# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table of Contents</td>
<td>i</td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Measure Conceptualization</td>
<td>4</td>
</tr>
<tr>
<td>Measure Specification</td>
<td>10</td>
</tr>
<tr>
<td>Measure Testing</td>
<td>15</td>
</tr>
<tr>
<td>Measure Implementation</td>
<td>19</td>
</tr>
<tr>
<td>Measure Use, Continuing Evaluation, and Maintenance</td>
<td>24</td>
</tr>
<tr>
<td>Conclusion</td>
<td>26</td>
</tr>
</tbody>
</table>
Introduction

ABOUT THE QUICKSTART

Designed with both experienced and novice measure developers in mind, the CMS MMS Blueprint QuickStart Guide (QuickStart) provides a start-to-finish overview of quality measure development, implementation, and maintenance steps and processes. Each section includes information about important steps associated with a given stage of the Measure Lifecycle, along with links to additional resources, templates, and references to specific chapters and supplemental materials of the Blueprint.

The QuickStart provides tables, checklists, and procedural graphics as a tool to guide measure developers through the mechanics of measure development, implementation, and maintenance. For a more comprehensive overview, visit the Blueprint. To take a deeper dive into specific topics, view the Blueprint supplemental materials.

CMS PRIORITIES & MEANINGFUL MEASURES

In nearly every setting of care, CMS is moving from paying for services to paying for value, not volume. CMS’s goal is to foster value by promoting the highest quality, safety, equity, and care experience with the most affordable, cost-efficient service possible for Americans. To do this, CMS develops quality measures (measures) that address healthcare priorities and goals and align with patient needs. Each measure focuses on a different aspect of health care, such as processes, patient health outcomes, patient perceptions, and organizational structure and/or systems.

The purpose of CMS measures is two-fold. First, measures promote quality and reduce waste in healthcare by incentivizing good performance and disincentivizing poor performance through public reporting and pay-for-reporting programs, and by allowing CMS and participating measured entities to track performance over time. Secondly, they improve patient decision-making by providing data through public reporting (e.g., Star Rating) to help patients, families, and caregivers make informed decisions about where to seek care that is not just based on cost. Given this critical role, measures must promote equity and be meaningful, robust,1 valid, feasible, based in scientific evidence, and well tested to ensure the measures do not lead to unintended negative consequences or burden for patients or measured entities.

CMS launched the comprehensive Meaningful Measures Initiative in 2017, which identifies high priority areas for quality measurement and improvement. The purpose of this initiative was to improve outcomes for patients, their families, and measured entities while also reducing burden and moving payment toward value through focusing everyone’s efforts on the same quality areas. The Meaningful Measures Initiative also helped to identify and close important gap areas of measures, align measures across the continuum of care and across payors, and spur innovation in new types of measures such as patient-reported measures and digital measures.

---

1 Throughout the document, “robust” refers to measures with the most vigorous quality action or guidance or as a descriptor to describe strong, vigorous, or thoroughly vetted components of a measure.
**MEASURE LIFECYCLE**

The Measure Lifecycle ensures measure developers create precisely specified, valid, reliable, and significant measures that directly link to CMS quality goals. The Measure Lifecycle graphic provides a high-level view of the major tasks involved in developing measures from the time of the initial concept through measure implementation and maintenance. While the stages follow a general sequence, the process is highly iterative and allows developers the flexibility to carry out stages concurrently. Additionally, measure developers perform cross-cutting activities such as information gathering, stakeholder engagement, and feasibility evaluation throughout the Measure Lifecycle.

![Measure Lifecycle Diagram](image)

*Figure 1. Key activities carried out by measure developers during the Measure Lifecycle.*

**EVALUATION CRITERIA**

Measure evaluation is not a single step in the Measure Lifecycle. Rather, measure developers should apply evaluation criteria throughout the development, implementation, and maintenance processes to identify weaknesses in the justification and provide an opportunity to revise and strengthen the measure. The more effectively the measure properties meet evaluation criteria, the more likely the measure will be robust and meaningful.

The evaluation criteria are:

- **Importance** of measure and report: evidence and performance gaps
- **Scientific acceptability** of measure properties: reliability and validity
- Feasibility
- **Usability and use**
- **Comparison to related** or **competing measures**

These criteria align with the National Quality Forum’s (NQF) evaluation criteria. Although NQF endorsement is not a requirement for use of a measure in a CMS program, CMS encourages seeking endorsement given that endorsement indicates a level of rigor in testing and evidence that CMS is seeking for its measures.
Further, CMS expects development and testing of measures that do not seek NQF endorsement to be in accordance with these evaluation criteria.

