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-017-10 **NQF Project:** Ambulatory Care - Additional Outpatient Measures 2010
MEASURE DESCRIPTIVE INFORMATION
De.1 **Measure Title:** Endoscopy/Poly Surveillance: Colonoscopy Interval for Patients with a History of Adenomatous Polyps- Avoidance of Inappropriate Use
De.2 **Brief description of measure***:* Percentage of patients aged 18 years and older receiving a surveillance colonoscopy, with a history of a prior colonic polyp in previous colonoscopy findings who had a follow-up interval of 3 or more years since their last colonoscopy documented in the colonoscopy report
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:** Overuse De.5 **IOM Quality Domain**: patient-centered, safety, timeliness, effectiveness De.6 **Consumer Care Need:** Getting Better

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. 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*)**:** A.3 **Measure Steward Agreement:** agreement signed and submitted	**A** **Y☐** **N☐**

A.4 Measure Steward Agreement attached:	
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☐ N☐
C. The intended use of the measure includes both public reporting and quality improvement. ►**Purpose:** public reporting, quality improvement Accountability	**C** Y☐ N☐
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.1Testing:** 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☐ N☐
(for NQF staff use) **Have all conditions for consideration been met?** **Staff Notes to Steward** (*if submission returned*)**:**	**Met** Y☐ N☐
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: high resource use, frequently performed procedure **1a.2** **1a.3 Summary of Evidence of High Impact:** Colorectal cancer is the 2nd leading cause of cancer death in the United States. Colonoscopy is the recommended method of surveillance after the removal of adenomatous polyps because it has been shown to significantly reduce subsequent Colorectal Cancer incidence. The time interval for the development of malignant changes in adenomatous polyps is estimated at 5 to 25 years (ICSI, 2006). Performing colonoscopy too often, however not only increases patients' exposure to procedural harm, but also drains limited resources that could be more effectively used to adequately screen those in need. **1a.4 Citations for Evidence of High Impact:** Institute for Clinical Systems Improvement (ICSI). Colorectal cancer screening. Bloomington (MN): Institute for Clinical Systems Improvement (ICSI); 2006 Jun. 50 p.	**1a** C☐ P☐ M☐ N☐
1b. Opportunity for Improvement **1b.1 Benefits (improvements in quality) envisioned by use of this measure:** Several published literature studies indicate that repeat colonoscopy is often over utilized and is not tied to clinical data on initial colonoscopy. The use of this measure is intended to increase physicians' adherence to the evidence based	**1b** C☐ P☐ M☐ N☐

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).

guideline and subsequently may reduce unnecessary tests, costs, and patient risk.

1b.2 Summary of data demonstrating performance gap (variation or overall poor performance) across providers:
A randomized controlled trial of 699 patients showed that after newly diagnosed adenomatous polyps have been removed by colonoscopy, follow-up colonoscopy at 3 years detects important colonic lesions as effectively as follow-up colonoscopy at both 1 and 3 years (ICSI, 2006). A recent pooled analysis of surveillance colonoscopy in 9167 patients from 8 studies confirmed the relative discrimination between high and low risk groups on the basis of patient features and endoscopic findings (Martinez, 2009) The index findings of advanced histology or greater size or number of polyps correlated with risk for subsequent advanced neoplasia. Hence, the timing of follow-up colonoscopy should be tailored to the number, size, and pathologic findings of the adenomatous polyps removed (Levin 2008). In a large multi-state community-based study of colorectal cancer surveillance, one third of those with only 1-2 low risk adenomas were noted to undergo relatively premature repeat colonoscopy within 4 years (median 3.1 yrs), despite recommendations for repeat examination in 5 years, and more recently, 5 to 10 years (Schoen, 2010). Similarly, evidence from 4 surveys indicated that postpolypectomy surveillance colonoscopy in the United States is frequently performed at intervals that are shorter than those recommended in guidelines (Rex et al, 2006). Some endoscopists in these studies performed colonoscopy in patients with only small hyperplastic polyps or a single tubular adenoma at 1 year. These surveys underscore the importance of measuring intervals between examinations in continuous quality improvement programs.

1b.3 Citations for data on performance gap:
Institute for Clinical Systems Improvement (ICSI). Colorectal cancer screening. Bloomington (MN): Institute for Clinical Systems Improvement (ICSI); 2006 Jun. 50 p.

Martinez ME, Baron JA, Lieberman DA, et al. A Pooled Analysis of Advanced Colorectal Neoplasia Diagnoses after Colonoscopic Polypectomy. Gastroenterology Vol 136, 2009.

Levin B, Lieberman DA, McFarland B et al. Screening and Surveillance for the Early Detection of Colorectal Cancer and Adenomatous Polyps, 2008: A Joint Guideline from the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology. CA Cancer J Clin Vol, No 58; 2008.

Schoen R, Pinsky PF, Weissfeld JL, et al. Utilization of Surveillance Colonoscopy in Community Practice. Gastroenterology Vol 138, 2010.

