's decisions are made public, a 30-day appeals period begins. Eligibility for an appeal must relate to procedural errors or new information.[85]

The following five criteria are considered throughout the CDP for all candidate measures:

- **Criterion 1: Importance to Measure and Report**: Extent to which the specific measure focus is evidence-based, important to making significant gains in healthcare quality, and improving health outcomes for a specific high-impact aspect of healthcare in which there is variation in or overall less-than-optimal performance.
- **Criterion 2: Reliability and Validity – Scientific Acceptability of Measure Properties**: Extent to which the measure produces consistent and credible results about the quality of care when implemented.
- **Criterion 3: Feasibility**: Extent to which the specifications, including measure logic, required data that are readily available or could be captured without undue burden and can be implemented for performance measurement.
- **Criterion 4: Usability and Use**: Extent to which potential audiences (e.g., consumers, purchasers, clinicians, and policymakers) are using or could use performance results for both accountability and performance improvement to achieve the goal of high quality, efficient healthcare for individuals or populations.
- **Criterion 5: Comparison to Related or Competing Measures:** If a measure meets all criteria and there are endorsed or new related or competing measures, the measures are compared to address harmonization and/or selection of the best measure.[22]

Appendix F: Technical Expert Panel Members, Federal Liaisons, and NQF Staff

Technical Expert Panel Members

* TEP members marked with an asterisk served in 2021; all others served 2021-2022

Catherine MacLean, MD, PhD *(Co-Chair)*
Chief Value Medical Officer, Rheumatologist, Hospital for Special Surgery
New York, New York

Sam Simon, PhD *(Co-Chair)*
Director, Health Program Improvement, Mathematica
Cambridge, Massachusetts

David Andrews, PhD*
Patient Advisor
Aiken, South Carolina

Katherine Ast, MSW, LCSW
Director, Quality & Research, American Academy of Hospice and Palliative Medicine
Glenview, Illinois

Rachel Brodie, BA
Senior Director, Measurement & Accountability, Pacific Business Group on Health
San Francisco, California

Zahid Butt, MD, FACG
President and Chief Executive Officer (CEO), Medisolv, Inc.
Columbia, Maryland

Collette Cole, BSN, RN, CPHQ
Clinical Measurement Developer, Minnesota Community Measurement
Minneapolis, Minnesota

Paula Farrell, BSN, RN, CPHQ, LSSGB*
Associate Project Director, Clinical Quality Measurement, Department of Quality Measurement, The Joint Commission
Oakbrook Terrace, Illinois

Mark Friedberg, MD, MPP
Senior Vice President, Performance Measurement & Improvement, Blue Cross Blue Shield of Massachusetts
Boston, Massachusetts

Debbie Gipson, MD, MS
Professor of Pediatrics, University of Michigan
Ann Arbor, Michigan

Ben Hamlin, DrPH
Senior Research Informaticist, Research & Performance Measurement, National Committee for Quality Assurance
Washington, District of Columbia

Janel Hanmer, MD, PhD
Medical Director of Patient Reported Outcomes, University of Pittsburgh Medical Center
Pittsburgh, Pennsylvania

Helen Haskell, MA
Founder and President, Mothers Against Medical Error
Columbia, South Carolina

Brian Hurley, MD, MBA, DFASAM
Director of Addiction Medicine, Los Angeles County Department of Health Services
Los Angeles, California

Christine Izui, MS
Health Policy Principle, Center for Transforming Health, MITRE Corporation
McLean, Virginia

Laura Jantos, LFHIMSS
Patient Advocate, Healthcare Technology/Digital Health
Mercer Island, Washington

Kirk Munsch
Patient Advocate, Rare Patient Voice
Colorado Springs, Colorado

Deborah Paone, DrPH, MHA
Performance Evaluation Lead and Policy Consultant, SNP Alliance
Washington, District of Columbia

Brenna Rabel, MPH
Lead Health Research Scientist, Battelle Memorial Institute
Arlington, Virginia

Nan Rothrock, PhD, MA
Clinical Psychologist, Feinberg School of Medicine, Northwestern University
Chicago, Illinois

Mike Sacca, AS
Vice President, Healthcare Quality and Data Analytics, RELI Group
Douglassville, Pennsylvania

Rachel Sisodia, MD*
Senior Medical Director for Specialty Care and Patient Reported Outcomes, Massachusetts General Brigham Health
Boston, Massachusetts

