

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-018-10	NQF Project: Ambulatory Care - Additional Outpatient Measures 2010
MEASURE DESCRIPTIVE INFORMATION	
De.1 Measure Title: Endoscopy/Polyp Surveillance: Comprehensive Colonoscopy Documentation	
De.2 Brief description of measure: Percentage of final colonoscopy reports for patients aged 18 years and older that include documentation of all of the following: pre-procedure risk assessment; depth of insertion; quality of the bowel prep; complete description of polyp(s) found, including location of each polyp, size, number and gross morphology; and recommendations for follow-up	
1.1-2 Type of Measure: process	
De.3 If included in a composite or paired with another measure, please identify composite or paired measure	
De.4 National Priority Partners Priority Area: population health	
De.5 IOM Quality Domain: patient-centered, effectiveness, safety, timeliness	
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:	NQF Staff A Y <input type="checkbox"/> N <input type="checkbox"/>
A. The measure is in the public domain or an intellectual property (measure steward agreement) is signed. <i>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.</i>	
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):	
A.3 Measure Steward Agreement: agreement signed and submitted	
A.4 Measure Steward Agreement attached:	

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	B Y <input type="checkbox"/> N <input type="checkbox"/>
C. The intended use of the measure includes both public reporting and quality improvement. ► Purpose: public reporting, quality improvement Accountability	C Y <input type="checkbox"/> N <input type="checkbox"/>
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	D Y <input type="checkbox"/> N <input type="checkbox"/>
(for NQF staff use) Have all conditions for consideration been met? Staff Notes to Steward (if submission returned):	Met Y <input type="checkbox"/> N <input type="checkbox"/>
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	
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) 1a. High Impact	Eval Rating
(for NQF staff use) Specific NPP goal:	
1a.1 Demonstrated High Impact Aspect of Healthcare: frequently performed procedure, high resource use, patient/societal consequences of poor quality, affects large numbers 1a.2 1a.3 Summary of Evidence of High Impact: Colonoscopy is an importance means of preventing and treating colorectal cancer, the second leading cause of death in the U.S. and a leading cause of death in Medicare beneficiaries. Incomplete colonoscopy reports that are missing the recommended elements have been documented in the published literature. As a result, patients may not be given an accurate diagnosis, which may lead to missed diagnoses or repeat exams in shorter time frames. 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). 1a.4 Citations for Evidence of High Impact: Lieberman, DA, Faigel, DO, Logan, J, Mattek, N, Holub, J, Eisen, G, Morris, C, Smith, R, Nadel, M. Assessment of the Quality of Colonoscopy Reports: Results from a multi-center consortium. Gastrointest Endosc Vol 69, 2009.	1a C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/>
1b. Opportunity for Improvement	1b C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/>
1b.1 Benefits (improvements in quality) envisioned by use of this measure: The goal of this measure is to ensure that a quality colonoscopy is completed including the patient's risk, findings, and recommendations. A fully comprehensive colonoscopy will help to better identify patients at risk and will minimize unnecessary follow-up. The use of this measure will contribute to the overall effectiveness of the	1b C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/>

Comment [KP1]: 1a. The measure focus addresses:
 • a specific national health goal/priority identified by NQF's National Priorities Partners; OR
 • a demonstrated high impact aspect of healthcare (e.g., affects large numbers, leading cause of morbidity/mortality, high resource use (current and/or future), severity of illness, and patient/societal consequences of poor quality).

Comment [KP2]: 1b. Demonstration of quality problems and opportunity for improvement, i.e., data demonstrating considerable variation, or overall poor performance, in the quality of care across providers and/or population groups (disparities in care).

colonoscopy, which influences the timing of repeat examinations, and reduces inappropriate colonoscopies and costs.

1b.2 Summary of data demonstrating performance gap (variation or overall poor performance) across providers:

Several studies have demonstrated significant gaps in the quality of colonoscopies, showing that colonoscopy reports often do not include the recommended components.

