NATIONAL QUALITY FORUM

Measure Evaluation 4.1
January 2010

This form contains the measure information submitted by stewards. Blank fields indicate no information was provided. Attachments also may have been submitted and are provided to reviewers. The sub-criteria and most of the footnotes from the evaluation criteria are provided in Word comments and will appear if your cursor is over the highlighted area (or in the margin if your Word program is set to show revisions in balloons). Hyperlinks to the evaluation criteria and ratings are provided in each section.

TAP/Workgroup (if utilized): Complete all yellow highlighted areas of the form. Evaluate the extent to which each sub-criterion is met. Based on your evaluation, summarize the strengths and weaknesses in each section.

Note: If there is no TAP or workgroup, the SC also evaluates the sub-criteria (yellow highlighted areas).

Steering Committee: Complete all pink highlighted areas of the form. Review the workgroup/TAP assessment of the sub-criterion, noting any areas of disagreement; then evaluate the extent to which each major criterion is met; and finally, indicate your recommendation for the endorsement. Provide the rationale for your ratings.

Evaluation ratings of the extent to which the criteria are met
C = Completely (unquestionably demonstrated to meet the criterion)
P = Partially (demonstrated to partially meet the criterion)
M = Minimally (addressed BUT demonstrated to only minimally meet the criterion)
N = Not at all (NOT addressed; OR incorrectly addressed; OR demonstrated to NOT meet the criterion)
NA = Not applicable (only an option for a few sub-criteria as indicated)

(for NQF staff use) NQF Review #: ACP-016-10
NQF Project: Ambulatory Care - Additional Outpatient Measures 2010

MEASURE DESCRIPTIVE INFORMATION

De.1 Measure Title: Endoscopy/Poly Surveillance: Appropriate follow-up interval for normal colonoscopy in average risk patients

De.2 Brief description of measure: Percentage of patients aged 50 years and older receiving a screening colonoscopy without biopsy or polypectomy who had a recommended follow-up interval of at least 10 years for repeat colonoscopy documented in their colonoscopy report.

De.3 Type of Measure: process

De.4 National Priority Partners Priority Area: Overuse

De.5 IOM Quality Domain: patient-centered, safety, timeliness, effectiveness

De.6 Consumer Care Need: Staying Healthy

CONDITIONS FOR CONSIDERATION BY NQF

Four conditions must be met before proposed measures may be considered and evaluated for suitability as voluntary consensus standards:

A. The measure is in the public domain or an intellectual property (measure steward agreement) is signed.

Public domain only applies to governmental organizations. All non-government organizations must sign a measure steward agreement even if measures are made publicly and freely available.

A.1 Do you attest that the measure steward holds intellectual property rights to the measure and the right to use aspects of the measure owned by another entity (e.g., risk model, code set)? Yes

A.2 Indicate if Proprietary Measure (as defined in measure steward agreement): Y

A.3 Measure Steward Agreement: agreement signed and submitted Y

A.4 Measure Steward Agreement attached: N

Rating: C=Completely; P=Partially; M=Minimally; N=Not at all; NA=Not applicable
B. The measure owner/steward verifies there is an identified responsible entity and process to maintain and update the measure on a schedule that is commensurate with the rate of clinical innovation, but at least every 3 years. Yes, information provided in contact section

C. The intended use of the measure includes both public reporting and quality improvement. Purpose: public reporting, quality improvement Accountability

D. The requested measure submission information is complete. Generally, measures should be fully developed and tested so that all the evaluation criteria have been addressed and information needed to evaluate the measure is provided. Measures that have not been tested are only potentially eligible for a time-limited endorsement and in that case, measure owners must verify that testing will be completed within 12 months of endorsement.

D.1 Testing: No, testing will be completed within 12 months

D.2 Have NQF-endorsed measures been reviewed to identify if there are similar or related measures? Yes

(for NQF staff use) Have all conditions for consideration been met?

Staff Notes to Steward (if submission returned):

Staff Notes to Reviewers (issues or questions regarding any criteria):

Staff Reviewer Name(s):

TAP/Workgroup Reviewer Name:

Steering Committee Reviewer Name:

1. IMPORTANCE TO MEASURE AND REPORT

<table>
<thead>
<tr>
<th>Extent to which the specific measure focus is important to making significant gains in health care quality (safety, timeliness, effectiveness, efficiency, equity, patient-centeredness) and improving health outcomes for a specific high impact aspect of healthcare where there is variation in or overall poor performance. Measures must be judged to be important to measure and report in order to be evaluated against the remaining criteria. (evaluation criteria)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a. High Impact</td>
</tr>
</tbody>
</table>

(for NQF staff use) Specific NPP goal:

<table>
<thead>
<tr>
<th>1a.1 Demonstrated High Impact Aspect of Healthcare: affects large numbers, frequently performed procedure, high resource use</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a.2</td>
</tr>
</tbody>
</table>

1a.3 Summary of Evidence of High Impact: Colorectal cancer is the 2nd leading cause of cancer death in the United States. Inappropriate interval recommendations can result in overuse of resources and can lead to significant patient harm. Performing colonoscopy too often not only increases patients’ exposure to procedural harm, but also drains resources that could be more effectively used to adequately screen those in need (Lieberman et al, 2009).