For more information on the evaluation criteria, see the Measure Evaluation Criteria and Guidance for Evaluating Measures for Endorsement.
Measure Conceptualization

The key components of measure conceptualization are information gathering, business case development, and stakeholder engagement.

### INFORMATION GATHERING

**DESCRIPTION**
Information gathering includes an environmental scan (e.g., review of literature, search for clinical practice guidelines, and existing measures, input from experts, and other related activities) and empirical data analysis. These activities yield information that will guide the prioritization of topics or conditions, gap analysis, business case development, and compilation of related and competing measures.

**PURPOSE**
Information gathering demonstrates the existence of a performance or measurement gap related to the topic of interest; it helps demonstrate measure importance and explore feasibility. This process should yield quality goals, strength of scientific evidence pertinent to the topics/conditions, and information with which to build a business case. It will also produce evidence of general agreement or conflicting views on the quality issues surrounding the topics/conditions.

### INFORMATION GATHERING CHECKLIST

- [ ] identify the healthcare quality issue and determine its priority area
- [ ] conduct an environmental scan
- [ ] analyze empirical data, as appropriate
- [ ] evaluate information collected during the environmental scan and empirical data analysis
- [ ] conduct a measurement gap analysis to identify areas for new measure development
- [ ] justify the creation of new measures
- [ ] apply measure evaluation criteria
- [ ] prepare an initial list of measures or measure topics

### HOW TO PERFORM AN ENVIRONMENTAL SCAN:

- Develop a series of unambiguous, structured questions to limit the search to a specific problem set.
- Determine the framework for relevant work, including literature databases and search engines, keywords and phrases, inclusion and exclusion criteria, and domain experts.
Assess the literature using qualitative techniques and quantitative metrics such as impact (e.g., number of times a paper is cited, number of page views), innovativeness, consistency with other works on the topic, recency of citations used in the work, seminality/originality, and quality of writing.

- Qualitatively evaluate and summarize the evidence. Evaluate the effectiveness and value of the data sources, sample sizes, data collection methods, statistical methods, periods, and research findings.
- Interpret findings by evaluating the similarities and differences among the findings; then, draw conclusions to inform data collection and analyses.
- Refine research questions and develop hypotheses. Generate a general analysis plan, including data sources and estimation procedures.

### BUSINESS CASE DEVELOPMENT

**DESCRIPTION**
The business case documents the information needed to assess the anticipated benefits of a measure against the resources and costs required to develop and implement a measure (burdens vs benefits).

**PURPOSE**
Facilitates decision-making about whether to invest resources in the development and implementation of a potential measure; helps demonstrate usability, importance, and feasibility.

The measure developer initiates a business case early in measure conceptualization, updates and enhances it throughout the measure lifecycle, and uses it to compare actual results during measure reevaluation and maintenance. The business case should demonstrate:

- why the measure is necessary
- the measure’s value and how it balances benefits, costs, and risks, including burden
- viability of the measure relative to the healthcare sector’s ability to respond
- whether the measure is sensitive to changes in behavior or policy such that improvements in measure performance reflect improvements in care delivery
- whether the costs of implementation are realistic
- whether the healthcare system has sufficient capacity to implement the measure

Key elements of the business case include:

- net benefit summary
- precise statement of need
- measure impact
- influencing factors
- resources required for measure implementation
- costs of clinical care

Table 1 outlines several research questions measure developers should ask when developing a business case and key areas for which the measure could provide benefits and decrease or increase costs.
Table 1. Research questions to pose during business case development.

<table>
<thead>
<tr>
<th>PRIMARY RESEARCH QUESTIONS</th>
<th>AREAS OF POTENTIAL COSTS, BENEFITS, AND SAVINGS</th>
</tr>
</thead>
<tbody>
<tr>
<td>How will this measure improve healthcare quality, processes of care delivery, outcomes of</td>
<td>• <strong>Patients</strong>: health outcomes, length of stay, readmissions, patient satisfaction, adverse events, medical</td>
</tr>
<tr>
<td>care, and/or decrease complications or untoward effects of care?</td>
<td>errors, trust of the healthcare system</td>
</tr>
<tr>
<td>How will the measure decrease variations of care across disadvantaged subgroups?</td>
<td>• <strong>Employee and organizational</strong>: workplace safety, staff time, staff turnover, sick time, training time,</td>
</tr>
<tr>
<td>How will implementing this measure decrease the cost of care or improve clinical efficiency?</td>
<td>turnover hiring costs, staff supervision costs</td>
</tr>
<tr>
<td>How does data collection for this measure affect clinical workflows?</td>
<td>• <strong>Liability</strong>: worker’s compensation claims, liability insurance premiums, litigation and judgment costs, fines</td>
</tr>
<tr>
<td>What are the far-reaching, long-term benefits of this measure? Does it address a gap in</td>
<td>• <strong>Materials</strong>: Product purchase, new technology or protocol, maintenance, storage, and disposal.</td>
</tr>
<tr>
<td>care not addressed in the literature?</td>
<td></td>
</tr>
</tbody>
</table>