1. The Clinical Outcomes Research Initiative (CORI) on Colonoscopy Quality Indicators Study of 53 gastroenterology practice sites in 24 states looked at all patients undergoing colonoscopy (n=438,521); in this study, documentation of risk assessment was measured. The ASA Classification field was not completed in 10.1% of reports. In 10 of 53 sites, completion rates were less than 90%. When completed, 7.0% of exams were performed in high-risk individuals with ASA class 3 or higher (Lieberman et al, 2009).
2. Numerous studies have shown that physicians routinely do not document the depth of insertion in the colonoscopy report. Quality evaluation of the colon consists of intubation of the entire colon and a detailed mucosal inspection. Cecal intubation improves sensitivity and reduces costs by eliminating the need for radiographic procedures or repeat colonoscopy to complete examination. Careful mucosal inspection is essential to effective colorectal cancer prevention and reduction of cancer mortality.
3. The Clinical Outcomes Research Initiative (CORI) on Colonoscopy Quality Indicators Study of 53 gastroenterology practice sites in 24 states looked at all patients undergoing colonoscopy (n=438,521); in this study, quality of bowel prep recorded was assessed. Findings indicated that 13.9% of reports did not have bowel prep quality reported and in 14 of 53 practices, over 20% did not have bowel prep quality (Lieberman et al, 2009).
4. A recent multi-center study looked at variations in practice and assessed the quality of colonoscopy procedures. Findings indicated that polyp size not recorded in 4.9% of polyps, polyp morphology (pedunculated, sessile, flat) was not reported in 14.7% of reported polyps, and polyp retrieval and submission to pathology was not documented in 4.5% of polyps (Lieberman et al, 2009). These gaps in the documentation of the description of the polyps removed during colonoscopy underscore the need to improve physician adherence to quality patient care.
5. Recent evidence suggests that surveillance colonoscopy for post-polypectomy patients in the United States is frequently performed at intervals that are shorter than those recommended in guidelines. In addition, many patient records do not have a recommended follow-up interval recorded. For example, in a 2006 study of 1282 colonoscopy reports, recommendations were consistent with contemporaneous guidelines in only 39.2% of cases and with current guidelines in 36.7% of cases. Correspondence from the endoscopist included no guidance on follow-up testing in 33.5% of cases (Krist et al, 2007).

1b.3 Citations for data on performance gap:

Faigel, DO, Pike, IM, Baron, TH, et al. ASGE/ACG Taskforce on Quality in Endoscopy. Quality Indicators For Gastrointestinal Endoscopic Procedures: An Introduction. Gastrointestinal Endoscopy. Volume 63, No. 4 : 2006.

Krist, AH, Jones, RM, Woolf, SH et al. Timing of Repeat Colonoscopy: Disparity Between Guidelines and Endoscopists' Recommendation. American Journal of Preventive Medicine. 2007.

Lieberman, DA, Faigel, DO, Logan, J, Mattek, N, Holub, J, Eisen, G, Morris, C, Smith, R, Nadel, M. Assessment of the Quality of Colonoscopy Reports: Results from a multi-center consortium. Gastrointest Endosc Vol 69, 2009.

1b.4 Summary of Data on disparities by population group:

At this time, we have not found any published literature/data on disparities by population group

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): The goal of this measure is

1c

C

P

M

Comment [k3]: 1 Examples of data on opportunity for improvement include, but are not limited to: prior studies, epidemiologic data, measure data from pilot testing or implementation. If data are not available, the measure focus is systematically assessed (e.g., expert panel rating) and judged to be a quality problem.

Comment [k4]: 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:
 oIntermediate outcome - evidence that the measured intermediate outcome (e.g., blood pressure, Hba1c) leads to improved health/avoidance of harm or cost/benefit.
 oProcess - 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).
 oStructure - 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.
 oPatient 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.
 oAccess - evidence that an association exists between access to a health service and the outcomes of, or experience with, care.
 oEfficiency - 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.

Comment [k5]: 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.

to ensure appropriate documentation of colonoscopy findings and recommendations, in order to improve physician adherence to quality patient care and decrease overuse of resources. Pre-procedure risk assessment is often used as a surrogate of co-morbidity; research has shown an association between higher class and adverse events. The need for cecal intubation is based on the continual finding that a substantial number of colorectal neoplasms are located in the proximal colon, including the cecum. Poor bowel preparation is a major impediment to the effectiveness of colonoscopy and impacts the ability to detect polyps and influences the timing of repeat examinations. Poor preparation prolongs cecal intubation time and withdrawal time and reduces detection of both small and large polyps (Faigel et al, 2006). The economic burden of repeating examinations because of inadequate bowel preparation is substantial. Accurate polyp descriptions are essential to assess disease progression and inform timing of repeat colonoscopy.