1b. Opportunity for Improvement

1b.1 Benefits (improvements in quality) envisioned by use of this measure: In the average-risk population, colonoscopy screening is recommended in all current guidelines at 10-year intervals. In recent years, colonoscopy has increasingly become a widely used procedure, often resulting in repeat screening.

Rating: C=Completely; P=Partially; M=Minimally; N=Not at all; NA=Not applicable
colonoscopies at intervals not consistent with clinical practice guidelines for patients who are not at risk. Inappropriate interval recommendations can result in overuse of resources and can lead to significant patient harm. Performing colonoscopy too often not only increases patients’ exposure to procedural harm, but also drains resources that could be more effectively used to adequately screen those in need (Lieberman et al., 2008).

1b.2 Summary of data demonstrating performance gap (variation or overall poor performance) across providers:
A recent community based multi-organ cancer screening study in 3627 patients noted that 49 % of low risk patients with adequate negative colonoscopic examinations underwent follow-up surveillance procedures within 7 years (median 3.1 yrs) of their first study, and 35% of low risk patients with two negative exams underwent a third study at a median of 3.3 years after the prior study, despite guidelines for repeat examination at 10 years (Schoen, 2010). Variations in the recommended time interval between colonoscopies also exist for patients with normal colonoscopy findings. In a 2006 study of 1282 colonoscopy reports, recommendations were consistent with current guidelines in only 36.7% of cases. (Krist et al., 2007).

1b.3 Citations for data on performance gap: 

1b.4 Summary of Data on disparities by population group:
We are not aware of any publications/evidence outlining disparities in this area.

1b.5 Citations for data on Disparities:
n/a

1c. Outcome or Evidence to Support Measure Focus

1c.1 Relationship to Outcomes (For non-outcome measures, briefly describe the relationship to desired outcome. For outcomes, describe why it is relevant to the target population): Inappropriate interval recommendations can result in overuse of resources and can lead to significant patient harm. The use of this measure will increase physician adherence to clinical practice guidelines; as such, reduce the patients’unnecessary risk/harm, decreasing the overall costs to patients and insurers, and encourage clinicians to screen only those patients who are at the greatest risk.

1c.2-3. Type of Evidence: evidence based guideline

1c.4 Summary of Evidence (as described in the criteria; for outcomes, summarize any evidence that healthcare services/care processes influence the outcome): Colonoscopy is recommended approximately every 10 years for average-risk individuals (Davila et al., 2006). The completeness of the examination and the quality of the preparation should be taken into account for the timing of subsequent examinations. After a good-quality colonoscopy examination without findings of colon cancer or adenomatous polyps is performed, further screening tests (eg, FOBT) should not be done for approximately 10 years (Davila et al., 2006).

1c.5 Rating of strength/quality of evidence (also provide narrative description of the rating and by whom):
1a-Randomized trials without important limitations
1c.6 Method for rating evidence: ASGE EVIDENCE CLASSIFICATIONS / RATING SCHEMES
Grade of Recommendation
1A. Randomized trials without important limitations
Strong recommendation: can be applied to most clinical settings

Rating: C=Completely; P=Partially; M=Minimally; N=Not at all; NA=Not applicable
1c.7 Summary of Controversy/Contradictory Evidence: n/a


1c.9 Quote the Specific guideline recommendation (including guideline number and/or page number): Colonoscopy is recommended approximately every 10 years for average-risk individuals (Davila et al, 2006). The completeness of the examination and the quality of the preparation should be taken into account for the timing of subsequent examinations. After a good-quality colonoscopy examination without findings of colon cancer or adenomatous polyps is performed, further screening tests (eg, FOBT) should not be done for approximately 10 years (Davila et al, 2006).

1c.10 Clinical Practice Guideline Citation: Davila, R, Rajan, E, Baron, T. American Society for Gastrointestinal Endoscopy. ASGE guideline: colorectal cancer screening and surveillance. Vol. 63. No.4; 2006.


1c.12 Rating of strength of recommendation (also provide narrative description of the rating and by whom):

**Strong Recommendation**

1c.13 Method for rating strength of recommendation (if different from USPSTF system, also describe rating and how it relates to USPSTF):

1A Clarity of Benefit: Clear
Methodologic strength/supporting evidence: Randomized trials without important limitations
Implications: Strong recommendation; can be applied to most clinical settings

1B Clarity of Benefit: Clear
Methodologic strength/supporting evidence: Randomized trials with important limitations (inconsistent results, nonfatal methodologic flaws)
Implications: Strong recommendation; likely to apply to most practice settings

1C Clarity of Benefit: Clear
Methodologic strength/supporting evidence: Overwhelming evidence from observational studies
Implications: Strong recommendation; can apply to most practice settings in most situations

1D Clarity of Benefit: Clear
Methodologic strength/supporting evidence: Observational studies
Implications: Intermediate-strength recommendation; may change when stronger evidence is available

2A Clarity of Benefit: Unclear
Methodologic strength/supporting evidence: Randomized trials without important limitations
Implications: Intermediate-strength recommendation; best action may differ depending on circumstances or patients' societal values