The cost savings analysis model is the method most commonly used to evaluate the business case. Regardless of the evaluation model, the business case should include a hypothesis that, at a minimum, states the measure’s effect over time. These details enable the measure developer to make cost-benefit determinations during measure use, continuing evaluation, and maintenance.

<table>
<thead>
<tr>
<th>STAKEHOLDER ENGAGEMENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>DESCRIPTION</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>PURPOSE</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
STAKEHOLDER ENGAGEMENT

**Technical Expert Panel (TEP)**
- **Description:** a group of experts and stakeholders, including patients, families, caregivers, representatives from marginalized communities, and others, who provide feedback to the measure developer during every stage of the Measure Lifecycle, from conceptualization through maintenance.
- **Purpose:** obtain guidance and thoughtful input from varied perspectives on what is important to measure and evaluate for a balanced quality measure useful to stakeholders.

**Person & Family Engagement (PFE)**
- **Description:** the process of meaningfully involving persons and/or family representatives throughout the Measure Lifecycle. Forms of involvement include informal conversations, focus groups, or TEPs. CMS uses the term “person” to reflect an individual’s identity as more than a patient while “family” broadly represents any participant in a person’s healthcare, such as caregivers, advocates, and advocacy groups. As a best practice, include at least two persons and/or family members on a TEP. Visit the CMS Person & Family Engagement Strategy and the Person and Family Engagement Toolkit for more information.
- **Purpose:** identify issues that are important and meaningful to persons and families, helping measure developers create high-quality measures that enable consumers to make informed healthcare decisions.

**Public Comment**
- **Description:** an opportunity for the widest array of interested parties to provide input on the measure.
- **Purpose:** solicit critical suggestions not previously considered by the measure developer or TEP, ensuring balance and transparency in measure development and maintenance.

**Engagement of Other Stakeholders**
- **Description:** targeted outreach to interested parties and subject matter experts to provide input on the measure.
- **Purpose:** solicit suggestions, potentially to answer specific questions, outside of the more formal processes of TEPs and public comment.
Figure 2. Stakeholder engagement and development activities accomplished during the Measure Lifecycle.

REFERENCES

BLUEPRINT CHAPTER
- Chapter 4: Measure Conceptualization

SUPPLEMENTAL MATERIALS
- Environmental Scans for Quality Measurement
- Technical Expert Panels (TEPs)
- Person and Family Engagement in Quality Measurement
REFERENCES

TEMPLATES

- Information Gathering Report Template
- Business Case Form & Instructions
Measure Specification

This stage of the Measure Lifecycle consists of developing the technical aspects of the measure specifications and harmonization with related measures. Measure developers revisit the measure specification process throughout the Measure Lifecycle to incorporate shifting measure concepts and specifications based on testing results and changes to standards.

STAKEHOLDER ENGAGEMENT

In addition to the TEP, the measure developer should engage stakeholders such as patients, caregivers, and clinicians at this stage to address the feasibility of data collection. The measure developer should also consider soliciting public comments. ONC Project Tracking System and the Measure Collaboration Workspace may post draft specifications for electronic clinical quality measures (eCQMs).

MEASURE SPECIFICATIONS

DESCRIPTION

Measure specifications are essentially the measure details; they include all the information required to define and calculate the measure. Development of specifications is an iterative process with testing.

PURPOSE

Ensures measure details are clear and unambiguous, creating a unique measure that is distinguishable from others and to support consistent implementation.

MEASURE SPECIFICATIONS CHECKLIST

- define the data source
- develop specifications and definitions
- specify the codes and code systems
- construct the data protocol
- document the measures

DEFINE THE DATA SOURCE: when identifying the source of data, the measure developer must consider the feasibility and method(s) of collecting data from that source. Types of sources include administrative data, claims data, paper patient medical records, electronic patient medical records, electronic clinical data, registries, and standardized patient assessments. Prioritize use of digital data sources and avoid paper data sources. Each source has its own benefits and limitations, such as time commitments and staff resources.

