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 #: 0034 **NQF Project:** Population Health: Prevention Project

(for Endorsement Maintenance Review)
Original Endorsement Date: Aug 10, 2009 **Most Recent Endorsement Date:** Aug 10, 2009

BRIEF MEASURE INFORMATION

De.1 **Measure Title:** Colorectal Cancer Screening

Co.1.1 **Measure Steward:** National Committee for Quality Assurance

De.2 **Brief Description of Measure:** The percentage of members 50–75 years of age who had appropriate screening for colorectal cancer.

2a1.1 **Numerator Statement:** One or more screenings for colorectal cancer. Appropriate screenings are defined by any one of the four criteria below:
•fecal occult blood test (FOBT) during the measurement year
•flexible sigmoidoscopy during the measurement year or the four years prior to the measurement year
•double contrast barium enema (DCBE) during the measurement year or the four years prior to the measurement year.
•Colonoscopy during the measurement year or the nine years prior to the measurement year

2a1.4 **Denominator Statement:** Patients 51–75 years of age as of December 31 of the measurement year.

2a1.8 **Denominator Exclusions:** Patients with a diagnosis of colorectal cancer or total colectomy. Look for evidence of colorectal cancer or total colectomy as far back as possible in the patient's history, through either administrative data or medical record review. Exclusionary evidence in the medical record must include a note indicating a diagnosis of colorectal cancer or total colectomy, which must have occurred by December 31 of the measurement year.

1.1 **Measure Type:** Process
2a1. 25-26 **Data Source:** Administrative claims, Electronic Clinical Data : Imaging/Diagnostic Study, Electronic Clinical Data : Laboratory, Paper Records
2a1.33 **Level of Analysis:** Clinician : Group/Practice, Clinician : Individual, Clinician : Team, Health Plan

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 : Colorectal, Prevention
De.5 **Cross Cutting Areas** (*Check all the areas that apply)*: Population Health

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)*:
An estimated 142,570 men and women were diagnosed with colon cancer in 2010. In the same year, 51,370 were estimated to have died from the disease, making colorectal cancer the second leading cause of cancer death in the United States (National Cancer Institute 2001; American Cancer Society 2011a).

Screening for colorectal cancer is extremely important as there are no signs or symptoms of the cancer in the early stages. Treatment in the disease's earliest stage is highly successful, with a five-year survival rate of 74 percent (American Cancer Society, 2011b).

Most colorectal cancers occur in people without a family history of colorectal cancer. While screening is extremely effective in detecting colorectal cancer, it remains underutilized (American Cancer Society, 2011c).

Studies have shown that the cost-effectiveness of colorectal cancer screening is $40,000 per life year gained, which by comparison is similar to; the cost-effectiveness of breast cancer mammography. (Hawk 2005)

Half of American adults do not receive the necessary colorectal cancer screening (Centers for Disease Control and Prevention, 2011). Fecal occult blood tests, colonoscopy, and flexible sigmoidoscopy are shown to be effective screening methods (American Cancer Society, 2011c). Colorectal screen of individuals with no symptoms can identify polyps whose removal can prevent more than 90 percent of colorectal cancers (Rozen, 2004).

1a.4 **Citations for Evidence of High Impact cited in 1a.3:**
1. National Cancer Institute. SEER Statistical Fact Sheets: Colon and Rectum. http://seer.cancer.gov/statfacts/html/colorect.html (May 2011).

2. American Cancer Society.: Cancer Facts and Figures 2010. http://www.cancer.org/acs/groups/content/@nho/documents/document/acspc-024113.pdf (May 2011).

3. American Cancer Society. Colorectal Cancer. What are the survival rates for colorectal cancer by state? http://www.cancer.org/Cancer/ColonandRectumCancer/DetailedGuide/colorectal-cancer-survival-rates. (May 2011).

