

NATIONAL QUALITY FORUM

Measure Evaluation Criteria and Guidance Summary Tables Effective for Projects Beginning after January 2011

5. Comparison to Related or Competing Measures [Definitions-Table11](#) [Guidance-Figure 1](#)

If a measure meets the above criteria and there are endorsed or new related measures (either the same measure focus or the same target population) or competing measures (both the same measure focus and the same target population), the measures are compared to address harmonization and/or selection of the best measure.

5a. The measure specifications are harmonized¹⁸ with related measures;

OR

the differences in specifications are justified. [Guidance-Table 13](#)

5b. The measure is superior to competing measures (e.g., is a more valid or efficient way to measure); [Guidance-Table 12](#)

OR

multiple measures are justified.

Note

18. Measure harmonization refers to the standardization of specifications for related measures with the same measure focus (e.g., *influenza immunization* of patients in hospitals or nursing homes); related measures with the same target population (e.g., eye exam and HbA1c for *patients with diabetes*); or definitions applicable to many measures (e.g., age designation for children) so that they are uniform or compatible, unless differences are justified (e.g., dictated by the evidence). The dimensions of harmonization can include numerator, denominator, exclusions, calculation, and data source and collection instructions. The extent of harmonization depends on the relationship of the measures, the evidence for the specific measure focus, and differences in data sources.

Guidance on Evaluating Related and Competing Measures

For more information, see full report: [Guidance for Measure Harmonization](#).

Table 11: Related versus Competing Measures

	Same concepts for measure focus—target process, condition, event, outcome	Different concepts for measure focus—target process, condition, event, outcome
Same target patient population	Competing measures—Select best measure from competing measures or justify endorsement of additional measure(s).	Related measures—Harmonize on target patient population or justify differences.
Different target patient population	Related measures—Combine into one measure with expanded target patient population or justify why different harmonized measures are needed.	Neither harmonization nor competing measure issue

Figure 1: Addressing Competing Measures and Harmonization of Related Measures in the NQF Evaluation Process

