eMEASURE FEASIBILITY ASSESSMENT

DRAFT REPORT (JANUARY 2013)

INTRODUCTION

As quality measurement shifts to measures derived from electronic health records (EHRs), additional clarity about the testing is needed to assure that eMeasures can be used for a range of accountability applications, which require both precision and results that are reliable and valid. While the concepts of reliability and validity apply equally to measures derived from EHRs, the electronic health record presents additional challenges related to measure testing: widespread EHR data are not yet available for measure development and testing; there is a lack of comparability across vendor products; and data elements needed for advanced measures currently may not be feasible to capture in EHRs.

Realizing the promise of EHRs as a tool for quality reporting will rest on the ability of providers, payers, certifiers, vendors and other users of the information to know that eMeasures provide valid and reliable data. During the public and member comment for the National Quality Forum's (NQF) eMeasure Review and Testing Proposal in early 2012, several organizations expressed support that eMeasure testing should incorporate the feasibility and clinical workflow requirements of data capture for the data elements utilized in addition to reliability and validity testing. This requirement is significantly hampered by the lack of clarity and definition in the field as to what constitutes feasibility assessment for EHRs.

PURPOSE

The purpose of the eMeasure Feasibility Assessment project was to assess the current state of feasibility assessment for new and retooled eMeasures and identify a set of principles and criteria for adequate feasibility assessment. Specifically, the goals of the project were to:

- conduct an environmental scan of approaches to feasibility assessment from measure developers, government contractors and EHR vendors to gather the current approaches used to assess measure feasibility; and
- review results of environmental scan and propose a set of feasibility recommendations, including a starter set of criteria for eMeasure feasibility assessment that will address the following considerations:
    - the timing of feasibility assessment in the course of measure development (e.g., iterative testing with development, feasibility testing of fully specified measures);
    - the potential differences in feasibility assessment requirements for de novo (new) eMeasures and re-specified eMeasures;
    - the number and diversity of EHRs used for testing and relation to comparability across EHRs in terms of data feasibility, validity, and reliability; and
    - the interrelationship of feasibility and validity testing of new measures.

This eMeasure feasibility report provides important guidance that can shape future measure development as well as product development and certification requirements.

To achieve these goals, NQF convened a 15-member Technical Expert Panel (TEP) which was comprised of measure developers, experts in measure development and testing, EHR vendors, and implementers. In addition to participating in several conference calls, the TEP also gathered for a one-day, in-person meeting in Washington, DC on December 7, 2012.

BACKGROUND

Several prior NQF activities provide the foundation for this present work on eMeasure feasibility assessment:

- NQF criteria for feasibility: Feasibility is one of four major criteria that NQF uses to evaluate measures for endorsement. Feasibility is defined as the *"extent to which the required data are readily available or could be captured without undue burden and can be implemented for performance measurement."* Sub-criteria further define how feasibility is evaluated:

  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., source, timing, frequency, sampling, patient confidentiality, costs associated with fees/licensing of proprietary measures) can be implemented (e.g., already in operational use, or testing demonstrates that it is ready to put into operational use). [1]

- *Health Information Technology Expert Panel (HITEP) Report (2008)* The HITEP 1 report identified common data types to be standardized for automation in EHRs and health information exchanges. This report described a framework to assess the quality of each data type at it then existed in EHRs. The data quality framework provided an initial assessment of the availability and quality of a given data type.[2]

- NQF Testing Task Force Report (2011) The Testing Task Force report provides the guidance for rating the level of evidence for reliability and validity of EHR measures. The testing criteria for all eMeasures (de novo and re-specified) include specification in accordance with the Quality Data Model (QDM); and empiric testing of validity of the data elements and measure score. [3]

- NQF's Draft Requirements for eMeasure Review and Testing (2012) During public and member comment on the draft document, several organizations expressed a need to incorporate feasibility assessment of data capture for the specified data elements. The Consensus Standards Approval Committee (CSAC) agreed that feasibility assessment should be required

for all eMeasures. The draft document concluded that further work is needed to define and standardize requirements for feasibility assessment.[4]

ENVIRONMENTAL SCAN

NQF conducted an environmental scan to identify approaches to feasibility assessment from measure developers, electronic health record (EHR) vendors, and providers. Each of the stakeholder groups was asked to respond to a series of questions. Up to nine organizations were queried in each of the three stakeholder groups.