4. American Cancer Society. Colorectal Cancer. What are the risk factors for colorectal cancer? http://www.cancer.org/Cancer/ColonandRectumCancer/DetailedGuide/colorectal-cancer-risk-factors. (May 2011).

5. Hawk ET, Levin B. Colorectal cancer prevention. J Clin Oncol. 2005;23:378-388.

6. Centers for Disease Control and Prevention. National Health Interview Survey. http://www.cdc.gov/nchs/nhis.htm (May 2011).

7. Rozen, P. 2004. Cancer of the gastrointestinal tract: early detection or early prevention? Eur J Cancer Prev 13(1):71-5.

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:
Reductions in the deaths associated with colorectal cancer. Decreases in medical costs associated with colorectal cancer.

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.***]**
Commercial
NQF #0034, COL - Reported Rate
Data Element; 2009; 2008; 2007
N; 243; 153; 351
MEAN; 60.7; 59.1; 56.3
STDEV; 9.81; 8.56; 9.62
STDERR; 0.63; 0.69; 0.51
MIN; 27.2; 32.8; 23.6
MAX; 81.3; 78.7; 77.9
P10; 47; 49.1; 43.9
P25; 56.1; 54.5; 50.2
P50; 61.3; 59.4; 56.3
P75; 67.4; 64; 63.3
P90; 72.3; 69.4; 69.3

Medicare
NQF #0034, COL - Reported Rate
Data Element; 2009; 2008; 2007
N; 296; 260; 234
MEAN; 54.9; 53.1; 50.4
STDEV; 14.4; 14.6; 14.7
STDERR; 0.84; 0.9; 0.96
MIN; 19.3; 16; 12.5
MAX; 90.5; 89.5; 85.6
P10; 34.8; 33.3; 31.3
P25; 45.5; 43.1; 39.7
P50; 54.7; 54.2; 50.9
P75; 65.3; 64.2; 61.9
P90; 73.8; 72.7; 69.1

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***]**
Section 1b.2 references data from the most recent three years of measurement for HEDIS. Some rates and measures are new, therefore data might only be available for one or two years. The data in section 1b.2 includes percentiles, mean, min, max, standard deviation and standard error. There were (Number from below) submissions for this measure/rate.

Rate	Frequency	Percent
COL - Reported Rate	1599	5.04

1b.4 Summary of Data on Disparities by Population Group: [*For Maintenance* –*Descriptive statistics for performance results for this measure by population group***]**
African Americans have the highest rate of colorectal cancer within the US and Jews of Eastern European descent have the highest rates among ethnic groups. (American Cancer Society 2008)
While colon and rectal cancer incidences for African Americans have leveled in the past 20 years, the disparity between African

Americans and white Americans has grown. Colorectal Cancer is the third leading cancer that affects African Americans. It is estimated that a little less than 50,000 people will die from colorectal cancer; of those deaths 7,070 will be African Americans. While these deaths rates have declined in the past years, they have done so at a slower rate than among white Americans. Disparities also exist in survival rates. This is partially due to diagnosis at later stages along with problems of access to and receipt of effective treatment (American Cancer Society 2008).

While not as severe as the disparities among African Americans and whites, there are disparities in screening rates among Hispanics and whites. Hispanics are more likely to be diagnosed at later stages of cancer and thus have a lower probability of survival. Late diagnosis is attributed to lower utilization of screening among Hispanics and less access to and usage of treatment. (American Cancer Society – Hispanics 2006)

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*]

American Cancer Society. Cancer Facts & Figures 2008. Atlanta: American Cancer Society; 2008. http://www.cancer.org/downloads/STT/2008CAFFfinalsecured.pdf

American Cancer Society. Cancer Facts & Figures for Hispanics/Latinos 2006-2008. Atlanta: American Cancer Society; 2006. http://www.cancer.org/downloads/STT/CAFF2006HispPWSecured.pdf

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)*:

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)*:
The United States Preventive Services Task Force (1)
The USPSTF recommends screening for colorectal cancer in adults, beginning at age 50 years and continuing until age 75 years. (A recommendation)

The USPSTF recommends against routine screening for colorectal cancer in adults age 76 to 85 years. There may be considerations that support colorectal cancer screening in an individual patient (C recommendation)

The USPSTF recommends against screening for colorectal cancer in adults older than age 85 years (D recommendation)

The American Cancer Society, The American College of Radiology, and the U.S. Multi-Society Task Force on Colorectal Cancer~ (2)
Beginning at age 50, both men and women at average risk for developing colorectal cancer should use one of the screening tests [noted].

