Print

NATIONAL QUALITY FORUM

Project Name	Cancer Project
Measure #	2267
Measure Title	test
Deadline Date	08.29.2013
Last Edit Date	06.27.2013 - 01:58 AM
Last Edit By	Mark Tobias
Measure Process	CDP
Status	Draft

Introduction

Thank you for your interest in submitting a measure to NQF for possible endorsement.

What criteria are used to evaluate measures? Measures are evaluated on four standardized criteria: importance to measure and report, scientific acceptability of measure properties, usability, and feasibility. For your measure to be evaluated against these measure evaluation criteria , you must complete the measure submission form.

Why do I have to complete a form? Due to the volume and/or complexity of proposed measures, NQF provides measure information to committee reviewers in a standardized format to facilitate their evaluation of whether the measure meets NQF's measure evaluation criteria. This form allows the measure steward to present information demonstrating that the proposed measure meets NQF's criteria.

What is on the form? The information requested in this form is directly related to NQF's measure evaluation criteria and is consistent with the data fields agreed upon in the *Common Data Fields Collaboration*.

Can't I just submit our files for consideration? No. Measures must be submitted through the online form to be considered. Requested information should be entered directly into this form.

Can I submit additional details and materials? Additional materials will be considered only as supplemental. Do NOT rely on material provided in attachments or in links to provide measure specifications or to demonstrate meeting the criteria. For example, definitions should be provided in the measure specification detail fields. Some examples of appropriate supplemental materials include code lists that exceed two pages, data collection tools, and methodology reports for complex measures. Even in these examples, the core information should be provided in the appropriate submission form fields. If supplemental materials are provided, a link to a web page is preferred over attached materials. Be sure to indicate specific page numbers or web page locations for the relevant information. Please contact **the designated project staff** regarding questions about submitting supplemental materials.

What do I do first? When you first start a new submission or click on "Begin Submission", you will be directed to the "NQF Conditions" tab, which asks questions about several conditions that must be met before your proposed measures may be considered and evaluated for suitability as NQF-endorsed voluntary consensus standards:

- Is there a signed Measure Steward Agreement (applicable to all non-government organizations)?

- Have you identified the entity and process that will be used to maintain and update the measure?
- Does the intended use include both public reporting and quality improvement?
- Is the measure fully specified and tested for reliability and validity?
- Have you addressed harmonization of related measures and issues with competing measures?
- Is the measure submission information complete with all requested information entered in the form?

Once you have agreed that the four conditions have been met by answering all questions marked with an asterisk, you can begin completing the measure submission form.

Can I come back later to complete a submission once I have started? Yes. You can return to your submission at your convenience to complete the form until the designated deadline for the specific project. To save and return, simply click on the save-draft option anytime during the submission process. When you want to continue, please login to the National Quality Forum website, go to your Dashboard, and click on submission.

Can I make changes to a form once I have submitted it? No. Once you submit your measure, you will NOT be able to return to this submission form to make further revisions.

What if I need additional help? Please contact the project director identified in the call for measures if you have questions regarding the information requested or submitting supplemental materials.

Please email us at web-help@qualityforum.org if you experience technical difficulties using the online submission form.

Thank you for your interest in submitting measures to NQF.

NQF Conditions

Conditions that must be met for consideration by NQF

Several conditions must be met before proposed measures may be considered and evaluated for suitability as voluntary consensus standards. If any of the conditions are not met, the measure will not be accepted for consideration.

A. The measure is in the public domain or a Measure Steward Agreement is signed. (All non-government organizations must sign a Measure Steward Agreement even if measures are made publicly and freely available.)

B. The measure owner/steward verifies there is an identified responsible entity and a process to maintain and update the measure on a schedule that is commensurate with the rate of clinical innovation, but at least every three years.

C. The intended use of the measure includes both accountability applications (including public reporting) and performance improvement to achieve high-quality, efficient healthcare.

D. The measure is fully specified and tested for reliability and validity.

E. The measure developer/steward attests that harmonization with related measures and issues with competing measures have been considered and addressed, as appropriate.

F. The requested measure submission information is complete and responsive to the questions so that all the information needed to evaluate all criteria is provided.

Do you agree to the condiitons?

☑ I have read and accept the conditions as specified above *

Composite

Composite Quality Construct and Rationale

1d.1. A composite performance measure is a combination of two or more component measures, each of which individually reflects quality of care, into a single performance measure with a single score.
For purposes of NQF measure submission, evaluation, and endorsement, the following will be considered composites:

- Measures with two or more individual performance measure scores combined into one score for an accountable entity.
- Measures with two or more individual component measures assessed separately for each patient and then aggregated into one score for an accountable entity:
  - all-or-none measures (e.g., all essential care processes received, or outcomes experienced, by each patient); or
  - any-or-none measures (e.g., any or none of a list of adverse outcomes experienced, or inappropriate or unnecessary care processes received, by each patient).

Please identify the composite measure construction

- two or more individual performance measure scores combined into one score
- all-or-none measures (e.g., all essential care processes received, or outcomes experienced, by each patient)
- any-or-none measures (e.g., any or none of a list of adverse outcomes experienced, or inappropriate or unnecessary care processes received, by each patient)

1d.2. Describe the quality construct, including:

- the overall area of quality
- included component measures and
- the relationship of the component measures to the overall composite and to each other.

1d.3. Describe the rationale for constructing a composite measure, including how the composite provides a distinctive or additive value over the component measures individually.

1d.4. Describe how the aggregation and weighting of the component measures are consistent with the stated quality construct and rationale.

Specifications

Descriptive Information

De.1. Measure Type *(Patient-reported outcomes include HRQoL/functional status, symptom/burden, experience with care, health-related behavior.)**

Composite

De.2. Measure Title*

test

De.3. Brief description of measure (*including type of score, measure focus, target population, timeframe, e.g., Percentage of adult patients aged 18-75 years receiving one or more HbA1c tests per year*)

De.4. IF PAIRED/GROUPED, what is the reason this measure must be reported with other measures to appropriately interpret results?

De.5. Subject/Topic Areas (Check all the areas that apply):

