NATIONAL QUALITY FORUM

Measure Submission and Evaluation Worksheet 5.0

This form contains the information submitted by measure developers/stewards, organized according to NQF's measure evaluation criteria and process. The evaluation criteria, evaluation guidance documents, and a blank online submission form are available on the submitting standards web page.

NQF #: 0392 NQF Project: Cancer Project

(for Endorsement Maintenance Review)
Original Endorsement Date: Jul 31, 2008 Most Recent Endorsement Date: Jul 31, 2008

BRIEF MEASURE INFORMATION

De.1 Measure Title: Colorectal Cancer Resection Pathology Reporting- pT category (primary tumor) and pN category (regional lymph nodes) with histologic grade

Co.1.1 Measure Steward: American Medical Association - Physician Consortium for Performance Improvement

De.2 Brief Description of Measure: Percentage of colon and rectum cancer resection pathology reports that include the pT category (primary tumor), the pN category (regional lymph nodes) and the histologic grade

2a1.1 Numerator Statement: Reports that include the pT category, the pN category and the histologic grade

2a1.4 Denominator Statement: All colon and rectum cancer resection pathology reports

2a1.8 Denominator Exclusions: Denominator Exclusion: Documentation of medical reason(s) for not including the pT category, the pN category or the histologic grade (eg; re-excision without residual tumor; non-carcinomasanal canal)

1.1 Measure Type: Process
2a1. 25-26 Data Source: Electronic Clinical Data, Electronic Clinical Data : Electronic Health Record, Electronic Clinical Data : Laboratory, Electronic Clinical Data : Registry, Paper Records
2a1.33 Level of Analysis: Clinician : Group/Practice, Clinician : Individual, Clinician : Team

1.2-1.4 Is this measure paired with another measure? No

De.3 If included in a composite, please identify the composite measure (*title and NQF number if endorsed*):

STAFF NOTES (*issues or questions regarding any criteria*)

Comments on Conditions for Consideration:

Is the measure untested? Yes☐ No☐ If untested, explain how it meets criteria for consideration for time-limited endorsement:

1a. Specific national health goal/priority identified by DHHS or NPP addressed by the measure (*check De.5*):
5. Similar/related endorsed or submitted measures (*check 5.1*):
Other Criteria:

Staff Reviewer Name(s):

1. IMPACT, OPPORTUITY, EVIDENCE - IMPORTANCE TO MEASURE AND REPORT

Importance to Measure and Report is a threshold criterion that must be met in order to recommend a measure for endorsement. All three subcriteria must be met to pass this criterion. See guidance on evidence.
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: H☐ M☐ L☐ I ☐
(*The measure directly addresses a specific national health goal/priority identified by DHHS or NPP, or some other high impact aspect of healthcare.*)

De.4 Subject/Topic Areas *(Check all the areas that apply)*: Cancer, Cancer : Colorectal
De.5 Cross Cutting Areas (*Check all the areas that apply)*:

1a.1 Demonstrated High Impact Aspect of Healthcare: Affects large numbers, A leading cause of morbidity/mortality, Patient/societal consequences of poor quality, Severity of illness

1a.2 If "Other," please describe:

1a.3 Summary of Evidence of High Impact *(Provide epidemiologic or resource use data)*:
According to NCCN, colorectal cancer is the fourth most frequently diagnosed cancer and the second leading cause of cancer death in the United States. In 2011 an estimated 101, 340 new cases of colon cancer and approximately 39,870 cases of rectal cancer will occur. Additionally, it is estimated that 49,380 people will die of colon and rectal cancer combined.

1a.4 Citations for Evidence of High Impact cited in 1a.3: NCCN Clinical Practice Guidelines in Oncology - Colon Cancer. Version 2, 2012. Available here: http://www.nccn.org/professionals/physician_gls/pdf/colon.pdf

1b. Opportunity for Improvement: H☐ M☐ L☐ I ☐
(*There is a demonstrated performance gap - variability or overall less than optimal performance*)

1b.1 Briefly explain the benefits (improvements in quality) envisioned by use of this measure:
A complete set of pathology descriptors is needed for therapeutic decisions regarding colorectal cancer management because of it being stage driven. Incomplete cancer resection pathology reports may result in misclassification of patients.

1b.2 Summary of Data Demonstrating Performance Gap *(Variation or overall less than optimal performance across providers)*:
[*For Maintenance – Descriptive statistics for performance results for this measure - distribution of scores for measured entities by quartile/decile, mean, median, SD, min, max, etc.*]
The CAP conducted a structured audit of colorectal cancer pathology report adequacy at 86 institutions. Overall, 21% of eligible reports were missing at least one of the ten CAP-recommended colorectal cancer elements.

CMS Physician Quality Reporting Initiative/System (PQRI/S)

This measure was used in the 2008 (claims), 2009 (claims and registry) and 2010 (claims and registry) CMS Physician Quality Reporting Initiative/System (PQRI/S) as NQF #100 Colorectal cancer resection pathology reporting- pT category and pN category with histologic grade. There is a gap in care as shown by this 2008 data; with 25.82% of patients reported on not receiving the optimal care.

10th percentile: 33.33%
25th percentile: 60.00%
50th percentile: 90.91%
75th percentile: 100.00%
90th percentile: 100.00%

1b.3 Citations for Data on Performance Gap: [*For Maintenance – Description of the data or sample for measure results reported in 1b.2 including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included*]
Idowu MO, Bekeris LG, et al. Adequacy of Surgical Pathology Reporting of Cancer: A College of American Pathologists Q-Probes Study of 86 institutions. Arch Pathol Lab Med-Vol 134, July 2010.