2B Clarity of Benefit: Unclear
Methodologic strength/supporting evidence: Randomized trials with important limitations (inconsistent results, nonfatal methodologic flaws)
Implications: Weak recommendation; alternative approaches may be better under some circumstances

2C Clarity of Benefit: Unclear
Methodologic strength/supporting evidence: Observational studies
Implications: Very weak recommendation; alternative approaches likely to be better under some circumstances

Comment [k7]: USPSTF grading system http://www.ahrq.gov/clinic/uspstf/grades.htm:
A - The USPSTF recommends the service. There is high certainty that the net benefit is substantial.
B - The USPSTF recommends the service. There is high certainty that the net benefit is moderate or there is moderate certainty that the net benefit is moderate to substantial.
C - The USPSTF recommends against routinely providing the service. There may be considerations that support providing the service in an individual patient. There is at least moderate certainty that the net benefit is small. Offer or provide this service only if other considerations support the offering or providing the service in an individual patient.
D - The USPSTF recommends against the service. There is moderate or high certainty that the service has no net benefit or that the harms outweigh the benefits.
I - The USPSTF concludes that the current evidence is insufficient to assess the balance of benefits and harms of the service. Evidence is lacking, of poor quality, or conflicting, and the balance of benefits and harms cannot be determined.
### 1c.14 Rationale for using this guideline over others:

It is the PCPI policy to use guidelines, which are evidence-based, applicable to physicians and other healthcare providers, and developed by a national specialty organization or government agency. In addition, the PCPI has now expanded what is acceptable as the evidence base for measures to include documented quality improvement (QI) initiatives or implementation projects that have demonstrated improvement in the quality of care.

### TAP/Workgroup: What are the strengths and weaknesses in relation to the sub-criteria for Importance to Measure and Report?

<table>
<thead>
<tr>
<th>Sub-criteria</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Importance to Measure and Report</td>
<td>1</td>
</tr>
</tbody>
</table>

### Steering Committee: Was the threshold criterion, Importance to Measure and Report, met?

**Rationale:**

**Y**

### 2. SCIENTIFIC ACCEPTABILITY OF MEASURE PROPERTIES

**Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. (evaluation criteria)**

### 2a. MEASURE SPECIFICATIONS

#### S.1 Do you have a web page where current detailed measure specifications can be obtained?

**2a. Precisely Specified**

<table>
<thead>
<tr>
<th>Numerator Statement (Brief, text description of the numerator - what is being measured about the target population, e.g. target condition, event, or outcome):</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients who had a recommended follow-up interval of at least 10 years for repeat colonoscopy documented in their colonoscopy report</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Numerator Time Window (The time period in which cases are eligible for inclusion in the numerator):</th>
</tr>
</thead>
<tbody>
<tr>
<td>Every procedure within the denominator time window</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Numerator Details (All information required to collect/calculate the numerator, including all codes, logic, and definitions):</th>
</tr>
</thead>
<tbody>
<tr>
<td>EHR Specifications for this measure are under development</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Denominator Statement (Brief, text description of the denominator - target population being measured):</th>
</tr>
</thead>
<tbody>
<tr>
<td>All patients aged 50 years and older receiving screening colonoscopy without biopsy or polypectomy</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Denominator population gender:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female, Male</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Denominator population age range:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients aged 50 years and older</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Denominator Time Window (The time period in which cases are eligible for inclusion in the denominator):</th>
</tr>
</thead>
<tbody>
<tr>
<td>Every procedure within the denominator time window</td>
</tr>
</tbody>
</table>

**Comment [KP8]:** 2a. The measure is well defined and precisely specified so that it can be implemented consistently within and across organizations and allow for comparability. The required data elements are of high quality as defined by NQF's Health Information Technology Expert Panel (HITEP).
2a.8 Denominator Details (All information required to collect/calculate the denominator - the target population being measured - including all codes, logic, and definitions):
EHR Specifications for this measure are under development

Claims Specifications:
Screening Colonoscopy Procedures
CPT code: 45378
OR
G-Code: G0121

CPT code with a modifier of -52, -53, -73 or -74 will not be included in the denominator of this measure

2a.9 Denominator Exclusions (Brief text description of exclusions from the target population):
Documentation of medical reason(s) for not recommending at least a 10 year follow-up interval (eg, above average risk patient, inadequate prep)

2a.10 Denominator Exclusion Details (All information required to collect exclusions to the denominator, including all codes, logic, and definitions):
Claims Specifications:
Append modifier to CPT Category II code: 0528F-1P

EHR Specifications for the exclusions for this measure are under development

2a.11 Stratification Details/Variables (All information required to stratify the measure including the stratification variables, all codes, logic, and definitions):
Stratification by insurance coverage (Commercial, Medicare and Medicaid) is recommended by some implementers

2a.12-13 Risk Adjustment Type: no risk adjustment necessary

2a.14 Risk Adjustment Methodology/Variables (List risk adjustment variables and describe conceptual models, statistical models, or other aspects of model or method):
N/A


2a.18-19 Type of Score: rate/proportion
2a.20 Interpretation of Score: better quality = higher score
2a.21 Calculation Algorithm (Describe the calculation of the measure as a flowchart or series of steps): See sample calculation algorithm attached

2a.22 Describe the method for discriminating performance (e.g., significance testing):
N/A