- [ ] Behavioral Health : Behavioral Health
- [ ] Behavioral Health : Alcohol, Substance Use/Abuse
- [ ] Behavioral Health : Attention Deficit Hyperactivity Disorder (ADHD)
- [ ] Behavioral Health : Depression
- [ ] Behavioral Health : Post-Traumatic Stress Disorder (PTSD)
- [ ] Behavioral Health : Screening
- [ ] Behavioral Health : Serious Mental Illness
- [ ] Behavioral Health : Suicide
- [ ] Behavioral Health : Tobacco Use
- [ ] Cancer : Cancer
- [ ] Cancer : Bladder
- [ ] Cancer : Breast
- [ ] Cancer : Colorectal
- [ ] Cancer : Gynecologic
- [ ] Cancer : Hematologic
- [ ] Cancer : Liver
- [ ] Cancer : Lung, Esophageal
- [ ] Cancer : Pancreatic
- [ ] Cancer : Prostate
- [ ] Cancer : Screening
- [ ] Cancer : Skin
- [ ] Cardiovascular : Cardiovascular
- [ ] Cardiovascular : Acute Myocardial Infarction
- [ ] Cardiovascular : Atrial Fibrillation
- [ ] Cardiovascular : Congestive Heart Failure
- [ ] Cardiovascular : Hyperlipidemia
- [ ] Cardiovascular : Hypertension
- [ ] Cardiovascular : Ischemic Heart Disease, Coronary Artery Disease
- [ ] Cardiovascular : Percutaneous Coronary Intervention (PCI)
- [ ] Cardiovascular : Screening
- [ ] Endocrine : Endocrine
- [ ] Endocrine : Diabetes
- [ ] Endocrine : Screening
- [ ] Endocrine : Thyroid Disorders
- [ ] Gastrointestinal (GI) : Gastrointestinal (GI)
- [ ] Gastrointestinal (GI) : Appendicitis
- [ ] Gastrointestinal (GI) : Cirrhosis
- [ ] Infectious Diseases : Hepatitis
- [ ] Infectious Diseases : Human Immunodeficiency Virus/Acquired Immune Deficiency Syndrome (HIV/AIDS)
- [ ] Infectious Diseases : Immunization
- [ ] Infectious Diseases : Respiratory
- [ ] Infectious Diseases : Sexually Transmitted
- [ ] Infectious Diseases : Tuberculosis
- [ ] Infectious Diseases : Screening
- [ ] Mental Health : Mental Health
- [ ] Mental Health : Alcohol, Substance Use/Abuse
- [ ] Mental Health : Depression
- [ ] Mental Health : Domestic Violence
- [ ] Mental Health : Serious Mental Illness
- [ ] Mental Health : Suicide
- [ ] Musculoskeletal : Musculoskeletal
- [ ] Musculoskeletal : Osteoarthritis
- [ ] Musculoskeletal : Rheumatoid Arthritis
- [ ] Musculoskeletal : Hip/Pelvic Fracture
- [ ] Musculoskeletal : Joint Surgery
- [ ] Musculoskeletal : Low Back Pain
- [ ] Musculoskeletal : Osteoporosis
- [ ] Neurology : Neurology
- [ ] Neurology : Brain Injury
- [ ] Neurology : Cognitive Impairment/Dementia
- [ ] Neurology : Delirium
- [ ] Neurology : Stroke/Transient Ischemic Attack (TIA)
- [ ] Perinatal and Reproductive Health : Perinatal and Reproductive Health
- [ ] Perinatal and Reproductive Health : Gynecology
- [ ] Perinatal and Reproductive Health : Newborn
- [ ] Perinatal and Reproductive Health : Perinatal
- [ ] Perinatal and Reproductive Health : Screening
- [ ] Prevention : Prevention
- [ ] Prevention : Development/Wellness
- [ ] Prevention : Immunization
- [ ] Prevention : Malnutrition
- [ ] Prevention : Obesity
- [ ] Prevention : Physical Activity
- [ ] Prevention : Screening

- [ ] Gastrointestinal (GI) : GI Bleeding
- [ ] Gastrointestinal (GI) : Gall Bladder Disease
- [ ] Gastrointestinal (GI) : Gastroenteritis
- [ ] Gastrointestinal (GI) : Gastro-Esophageal Reflux Disease (GERD)
- [ ] Gastrointestinal (GI) : Polyps
- [ ] Gastrointestinal (GI) : Screening
- [ ] Gastrointestinal (GI) : Peptic Ulcer
- [ ] GU/GYN : GU/GYN
- [ ] GU/GYN : Incontinence
- [ ] GU/GYN : Screening
- [ ] Head, Eyes, Ears, Nose, Throat (HEENT) : Head, Eyes, Ears, Nose, Throat (HEENT)
- [ ] Head, Eyes, Ears, Nose, Throat (HEENT) : Dental
- [ ] Head, Eyes, Ears, Nose, Throat (HEENT) : Ear Infection
- [ ] Head, Eyes, Ears, Nose, Throat (HEENT) : Hearing
- [ ] Head, Eyes, Ears, Nose, Throat (HEENT) : Pharyngitis
- [ ] Head, Eyes, Ears, Nose, Throat (HEENT) : Screening
- [ ] Head, Eyes, Ears, Nose, Throat (HEENT) : Vision
- [ ] Infectious Diseases : Infectious Diseases
- [ ] Prevention : Tobacco Use
- [ ] Pulmonary/Critical Care : Pulmonary/Critical Care
- [ ] Pulmonary/Critical Care : Asthma
- [ ] Pulmonary/Critical Care : Chronic Obstructive Pulmonary Disease (COPD)
- [ ] Pulmonary/Critical Care : Critical Care
- [ ] Pulmonary/Critical Care : Dyspnea
- [ ] Pulmonary/Critical Care : Pneumonia
- [ ] Pulmonary/Critical Care : Sleep Apnea
- [ ] Renal : Renal
- [ ] Renal : Chronic Kidney Disease (CKD)
- [ ] Renal : End Stage Renal Disease (ESRD)
- [ ] Surgery : Surgery
- [ ] Surgery : Cardiac Surgery
- [ ] Surgery : General Surgery
- [ ] Surgery : Perioperative
- [ ] Surgery : Thoracic Surgery
- [ ] Surgery : Vascular Surgery

De.6. Cross Cutting Areas (Check all the areas that apply):

- [ ] Health and Functional Status : Health and Functional Status
- [ ] Health and Functional Status : Development/Wellness
- [ ] Health and Functional Status : Functional Status
- [ ] Prevention : Prevention
- [ ] Prevention : Immunization
- [ ] Prevention : Nutrition
- [ ] Prevention : Obesity
- [ ] Prevention : Physical Activity
- [ ] Prevention : Screening
- [ ] Prevention : Social Determinants
- [ ] Access
- [ ] Care Coordination : Care Coordination
- [ ] Care Coordination : Readmissions
- [ ] Disparities
- [ ] Functional Status
- [ ] Overuse
- [ ] Palliative Care and End of Life Care
- [ ] Patient and Family Engagement
- [ ] Safety : Safety
- [ ] Safety : Complications
- [ ] Safety : Healthcare Associated Infections
- [ ] Safety : Medication Safety
- [ ] Safety : Venous Thromboembolism
- [ ] Safety : Readmissions
- [ ] Safety : Workforce

Measure Specifications

S.1. Measure-specific Web Page *(Provide a URL link to a web page specific for this measure that contains current detailed specifications including code lists, risk model details, and supplemental materials. Do not enter a URL linking to a home page or to general information.)* *

S.2a. If this is an eMeasure, HQMF specifications must be attached. Attach the output from the eMeasure authoring tool (MAT) - if the MAT was not used, contact staff. (Use the specification fields in this online form for the plain-language description of the specifications) *

- ☐ Available in attached file
- ☐ No HQMF specs

S.2b. Data Dictionary Code Table, *or value sets (and risk model codes and coefficients when applicable) must be attached. (Excel or csv file in the suggested format preferred - if not, contact staff)*

- ○ Available in attached Excel or csv file
- ○ No data dictionary/code table – all information provided in the submission form

S.3. For endorsement maintenance, please briefly describe any changes to the measure specifications since last endorsement date and explain the reasons.