NQF reached out to nine EHR vendors representing a cross section of inpatient and ambulatory electronic health record (EHR) companies. With the assistance of a technical expert panel (TEP) member, NQF worked with the HIMSS EHRA vendor group to distribute the questions to EHR vendors. Eight vendors responded. Four measure developers responded to NQF's request. All four are HHS contractors. Three develop inpatient measures and one develops ambulatory measures. The final group, providers, consisted of an ambulatory physician practice and two health systems which include inpatient and ambulatory sites.

Sample Environmental Scan Questions

- General approach to feasibility testing;
- Current efforts of collaboration or interrelationships with vendors, and/or measure developers and/or providers in regards to feasibility testing;
- How feasibility fits into their business cycle (vendor question)
- When testing occurs (measure developers)
- Impact to workflow to be addressed during development

NQF developed questions specific to each stakeholder group. Questions were based on their workflow and what their role was in the quality measure lifecycle. For example, it was important to ask EHR vendors how measure feasibility testing fit into their business development cycle, and to ask measure developers when feasibility testing occurs in their development cycle, and to ask providers to describe factors that impacted implementation and workflow issues that should be addressed or factored into the development of eMeasure. A full list of questions and a summary of responses from each stakeholder group can be found in Appendix 2.

Common themes emerged from each stakeholder group. Vendor responses had three common themes: 1) the need to assess measure requirements and analyze their EHR product for gaps, 2) the impact of measure requirements on EHR work flow, and 3) addressing identified gaps. Vendors also commented that not all data elements can be extracted from the EHR. Data may be stored in another system or is stored as text. There were reports of collaboration among vendors and measure developers but this was not the general rule.

Measure developer responses mentioned that feasibility testing is an evolving process and the approach to testing must be different for re-specified measures and de novo (new) measures. Workflow, test site experience and structure also impact the testing. Developers also suggested that when feasibility testing identifies a problem they need to look for alternatives for capturing the data or consider alternate representation for clinical concepts. One measure developer stated that feasibility testing is a continuum, it does not occur just once in the measure cycle, so they developed a five (5) point scale for testing and evaluating the feasibility. This concept of using a scale to rate feasibility has generated great discussion among the project TEP.

The final stakeholder group, providers, stated that testing of measures should occur with all major EHR vendors and in as many settings as possible: large, small, solo providers, multispecialty, academic, etc. They also suggested a test dataset to improve the process. Two other themes emerged from this group: the importance of workflow and the need to share data between disparate systems.

PRINCIPLES AND GUIDANCE

The following key principles were identified by the TEP that guided their recommendations for eMeasure feasibility assessment:

Definitions/Terminology

- **Clinical Quality Measures (CQM):** A mechanism used for assessing the degree to which a provider competently and safely delivers clinical services that are appropriate for the patient in an optimal timeframe. [5]
- **Electronic health record (EHR)** (also electronic patient record, electronic medical record, or computerized patient record): As defined by Healthcare Information Management and Systems Society (HIMSS), the electronic health record (EHR) is a longitudinal electronic record of patient health information generated by one or more encounters in any care delivery setting. Included in this information are patient demographics, progress notes, problems, medications, vital signs, past medical history, immunizations, laboratory data and radiology reports.[6]
- **eMeasure:** As defined by Health Level Seven (HL7), an eMeasure is a health quality measure encoded in the Health Quality Measures Format (HQMF) format. The HQMF is a standard for representing a health quality measure as an electronic document. Through standardization of a measure's structure, metadata, definitions, and logic, the HQMF provides for quality measure consistency and unambiguous interpretation.[7]
- ***De novo* eMeasure**: A new performance measure developed for use in EHRs; it is not re-specified from an existing measure based on other data sources

- **eMeasure feasibility**: The extent to which the required data are readily available or could be captured without undue burden and can be implemented for performance measurement in EHRs. A data element is highly feasible if the data is available in the current workflow without any additional input from the user required.