2a.23 Sampling (Survey) Methodology If measure is based on a sample (or survey), provide instructions for obtaining the sample, conducting the survey and guidance on minimum sample size (response rate):
N/A

2a.24 Data Source (Check the source(s) for which the measure is specified and tested)
Electronic administrative data/claims, Electronic clinical data, electronic Health/Medical Record, paper medical record/flowsheet, special or unique data

2a.25 Data source/data collection instrument (identify the specific data source/data collection instrument, e.g. name of database, clinical registry, collection instrument, etc.):

Comment [k9]: 11 Risk factors that influence outcomes should not be specified as exclusions.
12 Patient preference is not a clinical exception to eligibility and can be influenced by provider interventions.
2a.26-28 Data source/data collection instrument reference web page URL or attachment:

2a.29-31 Data dictionary/code table web page URL or attachment:

2a.32-35 Level of Measurement/Analysis (Check the level(s) for which the measure is specified and tested)
Clinicians: Individual, Clinicians: Group

2a.36-37 Care Settings (Check the setting(s) for which the measure is specified and tested)
Ambulatory Care: Ambulatory Surgery Center, Ambulatory Care: Office, Ambulatory Care: Clinic, Ambulatory Care: Hospital Outpatient

2a.38-41 Clinical Services (Healthcare services being measured, check all that apply)
Clinicians: Physicians (MD/DO), Clinicians: PA/NP/Advanced Practice Nurse

2b. Reliability testing

2b.1 Data/sample (description of data/sample and size):

2b.2 Analytic Method (type of reliability & rationale, method for testing):

2b.3 Testing Results (reliability statistics, assessment of adequacy in the context of norms for the test conducted):

2c. Validity testing

2c.1 Data/sample (description of data/sample and size):

2c.2 Analytic Method (type of validity & rationale, method for testing):

2c.3 Testing Results (statistical results, assessment of adequacy in the context of norms for the test conducted):

2d. Exclusions Justified

2d.1 Summary of Evidence supporting exclusion(s):
The PCPI supports the consideration of exceptions (or exclusions) on a measure by measure basis. There must be a clear rationale to permit an exception for a medical, patient, or system reason, based on whether or not that reason is significant and occurs frequently enough. The PCPI also advocates for the systematic review and analysis of each physician’s exceptions data to identify practice patterns and opportunities for quality improvement.

The decision to include an exception in an AHA PCPI measure is at the discretion of the individual Work Group. Work Groups are charged with deciding if there is a justifiable reason for excluding a patient from a measure, based on whether or not that reason is significant and occurs frequently enough. Denominator exceptions are included in this measure when physicians identify a patient for which a 10 year interval is not appropriate. For example, above average risk patients or patients who had an inadequate bowel prep may have recommendations other than a 10 year timeframe.

2d.2 Citations for Evidence:

Rating: C=Completely; P=Partially; M=Minimally; N=Not at all; NA=Not applicable
2d.3 Data/sample (description of data/sample and size):

2d.4 Analytic Method (type analysis & rationale):

2d.5 Testing Results (e.g., frequency, variability, sensitivity analyses):

2e. Risk Adjustment for Outcomes/Resource Use Measures

2e.1 Data/sample (description of data/sample and size):

2e.2 Analytic Method (type of risk adjustment, analysis, & rationale):

2e.3 Testing Results (risk model performance metrics):

2e.4 If outcome or resource use measure is not risk adjusted, provide rationale:

2f. Identification of Meaningful Differences in Performance

2f.1 Data/sample from Testing or Current Use (description of data/sample and size):

2f.2 Methods to identify statistically significant and practically/meaningfully differences in performance (type of analysis & rationale):

2f.3 Provide Measure Scores from Testing or Current Use (description of scores, e.g., distribution by quartile, mean, median, SD, etc.; identification of statistically significant and meaningfully differences in performance):

2g. Comparability of Multiple Data Sources/Methods

2g.1 Data/sample (description of data/sample and size):

2g.2 Analytic Method (type of analysis & rationale):

2g.3 Testing Results (e.g., correlation statistics, comparison of rankings):

2h. Disparities in Care

2h.1 If measure is stratified, provide stratified results (scores by stratified categories/cohorts): We are not aware of any existing research to indicate whether or not disparities in care exist regarding the implementation of this measure.

2h.2 If disparities have been reported/identified, but measure is not specified to detect disparities, provide follow-up plans: The PCPI and NCQA are currently developing a framework for stratifying measures to test for disparities.

TAP/Workgroup: What are the strengths and weaknesses in relation to the sub-criteria for Scientific Acceptability of Measure Properties?

Steering Committee: Overall, to what extent was the criterion, Scientific Acceptability of Measure Properties, met?