S.4. Numerator Statement *(Brief, narrative description of the measure focus or what is being measured about the target population, i.e., cases from the target population with the target process, condition, event, or outcome)*
IF an OUTCOME MEASURE, state the outcome being measured. Calculation of the risk-adjusted outcome should be described in the calculation algorithm.

S.5. Time Period for Data *(What is the time period in which data will be aggregated for the measure, e.g., 12 mo, 3 years, look back to August for flu vaccination? Note if there are different time periods for the numerator and denominator.)*

S.6. Numerator Details *(All information required to identify and calculate the cases from the target population with the target process, condition, event, or outcome such as definitions, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b)*
IF an OUTCOME MEASURE, describe how the observed outcome is identified/counted. Calculation of the risk-adjusted outcome should be described in the calculation algorithm.

S.7. Denominator Statement *(Brief, narrative description of the target population being measured)*

S.8. Target Population Category *(Check all the populations for which the measure is specified and tested if any):*

- ☐ Children's Health
- ☐ Maternal Health
- ☐ Populations at Risk : Individuals with multiple chronic conditions

☐ Populations at Risk : Populations at Risk ☐ Populations at Risk : Veterans

☐ Populations at Risk : Dual eligible beneficiaries ☐ Senior Care

S.9. Denominator Details *(All information required to identify and calculate the target population/denominator such as definitions, specific data collection items/responses , code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b)*

S.10. Denominator Exclusions *(Brief narrative description of exclusions from the target population)*

S.11. Denominator Exclusion Details *(All information required to identify and calculate exclusions from the denominator such as definitions, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b)*

S.12. Stratification Details/Variables *(All information required to stratify the measure results including the stratification variables, definitions, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format with at S.2b)*

S.13. Risk Adjustment Type (Select type. Provide specifications for risk stratification in S.12 and for statistical model in S.14-15)

S.14. Identify the statistical risk model method and variables *(Name the statistical method - e.g., logistic regression and list all the risk factor variables. Note - risk model development and testing should be addressed with measure testing under Scientific Acceptability)*

S.15. Detailed risk model specifications *(must be in attached data dictionary/code list Excel or csv file. Also indicate if available at measure-specific URL identified in S.1.)*
Note: Risk model details (including coefficients, equations, codes with descriptors, definitions), should be provided on a separate worksheet in the suggested format in the Excel or csv file with data dictionary/code lists at S.2b.

○ Available at measure-specific web page URL identified in S.1

○ Available in attached Excel or csv file

S.16. Type of score:

S.17. Interpretation of Score *(Classifies interpretation of score according to whether better quality is associated*

with a higher score, a lower score, a score falling within a defined interval, or a passing score)

S.18. Calculation Algorithm/Measure Logic *(Describe the calculation of the measure score as an ordered sequence of steps including identifying the target population; exclusions; cases meeting the target process, condition, event, or outcome; aggregating data; risk adjustment; etc.)*

S.19. Calculation Algorithm/Measure Logic Diagram URL or Attachment *(You also may provide a diagram of the Calculation Algorithm/Measure Logic described above at measure-specific Web page URL identified in S.1 OR in attached appendix)*

- ○ Available at measure-specific web page URL identified in S.1
- ○ Available in attached appendix
- ○ No diagram provided

S.20. Sampling *(If measure is based on a sample, provide instructions for obtaining the sample and guidance on minimum sample size.)*

IF a PRO-PM, identify whether (and how) proxy responses are allowed.

S.21. Survey/Patient-reported data *(If measure is based on a survey, provide instructions for conducting the survey and guidance on minimum response rate.)*

IF a PRO-PM, specify calculation of response rates to be reported with performance measure results.

S.22. Missing data (specify how missing data are handled, e.g., imputation, delete case.) Required for Composites and PRO-PMs.

S.23. Data Source *(Check ONLY the sources for which the measure is SPECIFIED AND TESTED). If other, please describe in 2a1.26.*

- ☐ Administrative claims
- ☐ Electronic Clinical Data : Electronic Clinical Data
- ☐ Electronic Clinical Data : Electronic Health Record
- ☐ Electronic Clinical Data : Imaging/Diagnostic Study
- ☐ Electronic Clinical Data : Laboratory
- ☐ Electronic Clinical Data : Pharmacy
- ☐ Electronic Clinical Data : Registry
- ☐ Healthcare Provider Survey
- ☐ Management Data
- ☐ Paper Medical Records
- ☐ Patient Reported Data/Survey
- ☐ Other

S.24. Data Source or Collection Instrument *(Identify the specific data source/data collection instrument e.g. name of database, clinical registry, collection instrument, etc.)*

IF a PRO-PM, identify the specific PROM(s); and standard methods, modes, and languages of administration.

S.25. Data Source or Collection Instrument *(available at measure-specific Web page URL identified in S.1 OR in attached appendix)*

- ○ Available at measure-specific web page URL identified in S.1
- ○ Available in attached appendix
- ○ No data collection instrument provided

S.26. Level of Analysis *(Check ONLY the levels of analysis for which the measure is SPECIFIED AND TESTED)*

- ☐ Clinician : Individual
- ☐ Clinician : Group/Practice
- ☐ Clinician : Team
- ☐ Facility
- ☐ Health Plan
- ☐ Integrated Delivery System
- ☐ Population : Community
- ☐ Population : County or City
- ☐ Population : National
- ☐ Population : Regional
- ☐ Population : State

S.27. Care Setting *(Check ONLY the settings for which the measure is SPECIFIED AND TESTED)*

- ☐ Ambulatory Care : Ambulatory Surgery Center (ASC)
- ☐ Ambulatory Care : Clinician Office/Clinic
- ☐ Ambulatory Care : Outpatient Rehabilitation
- ☐ Ambulatory Care : Urgent Care
- ☐ Behavioral Health/Psychiatric : Inpatient
- ☐ Behavioral Health/Psychiatric : Outpatient
- ☐ Dialysis Facility
- ☐ Emergency Medical Services/Ambulance
- ☐ Home Health
- ☐ Hospice
- ☐ Hospital/Acute Care Facility
- ☐ Imaging Facility
- ☐ Laboratory
- ☐ Pharmacy
- ☐ Post Acute/Long Term Care Facility : Nursing Home/Skilled Nursing Facility
- ☐ Post Acute/Long Term Care Facility : Inpatient Rehabilitation Facility
- ☐ Post Acute/Long Term Care Facility : Long Term Acute Care Hospital
- ☐ Other

S.28. COMPOSITE Performance Measure - Additional Specifications *(Use this section as needed for aggregation and weighting rules,or calculation of individual performance measures if not indvidually endorsed.)*

Importance

Importance to Measure and Report is a threshold criterion that must be met in order to recommend a measure for endorsement.

All three subcriteria must be met to pass this criterion. See guidance on evidence.

Opportunity for Improvement (Measure evaluation criterion 1a)

1a. Attach evidence submission form (Click here to download Evidence Submission Form Template)

1b.1. Briefly explain the rationale for this measure (e.g., the benefits or improvements in quality envisioned by use of this measure)
IF a COMPOSITE (e.g. combination of component measure scores, all-or-none, any-or-none), SKIP this question and provide rationale for composite in question 1d.3 on the composite tab.

