Measure Development Primer

An introduction to measure development from the PCPI
# Table of Contents

- Introduction ................................................................................................................. 3
- Measure Conceptualization ............................................................................................ 3
  - Evidence ....................................................................................................................... 3
  - Data Source .................................................................................................................. 4
  - Care Setting .................................................................................................................. 4
  - Level of Analysis .......................................................................................................... 4
- Stakeholder Input ............................................................................................................ 5
  - Expert Work Group ...................................................................................................... 5
  - Public Comment ............................................................................................................ 5
- Measure Specifications .................................................................................................... 5
- Measure Testing .............................................................................................................. 5
- Measure Implementation ............................................................................................... 6
- Measure Endorsement .................................................................................................... 6
- Citations ........................................................................................................................ 7
Introduction

The US healthcare system is experiencing a shift in the delivery of patient care and in the approach to payment for clinicians and medical facilities. The Centers for Medicare & Medicaid Services (CMS) is moving from a fee-for-service model that pays for volume to a model that focuses on value and quality. Clinical quality measures represent opportunities to improve the quality of patient care by assessing the processes, outcomes, patient perceptions, and organizational structure and/or systems that are associated with the ability to provide high-quality health care.1 CMS seeks to implement measures that are explicitly linked to the goals outlined within the CMS Quality Strategy (QS) to improve quality of care:

1. Make Care Safer
2. Strengthen Person and Family Engagement
3. Promote Effective Communication and Coordination of Care
4. Promote Effective Prevention and Treatment of Chronic Disease
5. Work with Communities to Promote Best Practices of Healthy Living
6. Make Care Affordable

Measure developers are encouraged to prioritize performance measures that address one or more of these elements, whether the measures are intended for use in CMS programs or other quality improvement initiatives. This framework outlines some of the key considerations and processes for the development and implementation of performance measures.

Measure Conceptualization

CMS refers to measure conceptualization as the first phase in the measure lifecycle. The measure lifecycle includes the processes involved in developing and maintaining measures, with precisely specified data elements, that are feasible, valid and reliable. Measure developers frequently incorporate the measure conceptualization processes to develop measures for use in CMS accountability programs as well as other relevant national or local quality improvement initiatives.

The measure conceptualization phase begins by delineating several elements required to develop a meaningful performance measure. These elements, in addition to the results of an initial environmental scan to identify related measures, will clarify the intended use and form the foundation for performance measures that drive improvements in quality.

Evidence

One of the first considerations when developing a performance measure is the supporting evidence. Measures typically focus on those clinical processes and outcomes with the highest levels of clinical evidence. Measure developers look to clinical practice guidelines, best practice statements, or other systematic reviews of the evidence to identify potential measure concepts with the highest degree of supporting evidence. Developing a measure with a dearth of evidence is not advisable for a measure intended to assess clinical performance and the quality of patient care.
In addition to having a strong evidence base, a performance measure must address a clinical process or outcome for which there is a meaningful opportunity to improve. In assessing the opportunity for improvement, measure developers look for data indicating overall suboptimal care or significant variations in care related to the focus of the performance measure. Some examples of opportunities for improvement include utilization, regional variation in care, variation in care by different provider types, disparities in care across gender, race, ethnicity or socioeconomic status, and inconsistent care over time. Additionally, measures can address gaps in the care of high- or multiple-risk patients (e.g., patients with diabetes or with heart disease who smoke).

Data Source
The data source is the origin of the data obtained for measurement. Consideration must be given to the data source to determine how the measure will be specified. Several data sources are available and include:

- Administrative claims
- Registries
- Electronic health record (EHR) systems

In recent years, measurement programs and developers have prioritized the use of electronic clinical data in performance measurement, either through an EHR or a registry. Understanding what types of data are available in an intended data source as well as how those data are captured is essential to developing a performance measure that can be feasibly implemented. A measure developer must also consider how the necessary data would be seamlessly captured within clinical workflow.

Care Setting
The care setting is the setting(s) in which the measure applies and is assessed. The care settings in which a measure is assessed include, but are not limited to, ambulatory care, hospitals/facilities, clinician offices or emergency departments. The care setting must be established early in the development process to determine what data elements are feasibly captured and the available data sources within the chosen setting.

Level of Analysis
The level of analysis is the level at which the measurement is assessed. Determining the level of analysis answers the question of whose performance the measure is trying to improve. Measures may be assessed at various levels including:

- Clinician (either individual or a group/practice)
- Facility
- Health Plan
- Integrated Delivery System
- Population (community, county/city, regional or state)

Many performance measures are intended to measure the performance of individual providers, while other measures address the performance of a hospital, health system, or health plan. It is important to align the level of measurement with the appropriate level where a change is needed to drive improvement in patient care. When considering measures for any topic area, it is vital to consider and determine the elements which are fundamental in the early phase of developing meaningful measures – strength of evidence, data source, care setting and level of analysis.
Stakeholder input

As the measure developer identifies potential measure concepts for further development, it is critically important to ensure stakeholder input at different stages in the measure development process. This is usually accomplished via two distinct approaches that include convening an expert work group to guide the development of the measure(s) and holding a public comment period.