N

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):

Before sedation is begun, a risk assessment is performed to stratify patients into higher or lower-risk-for-complications groups (particularly as pertains to sedation) [Grade 1C] (Faigel et al, 2006). The physician/nurse team should document the risk assessment. (Risk stratification systems commonly used are the ASA score and the Mallampati score). Visualization of the cecum by notation of landmarks and photo documentation of landmarks should be documented in every procedure (Grade 1C). Most important, these include the appendiceal orifice and the ileocecal valve. There should be documentation in the procedure note of the quality of the preparation of the bowel (Grade 2C) (Faigel et al, 2006). In clinical trials of bowel preparation, terms used to commonly characterize bowel preparation include "excellent," "good," "fair," and "poor." In clinical practice, these terms do not have standardized definitions. In clinical trials on the effectiveness of various regimens for bowel preparation, excellent is typically defined as no or minimal solid stool and only small amounts of clear fluid requiring suctioning. "Good" is typically no or minimal solid stool with large amounts of clear fluid requiring suctioning. "Fair" refers to collections of semisolid debris that are cleared with difficulty. "Poor" refers to solid or semisolid debris that cannot be effectively cleared. The endoscopist should be prepared to perform a total examination and remove all polyps found at the time of the first colonoscopy, although technical factors encountered during colonoscopy may limit completion of the procedure (Davila et al, 2006).

1c.5 Rating of strength/quality of evidence (also provide narrative description of the rating and by whom):

1 C, 2C

1c.6 Method for rating evidence: Grade of Recommendation: 1C

Clarity of Benefit: Clear

Methodologic strength/Supporting evidence: Observational studies

Implications: Intermediate-strength recommendation; may change when stronger evidence is available

Grade of Recommendation: 2C

Clarity of Benefit: Unclear

Methodologic strength/Supporting evidence: Observational studies

Implications: Very weak recommendation; alternative approaches likely to be better under some circumstances

1c.7 Summary of Controversy/Contradictory Evidence: n/a

1c.8 Citations for Evidence (other than guidelines): n/a

1c.9 Quote the Specific guideline recommendation (including guideline number and/or page number):

Before sedation is begun, a risk assessment is performed to stratify patients into higher or lower-risk-for-complications groups (particularly as pertains to sedation) [Grade 1C] (Faigel et al, 2006). The physician/nurse team should document the risk assessment. (Risk stratification systems commonly used are the ASA score and the Mallampati score). Visualization of the cecum by notation of landmarks and photo documentation of landmarks should be documented in every procedure (Grade 1C). Most important, these

Comment [k6]: 3 The strength of the body of evidence for the specific measure focus should be systematically assessed and rated (e.g., USPSTF grading system <http://www.ahrq.gov/clinic/uspstf07/methods/benefit.htm>). If the USPSTF grading system was not used, the grading system is explained including how it relates to the USPSTF grades or why it does not. However, evidence is not limited to quantitative studies and the best type of evidence depends upon the question being studied (e.g., randomized controlled trials appropriate for studying drug efficacy are not well suited for complex system changes). When qualitative studies are used, appropriate qualitative research criteria are used to judge the strength of the evidence.

include the appendiceal orifice and the ileocecal valve. There should be documentation in the procedure note of the quality of the preparation of the bowel (Grade 2C) (Faigel et al, 2006). In clinical trials of bowel preparation, terms used to commonly characterize bowel preparation include "excellent," "good," "fair," and "poor." In clinical practice, these terms do not have standardized definitions. In clinical trials on the effectiveness of various regimens for bowel preparation, excellent is typically defined as no or minimal solid stool and only small amounts of clear fluid requiring suctioning. "Good" is typically no or minimal solid stool with large amounts of clear fluid requiring suctioning. "Fair" refers to collections of semisolid debris that are cleared with difficulty. "Poor" refers to solid or semisolid debris that cannot be effectively cleared. The endoscopist should be prepared to perform a total examination and remove all polyps found at the time of the first colonoscopy, although technical factors encountered during colonoscopy may limit completion of the procedure (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.