Rationale:

Rating: C=Completely; P=Partially; M=Minimally; N=Not at all; NA=Not applicable
### 3. USABILITY

**Rating:** "C=Completely; P=Partially; M=Minimally; N=Not at all; NA=Not applicable"

<table>
<thead>
<tr>
<th>3a. Meaningful, Understandable, and Useful Information</th>
<th>Eval</th>
</tr>
</thead>
<tbody>
<tr>
<td>3a.1 Current Use: testing not yet completed</td>
<td>Rating</td>
</tr>
<tr>
<td>3a.2 Use in a public reporting initiative (disclosure of performance results to the public at large) (If used in a public reporting initiative, provide name of initiative(s), locations, Web page URL(s). If not publicly reported, state the plans to achieve public reporting within 3 years):</td>
<td></td>
</tr>
<tr>
<td>3a.3 If used in other programs/initiatives (If used in quality improvement or other programs/initiatives, name of initiative(s), locations, Web page URL(s). If not used for QI, state the plans to achieve use for QI within 3 years):</td>
<td></td>
</tr>
<tr>
<td>Testing of Interpretability (Testing that demonstrates the results are understood by the potential users for public reporting and quality improvement)</td>
<td></td>
</tr>
<tr>
<td>3a.4 Data/sample (description of data/sample and size):</td>
<td></td>
</tr>
<tr>
<td>3a.5 Methods (e.g., focus group, survey, QI project):</td>
<td></td>
</tr>
<tr>
<td>3a.6 Results (qualitative and/or quantitative results and conclusions):</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>3b/3c. Relation to other NQF-endorsed measures</th>
<th>Eval</th>
</tr>
</thead>
<tbody>
<tr>
<td>3b.1 NQF # and Title of similar or related measures:</td>
<td>Rating</td>
</tr>
<tr>
<td>(for NQF staff use) Notes on similar/related endorsed or submitted measures:</td>
<td></td>
</tr>
<tr>
<td>3b. Harmonization If this measure is related to measure(s) already endorsed by NQF (e.g., same topic, but different target population/setting/data source or different topic but same target population):</td>
<td></td>
</tr>
<tr>
<td>3b.2 Are the measure specifications harmonized? If not, why?</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>3c. Distinctive or Additive Value</th>
<th>Eval</th>
</tr>
</thead>
<tbody>
<tr>
<td>3c.1 Describe the distinctive, improved, or additive value this measure provides to existing NQF-endorsed measures:</td>
<td>Rating</td>
</tr>
<tr>
<td>5.1 Competing Measures If this measure is similar to measure(s) already endorsed by NQF (i.e., on the same topic and the same target population), describe why it is a more valid or efficient way to measure quality:</td>
<td></td>
</tr>
</tbody>
</table>

**TAP/Workgroup:** What are the strengths and weaknesses in relation to the sub-criteria for Usability?

**Steering Committee:** Overall, to what extent was the criterion, Usability, met?  
**Rationale:**

**Comment [KP22]:** 3a. Demonstration that information produced by the measure is meaningful, understandable, and useful to the intended audience(s) for both public reporting (e.g., focus group, cognitive testing) and informing quality improvement (e.g., quality improvement initiatives). An important outcome that may not have an identified improvement strategy still can be useful for informing quality improvement by identifying the need for and stimulating new approaches to improvement.

**Comment [KP23]:** 3b. The measure specifications are harmonized with other measures, and are applicable to multiple levels and settings.

**Comment [k24]:** Measure harmonization refers to the standardization of specifications for similar measures on the same topic (e.g., influenza immunization of patients in hospitals or nursing homes), or related measures for the same target population (e.g., eye exam and HbA1c for patients with diabetes), or definitions applicable to many measures (e.g., age designation for children) so that they are uniform or compatible, unless differences are dictated by the evidence. The dimensions of harmonization can include numerator, denominator, exclusions, and data source and collection instructions. The extent of harmonization depends on the relationship of the measures, the evidence for the specific measure focus, and differences in data sources.

**Comment [KP25]:** 3c. Review of existing endorsed measures and measure sets demonstrates that the measure provides a distinctive or additive value to existing NQF-endorsed measures (e.g., provides a more complete picture of quality for a particular condition or aspect of healthcare).

**Comment [k26]:** 5. Demonstration that the measure is superior to competing measures - new submissions and/or endorsed measures (e.g., is a more valid or efficient way to measure).
## 4. FEASIBILITY

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

<table>
<thead>
<tr>
<th>Eval Rating</th>
<th>4a. Data Generated as a Byproduct of Care Processes</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>How are the data elements that are needed to compute measure scores generated? (coding/abstraction performed by someone other than person obtaining original information)</td>
</tr>
</tbody>
</table>

4a.1-2

<table>
<thead>
<tr>
<th>Eval Rating</th>
<th>4a.1-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td></td>
</tr>
</tbody>
</table>

4b. Electronic Sources

4b.1 Are all the data elements available electronically? (elements that are needed to compute measure scores are in defined, computer-readable fields, e.g., electronic health record, electronic claims)

<table>
<thead>
<tr>
<th>Eval Rating</th>
<th>4b.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td></td>
</tr>
</tbody>
</table>

4b.2 If not, specify the near-term path to achieve electronic capture by most providers.

Electronic health record products are not uniform in ability to collect data in a standardized way at this time. Design decisions made by individual practices during the implementation of these measures can affect measure performance.

<table>
<thead>
<tr>
<th>Eval Rating</th>
<th>4b.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td></td>
</tr>
</tbody>
</table>

4c. Exclusions

4c.1 Do the specified exclusions require additional data sources beyond what is required for the numerator and denominator specifications?