1b.2. Provide performance scores on the measure as specified (**current and over time**) at the specified level of anlaysis. (This is required for endorsement maintenance. Include mean, std dev, min, max, interquartile range, scores by decile. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities include). This informationa also will be used to address the subcriterion on improvement (4b.1) under Usability and Use.

1b.3. If no or limited performance data on the measure as specified is reported in 1b2, then provide a summary of data from the literature that indicates opportunity for improvement or overall less than optimal performance **on the specific focus of measurement.**

1b.4. Provide disparities data from the measure as specified (**current and over time**) by population group, e.g., by race/ethnicity, gender, age, insurance status, socioeconomic status, and/or disability. (This is required for endorsement maintenance. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities include.) This information also will be used to address the subcriterion on improvement (4b.1) under Usability and Use.

1b.5. If no or limited data on disparities from the measure as specified is reported in 1b4, then provide a summary of data from the literature that addresses disparities **in care on the specific focus of measurement.** Include citations.

High Priority (Measure evaluation criterion 1c)

1c.1. Demonstrated high priority aspect of healthcare

- ☐ Affects large numbers
- ☐ Patient/societal consequences of poor quality
- ☐ A leading cause of morbidity/mortality
- ☐ Severity of illness
- ☐ Frequently performed procedure
- ☐ Other
- ☐ High resource use

1c.3. Provide epidemiologic or resource use data that demonstrates the measure addresses a high priority aspect of healthcare). List citations in 1a.4.

1c.4. Citations for data demonstrating high priority provided in 1a.3

1c.5. IF a PRO-PM (e.g. HRQoL/functional status, symptom/burden, experience with care, health-related behaviors), provide evidence that the target population values the measured PRO and finds it meaningful. (Describe how and from whom their input was obtained.)

Scientific Acceptability

Testing Attachment

2.1. Attach measure testing form (Click to here to download the Measure Testing Submission Form OR the Composite Measure Testing Form.)

Feasibility

Extent to which the required data are readily available, retrievable without undue burden, and can be implemented for performance measurement.

Data Elements Generated as Byproduct of Care Processes (Measure evaluation criterion 3a)

3a.1. How are the data elements needed to compute measure scores generated? (Check all that apply)

Data used in the measure are:

- ☐ Generated "or collected" by and used by healthcare personnel during the provision of care (e.g., blood pressure, lab value, diagnosis, "depression score")
- ☐ Coded by someone other than person obtaining original information (e.g., DRG, ICD-9 codes on claims)
- ☐ Abstracted from a record by someone other than person obtaining original information (e.g., chart abstraction for quality measure or registry)
- ☐ Other

Electronic Sources (Measure evaluation criterion 3b)

3b.1. To what extent are the specified data elements available electronically in defined fields (*i.e., data elements that are needed to compute the performance measure score are in defined, computer-readable fields*)

- ALL data elements are in defined fields in electronic health records (EHRs)
- ALL data elements are in defined fields in electronic claims
- ALL data elements are in defined fields in electronic clinical data (e.g., clinical registry, nursing home MDS, home health OASIS)
- ALL data elements are in defined fields in a combination of electronic sources
- Some data elements are in defined fields in electronic sources
- No data elements are in defined fields in electronic sources

3b.3. If this is an eMeasure, provide a summary of the feasibility assessment in an attached file or make available at a measure-specific URL.

- Available at measure-specific web page URL identified in S.1
- Available in attached file
- No feasibility assessment

Data Collection Strategy (Measure evaluation criterion 3c)

3c.1. Describe what you have learned/modified as a result of testing and/or operational use of the measure regarding data collection, availability of data, missing data, timing and frequency of data collection, sampling, patient confidentiality, time and cost of data collection, other feasibility/implementation issues. IF a PRO-PM, consider implications for both individuals providing PROM data (patients, service recipients, respondents) and those whose performance is being measured.

3c.2. Describe any fees, licensing, or other requirements to use any aspect of the measure as specified (e.g., value/code set, risk model, programming code, algorithm)?

Usability and Use

Extent to which intended audiences (e.g., consumers, purchasers, providers, policy makers) can understand the results of the measure and are likely to find them useful for decision making.

NQF-endorsed measures are expected to be used in at least one accountability application within 3 years and publicly reported within 6 years of initial endorsement in addition to performance improvement.

4.1. Current and Planned Use *(check all the current and planned uses; for any current uses that are checked, provide a program name and URL for the specific program)*

Use	Planned	Current	For current use, provide Program Name and URL
a. Public Reporting	○	○	

b. Public Health/Disease Surveillance	○	○	
c. Payment Program	○	○	
d. Regulatory and Accreditation Programs	○	○	
e. Professional Certification or Recognition Program	○	○	
f. Quality Improvement with Benchmarking (external benchmarking to multiple organizations)	○	○	
g. Quality Improvement (Internal to the specific organization)	○	○	
h. Not in use	○	○	

4a.1. For each CURRENT use, checked above, provide:

- **Name of program and sponsor**
- **Purpose**
- **Geographic area and number and percentage of accountable entities and patients included**

4a.2. If not currently publicly reported OR used in at least one other accountability application (e.g., payment program, certification, licensing) what are the reasons? (***e.g., Do policies or actions of the developer/steward or accountable entities restrict access to performance results or impede implementation?*** **)**

4a.3. If not currently publicly reported OR used in at least one other accountability application, provide a credible plan for implementation within the expected timeframes -- any accountability application within 3 years and publicly reported within 6 years of initial endorsement. (***Credible plan includes the specific program, purpose, intended audience, and timeline for implementing the measure within the specified timeframes. A plan for accountability applications addresses mechanisms for data aggregation and reporting.*** **)**

4b.1. Progress on Improvement. (Not required for initial endorsement unless available.)

Performance results on this measure (current and over time) should be provided in 1b.2 and 1b.4. Discuss:

- **Progress (trends in performance results, number and percentage of people receiving high-quality healthcare)**
- **Geographic area and number and percentage of accountable entities and patients included**

4b.2. If no improvement was demonstrated, what are the reasons? If not in use for performance improvement at the time of initial endorsement, provide a credible rationale that describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.

4c.1. Were any unintended negative consequences to individuals or populations identified during testing; OR has evidence of unintended negative consequences to individuals or populations been reported since implementation? If so, identify the negative unintended consequences and describe how benefits outweigh them or actions taken to mitigate them.

Related and Competing Measures

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.

Relation to Other NQF-endorsed® Measures (Measure evaluation criterion 5)

If there are related measures (conceptually, either same measure focus or target population) or competing measures (conceptually both the same measure focus and same target population), list the NQF # and title of all related and/or competing measures. (Can search and select measures.)

- Yes
- No

Harmonization (Measure evaluation criterion 5a)

5a.1. If this measure conceptually addresses EITHER the same measure focus OR the same target population as NQF-endorsed measure(s):

Are the measure specifications completely harmonized?

- Yes
- No

5a.2. If the measure specifications are not completely harmonized, identify the differences, rationale, and impact on interpretability and data collection burden.

Competing Measure(s) (Measure evaluation criterion 5b)

5b.1. If this measure conceptually addresses both the same measure focus and the same target population as NQF-endorsed measure(s):

Describe why this measure is superior to competing measures (e.g., a more valid or efficient way to measure quality); OR provide a rationale for the additive value of endorsing an additional measure. (Provide analyses when possible.)