Faigel, DO, Pike, IM, Baron, TH, et al. ASGE/ACG Taskforce on Quality in Endoscopy. Quality Indicators For Gastrointestinal Endoscopic Procedures: An Introduction. Gastrointestinal Endoscopy. Volume 63, No. 4 : 2006.

1c.11 National Guideline Clearinghouse or other URL:
http://www.guidelines.gov/summary/summary.aspx?doc_id=10162&nbr=5347#s24

1c.12 Rating of strength of recommendation *(also provide narrative description of the rating and by whom):*
 1C, 2C

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

1C 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' or 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

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.

<p>Implications: Very weak recommendation; alternative approaches likely to be better under some circumstances</p> <p>3 Clarity of Benefit: Unclear Methodologic strength/supporting evidence: Expert opinion only Implications: Weak recommendation; likely to change as data become available</p> <p>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 speciality organization or government agency. In addition, the PCPI has now expanded what is acceptable as the evidence base for measures to included documented quality improvement (QI) initiatives or implementation projects that have demonstrated improvement in the quality of care.</p>	
TAP/Workgroup: What are the strengths and weaknesses in relation to the sub-criteria for <i>Importance to Measure and Report</i> ?	1
Steering Committee: Was the threshold criterion, <i>Importance to Measure and Report</i> , met? Rationale:	1 Y <input type="checkbox"/> N <input type="checkbox"/>
2. SCIENTIFIC ACCEPTABILITY OF MEASURE PROPERTIES	
Extent to which the measure, <u>as specified</u> , produces consistent (reliable) and credible (valid) results about the quality of care when implemented. (evaluation criteria)	Eval Rating
2a. MEASURE SPECIFICATIONS	
S.1 Do you have a web page where current detailed measure specifications can be obtained? S.2 If yes, provide web page URL:	
2a. Precisely Specified	
<p>2a.1 Numerator Statement (<i>Brief, text description of the numerator - what is being measured about the target population, e.g. target condition, event, or outcome</i>): Final reports that include documentation of ALL of the following:</p> <ul style="list-style-type: none"> • Pre-procedure risk assessment (eg, ASA class, Mallampati score) • Depth of insertion (i.e., to cecum or other landmark) • Quality of the bowel prep (ie, prep was either adequate or inadequate) • Complete description of polyp(s) found, including location of each polyp, size, number and gross morphology • Recommendations for follow-up <p>2a.2 Numerator Time Window (<i>The time period in which cases are eligible for inclusion in the numerator</i>): Every procedure within the denominator time window</p> <p>2a.3 Numerator Details (<i>All information required to collect/calculate the numerator, including all codes, logic, and definitions</i>): EHR Specifications for this measure are under development</p> <p>Claims Specifications: 3018F- Pre-procedure risk assessment AND depth of insertion AND quality of the bowel prep AND complete description of polyp(s) found, including location of each polyp, size, number and gross morphology AND recommendations for follow-up in final colonoscopy report, documented</p>	<p>2a-specs C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/></p>

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).