Rex DK. Overuse of postpolypectomy surveillance colonoscopy. - - Rev Gastroenterol Disord Vol 6, No 3; 2006.

1b.4 Summary of Data on disparities by population group:
We are not aware of 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*)**:** Performing colonoscopy too often not only increases patients' exposure to procedural harm, but also drains limited resources that could be more effectively used to adequately screen those in need. This measure would therefore increase patient safety, decrease overuse of resources and
decrease the economic impact of colonoscopy,

1c
C☐
P☐
M☐
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*

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.

healthcare services/care processes influence the outcome)**:**
Patients with 1 to 2 small (1 cm) tubular adenomas with only low-grade dysplasia should undergo follow-up colonoscopy no earlier than 5 years later. Patients with advanced adenomatous lesions or >3 adenomas should have repeat colonoscopy in 3 years as long as all visualized polyps were completely removed, the colonoscopy was completed up to the cecum, and the colonic preparation was adequate. A shorter interval of follow-up is recommended in those patients with numerous adenomatous (>10) polyps and in those in whom the colonoscopy was incomplete or the preparation was inadequate. After a surveillance colonoscopy has normal results, repeat examinations should be done at 5-year intervals. Patients with large, sessile adenomatous lesions removed in a piecemeal fashion should have a repeat examination within 2 to 6 months to exclude and remove and remnant polypoid tissue (Grade 1a) (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: Grade of Recommendation: 1a
Clarity of Benefit: Clear
Methodologic strength/Supporting evidence: Randomized trials without important limitations
Strong Recommendation

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

1c.8 Citations for Evidence (*other than guidelines***):**
Davila , R, Rajan, E, Baron, T. American Society for Gastrointestinal Endoscopy. ASGE guideline: colorectal cancer screening and surveillance. Gastrointestinal Endoscopy Vol. 63. No.4; 2006.

1c.9 Quote the Specific guideline recommendation (*including guideline number and/or page number***):**
Patients with 1 to 2 small (1 cm) tubular adenomas with only low-grade dysplasia should undergo follow-up colonoscopy no earlier than 5 years later. Patients with advanced adenomatous lesions or >3 adenomas should have repeat colonoscopy in 3 years as long as all visualized polyps were completely removed, the colonoscopy was completed up to the cecum, and the colonic preparation was adequate. A shorter interval of follow-up is recommended
in those patients with numerous adenomatous (>10) polyps and in those in whom the colonoscopy was incomplete or the preparation was inadequate. After a surveillance colonoscopy has normal results, repeat examinations should be
done at 5-year intervals. Patients with large, sessile adenomatous lesions removed in a piecemeal fashion should have a repeat examination within 2 to 6 months to exclude and remove and remnant polypoid tissue (Grade 1a) (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.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***):**
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

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.

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.

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 Implications: Very weak recommendation; alternative approaches likely to be better under some circumstances 3 Clarity of Benefit: Unclear Methodologic strength/supporting evidence: Expert opinion only Implications: Weak recommendation; likely to change as data become available 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 expanced 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.	
TAP/Workgroup: What are the strengths and weaknesses in relation to the sub-criteria **for** *Importance to Measure and Report?*	**1**
Steering Committee: Was the threshold criterion, *Importance to Measure and Report*, **met?** **Rationale:**	**1** **Y☐** **N☐**
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)	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**	**2a-specs** **C☐** **P☐** **M☐**
2a.1 **Numerator Statement (***Brief, text description of the numerator - what is being measured about the*	

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) .

target population, e.g. target condition, event, or outcome)**:**
Patients who had an interval of 3 or more years since their last colonoscopy

2a.2 Numerator Time Window (*The time period in which cases are eligible for inclusion in the numerator*)**:**
Every procedure within the denominator time window

2a.3 Numerator Details (*All information required to collect/calculate the numerator, including all codes, logic, and definitions*)**:**
EHR Specifications for this measure are under development

Claims Specifications:
0529F- Interval of at least 3 or more years since patient's last colonoscopy, documented

2a.4 Denominator Statement (*Brief, text description of the denominator - target population being measured*)**:**
All patients aged 18 years and older receiving a surveillance colonoscopy with a history of a prior colonic polyp in a previous colonoscopy

2a.5 Target population gender: Female, Male
2a.6 Target population age range: Patients aged 18 years and older

2a.7 Denominator Time Window (*The time period in which cases are eligible for inclusion in the denominator*)**:**
12 month period

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:
ICD-9-CM diagnosis code: V12.72
AND
CPT codes or G-Codes: 44388, 44389, 44392, 44393, 44394, 45355, 45378, 45380, 45381, 45383, 45384, 45385, G0105

CPT codes 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*)**:**
Documentations of medical reason(s) for an interval of less than 3 years since the last colonoscopy (eg, last colonoscopy incomplete, last colonoscopy had inadequate prep, piecemeal removal of adenomas, or last colonoscopy found greater than 10 adenomas)

OR

Documentation of a system reason(s) for an interval of less than 3 years since the last colonoscopy (eg, unable to locate previous colonoscopy report, previous colonoscopy report was incomplete)

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: 0529F-1P
OR
Append modifier to CPT Category II code: 0529F-3P

EHR Specifications for the exclusions for this measure are under development

N☐

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.11 Stratification Details/Variables (*All information required to stratify the measure including the stratification variables, all codes, logic, and definitions***):**
Stratification by insurance coverage (Commerical, 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.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

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 adminstrative data/claims, paper medical record/flowsheet, electronic Health/Medical Record, 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.***):**

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, Hospital, 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

TESTING/ANALYSIS

2b. Reliability testing

2b.1 Data/sample *(description of data/sample and size)***:** This measure is used in the CMS PQRI program claims option for 2009 and 2010, and registry option for 2009 and 2010. Data from the 2009 PQRI program is not yet available.