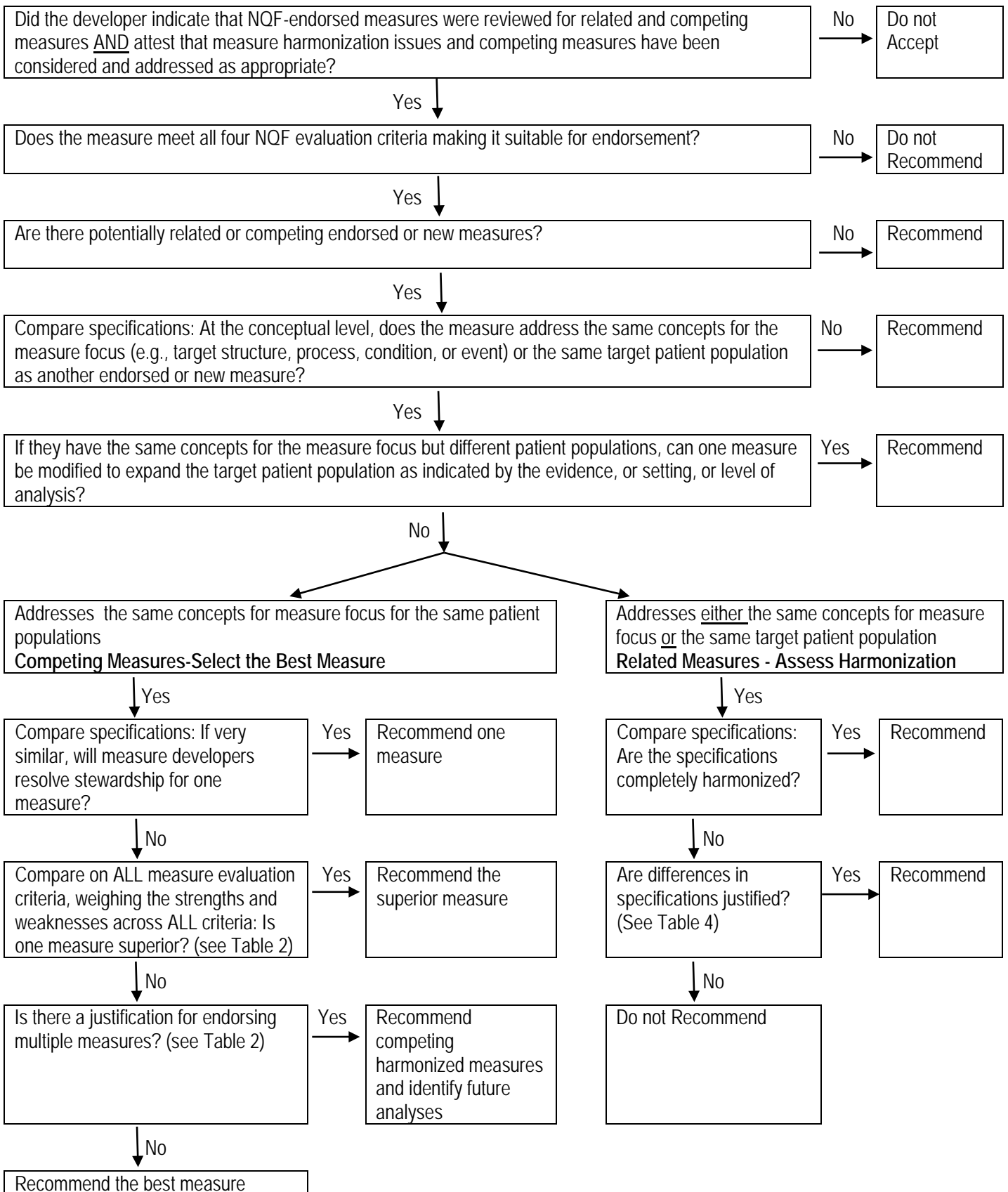


Table 12: Evaluating Competing Measures for Superiority or Justification for Multiple Measures

Steps	Evaluate Competing Measures
1. Determine if need to compare measures for superiority	Work through the steps in the algorithm (Figure 1) to determine if need to evaluate competing measures for superiority (i.e., two or more measures address the same concepts for measure focus for the same patient populations)
2. Assess Competing Measures for Superiority by weighing the strengths and weaknesses across ALL NQF evaluation criteria	<p>Because the competing measures have already been determined to have met NQF’s criteria for endorsement, the assessment of competing measures must include <u>weighing the strengths and weaknesses across ALL the criteria</u> and involves more than just comparing ratings. (For example, a decision is not based on just the differences in scientific acceptability of measure properties without weighing the evaluation of importance to measure and report, usability, and feasibility as well.)</p> <p>Impact, Opportunity, and Evidence—Importance to Measure and Report: Competing measures generally will be the same in terms of the measure focus addressing a high-impact aspect of healthcare (1a) and evidence for the focus of measurement (1c). However, due to differences in measure construction, they could differ on alignment with national health goals/priorities or opportunity for improvement.</p> <ul style="list-style-type: none"> • Compare measures on alignment with national health goals/priorities (1a) • Compare measures on opportunity for improvement (1b) <p>Reliability and Validity—Scientific Acceptability of Measure Properties:</p> <ul style="list-style-type: none"> • Compare evidence of reliability (2a1-2a2) • Compare evidence of validity, including threats to validity (2b1-2b6) <p>Untested measures cannot be considered superior to tested measures because there would be no empirical evidence on which to compare reliability and validity. (However, a new measure, when tested, could ultimately demonstrate superiority over an endorsed measure and the NQF endorsement maintenance cycles allow for regular submission of new measures.)</p> <p>Compare and identify differences in specifications <u>All else being equal on the criteria and subcriteria, the preference is for:</u></p> <ul style="list-style-type: none"> • Measures specified for the broadest application (target patient population as indicated by the evidence, settings, level of analysis) • Measures that address disparities in care when appropriate <p>Usability:</p> <ul style="list-style-type: none"> • Compare evidence of use and usefulness for public reporting, including availability of data for reporting performance results • Compare evidence of use and usefulness for quality improvement <p><u>All else being equal on the criteria and subcriteria, the preference is for:</u></p> <ul style="list-style-type: none"> • Measures that are publicly reported • Measures with the widest use (e.g., settings, numbers of entities reporting performance results) • Measures that are in use over those without evidence of use <p>Feasibility:</p> <ul style="list-style-type: none"> • Compare the ease of data collection/availability of required data • Compare the potential for inaccuracies, errors, and unintended consequences <p><u>All else being equal on the criteria and subcriteria, the preference is for:</u></p> <ul style="list-style-type: none"> • Measures based on data from electronic sources