Institute for Clinical Systems Improvement (3)
Routine screening for colorectal cancer. The patient meets the following criteria: 1) 50 years old, or if African American, 45 years old. 2) No personal history of polyps and/or colorectal cancer. 3) No personal history of inflammatory bowel disease. 4) No family history of colorectal cancer in: One first-degree relative diagnosed before age 60 or Two first-degree relatives diagnosed at any age; 5) No family history of adenomatous polyps in: One first-degree relative diagnosed before age 60(A single first-degree relative diagnosed with colorectal cancer after age 60 may put the patient at a slightly increased risk and may warrant starting colorectal cancer screening at age 40. A single first-degree relative with an adenomatous polyp diagnosed after age 60 may put the patient at a slightly increased risk and may also warrant starting colorectal cancer screening at age 40.

Kaiser Permanente Care Management Institute(4)
Colorectal cancer screening is strongly recommended for all asymptomatic, average-risk adults. (Evidence-based: A)
In the absence of sufficient evidence, the following ages at which to begin and end colorectal cancer screening in asymptomatic average-risk adults are recommended: 1) Initiation of screening is recommended at age 50. (Consensus-based) 2) Discontinuation of screening is generally recommended at age 80. The decision to discontinue screening should be based on physician judgment, patient preference, the increased risk of complications in older adults, and existing comorbidities. (Consensus-based)

1c.5 Quantity of Studies in the Body of Evidence *(Total number of studies, not articles)*: Refer to USPSTF http://www.uspreventiveservicestaskforce.org/uspstf08/colocancer/coloartwhit.htm

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*): High

1c.7 Consistency of Results across Studies *(Summarize the consistency of the magnitude and direction of the effect):*

1c.8 Net Benefit *(Provide estimates of effect for benefit/outcome; identify harms addressed and estimates of effect; and net benefit - benefit over harms)*:
The available evidence usually includes consistent results from well-designed, well-conducted studies in representative primary care populations. These studies assess the effects of the preventive service on health outcomes. This conclusion is therefore unlikely to be strongly affected by the results of future studies.

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:

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

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

1c.13 Grade Assigned to the Body of Evidence: A

1c.14 Summary of Controversy/Contradictory Evidence: Recently published conflicting guidelines from A) the United States Preventive Services Task Force (USPSTF) (1) and B) from the American Cancer Society, American College of Radiology and U.S. Multi-Society Task Force on Colorectal Cancer ("joint guideline") (2).
Conflicting guidelines over age range:
? Joint guideline: begin screening at age 50 (no explicit upper age limit)

? USPSTF: screen patients age 50-75, do not routinely screen patients age 76-80, and do not screen patients age >80.

In staying in line with the USPSTF recommendations, the age range has been changed to include appropriate screenings for people age 50-75 and not appropriate for those 76 and older. Though the joint guideline is silent on an upper age limit, the guidelines stress that screening should end at a point where curative therapy would not be offered due to life-limiting co-morbidity.
Conflicting guidelines over approved modalities:
? Joint guideline: approves use of fecal DNA (sDNA), computed tomographic colonography (CTC), and DCBE.
? USPSTF: evidence is insufficient to assess the benefits and harms of sDNA and CTC. The USPSTF did not review DCBE due to its substantially lower sensitivity relative to modern tests, lack of screening trials, and decreased usage.