Additional

Authorized Users

Steward	Developer	Username	First Name	Last Name	Organization
○	○	MTOBIAS	Mark	Tobias	NQF TEST

View / Edit My Account

Appendix-Attachment

A.1. Supplemental materials may be provided in an appendix.

All supplemental materials (such as data collection instrument or methodology reports) should be organized in one file with a table of contents or

bookmarks. If material pertains to a specific submission form number that should be indicated.

Requested information should be provided in the

submission form and measure testing attachment. There is no guarantee that supplemental materials will be reviewed.

- Available at measure-specific web page URL identified in S.1
- Available in attached file
- No appendix

Contact Information

Co.1. Steward Point of Contact

Co.1.1. Organization

Co.1.2. First Name

Co.1.3. Last Name

Co.1.4. Email Address

Co.1.5. Phone Number

( ) - ext.

Co.2. Developer Point of Contact

☐ **Same as Measure Steward Point of Contact**

Co.2.1. Organization

Co.2.2. First Name

Co.2.3. Last Name

Co.2.4. Email Address

Co.2.5. Phone Number

( ) - ext.

Additional Information

Ad.1. Workgroup/Expert Panel Involved in Measure Development

List the workgroup/panel members' names and organizations.

Describe the members' role in measure development.

Measure Developer/Steward Updates and Ongoing

Maintenance Ad.2. Year the Measure Was First Released

Ad.3. Month and Year of Most Recent Revision

Ad.4. What is your frequency for review/update of this measure?

Ad.5. When is your next scheduled review/update for this measure?

Ad.6. Copyright Statement

Ad.7. Disclaimers

Ad.8. Additional Information/Comments

NATIONAL QUALITY FORUM—Evidence (subcriterion 1a)

Measure Title: Click here to enter measure title
IF the measure is a component in a composite performance measure, provide the title of the Composite Measure here: Click here to enter composite measure title

Date of Submission: Click here to enter a date

Instructions

- *For composite performance measures:*
  - *A separate evidence form is required for each component measure unless several components were studied together.*
  - *If a component measure is submitted as an individual performance measure, attach the evidence form to the individual measure submission.*
- Respond to all questions as instructed with answers immediately following the question. All information needed to demonstrate meeting the evidence subcriterion (1a) must be in this form. An appendix of *supplemental* materials may be submitted, but there is no guarantee it will be reviewed.
- If you are unable to check a box, please highlight or shade the box for your response.
- Maximum of 10 pages (*incudes questions/instructions*; minimum font size 11 pt; do not change margins). ***Contact NQF staff if more pages are needed.***
- Contact NQF staff regarding questions. Check for resources at Submitting Standards webpage.

Note: The information provided in this form is intended to aid the Steering Committee and other stakeholders in understanding to what degree the evidence for this measure meets NQF's evaluation criteria.

Subcriterion 1a. Evidence to Support the Measure Focus

The measure focus is a health outcome or is evidence-based, demonstrated as follows:

- Health outcome:[3] a rationale supports the relationship of the health outcome to processes or structures of care.
- Intermediate clinical outcome, Process,[4] or Structure: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence[5] that the measure focus leads to a desired health outcome.
- Patient experience with care: evidence that the measured aspects of care are those valued by patients and for which the patient is the best and/or only source of information OR that patient experience with care is correlated with desired outcomes.
- Efficiency:[6] evidence for the quality component as noted above.

Notes

3. Generally, rare event outcomes do not provide adequate information for improvement or discrimination; however, serious reportable events that are compared to zero are appropriate outcomes for public reporting and quality improvement.

4. Clinical care processes typically include multiple steps: assess → identify problem/potential problem → choose/plan intervention (with patient input) → provide intervention → evaluate impact on health status. If the measure focus is one step in such a multistep process, the step with the strongest evidence for the link to the desired outcome should be selected as the focus of measurement.

5. The preferred systems for grading the evidence are the U.S. Preventive Services Task Force (USPSTF) grading definitions and methods, or Grading of Recommendations, Assessment, Development and Evaluation (GRADE) guidelines.

6. Measures of efficiency combine the concepts of resource use and quality (NQF's Measurement Framework: Evaluating Efficiency Across Episodes of Care; AQA Principles of Efficiency Measures).

1a.1.This is a measure of:
Outcome
- ☐ Health outcome: Click here to name the health outcome
  Health outcome includes patient-reported outcomes (PRO, i.e., HRQoL/functional status, symptom/burden, experience with care, health-related behaviors)
  - ☐ Intermediate clinical outcome: Click here to name the intermediate outcome
- ☐ Process: Click here to name the process
- ☐ Structure: Click here to name the structure
- ☐ Other: Click here to name what is being measured

HEALTH OUTCOME PERFORMANCE MEASURE *If not a health outcome, skip to 1a.3*

1a.2. Briefly state or diagram the linkage between the health outcome (or PRO) and the healthcare structures, processes, interventions, or services that influence it.

1a.2.1. State the rationale supporting the relationship between the health outcome (or PRO) and at least one healthcare structure, process, intervention, or service.

Note: For health outcome performance measures, no further information is required; however, you may provide evidence for any of the structures, processes, interventions, or service identified above.

INTERMEDIATE OUTCOME, PROCESS, OR STRUCTURE PERFORMANCE MEASURE

1a.3. Briefly state or diagram the linkages between structure, process, intermediate outcome, and health outcomes. Include all the steps between the measure focus and the health outcome.

1a.3.1. What is the source of the systematic review of the body of evidence that supports the performance measure?
- ☐ Clinical Practice Guideline recommendation – ***complete sections 1a.4, and 1a.7***
- ☐ US Preventive Services Task Force Recommendation – ***complete sections 1a.5 and 1a.7***
- ☐ Other systematic review and grading of the body of evidence (*e.g., Cochrane Collaboration, AHRQ Evidence Practice Center*) – ***complete sections 1a.6 and 1a.7***
- ☐ Other – ***complete section 1a.8***

Please complete the sections indicated above for the source of evidence. You may skip the sections that do not apply.

1a.4. CLINICAL PRACTICE GUIDELINE RECOMMENDATION

1a.4.1. Guideline citation (*including date*) and **URL for guideline** (*if available online*):

1a.4.2. Identify guideline recommendation number and/or page number and **quote verbatim, the specific guideline recommendation**.

1a.4.3. Grade assigned to the quoted recommendation with definition of the grade:

1a.4.4. Provide all other grades and associated definitions for recommendations in the grading system. (*Note: If separate grades for the strength of the evidence, report them in section 1a.7.*)

1a.4.5. Citation and URL for methodology for grading recommendations (*if different from 1a.4.1*)**:**

1a.4.6. If guideline is evidence-based (rather than expert opinion), are the details of the quantity, quality, and consistency of the body of evidence available (e.g., evidence tables)?

☐ Yes → ***complete section 1a.7***

☐ No → ***report on another systematic review of the evidence in sections 1a.6 and 1a.7; if another review does not exist, provide what is known from the guideline review of evidence in 1a.7***

1a.5. UNITED STATES PREVENTIVE SERVICES TASK FORCE RECOMMENDATION

1a.5.1. Recommendation citation (*including date*) and **URL for recommendation** (*if available online*):

1a.5.2. Identify recommendation number and/or page number and **quote verbatim, the specific recommendation**.