Questions for TEP:
•Does this definition of eMeasure feasibility meet the needs of stakeholders?
•Should the definition explicitly state that a feasible eMeasure is comprised of data elements that are highly feasible?

- **Data element, quality:** A quality data element is a single piece of information that is used in quality measures to describe part of the clinical care process, including both a clinical entity and its context of use (e.g., diagnosis, active). [8]
    - **Data element, critical** Among quality data elements those elements that contribute most to the computed measure score, that is, account for identifying the greatest proportion of the target condition, event, or outcome being measured (numerator); the target population (denominator); population excluded (exclusions); and, when applicable, risk factors with the largest contribution to variability in outcome. [9]
- **Quality Data Model (QDM):** Clinical data necessary to measure quality performance. The QDM framework contains three levels of information: standard elements, quality data elements, and data flow attributes. Standard elements (e.g., diagnosis) represent the atomic unit of data identified by a data element name, a code set, and a code list composed of one or more enumerated values. The quality data element includes the standard element plus quality data type or context (e.g., diagnosis active). Data flow attributes include source (originator), recorder, setting, and health record field.[10]
- **Health Quality Measures Format (HQMF**): A standard for representing a health quality measure as an electronic document.[11]
- **Quality Reporting Document Architecture (QRDA):** A standard for communicating health care quality measurement information.[12]

Overarching Principles:

- There is a spectrum of feasibility that depends on maturity of the various EHR systems and the context for the eMeasure implementation. EHR vendors indicate that all data elements can be ultimately integrated into an EHR; however, the important questions are the time required for development and deployment and whether the importance of the information justifies the cost.

Until agreement is reached that everyone must achieve the same level of interconnection and maturity, feasibility will be variable across all EHR implementations.

- The feasibility assessment is different than reliability and validity testing of an eMeasure. Feasibility should be assessed in several domains: data availability including heterogeneity across different EHR systems and mapping requirements; data accuracy and completeness; data standards (access to structured and coded data); workflow; measure logic; measure aggregation and reporting. Feasibility assessment of the data elements must precede testing for reliability and validity.

- The fundamental goal of performance measurement is to improve the quality of care delivered. As such, there are some aspects of care that are sufficiently important to merit changes in workflow, i.e., changing behaviors to align with best practices. The balance between feasibility and validity and reliability and the usefulness of the measure for care improvement is critical. The value of an eMeasure reflects a balance between the quality improvement potential in the eMeasure and the feasibility, including costs.

- In this dynamic and rapidly evolving environment feasibility assessment should not restrict the development of new quality measures that capitalize on the features of EHRs or address important areas of measures such as care coordination and patient reported outcomes that may be challenging at the present time. Data elements that seem less feasible in the near term can become feasible in 3-5 years with specific guidance to EHR vendors, certification bodies and other stakeholders. While there is a need for prioritization of expectations for EHR vendors, measure developers are encouraged to be forward thinking and develop new types of innovative measures that capture important, new aspects of quality care .

- A great deal of innovation is occurring in the local settings, particularly in high-performing organizations. Providers are using EHR products in creative ways that inform future product development. Collaborative measure development efforts should seek out innovation in the field and learn from high quality organizations. Providers are creating useful and meaningful measures for quality improvement that may be suitable for use as accountability measures.