John Spertus, MD, MPH, FACC, FAHA
Clinical Director of Cardiovascular Education and Outcomes Research, Mid America Heart Institute
Kansas City, Missouri

Ruth Wetta, PhD, MSN, MPH, RN
Lead Clinical Researcher, Population Health, Cerner Corporation
Kansas City, Missouri

Albert Wu, MD, MPH, FACP
Professor, Department of Health Policy and Management and Medicine, Johns Hopkins Bloomberg School of Public Health
Baltimore, Maryland

Federal Liaisons

Joel Andress
CMS, Center for Clinical Standards and Quality

Janis Grady, RHIT, FAC-COR III
CMS, Center for Clinical Standards and Quality

Sandra Mitchell, PhD, CRNP, FAAN
National Institutes of Health, National Cancer Institute

Ashley Wilder Smith, PhD, MPH
National Institutes of Health, National Cancer Institute

Rhona Limcango, PhD
Agency for Healthcare Research and Quality

Meaghan McHugh, PhD, MPH
Substance Abuse and Mental Health Services Administration

Clifford A. Smith, PhD, ABPP-Cn
US Department of Veterans Affairs, Office of Mental Health and Suicide Prevention

David Au, MD, MS
US Department of Veterans Affairs, Center of Innovation for Veteran-Centered and Value-Driven Care

Girma Alemu, MD, MPH
Health Resources and Services Administration, Office of Planning, Analysis, and Evaluation

NQF Staff

Kathleen Giblin, RN
Senior Vice President

Elizabeth Drye, MD, SM
Chief Scientific Officer

Alejandra Herr
Senior Managing Director

Maha Taylor, MHA, PMP
Managing Director

Chuck Amos, MBA
Senior Director

Teresa Brown, MHA, MA
Director

Deidra Smith, MBA, PMP
Senior Project Manager

Carina Smith, MPP, MBA
Manager

Evelyn Thomas, MPH, CHES
Senior Analyst

CMS Staff

LaWanda Burwell, ScD
Task Order COR

Gequincia Polk
IDIQ COR

Helen Dollar-Maples
Director, DPMS/QMVIG/CCSQ

Marsha Smith, MD, MPH, FAAP
Medical Officer, DPMS/QMVIG/CCSQ

Appendix G: Public Comments

The draft Environmental Scan Report for the Building a Roadmap initiative was posted on the NQF project webpage for public and NQF member comment in spring 2022. NQF offered six prompts to guide public commenters on key areas of interest. Eight comments from two organizations are grouped below by prompt, and the responses from NQF and the TEP are included beneath each comment. Unless otherwise noted, public comments are presented as they were received by NQF and have not been edited, except for minor updates to spacing, spelling, and punctuation.

How could the usefulness and clarity of the Executive Summary be improved?

Partners Health Management

COMMENT

None - it is excellent as written. Very clear, concise, yet informative.

RESPONSE

Thank you for your compliment on the executive summary and for your ongoing engagement in the Building a Roadmap initiative.

Phreesia

COMMENT

Overall, the Executive Summary is a useful and relatively clear document. It is clear that much thought and consideration went into its development.

RESPONSE

Thank you for your positive feedback on the executive summary and your interest in this work.

Which findings of the report need to be either expanded or clarified? What do you recommend to improve these findings?

Phreesia

COMMENT

At points in the document, it seemed as though there was some advocacy for the need to develop new PROMs. Adding new measures, though, will not fix the problem that we have few PRO-PMs. More fundamental to ask why there are so few PRO-PMs given all the existing PROMs currently.

Reviewing the Appendix, I find that the list of currently endorsed PRO-PMs includes both measures derived from patient-reported outcomes and patient experience measures. Might it be helpful to highlight this distinction between PROMs and PREMs. Does it make sense to label them all as PRO-PMs?

RESPONSE

Thank you for raising the concern about the unintended advocacy for development of new PROMs. We have reviewed the report and revised the language to clarify that the report advocates for the development of new PRO-PMs rather than new PROMs.

As for your comment on distinction between experience measures and outcome measures, the measures in the Appendix are grouped together as PRO-PMs in alignment with NQF's current approach to classifying quality measures. The TEP will continue to discuss how best to clearly describe and classify these types of performance measures for this and future reports. NQF has also raised your suggestion to its Consensus Development Process (CDP) and Maintenance teams.

The scan aims to describe the current state of digital measurement as it pertains to PRO-PMs. What is missing or unclear in the information on digital measurement and dQMs?