1a.5.3. Grade assigned to the quoted recommendation with definition of the grade:

1a.5.4. Provide all other grades and associated definitions for recommendations in the grading system. (*Note: the grading system for the evidence should be reported in section 1a.7.*)

1a.5.5. Citation and URL for methodology for grading recommendations (*if different from 1a.5.1*)**:**

Complete section 1a.7

1a.6. OTHER SYSTEMATIC REVIEW OF THE BODY OF EVIDENCE

1a.6.1. Citation (*including date*) and **URL** (*if available online*):

1a.6.2. Citation and URL for methodology for evidence review and grading (*if different from 1a.6.1*)**:**

Complete section 1a.7

1a.7. FINDINGS FROM SYSTEMATIC REVIEW OF BODY OF THE EVIDENCE SUPPORTING THE MEASURE

1a.7.1. What was the specific structure, treatment, intervention, service, or intermediate outcome addressed in the evidence review?

1a.7.2. Grade assigned for the quality of the quoted evidence with definition of the grade:

1a.7.3. Provide all other grades and associated definitions for strength of the evidence in the grading system.

1a.7.4. What is the time period covered by the body of evidence? (*provide the date range, e.g., 1990-2010*). Date range: Click here to enter date range

QUANTITY AND QUALITY OF BODY OF EVIDENCE

1a.7.5. How many and what type of study designs are included in the body of evidence? (*e.g., 3 randomized controlled trials and 1 observational study*)

1a.7.6. What is the overall quality of evidence across studies in the body of evidence? (*discuss the certainty or confidence in the estimates of effect particularly in relation to study factors such as design flaws, imprecision due to small numbers, indirectness of studies to the measure focus or target population*)

ESTIMATES OF BENEFIT AND CONSISTENCY ACROSS STUDIES IN BODY OF EVIDENCE

1a.7.7. What are the estimates of benefit—magnitude and direction of effect on outcome(s) across studies in the body of evidence? (*e.g., ranges of percentages or odds ratios for improvement/ decline across studies, results of meta-analysis, and statistical significance*)

1a.7.8. What harms were studied and how do they affect the net benefit (benefits over harms)?

UPDATE TO THE SYSTEMATIC REVIEW(S) OF THE BODY OF EVIDENCE

1a.7.9. If new studies have been conducted since the systematic review of the body of evidence, provide for each new study: 1) citation, 2) description, 3) results, 4) impact on conclusions of systematic review.

1a.8 OTHER SOURCE OF EVIDENCE

If source of evidence is NOT from a clinical practice guideline, USPSTF, or systematic review, please describe the evidence on which you are basing the performance measure.

1a.8.1 What process was used to identify the evidence?

1a.8.2. Provide the citation and summary for each piece of evidence.

NATIONAL QUALITY FORUM—Composite Measure Testing (subcriteria 2a2, 2b2-2b6, 2d)

Composite Measure Title: Click here to enter measure title

Date of Submission: Click here to enter a date

Composite Construction:

☐Two or more individual performance measure scores combined into one score

☐ All-or-none measures (e.g., all essential care processes received or outcomes experienced by each patient)

☐ Any-or-none measures (e.g., any or none of a list of adverse outcomes experienced, or inappropriate or unnecessary care processes received, by each patient)

Instructions: Please contact NQF staff before you begin.

- If a component measure is submitted as an individual performance measure, the non-composite measure testing form must also be completed and attached to the individual measure submission.
- Measures must be tested for all the data sources and levels of analyses that are specified. ***If there is more than one set of data specifications or more than one level of analysis, contact NQF staff*** about how to present all the testing information in one form.
- **For all composite measures, sections 1, 2a2, 2b2, 2b3, 2b5, and 2d must be completed.**
- **For composites with outcome and resource use measures**, section **2b4** also must be completed.
- If specified for **multiple data sources/sets of specificaitions** (e.g., claims and EHRs), section **2b6** also must be completed.
- Respond to all questions as instructed with answers immediately following the question. All information on testing to demonstrate meeting the subcriteria for reliability (2a2), validity (2b2-2b6), and composites (2d) must be in this form. An appendix for *supplemental* materials may be submitted, but there is no guarantee it will be reviewed.
- If you are unable to check a box, please highlight or shade the box for your response.
- Maximum of 25 pages (*incuding questions/instructions;* minimum font size 11 pt; do not change margins). ***Contact NQF staff if more pages are needed.***
- Contact NQF staff regarding questions. Check for resources at Submitting Standards webpage.

Note: The information provided in this form is intended to aid the Steering Committee and other stakeholders in understanding to what degree the testing results for this measure meet NQF's evaluation criteria for testing.

2a2. Reliability testing [10] demonstrates the measure data elements 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.

2b2. Validity testing [11] demonstrates that the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality.

2b3. Exclusions are supported by the clinical evidence; otherwise, they are supported by evidence of sufficient frequency of occurrence so that results are distorted without the exclusion; [12]

AND

If patient preference (e.g., informed decisionmaking) 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). [13]

2b4. For outcome measures and other measures when indicated (e.g., resource use):

- **an evidence-based risk-adjustment strategy** (e.g., risk models, risk stratification) is specified; is based on patient factors that influence the measured outcome (but not factors related to disparities in care or the quality of care) and are present at start of care; [14,15] and has demonstrated adequate discrimination and calibration

OR

- rationale/data support no risk adjustment/ stratification.

2b5. Data analysis of computed measure scores demonstrates that methods for scoring and analysis of the specified measure allow for **identification of statistically significant and practically/clinically meaningful [16] differences in performance**;
OR
there is evidence of overall less-than-optimal performance.

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

Composite 2d. For composite performance measures, empirical analyses support the composite construction approach and demonstrate that:

1) the component measures fit the quality construct and add value to the overall composite while achieving the related objective of parsimony to the extent possible; and
2) the aggregation and weighting rules are consistent with the quality construct and rationale while achieving the related objective of simplicity to the extent possible; and
3) the extent of missing data and how the specified handling of missing data minimizes bias (i.e., achieves scores that are an accurate reflection of quality).

Notes

10. Reliability testing applies to both the data elements and computed measure score. Examples of reliability testing for data elements include, but are not limited to: inter-rater/abstractor or intra-rater/abstractor studies; internal consistency for multi-item scales; test-retest for survey items. Reliability testing of the measure score addresses precision of measurement (e.g., signal-to-noise).

11. Validity testing applies to both the data elements and computed measure score. Validity testing of data elements typically analyzes agreement with another authoritative source of the same information. Examples of validity testing of the measure score include, but are not limited to: testing hypotheses that the measures scores indicate quality of care, e.g., measure scores are different for groups known to have differences in quality assessed by another valid quality measure or method; correlation of measure scores with another valid indicator of quality for the specific topic; or relationship to conceptually related measures (e.g., scores on process measures to scores on outcome measures). Face validity of the measure score as a quality indicator may be adequate if accomplished through a systematic and transparent process, by identified experts, and explicitly addresses whether performance scores resulting from the measure as specified can be used to distinguish good from poor quality.