No

<table>
<thead>
<tr>
<th>Eval Rating</th>
<th>4c.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td></td>
</tr>
</tbody>
</table>

4c.2 If yes, provide justification.

<table>
<thead>
<tr>
<th>Eval Rating</th>
<th>4c.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td></td>
</tr>
</tbody>
</table>

4d. Susceptibility to Inaccuracies, Errors, or Unintended Consequences

4d.1 Identify susceptibility to inaccuracies, errors, or unintended consequences of the measure and describe how these potential problems could be audited. If audited, provide results.

<table>
<thead>
<tr>
<th>Eval Rating</th>
<th>4d.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td></td>
</tr>
</tbody>
</table>

4e. Data Collection Strategy/Implementation

4e.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/frequency of data collection, patient confidentiality, time/cost of data collection, other feasibility/ implementation issues:

<table>
<thead>
<tr>
<th>Eval Rating</th>
<th>4e.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td></td>
</tr>
</tbody>
</table>

4e.2 Costs to implement the measure (costs of data collection, fees associated with proprietary measures):

<table>
<thead>
<tr>
<th>Eval Rating</th>
<th>4e.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td></td>
</tr>
</tbody>
</table>

4e.3 Evidence for costs:

<table>
<thead>
<tr>
<th>Eval Rating</th>
<th>4e.3</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td></td>
</tr>
</tbody>
</table>

4e.4 Business case documentation:

<table>
<thead>
<tr>
<th>Eval Rating</th>
<th>4e.4</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td></td>
</tr>
</tbody>
</table>

TAP/Workgroup: What are the strengths and weaknesses in relation to the sub-criteria for Feasibility?

<table>
<thead>
<tr>
<th>Eval Rating</th>
<th>TAP/Workgroup</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td></td>
</tr>
</tbody>
</table>

Steering Committee: Overall, to what extent was the criterion, Feasibility, met?

<table>
<thead>
<tr>
<th>Eval Rating</th>
<th>Steering Committee</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td></td>
</tr>
</tbody>
</table>

Rationale:

<table>
<thead>
<tr>
<th>Eval Rating</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td></td>
</tr>
</tbody>
</table>

Comment [KP27]: 4a. For clinical measures, required data elements are routinely generated concurrent with and as a byproduct of care processes during care delivery. (e.g., BP recorded in the electronic record, not abstracted from the record later by other personnel; patient self-assessment tools, e.g., depression scale; lab values, meds, etc.)

Comment [KP28]: 4b. The required data elements are available in electronic sources. If the required data are not in existing electronic sources, a credible, near-term path to electronic collection by most providers is specified and clinical data elements are specified for transition to the electronic health record.

Comment [KP29]: 4c. Exclusions should not require additional data sources beyond what is required for scoring the measure (e.g., numerator and denominator) unless justified as supporting measure validity.

Comment [KP30]: 4d. Susceptibility to inaccuracies, errors, or unintended consequences and the ability to audit the data items to detect such problems are identified.

Comment [KP31]: 4e. Demonstration that the data collection strategy (e.g., source, timing, frequency, sampling, patient confidentiality, etc.) can be implemented (e.g., already in operational use, or testing demonstrates that it is ready to put into operational use).
RECOMMENDATION

(for NQF staff use) Check if measure is untested and only eligible for time-limited endorsement.

Steering Committee: Do you recommend for endorsement?
Comments: Y N A

CONTACT INFORMATION

Co.1 Measure Steward (Intellectual Property Owner)
Co.1 Organization
American Medical Association | 515 N State St. | Chicago | Illinois | 60654

Co.2 Point of Contact
Mark Antman, DDS, MBA | mark.antman@ama-assn.org | 312-464-5056

Measure Developer if different from Measure Steward
Co.3 Organization
American Medical Association | 515 N State St. | Chicago | Illinois | 60654

Co.4 Point of Contact
Mark Antman, DDS, MBA | mark.antman@ama-assn.org | 312-464-5056

Co.5 Submitter if different from Measure Steward POC
Mark Antman, DDS, MBA | mark.antman@ama-assn.org | 312-464-5056 | American Medical Association

Co.6 Additional organizations that sponsored/participated in measure development
American Society for Gastrointestinal Endoscopy (ASGE)
American Gastroenterological Association (AGA)
National Committee for Quality Assurance

ADDITIONAL INFORMATION

Workgroup/Expert Panel involved in measure development
Ad.1 Provide a list of sponsoring organizations and workgroup/panel members’ names and organizations.
Describe the members’ role in measure development.
PCPI measures are developed through cross-specialty, multi-disciplinary work groups. All medical specialties and
other health care professional disciplines participating in patient care for the clinical condition or topic under
study must be equal contributors to the measure development process. In addition, the PCPI strives to include on
its work groups individuals representing the perspectives of patients, consumers, private health plans, and
employers. This broad-based approach to measure development ensures buy-in on the measures from all
stakeholders and minimizes bias toward any individual specialty or stakeholder group. All work groups have at
least two co-chairs who have relevant clinical and/or measure development expertise and who are responsible for
ensuring that consensus is achieved and that all perspectives are voiced.