Questions for the TEP

- **Should there be a principle that includes use of the QDM and MAT and/or the limits of QDM, HQMF, and QRDA?**
- **Should there be a principle on standardization and certification that timelines must be identified and rapid cycle development and integration must accommodate what is needed for eMeasures?**

RECOMMENDATIONS

RECOMMENDATION 1: Assess feasibility throughout eMeasure development.

Feasibility should be considered early in the eMeasure development process and assessed throughout the entire duration of measure development as an agile, iterative process. Greater collaboration among measure developers, EHR vendors, and providers at all stages of development would promote more rapid evolution in EHR functionality for current measures as well as to build capability for more complex, EHR enabled eMeasures.

RECOMMENDATION 2: Framework for eMeasure feasibility assessment

The TEP agreed that a framework for assessing eMeasure feasibility would provide a common language and provide decision-makers with valuable information about the technical feasibility of eMeasures. Building on the work from the HITEP 1 report, this proposed framework addresses assessment of technical feasibility of data collection only and does not address the value of the data element or eMeasure. Assessment of the feasibility of data elements while a measure is being conceptualized and specified can identify significant feasibility issues before the measure is field tested. A feasibility assessment should address both the data elements and subsequently the measure logic, aggregation and reporting.

2.1 Data element feasibility assessment

The goal of data element feasibility assessment is to guide and assist measure development to better understand what is more feasible in the near term or whether a longer timeframe is required to achieve changes in EHRs or workflow to capture the data. Measure developers can consider alternative approaches to specify a measure based on the feasibility assessment. Measure development sponsors and potential implementers can use the feasibility assessment to make decisions on proceeding with development of a new measure or re-specifying an existing measure or determine the likely timeframe when a measure might be ready for implementation. Providers and organization who would implement the eMeasure would have more information to provide meaningful input during measure development.

A report on the feasibility assessment of the data elements should address the following characteristics:

- Data availability – the extent to which the data is readily available in a structured format across EHR systems. EHR vendors can determine how often data is captured (data profiling) among current users. A major challenge is data exchange and the ability to link independent data sources such as inpatient and ambulatory data, inpatient and long-term care data, or emergency department and inpatient data. The capability for data exchange is quite variable. When difficulties with availability of critical data elements in valuable eMeasures are identified, the steps required to improve data availability should be proposed.

Questions for TEP:

- **How does data element availability affect validity?**
- **Will validity testing of data elements expose problems with data availability?**
- **Should there be a minimum number of EHR systems and a minimum number of care setting required for testing for reliability and validity?**

- Data accuracy - the extent to which the most accurate data source is used and/or the data is captured by the most appropriate healthcare professional or recorder. The HITEP 1 report stated "the authority and accuracy of data are often interrelated (e.g. lab results coming from a laboratory interface are both authoritative and accurate). While there are examples of data from an authoritative source (e.g. clinician) that are not always accurate (e.g. subjective historical findings) and vice versa, these two were considered an aggregate as they both assess the soundness of the data source."

  Questions for the TEP:
  - **How does data accuracy relate to data element validity (which is a sub-criterion of Scientific Acceptability rather than Feasibility)?**
  - **NQF requires testing for validity at either the level of the data element or the measures score or, preferably, both. How does testing for data element validity relate to data accuracy in the proposed feasibility assessment?**

- Data standards – the extent to which the data element is coded using a nationally accepted terminology standard. Standard data elements, associated definitions and code sets, and mapping to the Quality Data Model (QDM) are expected.

- Workflow - the extent to which capturing the data element fits the typical EHR workflow for that user. The HITEP 1 report indicated that "in order for quality data to be recorded at the point of care by authoritative sources, it needs to fit into the clinical workflow. For example, it is of little benefit to have the capability of capturing certain patient symptoms if it requires five clicks and three screens during a busy clinical encounter, for the end result will likely be missing data."