Confidential CMS PQRI 2008 Performance Information by Measure. Jan-Sept TAP file

1b.4 Summary of Data on Disparities by Population Group: [*For Maintenance –Descriptive statistics for performance results*

for this measure by population group]
We are not aware of any publications/evidence outlining disparities in documentation of cancer staging however the National Cancer Institute and AHRQ´s National Healthcare Disparities Report has shown that disparities exist in cancer incidence and deaths by race, ethnicity and socioeconomic status.

1b.5 Citations for Data on Disparities Cited in 1b.4: [*For Maintenance – Description of the data or sample for measure results reported in 1b.4 including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included*]
Harper S, Lynch J. Methods for Measuring Cancer Disparities: Using Data Relevant to Healthy People 2010 Cancer-Related Objectives. Cancer Control Monograph Series, No. 6. Bethesda, MD: National Cancer Institute; 2005. NIH publication 05-5777.

Agency for Healthcare Research and Quality. 2010 National Healthcare Disparities Report. http://www.ahrq.gov/qual/nhdr10/nhdr10.pdf. Published March 2011. Accessed January 3, 2011.

1c. Evidence (*Measure focus is a health outcome OR meets the criteria for quantity, quality, consistency of the body of evidence.*)
Is the measure focus a health outcome? Yes☐ No☐ If not a health outcome, rate the body of evidence.

Quantity: H☐ M☐ L☐ I ☐ Quality: H☐ M☐ L☐ I ☐ Consistency: H☐ M☐ L☐ I ☐

Quantity	Quality	Consistency	Does the measure pass subcriterion1c?
M-H	M-H	M-H	Yes☐
L	M-H	M	Yes☐ IF additional research unlikely to change conclusion that benefits to patients outweigh harms: otherwise No☐
M-H	L	M-H	Yes☐ IF potential benefits to patients clearly outweigh potential harms: otherwise No☐
L-M-H	L-M-H	L	No ☐

Health outcome – rationale supports relationship to at least one healthcare structure, process, intervention, or service	Does the measure pass subcriterion1c? Yes☐ IF rationale supports relationship

1c.1 Structure-Process-Outcome Relationship *(Briefly state the measure focus, e.g., health outcome, intermediate clinical outcome, process, structure; then identify the appropriate links, e.g., structure-process-health outcome; process- health outcome; intermediate clinical outcome-health outcome)*:
Inclusion of a complete set of pathology descriptors in a pathology report is optimal for complete patient care. Incomplete cancer care may result in missclassification of patients.

Idowu MO, Bekeris LG, et al. Adequacy of Surgical Pathology Reporting of Cancer: A College of American Pathologists Q-Probes Study of 86 institutions. Arch Pathol Lab Med-Vol 134, July 2010.

1c.2-3 Type of Evidence *(Check all that apply)*:
Clinical Practice Guideline

1c.4 Directness of Evidence to the Specified Measure *(State the central topic, population, and outcomes addressed in the body of evidence and identify any differences from the measure focus and measure target population)*:
Clinical practice guidelines for colorectal cancer recommend staging to allow for efficient identification of local treatment options, assists in identifying systemic treatment options, allows the comparison of outcomes results across institutions and clinical trials, and provides baseline prognotic information.

The measure focus is on Resection Pathology Reporting for Colorectal Cancer.

1c.5 Quantity of Studies in the Body of Evidence *(Total number of studies, not articles)*: The description of evidence review in the guideline did not address the overall quantity of studies in the body of evidence. However, 387 and 271 articles are cited in NCCN´s colon and rectal cancer guideline´s references section.

1c.6 Quality of Body of Evidence (*Summarize the certainty or confidence in the estimates of benefits and harms to patients across studies in the body of evidence resulting from study factors. Please address: a) study design/flaws; b) directness/indirectness of the evidence to this measure (e.g., interventions, comparisons, outcomes assessed, population included in the evidence); and c) imprecision/wide confidence intervals due to few patients or events*): The quality of the body of evidence supporting the guideline recommendation is summarized according to the NCCN categories of evidence and consensus as being based on "lower-level evidence". Lower-level evidence is later described as evidence that may include non-randomized trials; case series; or when other data are lacking, the clinical experience of expert physicians.

1c.7 Consistency of Results across Studies *(Summarize the consistency of the magnitude and direction of the effect):* Althought there is no explicit statement regarding the overall consistency of results across studies in the guidelines supporting the measures, the recommendation received uniform NCCN consensus that the intervention is appropriate.

1c.8 Net Benefit *(Provide estimates of effect for benefit/outcome; identify harms addressed and estimates of effect; and net benefit - benefit over harms)*:
A complete set of pathology descriptors is needed for therapeutic decisions regarding colorectal cancer management because of it being stage driven. Incomplete cancer resection pathology reports may result in misclassification of patients.

1c.9 Grading of Strength/Quality of the Body of Evidence. Has the body of evidence been graded? Yes

1c.10 If body of evidence graded, identify the entity that graded the evidence including balance of representation and any disclosures regarding bias: NCCN Guidelines Committee; The CAP protocol is revised by a multi-disciplinary team of experts.

1c.11 System Used for Grading the Body of Evidence: GRADE

1c.12 If other, identify and describe the grading scale with definitions:

1c.13 Grade Assigned to the Body of Evidence: 2A: Based upon lower-level evidence, there is uniform NCCN consensus that the intervention is appropriate; CAP is not available.

1c.14 Summary of Controversy/Contradictory Evidence: No controversy or contradictory evidence provided.