DEVELOP SPECIFICATIONS AND DEFINITIONS: the construction of measure specifications begins with the outline of the target/initial population, numerator, denominator, numerator, and
denominator exclusions, denominator exceptions, and measure logic. Then, the measure developer gives the measure concept increasing amounts of detail, including precisely defined data elements and the appropriate values, value sets, or direct reference codes. Every part of measure specification requires explicitly defined elements with accompanying analysis to identify constraints and criteria of the specification.

Table 2. Measure specification examples.

<table>
<thead>
<tr>
<th>COMPONENT</th>
<th>DESCRIPTION</th>
<th>EXAMPLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target/Initial Population</td>
<td>Refers to the cohort for selecting the denominator population. Some measures (e.g., ratio measures) will require multiple target/initial populations, one for the numerator and one for the denominator.</td>
<td>All patients aged 18 years and older with a diagnosis of major depressive disorder (MDD) (NQF 0104e) (CMIT Reference Number 432).</td>
</tr>
<tr>
<td>Denominator</td>
<td>Describes the population evaluated by the individual measure. The population defined by the denominator can be the same as the target/initial population or it can be a subset of the target/initial population to further constrain the population for the measure. Format: patients, age [age or age range], with [condition] in [setting] during [time frame]</td>
<td>All patients at least 18 years old as of the first day of the reporting month who are determined to be maintenance hemodialysis patients (in-center and home HD [hemodialysis]) for the complete reporting month at the same facility (NQF 2978) (CMIT Reference Number 5642).</td>
</tr>
<tr>
<td>Denominator Exclusion</td>
<td>Denominator exclusions refer to criteria that result in removal of patients or cases from the denominator before calculating the numerator. An exclusion means that the numerator event is not applicable to those covered by the exclusion. Limit use to just those that are absolutely necessary. Format: denominator-eligible patients who [have some additional characteristic, condition, procedure]</td>
<td>Patients with an active diagnosis for depression or a diagnosis of bipolar disorder (CMIT Reference Number 5824).</td>
</tr>
<tr>
<td>Numerator</td>
<td>Describes the process, condition, event, or outcome that satisfies the measure focus or intent. Format: patients who received/had [measure focus] (during [time frame] if different than for target population)</td>
<td>The number of adult patient-months in the denominator who were on maintenance hemodialysis using a catheter continuously for three months or longer as of the last hemodialysis session of the reporting month (NQF 2978) (CMIT Reference Number 5642).</td>
</tr>
<tr>
<td>COMPONENT</td>
<td>DESCRIPTION</td>
<td>EXAMPLE</td>
</tr>
<tr>
<td>----------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Denominator Exception</td>
<td>An exception permits the exercise of clinical judgment and implies the measured entity considered or offered treatment to each potentially eligible patient in the denominator. Exceptions are most appropriate when measurements of contraindications to drugs or procedures are relative. Only use a denominator exception in proportion measures. It is not appropriate for ratio or continuous variable measures.</td>
<td>Documentation of medical reason(s) for not prescribing beta-blocker therapy (e.g., low blood pressure, fluid overload, asthma, patients recently treated with an intravenous positive inotropic agent, allergy, intolerance, other medical reasons). Documentation of patient reason(s) for not prescribing beta-blocker therapy (e.g., patient declined, other patient reasons). Documentation of system reason(s) for not prescribing beta-blocker therapy (e.g., other reasons attributable to the healthcare system) (NQF 0083e) (CMIT Reference Number 5798).</td>
</tr>
<tr>
<td>Numerator Exclusion</td>
<td>Used only in ratio and proportion measures to define elements that should not be in the numerator data.</td>
<td>If the number of central line bloodstream infections per 1,000 catheter days were to exclude infections with a specific bacterium, that bacterium would be listed as a numerator exclusion.</td>
</tr>
<tr>
<td>Stratification Scheme</td>
<td>Measure developers may define a stratification scheme in lieu of risk adjustment by stratifying the population based on their risk for an outcome of a procedure.</td>
<td>Measure is to be stratified by a population type (i.e., race, ethnicity, age, social risk factors, income, region, gender, primary language, disability) (e.g., Chlamydia Screening for Women [NQF 0033] (CMIT Reference Number 2513).</td>
</tr>
</tbody>
</table>

**SPECIFY THE CODE AND/OR CODE SYSTEMS:** most measures rely at least in part on the use of various standardized codes or code systems for classifying healthcare provided in the United States. The measure developer should list all codes (plus the code system and the version that the codes came from) required for the measure and explicitly state the source of the codes and instructions pertaining to their use. Find more information in the [Codes, Code Systems, and Value Sets] supplemental material.