  Questions for the TEP:
  - **Which user(s) is the reference?**
  - **What is included in workflow ?**
    - **the data is usually captured by a provider during the patient encounter;**
    - **the interface does not require significant manipulation or levels (clicks) to document;**

  - **If a data element is provided through interfaces and does not require manual entry would it score the highest on workflow?**

DATA ELEMENT FEASIBILITY SCORECARD

A standard score card for assessing feasibility of data elements can provide a basic summary about overall feasibility of an eMeasure. Assessing feasibility requires identifying the assumptions made and the reference point against which feasibility is being assessed. For example, is the assessment made for EHRs that meet certification requirements, the average EHR, or 50% or 80% of EHRs? It may be necessary to compute the score against a matrix of different assumptions in terms of interconnectivity and maturity of the EHR.

The assessment would be made based on current implementation capabilities and future (3-5 years) implementation for data elements that have a current low score for feasibility:

> Current –Rate the characteristics of the data element using the 1-3 scale for current feasibility based on the assumptions and reference indicated.
>
> Future – Rate the characteristics of the data element using the 1-3 scale for feasibility in 3—5 years and indicate what is required to reach the future state if necessary.

The assessment should use quantitative methods whenever possible. For example, data profiling; structured interview surveys and questionnaires from providers in a variety of settings

> **Questions for the TEP:**
>
> **What further specifications for use of the scorecard are needed?**
>
> - **Where should the assessment occur – at the vendor level or point of care?**
> - **Is the measure developer ultimately responsible for completing the assessment in collaboration with others?**
> - **How many EHR systems should be assessed? What is the minimum? What is optimal?**
> - **How many settings of care should be assessed? What is the minimum? What is optimal?**

Data element:		
Assumption(s) when rating this data element:		
Reference Point/target audience:		
	Current (1-3)	Future* (1-3)
Data Availability – Is the data readily available in structured format across EHR systems? Scale: 3 One of the following: a. Data element is routinely collected in the vast majority of EHRs or b. Data element is currently in some EMRs and/or some health systems		

and the ability to collect data element is required for certified EHR 2 – Data element is currently not collected but the ability to collect data element is required for certified EMR 1 – Ability to collect data element is not required for certified EHR and is currently not widely collected		
Data Accuracy **–** What is the accuracy of the data element in EHRs? Are the data source and recorder specified? Scale: 3 – Highly accurate; very reliable 2 – generally accurate under usual circumstances 1 –(low) unreliable **Questions for TEP:** • **Should the rating scale refer to the accuracy of the data source and/or appropriateness of the recorder?** • **The NQF criterion for reliability refers to reproducibility of the data elements and the measure score. What is meant by "reliable" in the scale above?** • **How are accuracy and validity related?**		
Data Standards **–** Is the data element coded using a nationally accepted terminology standard? Scale: 3 – Usually coded in nationally accepted standard 2 – Standards available, but not widely adopted 1 – No standards available or typically entered as free text		
Workflow **–** Is the data captured during the course of care and it fit the typical EHR workflow for that user? Scale: 3 – Requires no additional data entry from clinician and no EMR user interface changes. Data element is available as a byproduct of routine care. Examples would be lab values, vital signs, referral orders, or problem list entry. 2 – Additional time and effort is required, but some perceived to be of benefit 1 – Additional time and effort is required but without immediate benefit to care		
DATA ELEMENT FEASIBILITY SCORE		

*For data elements that score low on current feasibility, indicate what is required to improve the feasibility score in 3-5 years.

Data element feasibility score

Questions for the TEP:

How should the data element feasibility score be determined?

- **total # of data elements;**
- **sum or weighted sum;**
- **average or weighted total average;**
- **weighted average of each component;**
- **list of data elements and components that scored 1;**
- **low feasibility if any domain is rated a "1".**

2.2 eMeasure feasibility assessment

The TEP emphasized that feasibility is not solely about the data elements. The measure specifications and the calculation logic are important to understanding the intent of the measure which can influence what data must be collected. The number of data elements increases complexity of the measure which is more costly for development and testing. Each structured data element increase implementation costs and the importance of the clinical component both to the provider and patient being measured should be weighed. For example, if the logic of an eMeasure is quite complex or some data elements score low on feasibility consideration of a different approach may be warranted. During the environmental scan some developers reported having studied the incremental effect of adding a data element. Such an approach to feasibility assessment requires more analysis and perhaps more quantitative assessment.