Computed tomography (CT) colonography: We are recommending to add CTC as an acceptable modality for screening in light of the Joint guideline's interpretation of the evidence and to encourage colorectal cancer screening, which is currentlly underused,

1c.15 Citations for Evidence other than Guidelines*(Guidelines addressed below)*:
Institute for Clinical Systems Improvement (ICSI). Colorectal cancer screening. Bloomington (MN): Institute for Clinical Systems Improvement (ICSI); 2010 May. 27 p. [57 references]

1c.16 Quote verbatim, <u>the specific guideline recommendation</u> *(Including guideline # and/or page #)*:
ACS/USMSTF/ACR (2008)
Testing Options for the Early Detection of Colorectal Cancer and Adenomatous Polyps for Asymptomatic Adults Aged 50 Years and Older
Tests that Detect Adenomatous Polyps and Cancer
- FSIG every 5 years
- Colonoscopy every 10 years
- DCBE every 5 years
- CTC every 5 years

Tests that Primarily Detect Cancer
- Annual gFOBT with high test sensitivity for cancer
- Annual FIT with high test sensitivity for cancer
- sDNA test with high sensitivity for cancer, interval uncertain

Screening Tests for the Detection of CRC
gFOBT — Conclusions and Recommendations. Annual screening with high-sensitivity gFOBT (such as Hemoccult SENSA) that have been shown in the published, peer-reviewed literature to detect a majority of prevalent CRC in an asymptomatic population is an acceptable option for colorectal screening in average-risk adults aged 50 years and older. Any positive test should be followed up with colonoscopy. Individuals should be informed that annual testing is necessary to achieve the fullest potential of this test and that they will need follow-up colonoscopy if test results are positive. Screening for CRC with gFOBT in the office following DRE or as part of a pelvic examination is not recommended and should not be done. Commonly used guaiac tests, with or without rehydration, that have not been shown in the literature to detect a majority of prevalent CRC at the time of testing are no longer recommended.
FIT — Conclusions and Recommendations. Annual screening with FIT that have been shown in the published, peer-reviewed literature to detect a majority of prevalent CRC in an asymptomatic population at the time of testing is an acceptable option for colorectal screening in average risk adults aged 50 years and older. Any positive test should be followed up with colonoscopy. Adults should be informed that annual testing is necessary to achieve the fullest potential of this test and that they will need follow-up colonoscopy if test results are positive.
sDNA — Conclusions and Recommendations. In previous assessments of the performance of sDNA, both the ACS and the USMSTF concluded that data were insufficient to recommend screening with sDNA for average-risk individuals. Based on the accumulation of evidence since the last update of these guidelines, the panel concluded that there now are sufficient data to include sDNA as an acceptable option for CRC screening. As noted above, testing stool for molecular markers is an evolving technology. New iterations of these tests, either technological enhancements of existing tests or completely new test variants, should be carefully evaluated in order to determine that they meet the criteria of detecting a majority of cancers at the time of screening but also have acceptable performance in a screening cohort. While the manufacturer of the one test that is commercially available currently is recommending a 5-year interval for routine screening between examinations with normal results, the panel concluded that there were insufficient data upon which to endorse this interval. Such an interval was judged by the committee to be appropriate only for a test that has very high sensitivity for both cancer and adenomatous polyps—a standard that has not been

documented for sDNA to date. At this time, further research is needed to determine the interval between negative sDNA exams. Based on current evidence, the appropriate interval is uncertain.