**CONSTRUCT DATA PROTOCOL:** the measure developer must explicitly identify the types of data and how to aggregate or link these data so that the measure calculation can be reliable and valid. The measure developer should proceed carefully when merging data from different sources or systems to prevent errors in assumptions.
HARMONIZATION

DESCRIPTION  Harmonization is the standardization of specifications for related measures. Measure harmonization may be based on shared focus, target population, or definitions applicable to many measures so that they are uniform or compatible (unless there is a compelling reason not to, i.e., dictated by the evidence).

PURPOSE  Harmonization helps to reduce burden associated with measure implementation and reporting at healthcare organizations. Harmonization efforts during development and maintenance help fulfill the NQF criterion of alignment with competing or existing measures. Find more information in the Measure Harmonization, Respecification, and Adoption supplemental material.
Table 3. Harmonization during measure development. N represents numerator and D represents denominator in the table.

<table>
<thead>
<tr>
<th>MEASURE</th>
<th>HARMONIZATION ISSUE</th>
<th>ACTION</th>
</tr>
</thead>
</table>
| N       | same measure focus  | competing measure | • use existing measure (adopted) or justify development of a new measure  
|         |                     |        | • a different data source will require new harmonized specifications (e.g., respecified) |
| D       | same target population |        | • harmonize on measure focus (i.e., respecified)  
|         |                     |        | • justify differences  
|         |                     |        | • respecify existing measure by expanding the target population |
| N       | same measure focus  | related measure | • harmonize on target population  
|         |                     |        | • justify differences |
| D       | different target populations |        | • proceed with new measure development |
| N       | different measure focus  | new measure | • proceed with new measure development |
| D       | different target population |        | • proceed with new measure development |

REFERENCES

**BLUEPRINT CHAPTER**
- Chapter 5: Measure Specification

**SUPPLEMENTAL MATERIALS**
- Codes, Code Systems, and Value Sets
- Composite Measures for Accountability Programs
- Cost and Resource Use Measures
- Electronic Clinical Quality Measures (eCQM) Specifications, Standards, and Tools
- Measure Harmonization, Respecification, and Adoption
- Multiple Chronic Condition Measures
- Patient-Reported Outcome Measures
- Risk Adjustment in Quality Measurement
Measure Testing

Key components of measure testing include developing a testing plan, alpha and beta testing, and measure evaluation.

**DESCRIPTION** Testing refers to all the data collection and analysis activities that contribute to the evaluation of the measure specifications.

**PURPOSE** Enables measure developers to assess the suitability of the technical specifications and acquire empirical evidence to help assess the strengths and challenges of a measure with respect to the evaluation criteria, especially scientific acceptability (reliability and validity) and feasibility. Testing also provides an opportunity to build upon earlier judgments about the measure’s importance and usability.

### STEPS TO PERFORM TESTING

☐ develop a plan for how to test the measure (ensure that planned methods will address evaluation criteria)

☐ implement the testing plan

☐ analyze the test results

☐ refine the measure; incorporate stakeholder inputs

☐ retest the refined measure

☐ document adherence to measure evaluation criteria:
  - ☐ prepare for NQF endorsement process (if applicable)
  - ☐ compile information to support measure selection (see “Measure Implementation”)

### ALPHA AND BETA TESTING

Table 4 provides the features of both alpha and beta testing; measure developers should consider both when developing a testing plan.