The ability to aggregate data and produce reports for internal and external use on eMeasure performance is integral. Existing communication standards such as Continuity of Care Document (CCD) and Quality Reporting Document Architecture (QRDA) can be used to assess additional aspects of feasibility related to reporting of quality data. Ensuring that data can be captured through QRDA is important because communications standards can, at times, lag behind data standards.

In addition to the scores for the individual data elements, the eMeasure feasibility assessment report should include:

- assessment of the feasibility of the aggregate data elements:
    - Do any critical data elements (those elements that contribute most to the computed measure score such as numerator, denominator) score low on feasibility?
    - How many data elements are required and are all data elements essential to the eMeasure? For examples, do data elements for exclusions impact the measure score significantly?
- assessment of the measure logic:
    - Does the calculation algorithm work in multiple EHR systems?
    - How complex is the logic?
    - How easy is it to explain to providers?

The feasibility assessment of the measure score (data aggregation) is naturally addressed by testing for reliability and validity at the level of the eMeasure score.

RECOMMENDATION 3: Validating the Data Element Feasibility Scoring

The proposed data element feasibility scoring system must be validated which is outside the scope of this project. The scoring system should be piloted by several measure developer/EHR vendor/provider collaboratives and results provided within 6-12 months.

Questions for the TEP:

- **How many eMeasures or data elements should be evaluated in the pilot to validate the scoring system?**
- **What is the source of "truth" about data element feasibility that should be used to judge the scoring system?**

RECOMMENDATION 4: Data element feasibility catalogue/repository

In order to leverage the work of others and avoid duplication of efforts, the results of the data feasibility assessments should be catalogued in a repository available to all stakeholders and reviewed annually. Measure developers would then be able to consult the catalogue to determine feasibility of previously used data elements and only expend resources to assess feasibility of new data elements or to provide new information on how the data element is performing. The catalogue would also allow standardization of how data elements are represented in the QDM.

RECOMMENDATION 5: NQF evaluation for endorsement

The results of the feasibility assessment(s) conducted during measure development would provide useful information to address NQF's endorsement criteria and sub-criteria, particularly in the areas of Scientific Acceptability and Feasibility. The eMeasure feasibility assessment and data element scores should accompany an eMeasure submitted to NQF for consideration for endorsement.

NQF's criteria and ratings for reliability and validity are described for *de novo* measures and re-specified measures:

Table 4: Evaluation of Reliability and Validity of Measures Specified for EHRs [13]