12. Examples of evidence that an exclusion distorts measure results include, but are not limited to: frequency of occurrence, variability of exclusions across providers, and sensitivity analyses with and without the exclusion.

13. Patient preference is not a clinical exception to eligibility and can be influenced by provider interventions.

14. Risk factors that influence outcomes should not be specified as exclusions.

15. Risk models should not obscure disparities in care for populations by including factors that are associated with differences/inequalities in care, such as race, socioeconomic status, or gender (e.g., poorer treatment outcomes of African American men with prostate cancer or inequalities in treatment for CVD risk factors between men and women). It is preferable to stratify measures by race and socioeconomic status rather than to adjust out the differences.

16. With large enough sample sizes, small differences that are statistically significant may or may not be practically or clinically meaningful. The substantive question may be, for example, whether a statistically significant difference of one percentage point in the percentage of patients who received smoking cessation counseling (e.g., 74 percent v. 75 percent) is clinically meaningful; or whether a statistically significant difference of $25 in cost for an episode of care (e.g., $5,000 v. $5,025) is practically meaningful. Measures with overall less-than-optimal performance may not demonstrate much variability across providers.

1. DATA/SAMPLE USED FOR ALL TESTING OF THIS MEASURE

Often the same data are used for all aspects of measure testing. In an effort to eliminate duplication, the first five questions apply to all measure testing. If there are differences by aspect of testing,(e.g., reliability vs. validity) be sure to indicate the specific differences in question 1.7.

1.1. What type of data was used for testing? (*Check all the sources of data identified in the measure specifications and data used for testing the measure. Testing must be provided for all the sources of data specified and intended for measure implementation. If different data sources are used for different components in the composite, indicate the component after the checkbox.*)

Measure Specified to Use Data From: (*must be consistent with data sources entered in S.23*)	**Measure Tested with Data From:**
☐ abstracted from paper record	☐ abstracted from paper record
☐ administrative claims	☐ administrative claims
☐ clinical database/registry	☐ clinical database/registry
☐ abstracted from electronic health record	☐ abstracted from electronic health record
☐ eMeasure (HQMF) implemented in EHRs	☐ eMeasure (HQMF) implemented in EHRs
☐ other: Click here to describe	☐ other: Click here to describe

1.2. If an existing dataset was used, identify the specific dataset (*the dataset used for testing must be consistent with the measure specifications for target population and healthcare entities being measured; e.g., Medicare Part A claims, Medicaid claims, other commercial insurance, nursing home MDS, home health OASIS, clinical registry*).

1.3. What are the dates of the data used in testing? Click here to enter date range

1.4. What levels of analysis were tested? (*testing must be provided for all the levels specified and intended for measure implementation, e.g., individual clinician, hospital, health plan*)

Measure Specified to Measure Performance of: (*must be consistent with levels entered in item S.26*)	**Measure Tested at Level of:**
☐ individual clinician	☐ individual clinician
☐ group/practice	☐ group/practice
☐ hospital/facility/agency	☐ hospital/facility/agency
☐ health plan	☐ health plan
☐ other: Click here to describe	☐ other: Click here to describe

1.5. How many and which measured entities were included in the testing and analysis (by level of analysis and data source)? (*identify the number and descriptive characteristics of measured entities included in the analysis (e.g., size, location, type); if a sample was used, describe how entities were selected for inclusion in the sample*)

1.6. How many and which patients were included in the testing and analysis (by level of analysis and data source)? (*identify the number and descriptive characteristics of patients included in the analysis (e.g., age, sex, race, diagnosis); if a sample was used, describe how patients were selected for inclusion in the sample*)

1.7. If there are differences in the data or sample used for different aspects of testing (e.g., reliability, validity, exclusions, risk adjustment), identify how the data or sample are different for each aspect of testing reported below.

_____

2a2. RELIABILITY TESTING

2a2.1. What level of reliability testing was conducted?

Note*: Current guidance for composite measure evaluation states that reliability must be demonstrated for the composite performance measure score.*

☒ **Performance measure score** (e.g., *signal-to-noise analysis*)

2a2.2. Describe the method of reliability testing and what it tests (*describe the steps—do not just name a method; what type of error does it test; what statistical analysis was used*)

2a2.3. What were the statistical results from reliability testing? (e.*g., percent agreement and kappa for the critical data elements; distribution of reliability statistics from a signal-to-noise analysis*)

2a2.4 What is your interpretation of the results in terms of demonstrating reliability? (i.*e., what do the results mean and what are the norms for the test conducted?*)

_____

2b2. VALIDITY TESTING

Note*: Current guidance for composite measure evaluation states that validity should be demonstrated for the composite performance measure score. If not feasible for initial endorsement, acceptable alternatives include assessment of content or face validity of the composite OR demonstration of validity for each component. Empirical validity testing of the composite measure score is expected by the time of endorsement maintenance.*

2b2.1. What level of validity testing was conducted?

☐ **Composite performance measure score**

- ☐ **Empirical validity testing**
- ☐ **Systematic assessment of face validity of performance measure score as an indicator** of quality or resource use (*i.e., is an accurate reflection of performance on quality or resource use and can distinguish good from poor performance*)
- ☐ **Systematic assessment of content validity**

☐ **Validity testing for component measures** *(check all that apply)*

Note*: applies to ALL component measures, unless already endorsed or are being submitted for individual endorsement.*

- ☐ **Endorsed (or submitted) as individual performance measures**
- ☐ **Critical data elements** (*data element validity must address ALL critical data elements*)
- ☐ **Empirical validity testing of the component measure score(s)**
- ☐ **Systematic assessment of face validity of component measure score(s) as an indicator** of quality or resource use (*i.e., is an accurate reflection of performance on quality or resource use and can distinguish good from poor performance*)

2b2.2. For each level checked above, describe the method of validity testing and what it tests (*describe the steps—do not just name a method; what was tested, e.g., accuracy of data elements compared to authoritative source, relationship to another measure as expected; what statistical analysis*

was used)

2b2.3. What were the statistical results from validity testing? (*e.g., correlation; t-test*)

2b2.4. What is your interpretation of the results in terms of demonstrating validity? (i.*e., what do the results mean and what are the norms for the test conducted?*)

_____

2b3. EXCLUSIONS ANALYSIS

Note: *Applies to the composite performance measure, as well all component measures unless they are already endorsed or are being submitted for individual endorsement.*

NA ☐ no exclusions — *skip to section 2b4*

2b3.1. Describe the method of testing exclusions and what it tests (*describe the steps—do not just name a method; what was tested, e.g., whether exclusions affect overall performance scores; what statistical analysis was used*)

2b3.2. What were the statistical results from testing exclusions? (*include overall number and percentage of individuals excluded, frequency distribution of exclusions across measured entities, and impact on performance measure scores*)

2b3.3. What is your interpretation of the results in terms of demonstrating that exclusions are needed to prevent unfair distortion of performance results? (*i.e., the value outweighs the burden of increased data collection and analysis. Note: **If patient preference is an exclusion**, the measure must be specified so that the effect on the performance score is transparent, e.g., scores with and without exclusion*)

_____

2b4. RISK ADJUSTMENT/STRATIFICATION FOR OUTCOME OR RESOURCE USE MEASURES

Note: *Applies to all outcome or resource use component measures, unless already endorsed or are being submitted for individual endorsement.*

If not an intermediate or health outcome, or PRO-PM, or resource use measure, skip to section 2b5.