*Table 4. Features of alpha and beta testing.*

<table>
<thead>
<tr>
<th></th>
<th>ALPHA TESTING</th>
<th>BETA TESTING</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Timing</strong></td>
<td>• conducted prior to finalization of technical specifications</td>
<td>• conducted after development of the measure developer’s detailed and precise technical specifications</td>
</tr>
<tr>
<td></td>
<td>• can conduct multiple times in quick succession</td>
<td></td>
</tr>
<tr>
<td><strong>Scale</strong></td>
<td>• requires only enough records to ensure the presence of all elements needed for the measure and identify common</td>
<td>• samples should be representative and of adequate size</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• may require data from multiple sites/settings, depending on the data</td>
</tr>
<tr>
<td></td>
<td>ALPHA TESTING</td>
<td>BETA TESTING</td>
</tr>
<tr>
<td>----------------------</td>
<td>------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td><strong>Scale (cont’d)</strong></td>
<td>occurrences or variation in the data</td>
<td>source (e.g., administrative, medical record)</td>
</tr>
<tr>
<td><strong>Sampling</strong></td>
<td>• convenience sampling</td>
<td>• sufficient to allow adequate testing of the measure’s scientific acceptability</td>
</tr>
<tr>
<td><strong>Specification</strong></td>
<td>• permits early detection of problems in technical specifications (e.g.,</td>
<td>• used to assess or revise the complexity of computations required to calculate the measure</td>
</tr>
<tr>
<td><strong>Refinement</strong></td>
<td>identification of additional inclusion and exclusion criteria)</td>
<td></td>
</tr>
<tr>
<td><strong>Importance</strong></td>
<td>• may help assess volume, frequency, or costs related to a measure topic (e.g., cost of treating the condition, costs related to procedures measured)</td>
<td>• includes empirical evaluation of performance thresholds, disparities in the population (e.g., by identifying variability among comparison groups, showing that the measure is not “topped-out”)</td>
</tr>
<tr>
<td></td>
<td>• establishes, on a preliminary basis, that the measure can identify gaps in care</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• provides support for further development of the measure</td>
<td></td>
</tr>
<tr>
<td><strong>Scientific</strong></td>
<td>• limited in scope if conducted during the formative stage</td>
<td>• empirically assesses measure reliability and validity, including analysis of exclusion criteria (if any used)</td>
</tr>
<tr>
<td><strong>Acceptability</strong></td>
<td>• may include preliminary assessment of face validity</td>
<td>• evaluates the risk adjustment model</td>
</tr>
<tr>
<td><strong>Feasibility</strong></td>
<td>• provides initial information about the feasibility of collecting required data and calculating measures using technical specifications</td>
<td>• provides enhanced information regarding feasibility, including greater determination of barriers and measured entity burden to implementation and costs associated with measurement</td>
</tr>
<tr>
<td></td>
<td>• identifies barriers to implementation</td>
<td>• evaluates feasibility of stratification factors based on occurrences of target events in the sample</td>
</tr>
<tr>
<td></td>
<td>• offers an initial estimate of costs or burden of data collection and analysis</td>
<td></td>
</tr>
<tr>
<td><strong>Usability and Use</strong></td>
<td>• no formal analytic testing at this stage; may use qualitative testing with patients and measured entities during alpha testing</td>
<td>• identifies unintended consequences, including susceptibility to inaccuracies and errors</td>
</tr>
<tr>
<td></td>
<td>• TEP may assess potential usability of the measure</td>
<td>• reports strategies to ameliorate unintended consequences</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• may consist of focus groups or similar means of assessing usefulness of the measure by consumers</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• can use the TEP to assess potential usability</td>
</tr>
</tbody>
</table>
DEVELOP THE TESTING PLAN: a testing plan (also referred to as a work plan for testing) should include enough information to explain how the proposed testing methodology will help meet the evaluation criteria. Testing plans for alpha testing may look a bit different than testing plans for beta testing; at a minimum, however, all testing plans should contain the elements listed.

### TESTING PLAN ELEMENTS

- name(s) of measure(s)
- type of testing
- study objective(s)
- timeline for testing and report completion
- data collection methodology
- description of test population (e.g., number of test sites/data sets)
- description of data elements to be collected
- sampling methods, if applicable
- description of strategy to recruit measured entities/test data sets
- planned analysis methods and description of test statistics
- description and forms documenting patient confidentiality and description of Institutional Review Board (IRB) compliance approval and/or steps to obtain data use agreements, if necessary
- methods to comply with the Paperwork Reduction Act (PRA), if relevant

ANALYZE THE TEST RESULTS: when the measure developer completes data gathering from the test sites, the measure developer conducts a series of analyses to characterize the evaluation criteria of the measures. The measure developer presents findings of testing analyses in a final summary report.

REFINE THE MEASURE: the measure developer may need to modify the measure specifications, data collection instructions, and calculation of measure results based on analysis of testing results. For example,

- following alpha testing, the measure developer may undertake measure respecification or efforts to overcome potential implementation barriers.
- following beta testing, changes in the definition of the population or adjustments to the comparison group definition may occur.

### STAKEHOLDER ENGAGEMENT OPPORTUNITY

If making changes to the measure, consult with the TEP prior to retesting the measure. View Figure 2. Stakeholder engagement and development activities accomplished during the Measure Lifecycle. graphic for more information.
RETEST THE MEASURE: measure testing is an iterative process. Continue to refine and retest the measure as deemed necessary.

EVALUATE THE MEASURE: throughout the Measure Lifecycle, especially through testing, the measure developer evaluates the measure to determine the degree to which it is consistent with the evaluation criteria. The measure developer uses resulting evaluation information to determine how to modify the measure to increase the importance, scientific acceptability, usability and use, and feasibility.