<p>2a.4 Denominator Statement (<i>Brief, text description of the denominator - target population being measured</i>): All final colonoscopy reports for patients aged 18 years and older</p>
<p>2a.5 Target population gender: Male, Female</p>
<p>2a.6 Target population age range: Ages 18 and over</p>
<p>2a.7 Denominator Time Window (<i>The time period in which cases are eligible for inclusion in the denominator</i>): 12 month period</p>
<p>2a.8 Denominator Details (<i>All information required to collect/calculate the denominator - the target population being measured - including all codes, logic, and definitions</i>): EHR Specifications for this measure are under development</p> <p>Claims Specifications: CPT Codes or G-Codes: 44388, 44389, 44390, 44391, 44392, 44393, 44394, 44397, 45355, 45378, 45379, 45380, 45381, 45382, 45383, 45384, 45385, 45386, 45387, 45391, 45392, G0105, G0121</p> <p>CPT codes with a modifier of -52, -53, -73 or -74 will not be included in the denominator of this measure</p>
<p>2a.9 Denominator Exclusions (<i>Brief text description of exclusions from the target population</i>): None</p>
<p>2a.10 Denominator Exclusion Details (<i>All information required to collect exclusions to the denominator, including all codes, logic, and definitions</i>): N/A</p>
<p>2a.11 Stratification Details/Variables (<i>All information required to stratify the measure including the stratification variables, all codes, logic, and definitions</i>): Stratification by insurance coverage (Commerical, Medicare and Medicaid) is recommended by implementers</p>
<p>2a.12-13 Risk Adjustment Type: no risk adjustment necessary</p>
<p>2a.14 Risk Adjustment Methodology/Variables (<i>List risk adjustment variables and describe conceptual models, statistical models, or other aspects of model or method</i>): N/A</p>
<p>2a.15-17 Detailed risk model available Web page URL or attachment: URL http://www.ama-assn.org/ama1/pub/upload/mm/370/endoscopy-ms.pdf</p>
<p>2a.18-19 Type of Score: rate/proportion</p>
<p>2a.20 Interpretation of Score: better quality = higher score</p>
<p>2a.21 Calculation Algorithm (<i>Describe the calculation of the measure as a flowchart or series of steps</i>): See sample calculation algorithm attached</p>
<p>2a.22 Describe the method for discriminating performance (<i>e.g., significance testing</i>): N/A</p>
<p>2a.23 Sampling (Survey) Methodology <i>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):</i> N/A</p>
<p>2a.24 Data Source (<i>Check the source(s) for which the measure is specified and tested</i>) paper medical record/flowsheet, Electronic adminstrative data/claims, electronic Health/Medical Record, special or unique data</p>
<p>2a.25 Data source/data collection instrument (<i>Identify the specific data source/data collection instrument, e.g. name of database, clinical registry, collection instrument, etc.</i>):</p>
<p>2a.26-28 Data source/data collection instrument reference web page URL or attachment:</p>

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.

<p>2a.29-31 Data dictionary/code table web page URL or attachment:</p> <p>2a.32-35 Level of Measurement/Analysis (<i>Check the level(s) for which the measure is specified and tested</i>) Clinicians: Individual, Clinicians: Group</p> <p>2a.36-37 Care Settings (<i>Check the setting(s) for which the measure is specified and tested</i>) Ambulatory Care: Ambulatory Surgery Center, Ambulatory Care: Office, Ambulatory Care: Clinic, Hospital, Ambulatory Care: Hospital Outpatient</p> <p>2a.38-41 Clinical Services (<i>Healthcare services being measured, check all that apply</i>) Clinicians: Physicians (MD/DO), Clinicians: PA/NP/Advanced Practice Nurse</p>	
TESTING/ANALYSIS	
<p>2b. Reliability testing</p> <p>2b.1 Data/sample (<i>description of data/sample and size</i>):</p> <p>2b.2 Analytic Method (<i>type of reliability & rationale, method for testing</i>):</p> <p>2b.3 Testing Results (<i>reliability statistics, assessment of adequacy in the context of norms for the test conducted</i>):</p>	<p>2b</p> <p>C <input type="checkbox"/></p> <p>P <input type="checkbox"/></p> <p>M <input type="checkbox"/></p> <p>N <input type="checkbox"/></p>
<p>2c. Validity testing</p> <p>2c.1 Data/sample (<i>description of data/sample and size</i>):</p> <p>2c.2 Analytic Method (<i>type of validity & rationale, method for testing</i>): It is the consensus of the PCPI Measures Implementation and Evaluation Committee that face and content validity of PCPI measures can be assumed to be established once they have progressed beyond the Public Comment period by virtue of the specialized expertise of the PCPI work group members who are involved in identifying and drafting performance measures within a topical domain as well, as the rigorous, structured discussions that are prescribed according to PCPI protocols for work group conduct.</p> <p>2c.3 Testing Results (<i>statistical results, assessment of adequacy in the context of norms for the test conducted</i>):</p>	<p>2c</p> <p>C <input type="checkbox"/></p> <p>P <input type="checkbox"/></p> <p>M <input type="checkbox"/></p> <p>N <input type="checkbox"/></p>
<p>2d. Exclusions Justified</p> <p>2d.1 Summary of Evidence supporting exclusion(s):</p> <p>2d.2 Citations for Evidence:</p> <p>2d.3 Data/sample (<i>description of data/sample and size</i>):</p> <p>2d.4 Analytic Method (<i>type analysis & rationale</i>):</p> <p>2d.5 Testing Results (<i>e.g., frequency, variability, sensitivity analyses</i>):</p>	<p>2d</p> <p>C <input type="checkbox"/></p> <p>P <input type="checkbox"/></p> <p>M <input type="checkbox"/></p> <p>N <input type="checkbox"/></p> <p>NA <input type="checkbox"/></p>
<p>2e. Risk Adjustment for Outcomes/ Resource Use Measures</p>	<p>2e</p> <p>C <input type="checkbox"/></p>

Comment [KP10]: 2b. Reliability testing demonstrates the measure results are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period.