	New Measure Specified for EHR	

Rating	Reliability Description and Evidence	Validity Description and Evidence	Modifications for Endorsed Measure *Re-specified* for EHRs
High	All EHR measure specifications are unambiguous[+] and include only data elements from the Quality Data Model (QDM)***** including quality data elements, code lists, and measure logic; **OR** new data elements are submitted for inclusion in the QDM; **AND** Empirical evidence of reliability of both data element **AND** measure score within acceptable norms: • Data element: reliability (repeatability) assured with computer programming—**must test data element validity** **AND** • Measure score: appropriate method, scope, and reliability statistic within acceptable norms	The measure specifications (numerator, denominator, exclusions, risk factors) reflect the quality of care problem (1a,1b) and evidence cited in support of the measure focus (1c) under *Importance to Measure and Report*; **AND** Empirical evidence of validity of both data elements **AND** measure score within acceptable norms: • Data element: validity demonstrated by analysis of agreement between data elements electronically extracted and data elements visually abstracted from the entire EHR with statistical results within acceptable norms; **OR** complete agreement between data elements and computed measure scores obtained by applying the EHR measure specifications to a simulated test EHR data set with known values for the critical data elements; **AND** • Measure score: appropriate method, scope, and validity testing result within acceptable norms; **AND** Identified threats to validity (lack of risk adjustment/stratification, multiple data types/methods, systematic missing or "incorrect" data) are empirically assessed and adequately addressed so that results are not biased	The EHR measure specifications use only data elements from the Quality Data Model (QDM)***** and include quality data elements, code lists, and measure logic; **AND** Crosswalk of the EHR measure specifications (QDM quality data elements, code lists, and measure logic) to the endorsed measure specifications demonstrates that they represent the original measure, which was judged to be a valid indicator of quality; **AND** Analysis of comparability of scores produced by the retooled EHR measure specifications with scores produced by the original measure specifications demonstrated similarity within tolerable error limits
Moder-ate	All EHR measure specifications are unambiguous[+] and include only data elements from the QDM;* **OR** new data elements are submitted for inclusion in the QDM; **AND** Empirical evidence of reliability within acceptable norms for either data elements **OR** measure score as noted above	The measure specifications reflect the evidence cited under *Importance to Measure and Report* as noted above; **AND** Empirical evidence of validity within acceptable norms for either data elements **OR** measure score as noted above; **OR** Systematic assessment of face validity of measure score as a quality indicator (as described in Table A-3) explicitly addressed and found substantial agreement that ***the scores obtained from the measure as specified will provide an accurate reflection of quality and can be used to distinguish good and poor quality*** **AND** Identified threats to validity noted above are empirically assessed and adequately addressed so that results are not biased	The EHR measure specifications use only data elements from the QDM as noted above **AND** Crosswalk of the EHR measure specifications as noted above demonstrates that they represent the original measure **AND** For measures with time-limited status, testing of the original measure and evidence ratings of moderate for reliability and validity as described in Table 2.
Low	One or more EHR measure specifications are ambiguous[+] or do not use data elements from the QDM*; **OR** Empirical evidence of unreliability for either data elements **OR** measure score—i.e., statistical results outside of acceptable norms	The EHR measure specifications do not reflect the evidence cited under *Importance to Measure and Report* as noted above; **OR** Empirical evidence (using appropriate method and scope) of invalidity for either data elements **OR** measure score— i.e., statistical results outside of acceptable norms **OR** Identified threats to validity noted above are empirically assessed and determined to bias results	The EHR measure specifications do not use only data elements from the QDM; **OR** Crosswalk of the EHR measure specifications as noted above identifies that they do not represent the original measure **OR** For measures with time-limited status, empirical evidence of low reliability or validity for original time-limited measure
Insuffi	Inappropriate method or scope	Inappropriate method or scope of validity testing	Crosswalk of the EHR measure

	New Measure Specified for EHR		Modifications for Endorsed Measure *Re-specified* for EHRs
Rating	**Reliability Description and Evidence**	**Validity Description and Evidence**	
cient evidence	of reliability testing	(including inadequate assessment of face validity as noted above) **OR** Threats to validity as noted above are likely and are NOT empirically assessed	specifications as noted above was not completed OR For measures with time-limited status, inappropriate method or scope of reliability or validity testing for original time-limited measure

[+]Specifications are considered unambiguous if they are likely to consistently identify who is included and excluded from the target population and the process, condition, event, or outcome being measured; how to compute the score, etc.
*QDM elements should be used when available. When quality data elements are needed but are not yet available in the QDM, they will be considered for addition to the QDM.

Questions for TEP:

During NQF's evaluation for endorsement, reliability and validity is a "must pass" criterion and precedes consideration of feasibility. How does the feasibility assessment inform the evaluation of reliability and validity?

How does missing data affect validity? Should NQF require an analysis of missing data for eMeasures?

Should the criteria for validity include evaluation of data element accuracy, such as whether the most accurate data source and the most appropriate recorder of the information is used?