1c.15 Citations for Evidence other than Guidelines*(Guidelines addressed below)*:

1c.16 Quote verbatim, the specific guideline recommendation *(Including guideline # and/or page #)*:
Surgical resection remains the most effective therapy for colorectal carcinoma, and the best estimation of prognosis is derived from the pathologic findings on the resection specimen. The anatomic extent of disease is by far the most important prognostic factor in colorectal cancer. The protocol recommends the TNM staging system of the American Joint Committee on Cancer (AJCC) and the International Union Against Cancer(UICC)1 but does not preclude the use of other staging systems. By AJCC/UICC convention, the designation “T” refers to a primary tumor that has not been previously treated. The symbol “p” refers to the pathologic classification of the TNM, as opposed to the clinical classification, and is based on gross and microscopic examination. pT entails a resection of the primary tumor or biopsy adequate to evaluate the highest pT category, pN entails removal or biopsy of nodes adequate to validate lymph node metastasis, and pM implies microscopic examination of distant lesions.

Colorectal cancers are usually staged after surgical exploration of the abdomen and pathologic examination of the surgical specimen. Some of the criteria that should be included in the report of the pathologic evluation include the following: grade of the cancer; depth of penetration and extension to adjacent structures (T); number of regional lymph nodes evaluated; number of positive regional lymph nodes (N); an assessment of the presence of distant metasteses to other organs, the peritoneum of an abdominal structure, or in non-regional lymph nodes (M); the status of proximal, distal and radial margins; lymphovascular invasion, perineurial invasion and extra-nodal tumor deposits (NCCN).

1c.17 Clinical Practice Guideline Citation: Colon and Rectum, Protocol applies to all invasive carcinomas of the colon and rectum. College of American Pathologists. Revised Febuary 2011. Available at:

http://www.cap.org/apps/docs/committees/cancer/cancer_protocols/2011/Colon_09protocol.pdf

NCCN Clinical Practice Guidelines in Oncology - Colon Cancer. Version 2, 2012. Available here: http://www.nccn.org/professionals/physician_gls/pdf/colon.pdf

1c.18 National Guideline Clearinghouse or other URL:

1c.19 Grading of Strength of Guideline Recommendation. Has the recommendation been graded? Yes

1c.20 If guideline recommendation graded, identify the entity that graded the evidence including balance of representation and any disclosures regarding bias: Guidelines Committee

1c.21 System Used for Grading the Strength of Guideline Recommendation: Other

1c.22 If other, identify and describe the grading scale with definitions: 2A; CAP not available

1c.23 Grade Assigned to the Recommendation: 2A; CAP not available

1c.24 Rationale for Using this Guideline Over Others: It is the PCPI policy to use guidelines, which are evidence-based, applicable to physicians and other health-care providers, and developed by a national specialty organization or government agency. In addition, the PCPI has now expanded what is acceptable as the evidence base for measures to include documented quality improvement (QI) initiatives or implementation projects that have demonstrated improvement in quality of care.

Based on the NQF descriptions for rating the evidence, what was the developer's assessment of the quantity, quality, and consistency of the body of evidence?
1c.25 Quantity: Moderate 1c.26 Quality: Moderate1c.27 Consistency: Moderate

Was the threshold criterion, *Importance to Measure and Report*, met?
(*1a & 1b must be rated moderate or high and 1c yes*) Yes☐ No☐
Provide rationale based on specific subcriteria:

For a new measure if the Committee votes NO, then STOP.
For a measure undergoing endorsement maintenance, if the Committee votes NO because of 1b. (no opportunity for improvement), it may be considered for continued endorsement and all criteria need to be evaluated.

2. RELIABILITY & VALIDITY - 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)
Measure testing must demonstrate adequate reliability and validity in order to be recommended for endorsement. Testing may be conducted for data elements and/or the computed measure score. Testing information and results should be entered in the appropriate field. Supplemental materials may be referenced or attached in item 2.1. See guidance on measure testing.

S.1 Measure Web Page *(In the future, NQF will require measure stewards to provide a URL link to a web page where current detailed specifications can be obtained*). Do you have a web page where current detailed specifications for this measure can be obtained? No

S.2 If yes, provide web page URL:

2a. RELIABILITY. Precise Specifications and Reliability Testing: H☐ M☐ L☐ I☐

2a1. Precise Measure Specifications. (*The measure specifications precise and unambiguous.*)

2a1.1 Numerator Statement *(Brief, narrative description of the measure focus or what is being measured about the target population, e.g., cases from the target population with the target process, condition, event, or outcome)*:
Reports that include the pT category, the pN category and the histologic grade

2a1.2 Numerator Time Window *(The time period in which the target process, condition, event, or outcome is eligible for inclusion)*:
Each final report during measurement period

2a1.3 Numerator Details *(All information required to identify and calculate the cases from the target population with the target process, condition, event, or outcome such as definitions, codes with descriptors, and/or specific data collection items/responses*:
For EHR:

eSpecification currently under development. Data elements (using Quality Data Model) required for the measure attached.

For Claims/Administrative:

CPT Category II code 3260F: pT (primary tumor), pN (regional lymph node), and histologic grade documented in pathology report

2a1.4 Denominator Statement *(Brief, narrative description of the target population being measured)*:
All colon and rectum cancer resection pathology reports

2a1.5 Target Population Category *(Check all the populations for which the measure is specified and tested if any)*: Adult/Elderly Care

2a1.6 Denominator Time Window *(The time period in which cases are eligible for inclusion)*:
12 consecutive months

2a1.7 Denominator Details (*All information required to identify and calculate the target population/denominator such as definitions, codes with descriptors, and/or specific data collection items/responses)*:
For EHR:

eSpecification currently under development. Data elements (using Quality Data Model) required for the measure attached.