Phreesia

COMMENT

Overall, the document does a good job describing the current state of digital PRO-PMs. One question that was left unanswered for me, though, was what is known about the usage of the existing PRO-PMs. Yes, of course we can increase the numbers of PRO-PMs available, but what do we know about the ones that are already in play? Do we have clarity on the known barriers of PRO-PM utilization?

RESPONSE

We agree that this information would be valuable, but detailed data on the utilization of existing PRO-PMs are not readily available. Appendix D (specifically Table 3) does include information about PRO-PMs that are used in existing CMS programs.

How can we improve the section titled, "Challenges and Barriers of Developing PRO-PMs"?

Phreesia

COMMENT

Whether in this section or another, it would be helpful to delineate the challenges of developing high-quality PRO-PM (in any format) and the incremental challenges of digitizing those PRO-PMs. What are the known challenges or burdens to digitizing PRO-PM, above and beyond measure development?

Within the "challenges due to patient burden" section, the framing seems to place blame for suboptimal PROM completion on the patient. Please consider that part of the problem may be related to how the rationale or justification for PROM data collection was presented to the patient. Consider that the problem manifests in patients, but wasn't caused by them.

RESPONSE

In response to your suggestion, we added details to the "Challenges and Barriers of Developing PRO-PMs" section. The Technical Guidance Report will delineate challenges of developing high-quality PRO-PMs in greater detail, along with potential strategies and solutions to overcoming these challenges. While digitizing existing measures is an important consideration, it is beyond the scope of this initiative except at a very high level, so we have updated the "Limitations" section to better reflect this scope.

Thank you for raising the concern around language that could be perceived as blaming the patient for low response rates. We have revised the language under "Challenges due to Patient Burden."

What important practical information about the development and use of PROM-based PRO-PMs is not addressed within the scan?

Phreesia

COMMENT

Throughout the report, there is reference to the concept of a "high quality" PROM, but the criteria to identify such a measure is not well delineated. That may be helpful guidance to summarize or point to for those interested in developing PRO-PM.

It may also be helpful for readers to have clarity on the qualities of PROMs that may be ideal for clinical care vs for performance measurement. If that distinction isn't perceived to be meaningful, a straightforward statement to that end would also be helpful.

Reading through the document, I was also curious how many PRO-PM are currently under review for potential endorsement by the NQF.

While the document correctly outlines that there is very little specific guidance on how best to create a PRO-PM, I wonder if CMS or another entity will have the appetite to develop such guidance? FDA, for e.g., provides guidance for development of PROMs for use in outcomes assessment for use in drug labelling claims. Perhaps the report should be more forceful in calling for such guidance to be developed for PRO-PM to help advance this field.

RESPONSE

We have added this initiative's definition of a "high-quality" PROM to the "Introduction" section, and the Interim Report details the attributes of a high quality PROM for use in a PRO-PM.

The TEP leaned against making a distinction between PROM data for clinical care and performance measurement, noting that even when PROM data cannot be used in treatment decisions for the individual who completed the PROM, they can shape clinical decisions for future patients.

We cannot quantify the number of PRO-PMs under review for endorsement since this number constantly changes and would become outdated by the time the report is published.

We have added language to the "Implications of Pending Advances in Interoperability for PRO-PMs" section of the report to highlight how all stakeholders, including federal agencies, need to be involved in fostering an environment that values interoperability and elevating the patient voice. Notably, the Technical Guidance Report that is part of the CMS-funded Building a Roadmap initiative is a tangible example of a federal agency supporting new guidance that advances the field of PRO-PMs.

What general comments do you have that would improve the Environmental Scan?

Partners Health Management

COMMENT

General comment: The position of the page numbering (upper left) does create a bit of visual conflict - in particular when the reader gets to the References (starting on page 26). Sharing as information only

as I believe the overall draft document is a wonderful resource and tool and is very reader friendly providing helpful hyperlinks and references.

RESPONSE

Thank you for your comment and for Partners Health Management's engagement in this important work. We appreciate your support for the report, and we recognize that formatting decisions (e.g., page numbers) can change the presentation of content. The page numbers are consistent with the NQF style guide, but we will share your feedback with our editorial team for future consideration.

Phreesia

COMMENT

Again, we would like to reiterate that this Environmental Scan summarizes an impressive amount of thought and work to date. We offer some pointed feedback in the hopes that it may make the document even more clear and useful.

RESPONSE

Thank you for your positive comment and for your feedback on informing the improvements of the environmental scan. We appreciate your engagement in this work, and we believe the report is stronger thanks to your input.