Comment [k11]: 8 Examples of reliability testing 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 may address the data items or final measure score.

Comment [KP12]: 2c. Validity testing demonstrates that the measure reflects the quality of care provided, adequately distinguishing good and poor quality. If face validity is the only validity addressed, it is systematically assessed.

Comment [k13]: 9 Examples of validity testing include, but are not limited to: determining if measure scores adequately distinguish between providers known to have good or poor quality assessed by another valid method; correlation of measure scores with another valid indicator of quality for the specific topic; ability of measure scores to predict scores on some other related valid measure; content validity for multi-item scales/tests. Face validity is a subjective assessment by experts of whether the measure reflects the quality of care (e.g., whether the proportion of patients with BP < 140/90 is a marker of quality). If face validity is the only validity addressed, it is systematically assessed (e.g., ratings by relevant stakeholders) and the measure is judged to represent quality care for the specific topic and that the measure focus is the most important aspect of quality for the specific topic.

Comment [KP14]: 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 ca... [1])

Comment [k15]: 10 Examples of evidence that an exclusion distorts measure results include, but are not limited to: frequency of occurrence, sensitivity analyses with and without the exclusion, and variability of exclusions across providers.

Comment [KP16]: 2e. For outcome measures and other measures (e.g., resource use) when indicated:
 •an evidence-based risk-adjustment strategy (e.g., risk models, risk stratification) is specified and is based on patient clinical factors that influence the measured outcome (but not disparities in care) and are present at start of care; Error! Bookmark not defined. OR ... [2]

2e.1 Data/sample (description of data/sample and size):	P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/> NA <input type="checkbox"/>
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):	2f C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/>
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):	2g C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/> NA <input type="checkbox"/>
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.	2h C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/> NA <input type="checkbox"/>
TAP/Workgroup: What are the strengths and weaknesses in relation to the sub-criteria for Scientific Acceptability of Measure Properties?	2
Steering Committee: Overall, to what extent was the criterion, Scientific Acceptability of Measure Properties, met? Rationale:	2 C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/>
3. USABILITY	
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. (evaluation criteria)	Eval Rating
3a. Meaningful, Understandable, and Useful Information	
3a.1 Current Use: testing not yet completed	3a C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/>
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	

Comment [k17]: 13 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, gender (e.g., poorer treatment outcomes of African American men with prostate cancer, inequalities in treatment for CVD risk factors between men and women). It is preferable to stratify measures by race and socioeconomic status rather than adjusting out differences.

Comment [KP18]: 2f. Data analysis demonstrates that methods for scoring and analysis of the specified measure allow for identification of statistically significant and practically/clinically meaningful differences in performance.

Comment [k19]: 14 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% v. 75%) 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 poor performance may not demonstrate much variability across providers.

Comment [KP20]: 2g. If multiple data sources/methods are allowed, there is demonstration they produce comparable results.

Comment [KP21]: 2h. If disparities in care have been identified, measure specifications, scoring, and analysis allow for identification of disparities through stratification of results (e.g., by race, ethnicity, socioeconomic status, gender); OR rationale/data justifies why stratification is not necessary or not feasible.

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.