For Claims/Administrative:

ICD-9-CM diagnosis codes: 153.0, 153.1, 153.2, 153.3, 153.4, 153.5, 153.6, 153.7, 153.8, 153.9, 154.0, 154.1, 154.8
ICD-10-CM diagnosis codes: C18.0, C18.1, C18.2, C18.3, C18.4, C18.5, C18.6, C18.7, C18.8, C18.9, C19, C20, C21.2, C21.8
AND
CPT Codes: 88309

2a1.8 Denominator Exclusions *(Brief narrative description of exclusions from the target population)*:
Denominator Exclusion: Documentation of medical reason(s) for not including the pT category, the pN category or the histologic grade (eg; re-excision without residual tumor; non-carcinomasanal canal)

2a1.9 Denominator Exclusion Details *(All information required to identify and calculate exclusions from the denominator such as definitions, codes with descriptors, and/or specific data collection items/responses)*:
The PCPI methodology uses three categories of reasons for which a patient may be excluded from the denominator of an individual measure. These measure exception categories are not uniformly relevant across all measures; for each measure, there must be a clear rationale to permit an exception for a medical, patient, or system reason. Examples are provided in the measure exception language of instances that may constitute an exception and are intended to serve as a guide to clinicians. For this measure exceptions may include Documentation of medical reason(s) for not including the pT category, the pN category or the histologic grade. Where examples of exceptions are included in the measure language, these examples are coded and included in the eSpecifications. Although this methodology does not require the external reporting of more detailed exception data, the PCPI recommends that physicians document the specific reasons for exception in patients' medical records for purposes of optimal patient management and audit-readiness. The PCPI also advocates the systematic review and analysis of each physician's exceptions data to identify practice patterns and opportunities for quality improvement. For example, it is possible for implementers to calculate the percentage of patients that physicians have identified as meeting the criteria for exception. Additional details by

data source are as follows:

For EHR:

eSpecification currently under development. Data elements (using Quality Data Model) required for the measure attached.

For Claims/Administrative:

Documentation of medical reason(s) for not including the pT category, the pN category or the histologic grade
- Append modifier to CPT Category II code: 3260F-1P

OR
If the specimen is not primary breast tissue (e.g., liver, lung) report:
CPT II 3250F: Specimen site other than anatomic location of primary tumor

2a1.10 Stratification Details/Variables *(All information required to stratify the measure results including the stratification variables, codes with descriptors, definitions, and/or specific data collection items/responses)*:
We encourage the results of this measure to be stratified by race, ethnicity, gender, and primary language, and have included these variables as recommended data elements to be collected.

2a1.11 Risk Adjustment Type *(Select type. Provide specifications for risk stratification in 2a1.10 and for statistical model in 2a1.13)*: No risk adjustment or risk stratification 2a1.12 If "Other," please describe:

2a1.13 Statistical Risk Model and Variables *(Name the statistical method - e.g., logistic regression and list all the risk factor variables. Note - risk model development should be addressed in 2b4.)*:

2a1.14-16 Detailed Risk Model Available at Web page URL (or attachment). Include coefficients, equations, codes with descriptors, definitions, and/or specific data collection items/responses. Attach documents only if they are not available on a webpage and keep attached file to 5 MB or less. NQF strongly prefers you make documents available at a Web page URL. Please supply login/password if needed:

2a1.17-18. Type of Score:

2a1.19 Interpretation of Score *(Classifies interpretation of score according to whether better quality is associated with a higher score, a lower score, a score falling within a defined interval, or a passing score)*: Better quality = Higher score

2a1.20 Calculation Algorithm/Measure Logic*(Describe the calculation of the measure score as an ordered sequence of steps including identifying the target population; exclusions; cases meeting the target process, condition, event, or outcome; aggregating data; risk adjustment; etc.)*:
To calculate performance rates:
1) Find the patients who meet the initial patient population (ie, the general group of patients that the performance measure is designed to address).
2) From the patients within the initial patient population criteria, find the patients who qualify for the denominator (ie, the specific group of patients for inclusion in a specific performance measure based on defined criteria). Note: in some cases the initial patient population and denominator are identical.
3) From the patients within the denominator, find the patients who qualify for the Numerator (ie, the group of patients in the denominator for whom a process or outcome of care occurs). Validate that the number of patients in the numerator is less than or equal to the number of patients in the denominator
4) From the patients who did not meet the numerator criteria, determine if the physician has documented that the patient meets any criteria for denominator exception when exceptions have been specified [For this measure: documentation of medical

reason(s) for not including the pT category, the pN category or the histologic grade]. If the patient meets any exception criteria, they should be removed from the denominator for performance calculation. --Although the exception cases are removed from the denominator population for the performance calculation, the number of patients with valid exceptions should be calculated and reported along with performance rates to track variations in care and highlight possible areas of focus for QI.

If the patient does not meet the numerator and a valid exception is not present, this case represents a quality failure.

Calculation algorithm is included in data dictionary/code table attachment 2a1.30.

2a1.21-23 Calculation Algorithm/Measure Logic Diagram URL or attachment:
Attachment
AMA-PCPI_Measure Calculation.pdf

2a1.24 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):
Not applicable. The measure does not require sampling or a survey.