Steps	Evaluate Competing Measures
	<ul style="list-style-type: none"> • Clinical data from EHRs • Measures that are freely available <p>After weighing the strengths and weaknesses across ALL criteria, identify if one measure is clearly superior and provide the rationale based on the NOF criteria.</p>
<p>3.If a competing measure does not have clear superiority, assess justification for multiple measures</p>	<p>If a competing measure does not have clear superiority, is there a justification for endorsing multiple measures? Does the added value offset any burden or negative impact?</p> <p>Identify the value of endorsing competing measures Is an additional measure necessary?</p> <ul style="list-style-type: none"> • to change to EHR-based measurement; • to have broader applicability (if one measure cannot accommodate all patient populations; settings, e.g., hospital, home health; or levels of analysis, e.g., clinician, facility; etc.); • to increase availability of performance results (if one measure cannot be widely implemented, e.g., if measures based on different data types increase the number of entities for whom performance results are available) <p>Note: Until clinical data from electronic health records (EHRs) are widely available for performance measurement, endorsement of competing measures based on different data types (e.g., claims and EHRs) may be needed to achieve the dual goals of 1) advocating widespread access to performance data and 2) migrating to performance measures based on EHRs. EHRs are the preferred source for clinical record data, but measures based on paper charts or data submitted to registries may be needed in the transition to EHR-based measures.</p> <p>Is an additional measure unnecessary?</p> <ul style="list-style-type: none"> • primarily for unique developer preferences <p>Identify the burden of endorsing competing measures Do the different measures affect interpretability across measures? Does having more than one endorsed measure increase the burden of data collection?</p> <p>Determine if the added value of endorsing competing measures offsets any burden or negative impact?</p> <ul style="list-style-type: none"> • If yes, recommend competing measures for endorsement (if harmonized) and provide the rationale for recommending endorsement of multiple competing measures. Also, identify analyses needed to conduct a rigorous evaluation of the use and usefulness of the measures at the time of endorsement maintenance. • If no, recommend the best measure for endorsement and provide rationale.

Table 13: Sample Considerations to Justify Lack of Measure Harmonization

Related Measures	Lack of Harmonization	Assess Justification for Conceptual Differences	Assess Justification for Technical Differences
Same measure focus (numerator); different target population (denominator)	Inconsistent measure focus (numerator)	The evidence for the measure focus is different for the different target population so that one measure cannot accommodate both target populations. Evidence should always guide measure specifications.	<ul style="list-style-type: none"> • Differences in the available data drive differences in the technical specifications for the measure focus. • Effort has been made to reconcile the differences across measures but important differences remain.
Same target population (denominator); different measure focus (numerator)	Inconsistent target population (denominator) and/or exclusions	The evidence for the different measure focus necessitates a change in the target population and/or exclusions. Evidence should always guide measure specifications.	<ul style="list-style-type: none"> • Differences in the available data drive differences in technical specifications for the target population. • Effort has been made to reconcile the differences across measures but important differences remain.
For any related measures	Inconsistent scoring/ computation	The difference does not affect interpretability or burden of data collection. If it does, it adds value that outweighs any concern regarding interpretability or burden of data collection.	The difference does not affect interpretability or burden of data collection. If it does, it adds value that outweighs any concern regarding interpretability or burden of data collection.