<u>publicly reported</u> , state the plans to achieve public reporting within 3 years):	
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):	
Testing of Interpretability (Testing that demonstrates the results are understood by the potential users for public reporting and quality improvement)	
3a.4 Data/sample (description of data/sample and size):	
3a.5 Methods (e.g., focus group, survey, QI project):	
3a.6 Results (qualitative and/or quantitative results and conclusions):	
3b/3c. Relation to other NQF-endorsed measures	
3b.1 NQF # and Title of similar or related measures:	
(for NQF staff use) Notes on similar/related endorsed or submitted measures:	
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): 3b.2 Are the measure specifications harmonized? If not, why?	3b C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/> NA <input type="checkbox"/>
3c. Distinctive or Additive Value 3c.1 Describe the distinctive, improved, or additive value this measure provides to existing NQF-endorsed measures:	
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:	3c C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/>
TAP/Workgroup: What are the strengths and weaknesses in relation to the sub-criteria for Usability?	3
Steering Committee: Overall, to what extent was the criterion, Usability, met? Rationale:	3 C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/>
4. FEASIBILITY	
Extent to which the required data are readily available, retrievable without undue burden, and can be implemented for performance measurement. (evaluation criteria)	Eval Rating
4a. Data Generated as a Byproduct of Care Processes	4a C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/>
4a.1-2 How are the data elements that are needed to compute measure scores generated?	
4b. Electronic Sources	4b

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

Comment [k24]: 16 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).

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.

<p>4b.1 Are all the data elements available electronically? (<i>elements that are needed to compute measure scores are in defined, computer-readable fields, e.g., electronic health record, electronic claims</i>) No</p> <p>4b.2 If not, specify the near-term path to achieve electronic capture by most providers. <i>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.</i></p>	<p>C <input type="checkbox"/></p> <p>P <input type="checkbox"/></p> <p>M <input type="checkbox"/></p> <p>N <input type="checkbox"/></p>
<p>4c. Exclusions</p> <p>4c.1 Do the specified exclusions require additional data sources beyond what is required for the numerator and denominator specifications?</p> <p>4c.2 If yes, provide justification.</p>	<p>4c</p> <p>C <input type="checkbox"/></p> <p>P <input type="checkbox"/></p> <p>M <input type="checkbox"/></p> <p>N <input type="checkbox"/></p> <p>NA <input type="checkbox"/></p>
<p>4d. Susceptibility to Inaccuracies, Errors, or Unintended Consequences</p> <p>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.</p>	<p>4d</p> <p>C <input type="checkbox"/></p> <p>P <input type="checkbox"/></p> <p>M <input type="checkbox"/></p> <p>N <input type="checkbox"/></p>
<p>4e. Data Collection Strategy/Implementation</p> <p>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:</p> <p>4e.2 Costs to implement the measure (<i>costs of data collection, fees associated with proprietary measures</i>):</p> <p>4e.3 Evidence for costs:</p> <p>4e.4 Business case documentation:</p>	<p>4e</p> <p>C <input type="checkbox"/></p> <p>P <input type="checkbox"/></p> <p>M <input type="checkbox"/></p> <p>N <input type="checkbox"/></p>
<p>TAP/Workgroup: What are the strengths and weaknesses in relation to the sub-criteria for <i>Feasibility</i>?</p>	<p>4</p>
<p>Steering Committee: Overall, to what extent was the criterion, <i>Feasibility</i>, met? Rationale:</p>	<p>4</p> <p>C <input type="checkbox"/></p> <p>P <input type="checkbox"/></p> <p>M <input type="checkbox"/></p> <p>N <input type="checkbox"/></p>
<p>RECOMMENDATION</p>	
<p>(for NQF staff use) Check if measure is untested and only eligible for time-limited endorsement.</p>	<p>Time-limited</p> <p><input type="checkbox"/></p>
<p>Steering Committee: Do you recommend for endorsement? Comments:</p>	<p>Y <input type="checkbox"/></p> <p>N <input type="checkbox"/></p> <p>A <input type="checkbox"/></p>
<p>CONTACT INFORMATION</p>	

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).