2a1.25 Data Source *(Check all the sources for which the measure is specified and tested*). If other, please describe:
Electronic Clinical Data, Electronic Clinical Data : Electronic Health Record, Electronic Clinical Data : Laboratory, Electronic Clinical Data : Registry, Paper Records

2a1.26 Data Source/Data Collection Instrument (*Identify the specific data source/data collection instrument, e.g. name of database, clinical registry, collection instrument, etc.*):

2a1.27-29 Data Source/data Collection Instrument Reference Web Page URL or Attachment:

2a1.30-32 Data Dictionary/Code Table Web Page URL or Attachment:
Attachment
AMA-PCPI_0392_PATH ColorectalCancerResectionPathology_DataElements_1 2012.pdf

2a1.33 Level of Analysis (*Check the levels of analysis for which the measure is specified and tested)*: Clinician : Group/Practice, Clinician : Individual, Clinician : Team

2a1.34-35 Care Setting *(Check all the settings for which the measure is specified and tested)*: Ambulatory Care : Ambulatory Surgery Center (ASC), Laboratory

2a2. Reliability Testing. (*Reliability testing was conducted with appropriate method, scope, and adequate demonstration of reliability*.)

2a2.1 Data/Sample *(Description of the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included)*:
AMA-PCPI Testing Project
- The data sample came from four sites representing various types, locations and sizes
  - Three of the practices were urban, and one was more rural; each located in a different state
- The sample consisted of 25 colorectal cancer pathology reports for three of the four sites and 19 at the fourth site, for a total of 94 patient records
- Data collected from patients seen between January 1, 2009 and December 31, 2009.
- Data abstraction performed between June 10, 2010 and December 8, 2010.

2a2.2 Analytic Method *(Describe method of reliability testing & rationale)*:
Data abstracted from randomly sampled patient records were used to calculate inter-rater reliability for the measure.

Patients were randomly selected from visits for a diagnosis of colorectal cancer.
Data analysis included:

- Percent agreement
- Kappa statistic to adjust for chance agreement

2a2.3 Testing Results *(Reliability statistics, assessment of adequacy in the context of norms for the test conducted)*:
Overall, this measure is highly reliable.

Reliability (N, % Agreement, Kappa)
Numerator (94, 98.3%, 1.00)
Denominator (94, 100%, kappa not calculable *)
Exceptions (94, 100%, kappa not calculable *)
Overall (94, 100%, kappa not calculable *)

* Kappa statistics cannot be calculated because of complete agreement. Confidence intervals cannot be calculated because to do so would involve dividing by zero which cannot be done.

2b. VALIDITY. Validity, Testing, including all Threats to Validity: H☐ M☐ L☐ I ☐

2b1.1 Describe how the measure specifications *(measure focus, target population, and exclusions)* are consistent with the evidence cited in support of the measure focus (*criterion 1c)* and identify any differences from the evidence:

2b2. Validity Testing. (*Validity testing was conducted with appropriate method, scope, and adequate demonstration of validity.)*

2b2.1 Data/Sample *(Description of the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included)*:
An expert panel was used to assess face validity of the measure. This panel consisted of the following # members, with representation from the following specialties:

David L. Witte, MD, PhD, FCAP (Co-Chair, pathology)
Susan R. Snyder, PhD, MBA (Co-Chair, methodology)
Nancy Baxter, MD, PhD (colorectal surgery)
Joel V. Brill, MD, AGAF, FACG, CHCQM (gastroenterology)
Patrick Fitzgibbons, MD, FCAP (pathology)
M. Kay Washington, MD, PhD, FCAP (pathology)
Mario Gonzalez, MD, FCAP, FASCP (pathology)
Richard M Gore, MD, FACR (diagnostic radiology)
Dana Marie Grzybicki, MD, PhD (pathology)
Harvey W. Kaufman, MD, FCAP(pathology)
Jonathon Myles, MD, FCAP (pathology)
Raouf E. Nakhleh, MD, FCAP (pathology)
Felicia Nicholson, RN, BSN (health plan representative)
Omar Yousef, MD, FCAP (pathology)

2b2.2 Analytic Method *(Describe method of validity testing and rationale; if face validity, describe systematic assessment)*:
All PCPI performance measures are assessed for content validity by a panel of expert work group members during the development process. Additional input on the content validity of draft measures is obtained through a 30-day public comment period and by also soliciting comments from a panel of consumer, purchaser, and patient representatives convened by the PCPI specifically for this purpose. All comments received are reviewed by the expert work group and the measures adjusted as needed. Other external review groups (eg, focus groups) may be convened if there are any remaining concerns related to the content validity of the measures.

The expert panel was used to assess face validity of the measure. This panel consisted of X members, with representation from the

following specialties: pathology, colorectal surgery, gastroenterology and diagnostic radiology.

The aforementioned panel was asked to rate their agreement with the following statement:

The scores obtained from the measure as specified will accurately differentiate quality across providers.

Scale 1-5, where 1=Strongly Disagree; 3=Neither Disagree nor Agree; 5=Strongly Agree

2b2.3 Testing Results *(Statistical results, assessment of adequacy in the context of norms for the test conducted; if face validity, describe results of systematic assessment)*:
The results of the expert panel rating of the validity statement were as follows: N = 12; Mean rating = 3.92
Frequency Distribution of Ratings
1 - 1 (Disagree)
2 – 0
3 - 3 (Neither Disagree nor Agree)
4 - 3
5 - 5 (Agree)

POTENTIAL THREATS TO VALIDITY. (*All potential threats to validity were appropriately tested with adequate results.*)

2b3. Measure Exclusions. (*Exclusions were supported by the clinical evidence in 1c or appropriately tested with results demonstrating the need to specify them.*)

2b3.1 Data/Sample for analysis of exclusions *(Description of the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included)*:
There are no documented exceptions for this measure.

2b3.2 Analytic Method (*Describe type of analysis and rationale for examining exclusions, including exclusion related to patient preference)*:
There are no documented exceptions for this measure.