If an eMeasure meets the criteria (either high or moderate) for reliability and validity testing, what questions remain about feasibility of eMeasure remain?

The eMeasure feasibility framework addresses the sub-criteria of NQF's evaluation criterion of feasibility:

NQF Measure Evaluation Criteria
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., source, timing, frequency, sampling, patient confidentiality, [17] costs associated with fees/licensing of proprietary measures) can be implemented (e.g., already in operational use, or testing demonstrates that it is ready to put into operational use).

Note
17. All data collection must conform to laws regarding protected health information. Patient confidentiality is of particular concern with measures based on patient surveys and when there are small numbers of patients.

Question for the TEP:

Are there additional criteria specific to eMeasures that should be included in the evaluation for feasibility for eMeasures under consideration for NQF endorsement?

RECOMMENDATION 6. Composite measures

NQF is currently updating the guidance for composite measures in the Composite Evaluation Guidance Reassessment Project. Since the growing interest in composite measures extends to eMeasures , the guidance should assist eMeasure developers also. The composite guidance notes that "feasibility of the composite measure will be influenced by the least feasible of the component measures." The data element feasibility scoring system proposed here will assist development of composite measures by assessing the feasibility of the component measures and all of the data elements. A critical data element with low feasibility will affect the overall feasibility of the composite as well as the individual component measure.

RECOMMENDATION 7. Greater collaborative efforts are needed for eMeasure development and implementation

The eMeasure Feasibility Assessment recommendations emphasize the need for greater collaborative efforts among all stakeholders. Collaborative efforts such as the eMeasure Learning Collaborative that "seeks to create a learning environment for advancing knowledge and promoting best practices related to developing and implementing eMeasures" provides an opportunity for stakeholders to work together. All stakeholders must play an active role in the development and implementation of eMeasures. In addition to measure developers and EHR vendors, providers must be involved at all stages of eMeasure development to provide critical input on feasibility. Working through their health systems or professional communities, providers should take a more proactive role in the development of eMeasures. During measure development there is a constant need for testing partners and pilot participants that could be facilitated by ongoing collaboratives focused on developing and implementing eMeasures.

NOTES

1. National Quality Forum *Measure Evaluation Criteria* November 2012 http://www.qualityforum.org/docs/measure_evaluation_criteria.aspx
2. National Quality Forum *Health Information Technology Expert Panel (HITEP) Report (2008) http://www.qualityforum.org/Projects/h/Health_IT_Expert_Panel_I/Health_IT_Expert_Panel_I_(HITEP_I).aspx*
3. National Quality Forum *Guidance for Measure Testing and Evaluating Scientific Acceptability of Measure Properties (January 2011)*
4. National Quality Forum *NQF's Draft Requirements for eMeasure Review and Testing (2012)*
5. *http://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/Downloads/CQM_Webinar_10-25-2011.pdf*
6. Healthcare Information Management and Systems Society, Electronic Health Record, HIMSS EHR Web Page, 2012. *Available at: www.himss.org/ASP/topics_ehr.asp.* Last accessed December 2012.
7. Health Level Seven, 1.1 What is the HQMF, and what is an eMeasure?, HL7 Version 3 Standard: Representation of the Health Quality Measures Format (eMeasure), Release 1 Last Ballot: Draft Standard for Trial Use - March 2010, March 2010. *Available at: www.hl7.org/v3ballot/html/domains/uvqm/UVQM.html#WhatisHQMF .* Last accessed December 2012.
8. National Quality Forum, *Health Information Technology Expert Panel II - Health IT Enablement of Quality Measurement,* Washington, DC: NQF, 2009.
9. *Testing TF Report*
10. National Quality Forum*, Health Information Technology Expert Panel II - Health IT Enablement of Quality Measurement,* Washington, DC: NQF, 2009.
11. *HIT*
12. *HIT*
13. *TestingTF Report*

APPENDICES

1. TEP roster
2. Environmental scan details
3. Testing TF section on EHRs