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*

2b
C ☐
P ☐
M ☐
N ☐

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.

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)*: 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. 2c.3 **Testing Results** *(statistical results, assessment of adequacy in the context of norms for the test conducted)*:	**2c** C☐ P☐ M☐ N☐
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 AMA 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 because there are medical instances when a patient would need a repeat colonoscopy in a shorter time frame. For example, patient's whose last colonoscopy was incomplete, whose last colonoscopy had inadequate bowel prep, who had piecemeal removal of adenomas, or patient's whose last colonoscopy found greater than 10 adenomas would be excluded from this measure. Additionally, for those instances for which the clinician was unable to locate a prior colonoscopy report or the previous colonoscopy report was incomplete would be excluded from this measure. 2d.2 **Citations for Evidence:** 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)*:	**2d** C☐ P☐ M☐ N☐ NA☐
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:**	**2e** C☐ P☐ M☐ N☐ NA☐

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 rationale/data support no risk adjustment.

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.

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☐ P☐ M☐ N☐
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☐ P☐ M☐ N☐ NA☐
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☐ P☐ M☐ N☐ NA☐
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☐ P☐ M☐ N☐
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:** in use **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)*: This measure is used in the CMS PQRI program claims option for 2009 and 2010, and registry option for 2009 and 2010. Proposed use in Centers for Medicare & Medicaid Services (CMS) proposed rule for the adoption and meaningful use of electronic health record systems (EHRs) http://www.cms.hhs.gov/Recovery/11_HealthIT.asp	**3a** C☐ P☐ M☐ N☐

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.

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☐ P☐ M☐ N☐ NA☐
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☐ P☐ M☐ N☐
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☐ P☐ M☐ N☐
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.1-2 **How are the data elements that are needed to compute measure scores generated?**	**4a** C☐ P☐ M☐ N☐
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*)	**4b** C☐ P☐ M☐

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.

No 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.	N☐
4c. Exclusions 4c.1 **Do the specified exclusions require additional data sources beyond what is required for the numerator and denominator specifications?** 4c.2 **If yes, provide justification.**	**4c** C☐ P☐ M☐ N☐ NA☐
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.**	**4d** C☐ P☐ M☐ N☐
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:** 4e.2 **Costs to implement the measure** (*costs of data collection, fees associated with proprietary measures*)**:** 4e.3 **Evidence for costs:** 4e.4 **Business case documentation:**	**4e** C☐ P☐ M☐ N☐
TAP/Workgroup: What are the strengths and weaknesses in relation to the sub-criteria **for** *Feasibility*?	**4**
Steering Committee: Overall, to what extent was the criterion, *Feasibility*, **met?** **Rationale:**	**4** C☐ P☐ M☐ N☐
RECOMMENDATION	
(for NQF staff use) **Check if measure is untested and only eligible for time-limited endorsement.**	**Time-limited** ☐
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	

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).

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, FACRS, FACS (Colon and Rectal Surgery)
Stephen Bickston, MD, AGAF (Gastroenterology)
Joel V. Brill, MD, AGAF, FASGE, FACG, CHCQM (Gastroenterology)
Kirk Brandon, MBA (Business Administration/Coding)
Jason A. Dominitz, MD, MHS, AGAF (Gastroenterology)
Ira L. Flax, MD, FACG (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, FACG, FASGE, FACPE (Gastroenterology)
Bret Petersen, MD, FASGE (Gastroenterology)
Irving M. Pike, MD, FACG (Gastroenterology)
Bart Pope, MD (Family Medicine)
Harry Sarles, MD, FACG (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)

American Gastroenterological Association
Debbie Robin, MSN, RN, CHCQM

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 HHSM-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-634008798221354344.doc

Date of Submission (*MM/DD/YY*)**:** 02/17/2010

Page 8: [1] Comment [KP14] **Karen Pace** **10/5/2009 8:59:00 AM**

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

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

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

$$\frac{C_1 \text{ (\# patients with medical reason)}}{\text{PD (\# patients in denominator)}} \qquad \frac{C_2 \text{ (\# patients with patient reason)}}{\text{PD (\# patients in denominator)}} \qquad \frac{C_3 \text{ (\# patients with system reason)}}{\text{PD (\# patients in denominator)}}$$

© 2009. American Medical Association. All Rights Reserved.