2b4.1. What method of controlling for differences in case mix is used? *(check all that apply)*

☐ **Endorsed (or submitted) as individual performance measures**
☐ **No risk adjustment or stratification**
☐ **Statistical risk model**
☐ **Stratification by risk categories**
☐ **Other,** Click here to enter description

2b4.2. If an outcome or resource use component measure is not risk adjusted or stratified, provide rationale and analyses to demonstrate that controlling for differences in patient characteristics (case mix) is not needed to achieve fair comparisons across measured entities.

2b4.3. Describe the conceptual/clinical and statistical methods and criteria used to select patient factors used in the statistical risk model or for stratification by risk (*e.g., potential factors identified in the literature and/or expert panel; regression analysis; statistical significance of p<0.10; correlation of x or higher; patient factors should be present at the start of care and not related to disparities*)

2b4.4. What were the statistical results of the analyses used to select risk factors?

2b4.5. Describe the method of testing/analysis used to develop and validate the adequacy of the statistical model or stratification approach (*describe the steps—do not just name a method; what statistical analysis was used*)

Provide the statistical results from testing the approach to controlling for differences in patient characteristics (case mix) below.
if stratified, skip to 2b4.9

2b4.6. Statistical Risk Model Discrimination Statistics (*e.g., c-statistic, R-squared*)**:**

2b4.7. Statistical Risk Model Calibration Statistics (*e.g., Hosmer-Lemeshow statistic*):

2b4.8. Statistical Risk Model Calibration – Risk decile plots or calibration curves:

2b4.9. Results of Risk Stratification Analysis:

2b4.10. What is your interpretation of the results in terms of demonstrating adequacy of controlling for differences in patient characteristics (case mix)? (i.*e., what do the results mean and what are the norms for the test conducted?*)

***2b4.11. Optional Additional Testing for Risk Adjustment** (*not required, but would provide additional support of adequacy of risk model, e.g., testing of risk model in another data set; sensitivity analysis for missing data; other methods*)

_____

2b5. IDENTIFICATION OF STATISTICALLY SIGNIFICANT & MEANINGFUL DIFFERENCES IN PERFORMANCE
Note: *Applies to the composite performance measure.*

2b5.1. Describe the method for determining if statistically significant and clinically/practically meaningful differences in performance measure scores among the measured entities can be identified (*describe the steps—do not just name a method; what statistical analysis was used? Do not just repeat the information provided related to performance gap in 1b)*

2b5.2. What were the statistical results from testing the ability to identify statistically significant and/or clinically/practically meaningful differences in performance measure scores across measured entities? (e.g., *number and percentage of entities with scores that were statistically significantly different from mean or some benchmark, different from expected; how was meaningful difference defined*)

2b5.3. What is your interpretation of the results in terms of demonstrating the ability to identify statistically significant and/or clinically/practically meaningful differences in performance across measured entities? (i.*e., what do the results mean in terms of statistical and meaningful differences?*)

2b6. COMPARABILITY OF PERFORMANCE SCORES WHEN MORE THAN ONE SET OF SPECIFICATIONS
Note: *Applies to all component measures, unless already endorsed or are being submitted for individual endorsement.*
If only one set of specifications for each component, this section can be skipped.

Note*: This criterion is directed to measures with more than one set of specifications/instructions (e.g., one set of specifications for how to identify and compute the measure from medical record abstraction and a different set of specifications for claims or eMeasures). It does not apply to measures that use more than one source of data in one set of specifications/instructions (e.g., claims data to identify the denominator and medical record abstraction for the numerator).* ***If comparability is not demonstrated, the different specifications should be submitted as separate measures.***

2b6.1. Describe the method of testing conducted to demonstrate comparability of performance scores for the same entities across the different data sources/specifications (*describe the steps—do not just name a method; what statistical analysis was used*)

2b6.2. What were the statistical results from testing comparability of performance scores for the same entities when using different data sources/specifications? (*e.g., correlation, rank order*)

2b6.3. What is your interpretation of the results in terms of demonstrating comparability of performance measure scores for the same entities across the different data sources/specifications? (i.*e., what do the results mean and what are the norms for the test conducted?*)

2d. EMPIRICAL ANALYSIS TO SUPPORT COMPOSITE CONSTRUCTION APPROACH
Note: *If empirical analyses do not provide adequate results—or are not conducted—justification must be provided and accepted in order to meet the must-pass criterion of Scientific Acceptability of Measure Properties. Each of the following questions has instructions if there is no empirical analysis.*

2d1. Empirical analysis demonstrating that the component measures fit the quality construct, add value to the overall composite, and achieve the object of parsimony to the extent possible.

2d1.1 Describe the method used (*describe the steps—do not just name a method; what statistical analysis was used; if no empirical analysis, provide justification)*

2d1.2. What were the statistical results obtained from the analysis of the components? (e.g., *correlations, contribution of each component to the composite score, etc*.; *if no empirical analysis, identify the components that were considered and the pros and cons of each*)

2d1.3. What is your interpretation of the results in terms of demonstrating that the components included in the composite are consistent with the described quality construct and add value to the overall composite? (i.*e., what do the results mean in terms of supporting inclusion of the components; if no empirical analysis, provide rationale for the components that were selected)*

2d2. Empirical analysis demonstrating that the aggregations and weighting rules are consistent with the quality construct and achieve the objective of simplicity to the extent possible

2d2.1 Describe the method used (*describe the steps—do not just name a method; what statistical analysis was used; if no empirical analysis, provide justification)*

2d2.2. What were the statistical results obtained from the analysis of the aggregation and weighting rules? (e.g., *results of sensitivity analysis of effect of different aggregations and/or weighting rules; if no empirical analysis, identify the aggregation and weighting rules that were considered and the pros and cons of each*)

2d2.3. What is your interpretation of the results in terms of demonstrating the aggregation and weighting rules are consistent with the described quality construct? (i.*e., what do the results mean in terms of supporting the selected rules for aggregation and weighting; if no empirical analysis, provide rationale for the selected rules for aggregation and weighting*)

2d3. Empirical analysis demonstrating that the approach for handling missing data minimizes bias (*i.e., achieves scores that are an accurate reflection of quality*)**.**
Note: *Applies to the overall composite measure; the focus is on missing data rather than exclusions, which are considered in 2b3.*

2d3.1. What is the overall frequency of missing data and the distribution of missing data across providers?

2d3.2. Describe the method used to compare approaches for handling missing data (*describe the steps—do not just name a method; what statistical analysis was used; if no empirical analysis, provide justification)*

2d3.3. What were the statistical results obtained from the analysis of missing data? (e.g., *results of sensitivity analysis of effect of various rules for missing data; if no empirical analysis, identify the approaches for handling missing data that were considered and pros and cons of each*)

2d3.4. What is your interpretation of the results in terms of demonstrating that the approach used for missing data minimizes bias? (i.*e., what do the results mean in terms of supporting the selected approach for missing data; if no empirical analysis, provide rationale for the selected approach for missing data*)