NATIONAL QUALITY FORUM

Cost and Efficiency Measure Evaluation Criteria

Importance to Measure and Report

1a. The measure focus addresses:
– a demonstrated high-impact aspect of healthcare (e.g., affects large numbers, leading cause of morbidity/mortality, high resource use [current and/or future], severity of illness, and patient/societal consequences of poor quality).

AND

Performance Gap
1b. Demonstration of resource use or cost problems and opportunity for improvement, i.e., data demonstrating
• Considerable variation cost or resource across providers; and/or
• Disparities in care across population groups

2. Scientific acceptability of the measure properties

Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the cost or resources used to deliver care. Measures must be judged to meet the sub criteria for both reliability and validity to pass this criterion and be evaluated against the remaining criteria.

2a. Reliability
2a1. The measure is well defined and precisely specified so that it can be implemented consistently within and across organizations and allow for comparability.

- All measures that use the ICD classification system must use ICD-10-CM.
- eMeasures should be specified in the Health Quality Measures Format (HQMF) and must use the Quality Data Model (QDM) and value sets vetted and published through the National Library of Medicine's Value Set Authority Center (VASC).

2a2. Reliability testing demonstrates that the measure results are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period, and/or that the measure score is precise.

2b. Validity
2b1. The measure specifications are consistent with the measure intent and captures the most inclusive target population.

2b2. Validity testing demonstrates that the measure data elements are correct and/or the measure score correctly reflects the cost of care or resources provided.

2b3. Exclusions are supported by the clinical evidence

AND/OR

There is a rationale or analysis demonstrating that the measure results are sufficiently distorted due to the magnitude and/or frequency of then on-clinical exclusions;
AND

− Measure specifications for scoring include computing exclusions so that the effect on the measure is transparent (i.e., impact clearly delineated, such as number of cases excluded, exclusion rates by type of exclusion);
AND

− If patient preference (e.g., informed decision-making) is a basis for exclusion, there must be evidence that the exclusion impacts performance on the measure; in such cases, the measure must be specified so that the information about patient preference and the effect on the measure is transparent (e.g., numerator category computed separately, denominator exclusion category computed separately).

2b4. For resource use measures and other measures when indicated:
− an evidence-based risk-adjustment strategy is specified and is based on patient factors (including clinical and sociodemographic risk factors) that influence the measured outcome and are present at start of care, and has demonstrated adequate discrimination and calibration
OR
− rationale/data support no risk-adjustment/-stratification.

2b5. Data analysis demonstrates that methods for scoring and analysis of the specified measure allow for identification of statistically significant and practically/ clinically meaningful differences in performance.

2b6. If multiple data sources/methods are specified, there is demonstration that they produce comparable results.

2c. If disparities in care have been identified, measure specifications, scoring, and analysis allow for identification of disparities through stratification of results (e.g., by race, ethnicity, socioeconomic status, gender)
OR
rationale/data justifies why stratification is not necessary or not feasible.

3. Feasibility

Extent to which the required data are readily available or could be captured without undue burden, and can be implemented for performance measurement.
3a. For clinical measures, the required data elements are routinely generated and used during care delivery (e.g., blood pressure, lab test, diagnosis, medication order).

3b. The required data elements are available in electronic health records or other electronic sources. If the required data are not in electronic health records or existing electronic sources, a credible, near-term path to electronic collection is specified.

3c. Demonstration that the data collection strategy (e.g., data source/availability, timing, frequency, sampling, patient confidentiality, costs associated with fees/licensing of proprietary measures) or elements such as risk model, grouper, instrument) can be implemented (e.g., already in operational use, or testing demonstrates that it is ready to put into operational use).

4. Usability and use

Extent to which potential audiences (e.g., consumers, purchasers, providers, 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.

4a. Accountability and Transparency
Performance results are used in at least one accountability application one within three years after initial endorsement and are publicly reported within six years after initial endorsement (or the data on performance results are available). If not in use at the time of initial endorsement, then a credible plan for implementation within the specified timeframes is provided. AND

4b. Improvement
Progress toward achieving the goal of high-quality, efficient healthcare for individuals or populations is demonstrated. If not in use for performance improvement at the time of initial endorsement, then a credible rationale describes how the performance results could be used to further the goal of high quality, efficient healthcare for individuals or populations. AND

4d. Data and result detail are maintained such that the resource use measure, including the clinical and construction logic for a defined unit of measurement can be deconstructed to facilitate transparency and understanding.