Co-chairs
John Allen, MD, MBA, AGAF (Gastroenterology)
Doug Faigel, MD (Gastroenterology)

Work Group Members
Nancy Baxter, MD, PhD, FACS, FACS (Colon and Rectal Surgery)
Stephen Bickston, MD, AGAF (Gastroenterology)
Joel V. Brill, MD, AGAF, FASGE, FACC, CHCQM (Gastroenterology)
Kirk Brandon, MBA (Business Administration/Coding)
Jason A. Dominitz, MD, MHS, AGAF (Gastroenterology)
Ira L. Fiax, MD, FACC (Gastroenterology)
Karen E. Hall, MD, PhD (Geriatrics)
Robert Haskey, MD, FACS (General Surgery, Health Plan representative)
Brian C. Jacobson, MD, MPH (Gastroenterology)
<table>
<thead>
<tr>
<th>Name</th>
<th>Specialty</th>
</tr>
</thead>
<tbody>
<tr>
<td>David Lieberman, MD</td>
<td>Gastroenterology</td>
</tr>
<tr>
<td>Klaus Mergener, MD, PhD, FACP, FACG, FASGE, FACPE</td>
<td>Gastroenterology</td>
</tr>
<tr>
<td>Bret Petersen, MD, FASGE</td>
<td>Gastroenterology</td>
</tr>
<tr>
<td>Irving M. Pike, MD, FACG</td>
<td>Gastroenterology</td>
</tr>
<tr>
<td>Bart Pope, MD (Family Medicine)</td>
<td></td>
</tr>
<tr>
<td>Harry Sarles, MD, FACG</td>
<td>Gastroenterology</td>
</tr>
<tr>
<td>Kay Schwebke, MD, MPH (Specialty: Internal Medicine, Infectious</td>
<td>Diseases &amp; Medical Informatics)</td>
</tr>
<tr>
<td>Tom Lynn, MD (Medical Informatics, Methodology)</td>
<td></td>
</tr>
<tr>
<td>Emily E. Voik, MD, FCAP</td>
<td>Pathology</td>
</tr>
<tr>
<td>Michael Weinstein, MD (Specialty: Gastroenterology)</td>
<td></td>
</tr>
<tr>
<td>Debbie Robin, MSN, RN, CHCQM</td>
<td></td>
</tr>
<tr>
<td>American Gastroenterological Association</td>
<td></td>
</tr>
<tr>
<td>Jill Blim</td>
<td></td>
</tr>
<tr>
<td>Chris Recker, RN, MPH</td>
<td></td>
</tr>
<tr>
<td>Martha Espronceda</td>
<td></td>
</tr>
<tr>
<td>American Society for Gastrointestinal Endoscopy</td>
<td></td>
</tr>
<tr>
<td>Julie Cantor-Weinberg, MPP</td>
<td></td>
</tr>
<tr>
<td>American College of Gastroenterology</td>
<td></td>
</tr>
<tr>
<td>Joseph Gave, MPH</td>
<td></td>
</tr>
<tr>
<td>Karen Kmetik, PhD</td>
<td></td>
</tr>
<tr>
<td>Shannon Sims, MD, PhD</td>
<td></td>
</tr>
<tr>
<td>Beth Tapper, MA</td>
<td></td>
</tr>
<tr>
<td>Consortium Consultants</td>
<td></td>
</tr>
<tr>
<td>Rebecca Kresowik</td>
<td></td>
</tr>
<tr>
<td>Timothy Kresowik</td>
<td></td>
</tr>
<tr>
<td><strong>Ad.2 If adapted, provide name of original measure:</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Ad.3-5 If adapted, provide original specifications URL or attachment:</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Measure Developer/Steward Updates and Ongoing Maintenance</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Ad.6 Year the measure was first released:</strong></td>
<td>2008</td>
</tr>
<tr>
<td><strong>Ad.7 Month and Year of most recent revision:</strong></td>
<td>2008-08</td>
</tr>
<tr>
<td><strong>Ad.8 What is your frequency for review/update of this measure?</strong></td>
<td>Every 3-4 years or as new evidence becomes</td>
</tr>
<tr>
<td>available that materially affects the measures</td>
<td></td>
</tr>
<tr>
<td><strong>Ad.9 When is the next scheduled review/update for this measure?</strong></td>
<td>2011-08</td>
</tr>
<tr>
<td><strong>Ad.10 Copyright statement/disclaimers:</strong></td>
<td>Physician Performance Measures (Measures) and related data specifications developed by the American Medical Association (AMA) in collaboration with the Physician Consortium for Performance Improvement ® (PCPI) and the National Committee for Quality Assurance (NCQA), pursuant to government sponsorship under Subcontract No. 6414-07-089 with Mathematica Policy Research under Contract HHSN-500-2005-000251(0004) with Centers for Medicare and Medicaid Services.</td>
</tr>
<tr>
<td>These performance Measures are not clinical guidelines and do not establish a standard of medical care, and have not been tested for all potential applications. The Measures, while copyrighted, can be reproduced and distributed, without modification, for noncommercial purposes, e.g., use by health care providers in connection with their practices. Commercial use is defined as the sale, license, or distribution of the Measures for commercial gain, or incorporation of the Measures into a product or service that is sold, licensed or distributed for commercial gain. Commercial uses of the Measures require a license agreement between the user and the AMA, on behalf of the PCPI or NCQA. Neither the AMA, NCQA, PCPI nor its members shall be responsible for any use of the Measures.</td>
<td></td>
</tr>
<tr>
<td><strong>THE MEASURES AND SPECIFICATIONS ARE PROVIDED “AS IS” WITHOUT WARRANTY OF ANY KIND.</strong></td>
<td></td>
</tr>
<tr>
<td>© 2008 American Medical Association and National Committee for Quality Assurance. All Rights Reserved.</td>
<td></td>
</tr>
</tbody>
</table>
Limited proprietary coding is contained in the Measure specifications for convenience. Users of the proprietary code sets should obtain all necessary licenses from the owners of these code sets. The AMA, NCQA, the PCPI and its members disclaim all liability for use or accuracy of any Current Procedural Terminology (CPT®) or other coding contained in the specifications.