<p>Co.1 Measure Steward (Intellectual Property Owner) Co.1 Organization American Medical Association 515 N State St. Chicago Illinois 60654</p> <p>Co.2 Point of Contact Mark Antman, DDS, MBA mark.antman@ama-assn.org 312-464-5056</p>
<p>Measure Developer If different from Measure Steward Co.3 Organization American Medical Association 515 N State St. Chicago Illinois 60654</p> <p>Co.4 Point of Contact Mark Antman, DDS, MBA mark.antman@ama-assn.org 312-464-5056</p>
<p>Co.5 Submitter If different from Measure Steward POC Mark Antman, DDS, MBA mark.antman@ama-assn.org 312-464-5056- American Medical Association</p>
<p>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</p>
<p>ADDITIONAL INFORMATION</p>
<p>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.</p> <p>Co-chairs John Allen, MD, MBA, AGAF (Gastroenterology) Doug Faigel, MD (Gastroenterology)</p> <p>Work Group Members Nancy Baxter, MD, PhD, FACRS, FACS (Colon and Rectal Surgery) Stephen Bickston, MD, AGAF (Gastroenterology) Joel V. Brill, MD, AGAF, FASGE, FACP, CHCOM (Gastroenterology) Kirk Brandon, MBA (Business Administration/Coding) Jason A. Dominitz, MD, MHS, AGAF (Gastroenterology) Ira L. Flax, MD, FACP (Gastroenterology) Karen E. Hall, MD, PhD (Geriatrics) Robert Haskey, MD, FACS (General Surgery, Health Plan representative) Brian C. Jacobson, MD, MPH (Gastroenterology) David Lieberman, MD (Gastroenterology) Klaus Mergener, MD, PhD, CPE, FACP, FACP, FASGE, FACPE (Gastroenterology) Bret Petersen, MD, FASGE (Gastroenterology) Irving M. Pike, MD, FACP (Gastroenterology) Bart Pope, MD (Family Medicine) Harry Sarles, MD, FACP (Gastroenterology) Kay Schwebke, MD, MPH (Specialty: Internal Medicine, Infectious Diseases & Medical Informatics) Tom Lynn, MD (Medical Informatics, Methodology) Emily E. Volk, MD, FCAP (Pathology) Michael Weinstein, MD Specialty: Gastroenterology)</p>

American Gastroenterological Association
Debbie Robin, MSN, RN, CHCOM

American Society for Gastrointestinal Endoscopy
Jill Blim
Chris Recker, RN, MPH
Martha Espronceda

American College of Gastroenterology
Julie Cantor-Weinberg, MPP

American Medical Association
Joseph Gave, MPH
Karen Kmetik, PhD
Shannon Sims, MD, PhD
Beth Tapper, MA

Consortium Consultants
Rebecca Kresowik
Timothy Kresowik, MD

Ad.2 If adapted, provide name of original measure:
Ad.3-5 If adapted, provide original specifications URL or attachment

Measure Developer/Steward Updates and Ongoing Maintenance

Ad.6 Year the measure was first released: 2008

Ad.7 Month and Year of most recent revision: 2008-08

Ad.8 What is your frequency for review/update of this measure? Every 3-4 years or as new evidence becomes available that materially affects the measures

Ad.9 When is the next scheduled review/update for this measure? 2011-08

Ad.10 Copyright statement/disclaimers: 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 HHS-500-2005-000251(0004) with Centers for Medicare and Medicaid Services.

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.

THE MEASURES AND SPECIFICATIONS ARE PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND.

© 2008 American Medical Association and National Committee for Quality Assurance. All Rights Reserved.

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.

Ad.11 -13 Additional Information web page URL or attachment: [Attachment Sample Calculation Algorithm-634008804241472065.doc](#)

Date of Submission (MM/DD/YY): [02/17/2010](#)

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).

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

- an evidence-based risk-adjustment strategy (e.g., risk models, risk stratification) is specified and is based on patient clinical factors that influence the measured outcome (but not disparities in care) and are present at start of care;^{Error! Bookmark not defined.} OR

rationale/data support no risk adjustment.

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

$$\frac{\text{A (\# of patients meeting numerator criteria)}}{\text{PD (\# patients in denominator) - C (\# patients with valid denominator exclusions)}}$$

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

$$\frac{\text{A (\# of patients meeting measure criteria)}}{\text{PD (\# of patients in denominator)}}$$

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

$$\frac{\text{C (\# of patients with any valid exclusion)}}{\text{PD (\# patients in denominator)}}$$

OR

Exclusion Calculation by Type

$$\frac{\text{C}_1 (\# \text{ patients with medical reason})}{\text{PD (\# patients in denominator)}}$$

$$\frac{\text{C}_2 (\# \text{ patients with patient reason})}{\text{PD (\# patients in denominator)}}$$

$$\frac{\text{C}_3 (\# \text{ patients with system reason})}{\text{PD (\# patients in denominator)}}$$