2b3.3 Results *(Provide statistical results for analysis of exclusions, e.g., frequency, variability, sensitivity analyses):*
There are no documented exceptions for this measure.

2b4. Risk Adjustment Strategy. (*For outcome measures, adjustment for differences in case mix (severity) across measured entities was appropriately tested with adequate results.*)

2b4.1 Data/Sample *(Description of the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included)*:
This measure is not risk adjusted.

2b4.2 Analytic Method (*Describe methods and rationale for development and testing of risk model or risk stratification including selection of factors/variables)*:
This measure is not risk adjusted.

2b4.3 Testing Results *(Statistical risk model: Provide quantitative assessment of relative contribution of model risk factors; risk model performance metrics including cross-validation discrimination and calibration statistics, calibration curve and risk decile plot, and assessment of adequacy in the context of norms for risk models. Risk stratification: Provide quantitative assessment of relationship of risk factors to the outcome and differences in outcomes among the strata)*:
Not Applicable

2b4.4 If outcome or resource use measure is not risk adjusted, provide rationale and analyses to justify lack of adjustment:

2b5. Identification of Meaningful Differences in Performance. (*The performance measure scores were appropriately analyzed and discriminated meaningful differences in quality.*)

2b5.1 Data/Sample *(Describe the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included)*:

2b5.2 Analytic Method *(Describe methods and rationale to identify statistically significant and practically/meaningfully differences in performance)*:

2b5.3 Results *(Provide measure performance results/scores, e.g., distribution by quartile, mean, median, SD, etc.; identification of statistically significant and meaningfully differences in performance)*:
CMS Physician Quality Reporting Initiative 2008:
This measure was used in the 2008 (claims), 2009 (claims and registry) and 2010 (claims and registry) CMS Physician Quality Reporting Initiative/System (PQRI/S) as NQF #100 Colorectal cancer resection pathology reporting- pT category and pN category with histologic grade. There is a gap in care as shown by this 2008 data; with 25.82% of patients reported on not receiving the optimal care.

10th percentile: 33.33%
25th percentile: 60.00%
50th percentile: 90.91%
75th percentile: 100.00%
90th percentile: 100.00%

The inter-quartile range (IQR) provides a measure of the dispersion of performance. The IQR is 40, and indicates that 50% of physicians have performance on this measure ranging from 60.00% and 100.00%.

Confidential CMS PQRI 2008 Performance Information by Measure. Jan-Sept TAP file.

2b6. Comparability of Multiple Data Sources/Methods. (*If specified for more than one data source, the various approaches result in comparable scores.*)

2b6.1 Data/Sample *(Describe the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included)*:
This test was not performed for this measure.

2b6.2 Analytic Method *(Describe methods and rationale for testing comparability of scores produced by the different data sources specified in the measure)*:
This test was not performed for this measure.

2b6.3 Testing Results *(Provide statistical results, e.g., correlation statistics, comparison of rankings; assessment of adequacy in the context of norms for the test conducted)*:
This test was not performed for this measure.

2c. Disparities in Care: H☐ M☐ L☐ I ☐ NA☐ (*If applicable, the measure specifications allow identification of disparities.*)

2c.1 If measure is stratified for disparities, provide stratified results *(Scores by stratified categories/cohorts)*: We encourage the results of this measure to be stratified by race, ethnicity, gender, and primary language, and have included these variables as recommended data elements to be collected.

2c.2 If disparities have been reported/identified (e.g., in 1b), but measure is not specified to detect disparities, please explain:
The PCPI advocates that performance measure data should, where possible, be stratified by race, ethnicity, and primary language to assess disparities and initiate subsequent quality improvement activities addressing identified disparities, consistent with recent national efforts to standardize the collection of race and ethnicity data. A 2008 NQF report endorsed 45 practices including stratification by the aforementioned variables.(1) A 2009 IOM report "recommends collection of the existing Office of Management and Budget (OMB) race and Hispanic ethnicity categories as well as more fine-grained categories of ethnicity(referred to as

granular ethnicity and based on one's ancestry) and language need (a rating of spoken English language proficiency of less than very well and one's preferred language for health-related encounters)."(2)

References:
(1)National Quality Forum Issue Brief (No.10). Closing the Disparities Gap in Healthcare Quality with Performance Measurement and Public Reporting. Washington, DC: NQF, August 2008.

(2)Race, Ethnicity, and Language Data: Standardization for Health Care Quality Improvement. March 2010. AHRQ Publication No. 10-0058-EF. Agency for Healthcare Research and Quality, Rockville, MD. Available at: http://www.ahrq.gov/research/iomracereport. Accessed May 25, 2010.

2.1-2.3 Supplemental Testing Methodology Information:

Steering Committee: Overall, was the criterion, *Scientific Acceptability of Measure Properties*, met?
(*Reliability and Validity must be rated moderate or high*) Yes☐ No☐
Provide rationale based on specific subcriteria:

If the Committee votes No, STOP

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)

C.1 Intended Purpose/ Use *(Check all the purposes and/or uses for which the measure is intended)*: Public Reporting, Quality Improvement (Internal to the specific organization)

3.1 Current Use *(Check all that apply; for any that are checked, provide the specific program information in the following questions)*:

3a. Usefulness for Public Reporting: H☐ M☐ L☐ I ☐
(*The measure is meaningful, understandable and useful for public reporting.*)

3a.1. Use in Public Reporting - disclosure of performance results to the public at large *(If used in a public reporting program, provide name of program(s), locations, Web page URL(s)).* If not publicly reported in a national or community program, state the reason AND plans to achieve public reporting, potential reporting programs or commitments, and timeline, e.g., within 3 years of endorsement: [*For Maintenance – If not publicly reported, describe progress made toward achieving disclosure of performance results to the public at large and expected date for public reporting; provide rationale why continued endorsement should be considered.*]
The PCPI believes that the reporting of participation information is a beneficial first step on a trajectory toward the public reporting of performance results, which is appropriate since the measure has been tested and the reliability of the performance data has been validated. Continued NQF endorsement will facilitate our ongoing progress toward this public reporting objective.