REFERENCES

BLUEPRINT CHAPTERS

• Chapter 6: Measure Testing

SUPPLEMENTAL MATERIALS

• Composite Measures for Accountability Programs
• Patient-Reported Outcome Measures
• Cost and Resource Use Measures
• Electronic Clinical Quality Measures (eCQM) Specifications, Standards, and Tools

TEMPLATES

• Measure Evaluation Report Template
Measure Implementation

This stage of the Measure Lifecycle includes all activities associated with taking a measure from a development state to an active, in-use state, which includes—but not limited to—consensus endorsement processes, measure selection processes, and measure rollout.

DESCRIPTION

The implementation process that measures undergo varies significantly on several factors, which may include:
- scope of measure implementation
- measured entity being measured
- data collection processes
- ultimate use of the measure (e.g., quality improvement, public reporting, pay-for-reporting, or value-based purchasing)
- program into which the measure is added

The scope of implementation could entail a measure or measure set:
- implemented in a new program
- added to an existing program

PURPOSE

Implementation processes ensure careful review of all new and respecified measures to ensure the selection of only high-quality, meaningful measures meeting the needs of CMS programs.

NQF ENDORSEMENT: NQF currently serves as the consensus-based entity (CBE) regarding performance measurement for the Department of Health and Human Services (HHS). To the extent feasible, CMS uses NQF-endorsed measures in CMS public reporting and value-based purchasing programs. Measure developers must consider several items to facilitate the NQF submission process as well as minimize rework.

CONSIDERATIONS TO FACILITATE NQF SUBMISSION

☐ provide clear, concise, and substantive answers to all sections of the NQF submission form
☐ ensure NQF submission form can be understood as a standalone document. All attachments, uniform resource locators (URLs), and references must include specific page numbers or table number references
☐ provide the code list, risk adjustment methodology, and calculation algorithm as attachments or URLs
☐ for eCQMs, include the Measure Authoring Tool (MAT)-exported human-readable hypertext markup language (HTML) and executable files and URLs to value sets and direct reference codes in the attachments
☐ include pilot testing information, as applicable
☐ give clear rationales for decisions related to measure specifications, including use of numerator and denominator exclusions and denominator exceptions
CMS measure selection criteria help to ensure that each measure
- is responsive to specific program goals and statutory requirements
- addresses an important condition or topic with a performance gap and has a strong scientific evidence base to demonstrate that the measure, when implemented, can lead to the desired outcomes and more affordable care. This requirement corresponds to NQF’s importance criterion.
- addresses one of the Meaningful Measure areas
- promotes alignment with CMS program attributes and across HHS programs
- is feasible to report, fully developed, and tested. At a minimum, the measure developer should test measures for reliability and validity.
- offers results and performance that identifies opportunities for improvement. CMS will not select measures when evidence already identifies high levels of performance with minimal opportunity for improvement (i.e., topped out measures).
- avoids negative unintended consequences
- does not duplicate another measure currently implemented in one or more programs

If it is an eCQM, the measure developer must create using the MAT and express in Health Quality Measure Format (HQMF) using the Quality Data Model (QDM) and Clinical Quality Language.

CONSIDERATIONS TO FACILITATE NQF SUBMISSION
- provide explanations of controversies about the science behind the measure
- confirm points of contact on the Measure Submission Form are accurate

MEASURE SELECTION PROCESSES

DESCRIPTION
Different CMS programs have different paths a measure can take for selection and implementation. In general, measures undergo identification and finalization during a rigorous public process.

PROCESS
For measures not subject to pre-rulemaking or rulemaking, the measure selection process includes the following steps:
- CMS issues a call for measures to solicit measures and/or identify measures considered for removal
- submitted measures follow the HHS clearance process
- cleared measures go through a consensus development process, which might include the Measure Applications Partnership (MAP) process
PRE-RULEMAKING & RULEMAKING

DESCRIPTION  Pre-rulemaking and rulemaking represent one specific pathway for measure selection. The programs that participate in the CMS pre-rulemaking and rulemaking process include those identified under the Patient Protection and Affordable Care Act (ACA) Section 3014. Measure developers submit measures for potential inclusion in the Measures Under Consideration (MUC) list for these programs. The MUC list, which is publicly posted, is a list of the measures HHS is considering adopting through the federal rulemaking process for use in several Medicare quality and payment programs.

PURPOSE  Maximize transparency and rigor in the measure identification and selection process

In Figure 3, the gray boxes provide an overview of the pre-rulemaking process through the publication of the MUC list. Below the arrow are measure developer activities that occur at various points in the pre-rulemaking process.

Figure 3. Pre-rulemaking timeline.