Expert Work Group

An expert work group, also known as a technical expert panel, is convened to provide guidance and input during the measure development process. Expert work groups provide input from the clinical perspective to ensure measure validity, feasibility, and adherence to current evidence and practice. However, other perspectives are equally important and therefore it is imperative to include in an expert work group individuals with a consumer, patient, caregiver and purchaser perspective in addition to experts on performance measurement and quality improvement.

Public Comment

Another critical phase of the measure development process is public comment. Once the expert work group helps shape a measure, the draft measure is submitted for public comment to ensure a transparent process with balanced input from stakeholders and other interested parties. Public comment helps ensure that groups not represented in the expert work group or with differing perspectives provide input that may not have been considered during the measure development process. Solicitation of public comments is disseminated as widely as possible to invite a wide range of perspectives to help inform the finalization of a measure.

Measure Specifications

Measure specifications are the precise technical instructions for capturing required data elements including measure numerator, denominator, exclusions and exceptions. Measure specifications also provide the instructions for calculating a measure, including data source, data collection timing and methodology and risk stratification and adjustment, when necessary. CMS refers to measure specifications as “the blueprint that tells the user how to properly implement the measure within their organization.” The measure developer will need to specify a measure to capture the required data elements and calculate the measure using the appropriate code sets and terminologies for each data element and data source.

Measure Testing

Measure testing is part of full measure development and determines whether a measure is working as intended. The goal of measure testing is to ensure that performance measures are feasible to implement, valid, and reliable:

- The feasibility of a performance measure/measure set refers to the extent to which a clinical practice can interpret measure definitions and technical specifications and integrate them into existing workflows and health information systems to collect, manage, and manipulate data elements, compute performance measures, and generate performance reports.
- The validity of a performance measure refers to the extent to which it truly measures that which it is intended and designed to measure. Four common types of validity that are relevant to the evaluation of a performance measure are face, content, construct, and predictive validity.
• Validity testing demonstrates that the measure data elements are correct and/or the measure score correctly reflects the quality of care provided. This can be assessed empirically or through a survey-based face validity process
• Reliability refers to the stability of a set of observations generated by an indicator under a fixed set of conditions, regardless of who collects the observations or of when or where they are collected

Measure testing results can provide vital information to a measure developer regarding clarifications that would improve a measure. The measure testing results can also provide guidance to physicians and organizations who would like to implement the measures in the future. Properly conducting measure testing and analysis is critical to approval of a measure by CMS and endorsement by the National Quality Forum (NQF).

Measure Implementation

Performance measures can be developed for a number of applications. Some measures are designed to measure the success of a quality improvement initiative, while other measures are designed to be used in accountability and payment programs for providers or hospitals. Typically, measures that are developed for use in accountability and payment programs are held to a higher standard than those that are developed for quality improvement. Accountability measures are expected to have a rigorous evidence base and to demonstrate high reliability and validity.

Many performance measures are used in one or more federal reporting and/or payment programs, which generally apply to either individual providers or facilities in a specific care setting. Federal reporting programs place a high priority on measures that address outcomes of care, with a special interest in patient-reported outcomes. Currently, for a measure to be included in a federal program, it must be submitted to the Measures Under Consideration (MUC) list which is reviewed by the NQF’s Measures Applications Partnership (MAP) by the summer of each year. The MAP conducts an annual review of quality measures and then makes recommendations to the federal government regarding quality measures for inclusion in more than 20 federal programs.

Measure developers must also determine the appropriate reporting mechanism for each application. For example, CMS uses different types of reporting mechanisms within their quality programs, including claims-based, EHR, qualified registry and Qualified Clinical Data Registry (QCDR) reporting. CMS defines a QCDR as a CMS-approved entity, such as a specialty society, certification board, or regional health collaborative, that collects medical and/or clinical data for the purpose of patient and disease tracking to foster improvement in the quality of care provided to patients. A QCDR will complete the collection and submission of PQRS quality measures data on behalf of eligible professionals.

Measure Endorsement

The NQF is an organization that reviews performance measures for national endorsement based on a set of defined evaluation criteria. CMS highly prioritizes NQF-endorsed measures for use in their reporting and pay-for-performance programs. Endorsement of measures by the NQF is intended to be an important indicator that the measure has undergone a rigorous review by a range of stakeholders. Each measure reviewed by NQF is evaluated for their suitability for endorsement based on NQF’s measure evaluation criteria in the following order:

1. Importance to Measure and Report
2. Scientific Acceptability of Measure Properties
3. Feasibility
4. Usability and Use
5. Related and Competing Measures

With the increased role of performance measures in various programs, there has been increased emphasis from the NQF and elsewhere to reduce redundancy among performance measures. When developing a performance measure, it is important to conduct an environmental scan to avoid duplication and determine what measures exist that are related to the intended measure focus. When measures exist that address a similar topic, it is important to harmonize where possible such that common definitions and structures are shared across the measures to allow for easier implementation.

Citations