CMS Physician Quality Reporting Initiative/System (PQRI/S)

This measure was used in the 2008 (claims), 2009 (claims and registry) and 2010 (claims and registry) CMS Physician Quality Reporting Initiative/System (PQRI/S) as NQF #100 Colorectal cancer resection pathology reporting- pT category and pN category with histologic grade. There is a gap in care as shown by this 2008 data; with 25.82% of patients reported on not receiving the optimal care.

10th percentile: 33.33%
25th percentile: 60.00%

50th percentile: 90.91%
75th percentile: 100.00%
90th percentile: 100.00%

3a.2.Provide a rationale for why the measure performance results are meaningful, understandable, and useful for public reporting. If usefulness was demonstrated (e.g., focus group, cognitive testing), describe the data, method, and results: The PCPI believes that the reporting of participation information is a beneficial first step on a trajectory toward the public reporting of performance results, which is appropriate since the measure has been tested and the reliability of the performance data has been validated. Continued NQF endorsement will facilitate our ongoing progress toward this public reporting objective.

3.2 Use for other Accountability Functions (payment, certification, accreditation). If used in a public accountability program, provide name of program(s), locations, Web page URL(s): This measure may be used in a Maintenance of Certification program.

3b. Usefulness for Quality Improvement: H☐ M☐ L☐ I ☐
(*The measure is meaningful, understandable and useful for quality improvement.*)

3b.1. Use in QI. If used in quality improvement program, provide name of program(s), locations, Web page URL(s):
[*For Maintenance – If not used for QI, indicate the reasons and describe progress toward using performance results for improvement*].
All PCPI measures are suitable for use in quality improvement initiatives and are made freely available on the PCPI website and through the implementation efforts of medical specialty societies and other PCPI members. The PCPI strongly encourages the use of its measures in QI initiatives and seeks to provide information on such initiatives to PCPI members.

3b.2. Provide rationale for why the measure performance results are meaningful, understandable, and useful for quality improvement. If usefulness was demonstrated *(e.g., QI initiative),* describe the data, method and results:
The PCPI believes that the use of PCPI measures in quality improvement initiatives is a beneficial way to gather scientific data with which to improve physician performance. This is appropriate since the measure has been tested and the reliability of the performance data has been validated. NQF endorsement will facilitate our ongoing progress toward this quality improvement objective.

Overall, to what extent was the criterion, *Usability*, met? H☐ M☐ L☐ I ☐
Provide rationale based on specific subcriteria:

4. FEASIBILITY

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

4a. Data Generated as a Byproduct of Care Processes: H☐ M☐ L☐ I ☐

4a.1-2 How are the data elements needed to compute measure scores generated? *(Check all that apply).*
Data used in the measure are:
generated by and used by healthcare personnel during the provision of care, e.g., blood pressure, lab value, medical condition

4b. Electronic Sources: H☐ M☐ L☐ I ☐

4b.1 Are the data elements needed for the measure as specified available electronically *(Elements that are needed to compute measure scores are in defined, computer-readable fields)*: ALL data elements in electronic health records (EHRs)

4b.2 If ALL data elements are not from electronic sources, specify a credible, near-term path to electronic capture, OR provide a rationale for using other than electronic sources:

4c. Susceptibility to Inaccuracies, Errors, or Unintended Consequences: H☐ M☐ L☐ I ☐

4c.1 Identify susceptibility to inaccuracies, errors, or unintended consequences of the measurement identified during testing and/or operational use and strategies to prevent, minimize, or detect. If audited, provide results:
We are not aware of any unintended consequences related to this measurement.

4d. Data Collection Strategy/Implementation: H☐ M☐ L☐ I ☐

A.2 Please check if either of the following apply (*regarding proprietary measures*):
4d.1 Describe what you have learned/modified as a result of testing and/or operational use of the measure regarding data collection, availability of data, missing data, timing and frequency of data collection, sampling, patient confidentiality, time and cost of data collection, other feasibility/implementation issues *(e.g., fees for use of proprietary measures)*:
This measure was found to be reliable and feasible for implementation.

Overall, to what extent was the criterion, *Feasibility*, met? H☐ M☐ L☐ I ☐
Provide rationale based on specific subcriteria:

OVERALL SUITABILITY FOR ENDORSEMENT

Does the measure meet all the NQF criteria for endorsement? Yes☐ No☐
Rationale:

If the Committee votes No, STOP.
If the Committee votes Yes, the final recommendation is contingent on comparison to related and competing measures.

5. COMPARISON TO RELATED AND COMPETING MEASURES

If a measure meets the above criteria and there are endorsed or new related measures (either the same measure focus or the same target population) or competing measures (both the same measure focus and the same target population), the measures are compared to address harmonization and/or selection of the best measure before a final recommendation is made.

5.1 If there are related measures *(either same measure focus or target population)* or competing measures *(both the same measure focus and same target population*), list the NQF # and title of all related and/or competing measures:

5a. Harmonization

5a.1 If this measure has EITHER the same measure focus OR the same target population as NQF-endorsed measure(s): Are the measure specifications completely harmonized?