CMS provides the finalized MUC list to the MAP. MAP input is due by February 1. CMS encourages measure developers to attend the MAP meeting to be fully involved in the process. After CMS receives the MAP input, CMS begins a deliberation process to determine which measures to include in the federal rulemaking process. The measure selection criteria used during development of the MUC list are the same criteria used for federal rulemaking. HHS must consider MAP input and publish the rationale for selecting any measure for use in a CMS program—in proposed or final rules—not endorsed by NQF.
After CMS completes the pre-rulemaking process and selects measures for potential inclusion in rulemaking, the next steps in the cycle are

1. **PROPOSED RULES:** CMS develops the proposed rules and publishes them in the Federal Register. A proposed rule is generally available for public comment for 60 days.

2. **FINAL RULES:** CMS considers the comments received and publishes the final rules in the Federal Register.

Measure rollout occurs after a program adopts the measure. The rollout process may include collection of data for a dry run from all relevant measured entities across the country and sharing of calculated rates with the measured entities. CMS does not use dry run data for payment but may use them as a baseline for future payment years.

When communicating and coordinating with all parties involved in the rollout, the measure developer must consider the timelines of other processes (for example, rulemaking, NQF projects, and quality alliances). The measure developer prepares and presents education for the end users on what is being measured and how to interpret the results.

The measure developer also documents the results of any educational activities and assesses whether the activities were adequate to meet the needs of the end users of the measures. For example, the measure developer should report on the number of events, including the attendance at each:

- conference call and call recording
- web-based presentation and presentation recording
- workshop at conference or scientific society meeting
- train-the-trainer event

A measure steward is the individual or organization that owns a measure and is responsible for maintaining it. Sometimes the measure steward is also the measure developer. If an existing measure in a program undergoes a substantive change during these any of these updates, it is the responsibility of the measure steward to resubmit the measure to restart the measure selection process.
REFERENCES

BLUEPRINT SECTION
- Chapter 7: Measure Implementation

SUPPLEMENTAL MATERIAL
- National Quality Forum (NQF) Endorsement and Maintenance
Measure Use, Continuing Evaluation, and Maintenance

To help CMS ensure the continued soundness of a measure, the measure developer must provide strong evidence that a measure currently in use continues to add value to quality reporting and incentive programs and that its construction continues to be sound throughout its lifecycle.

**DESCRIPTION**
The processes for continuous review and evaluation for appropriateness of measures currently used in programs

**PURPOSE**
Maintenance and continuing evaluation ensure that the measure continues to meet criteria of importance, feasibility, scientific acceptability and usability and use

**CONTINUING EVALUATION**

**DESCRIPTION**
Continuing evaluation refers to the process through which the measure developer updates measure specifications to demonstrate the measure’s continued suitability for use.

**STEPS FOR CONTINUING EVALUATION**
- conduct data collection and ongoing surveillance
- respond to questions about the measure
- produce preliminary reports
- report measure results
- monitor and analyze the measure rates and audit findings
- perform measure maintenance or early maintenance review, when appropriate
- provide information that CMS can use in measure priorities planning

**MAINTENANCE**

**DESCRIPTION**
Measure maintenance is a multi-step review process that includes annual updates, comprehensive reevaluations, and early maintenance reviews.
**STEPS TO MAINTAIN A MEASURE**

- ☐ conduct data collection and ongoing surveillance
- ☐ respond to questions about the measure
- ☐ produce preliminary reports
- ☐ report measure results
- ☐ monitor and analyze measure rates and audit findings
- ☐ perform measure maintenance or early maintenance review, when appropriate
- ☐ provide information that CMS can use in measure priorities planning

**REFERENCES**

**BLUEPRINT SECTION**
- Chapter 8: Measure Use, Continuing Evaluation, and Maintenance

**SUPPLEMENTAL MATERIALS**
- Measure Maintenance Reviews
- National Quality Forum (NQF) Endorsement and Maintenance
Conclusion

The quality and efficiency measures developed for use in CMS programs have a real-world effect on patients, families, and measured entities. Implementation and use of thoughtfully developed measures has positively influenced several critically important metrics at the national level, including hospital readmission rates, central line-associated blood stream infections (CLABSI), surgical site infections, early elective deliveries, and ventilator-associated pneumonia.

For more in-depth information about the Measure Lifecycle, access the Blueprint and Blueprint supplemental materials, which provide greater context and additional detail on the topics highlighted in this guide. Visit the CMS MMS website for educational presentations or to download fillable templates.

Contact the MMS Support Team at MMSsupport@battelle.org if you have additional questions.