CPT® contained in the Measures specifications is copyright 2007 American Medical Association. LOINC® copyright 2004 Regenstrief Institute, Inc. SNOMED CLINICAL TERMS (SNOMED CT®) copyright 2004 College of American Pathologists (CAP). All Rights Reserved. Use of SNOMED CT® is only authorized within the United States.

<table>
<thead>
<tr>
<th>Ad.11 -13 Additional Information web page URL or attachment:</th>
<th>Attachment Sample Calculation Algorithm-634008794025999947.doc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date of Submission (MM/DD/YY):</td>
<td>02/17/2010</td>
</tr>
</tbody>
</table>
1c. The measure focus is:

• an outcome (e.g., morbidity, mortality, function, health-related quality of life) that is relevant to, or associated with, a national health goal/priority, the condition, population, and/or care being addressed; OR

• if an intermediate outcome, process, structure, etc., there is evidence that supports the specific measure focus as follows:
  - Intermediate outcome - evidence that the measured intermediate outcome (e.g., blood pressure, Hba1c) leads to improved health/avoidance of harm or cost/benefit.
  - Process - evidence that the measured clinical or administrative process leads to improved health/avoidance of harm and if the measure focus is on one step in a multi-step care process, it measures the step that has the greatest effect on improving the specified desired outcome(s).
  - Structure - evidence that the measured structure supports the consistent delivery of effective processes or access that lead to improved health/avoidance of harm or cost/benefit.
  - Patient experience - evidence that an association exists between the measure of patient experience of health care and the outcomes, values and preferences of individuals/ the public.
  - Access - evidence that an association exists between access to a health service and the outcomes of, or experience with, care.
  - Efficiency - demonstration of an association between the measured resource use and level of performance with respect to one or more of the other five IOM aims of quality.

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 multi-step process, the step with the greatest effect on the desired outcome should be selected as the focus of measurement. For example, although assessment of immunization status and recommending immunization are necessary steps, they are not sufficient to achieve the desired impact on health status - patients must be vaccinated to achieve immunity. This does not preclude consideration of measures of preventive screening interventions where there is a strong link with desired outcomes (e.g., mammography) or measures for multiple care processes that affect a single outcome.

2d. Clinically necessary measure exclusions are identified and must be:

• supported by evidence of sufficient frequency of occurrence so that results are distorted without the exclusion; AND

• a clinically appropriate exception (e.g., contraindication) to eligibility for the measure focus; AND

• precisely defined and specified:
  - if there is substantial variability in exclusions across providers, the measure is specified so that exclusions are computable and the effect on the measure is transparent (i.e., impact clearly delineated, such as number of cases excluded, exclusion rates by type of exclusion);

if patient preference (e.g., informed decision-making) is a basis for exclusion, there must be evidence that it strongly impacts performance on the measure and 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).
Sample PCPI Calculation Algorithm

Calculation for Performance
For performance purposes, a measure is calculated by creating a fraction with the following components: Numerator, Denominator, and Denominator Exclusions.

Numerator (A) Includes:
Number of patients meeting numerator criteria

Denominator (PD) Includes:
Number of patients meeting criteria for denominator inclusion

Denominator Exclusions (C) Include:
Number of patients with valid medical, patient or system exclusions (where applicable; will differ by measure)

Performance Calculation

\[
\frac{A}{PD - C}
\]

Where:
- \( A \) is the number of patients meeting numerator criteria
- \( PD \) is the number of patients in denominator
- \( C \) is the number of patients with valid denominator exclusions

If a measure does not allow for exclusion(s), it is calculated by creating a fraction with the following components: Numerator and Denominator.

Numerator (A) Includes:
Number of patients meeting numerator criteria

Denominator (PD) Includes:
Number of patients meeting criteria for denominator inclusion

\[
\frac{A}{PD}
\]

It is also possible to calculate the percentage of patients excluded overall, or excluded by medical, patient, or system reason where applicable:

Overall Exclusion Calculation

\[
\frac{C}{PD}
\]

Where:
- \( C \) is the number of patients with any valid exclusion
- \( PD \) is the number of patients in denominator

OR

Exclusion Calculation by Type

\[
\frac{C_1}{PD}, \frac{C_2}{PD}, \frac{C_3}{PD}
\]

Where:
- \( C_1 \) is the number of patients with medical reason
- \( C_2 \) is the number of patients with patient reason
- \( C_3 \) is the number of patients with system reason

PD is the number of patients in denominator.