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

5b. Competing Measure(s)

5b.1 If this measure has both the same measure focus and the same target population as NQF-endorsed measure(s): Describe why this measure is superior to competing measures *(e.g., a more valid or efficient way to measure quality);* OR provide a rationale for the additive value of endorsing an additional measure. *(Provide analyses when possible)*:

CONTACT INFORMATION

Co.1 Measure Steward (Intellectual Property Owner): American Medical Association - Physician Consortium for Performance Improvement, 515 N State Street, Chicago, Illinois, 60654

Co.2 Point of Contact: Mark S., Antman, DDS, MBA, Director, Measure Development Operations Performance Improvement, mark.antman@ama-assn.org, 312-464-5056-

Co.3 Measure Developer if different from Measure Steward: American Medical Association - Physician Consortium for Performance Improvement, 515 N State Street, Chicago, Illinois, 60654

Co.4 Point of Contact: Mark S., Antman, DDS, MBA, Director, Measure Development Operations Performance Improvement, mark.antman@ama-assn.org, 312-464-5056-

Co.5 Submitter: Molly, Siegel, molly.siegel@ama.assn-org, 312-464-4901-, American Medical Association - Physician Consortium for Performance Improvement

Co.6 Additional organizations that sponsored/participated in measure development:
College of American Pathologists

Co.7 Public Contact: Mark, Antman, mark.antman@ama-assn.org, 312-464-5056-, American Medical Association - Physician Consortium for Performance Improvement

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.
David L. Witte, MD, PhD, FCAP (Co-Chair, pathology)
Susan R. Snyder, PhD, MBA (Co-Chair, methodology)
Nancy Baxter, MD, PhD (colorectal surgery)
Joel V. Brill, MD, AGAF, FACG, CHCQM (gastroenterology)
Patrick Fitzgibbons, MD, FCAP (pathology)
M. Kay Washington, MD, PhD, FCAP (pathology)
Mario Gonzalez, MD, FCAP, FASCP (pathology)
Richard M Gore, MD, FACR (diagnostic radiology)
Dana Marie Grzybicki, MD, PhD (pathology)
Harvey W. Kaufman, MD, FCAP(pathology)
Jonathon Myles, MD, FCAP (pathology)
Raouf E. Nakhleh, MD, FCAP (pathology)
Felicia Nicholson, RN, BSN (health plan representative)
Omar Yousef, MD, FCAP (pathology)

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.

Ad.2 If adapted, provide title of original measure, NQF # if endorsed, and measure steward. Briefly describe the reasons for adapting the original measure and any work with the original measure steward:

Measure Developer/Steward Updates and Ongoing Maintenance
Ad.3 Year the measure was first released: 2007
Ad.4 Month and Year of most recent revision: 09, 2010
Ad.5 What is your frequency for review/update of this measure? Usually on a three-year cycle, when feasible
Ad.6 When is the next scheduled review/update for this measure? 12, 2012

Ad.7 Copyright statement:

Ad.8 Disclaimers:

Ad.9 Additional Information/Comments:

Date of Submission (*MM/DD/YY*): 10/03/2011

Basic Measure Calculation:

$$\frac{(N)}{(D) - (E)} = \%$$

The PCPI strongly recommends that exception rates also be computed and reported alongside performance rates as follows:

Exception Calculation:

$$\frac{(E)}{(D)} = \%$$

Exception Types:
E= E1 (Medical Exceptions) + E2 (Patient Exceptions) + E3 (System Exceptions)
For patients who have more than one valid exception, only one exception should be be counted when calculating the exception rate

Initial Patient Population (IPP)	Denominator (D)	Numerator (N)	Denominator Exceptions (E)
Definition: The initial patient population identifies the general group of patients that the performance measureis designed to address; usually focused on a specific clinical condition (e.g., coronary artery disease, asthma). For example, a patient aged 18 years and older with a diagnosis of CADwho has at least 2 Visits during the measurement period.	**Definition:** The denominator defines the specific group of patients for inclusion in a specific performance measure based on specific criteria (e.g., patient's age, diagnosis, prior MI). In some cases, the denominator may be I dentical to the initial patient population.	**Definition:** The numerator defines the group of patients in the denominator for whom a process or outcome of care occurs (e.g., flu vaccine received).	**Definition:** Denominator exceptions are the valid reasons why patients who are included in the denominator population did not receive a process or outcome of care (described in the numerator). Patients may have Denominator Exceptions for medical reasons (e.g., patient has an egg allergy so they did not receive flu vaccine); patient reasons (e.g., patient declined flu vaccine); or system reasons (e.g., patient did not receive flu Vaccine due to vaccine shortage). These cases are removed from the denominator population for the performance calculation, however the number of patients with valid exceptions should be calculated and reported. This group of patients constitutes the Denominator Exception reporting population – patients for whom the numerator was not achieved and a there is a valid Denominator Exception.
Find the patients who meet the Initial Patient Population criteria (IPP)	Find the patients who qualify for the denominator (D): o From the patients within the Patient Population criteria (IPP) select those people who meet Denominator selection criteria. (In some cases the IPP and D are identical).	Find the patients who qualify for the Numerator (N): o From the patients within the Denominator (D) criteria, select those people who meet Numerator selection criteria. o Validate that the number of patients in the numerator is less than or equal to the number of patients in the denominator	From the patients who did not meet the Numerator criteria, determine if the patient meets any criteria for the Denominator Exception (E1 + E2+E3). If they meet any criteria, they should be removed from the Denominator for performance calculation. As a point of reference, these cases are removed from the denominator population for the performance calculation, however the number of patients with valid exceptions should be calculated and reported.