Tests for the Detection of Adenomas and CRC

FSIG — Conclusion and Recommendations. FSIG can result in the identification of the majority of prevalent CRC at the time of screening, when the examination reaches the splenic flexure or beyond 40 cm as a reasonable target for insertion and when adenomas in the distal colon are used as an indication for the need for colonoscopy. Although the appropriate interval between normal examinations is uncertain, FSIG is recommended to be performed for screening every 5 years in most clinical settings due to concerns about exam quality and completeness. FSIG can be performed alone, or consideration can be given to combining FSIG performed every 5 years with a highly sensitive gFOBT or FIT performed annually. In high-quality centers (such as the program operated by Kaiser Permanente in California) where procedures are conducted by properly trained and experienced endoscopists who document regular insertion beyond 40 cm with a good bowel preparation, a 10-year interval between negative exams may be reasonable. Individuals should be informed about the limitations of FSIG, including the fact that it examines only the distal colon; that there is a risk, albeit small, of perforation; and that they may experience discomfort during and after the examination. Patients should also understand that the examination achieves higher quality when bowel cleansing follows the same protocol as that for colonoscopy. Finally, patients should be informed that positive test findings will need to be followed up with colonoscopy.

Colonoscopy—Conclusions and Recommendations. The evidence base to support screening colonoscopy, though indirect, is substantial. The appropriate interval between negative colonoscopy screening exams is uncertain because of lack of long-term follow-up data. At present, colonoscopy every 10 years is an acceptable option for CRC screening in average-risk adults beginning at age 50 years. Individuals should be informed about the limitations of colonoscopy, including the fact that it may miss some cancers and significant adenomas and that there is a risk, albeit small, of perforation, hemorrhage (following polypectomy), subsequent hospitalization, and in very rare circumstances, more serious harms. A full bowel cleansing is necessary prior to colonoscopy. Sedation usually is used to minimize discomfort during the examination, and thus a chaperone is required to provide transportation after the examination.

Imaging Examinations of the Colon and Rectum—DCBE and Computed Tomography

DCBE — Conclusions and Recommendations. DCBE every 5 years is an acceptable option for CRC screening in average-risk adults aged 50 years and older. Discussions with patients should include a description of the test characteristics, the importance of adherence to a thorough colon cleansing, test accuracy, the likelihood of a positive test, and the need for subsequent colonoscopy if the test is abnormal. The choice of DCBE for screening can be made on an individual basis, depending on factors such as personal preference, cost, and the local availability of trained radiologists able to offer a high-quality examination.

CTC — Conclusions and Recommendations. In terms of detection of colon cancer and advanced neoplasia, which is the primary goal of screening for CRC and adenomatous polyps, recent data suggest CTC is comparable to optical colonoscopy for the detection of cancer and polyps of significant size when state-of-the-art techniques are applied. In previous assessments of the performance of CTC, the ACS concluded that data were insufficient to recommend screening with CTC for average-risk individuals. Based on the accumulation of evidence since that time, the expert panel concludes that there are sufficient data to include CTC as an acceptable option for CRC screening.

Screening of average-risk adults with CTC should commence at age 50 years. The interval for repeat exams after a negative CTC has not been studied and is uncertain. However, if current studies confirm the previously reported high sensitivity for detection of cancer and of polyps 6 mm, it would be reasonable to repeat exams every 5 years if the initial CTC is negative for significant polyps until further studies are completed and are able to provide additional guidance. Until there is more research on the safety of observation, colonoscopy should be offered to patients whose largest polyp is 6 mm or greater. CTC surveillance could be offered to those patients who would benefit from screening but either decline colonoscopy or who are not good candidates for colonoscopy for one or more reasons. However, if colonoscopy is contraindicated because the patient is not likely to benefit from screening due to life-limiting comorbidity, then neither CTC nor any other CRC screening test would be appropriate.

KPCMI (2008)

Recommendation: Effectiveness of Colorectal Cancer Screening Tests

A. CRC screening is strongly recommended for all asymptomatic, average-risk adults. (Evidence-based: A)

B. Any of the following tests are acceptable for CRC screening in asymptomatic, average-risk adults:*

- High-sensitivity fecal occult blood test. (Consensus-based)
- Immunochemical fecal occult blood test (iFOBT/FIT).** (Consensus-based)
- FSIG. (Evidence-based: B)
- Colonoscopy.** (Consensus-based)
- A combination of high-sensitivity gFOBT test and FSIG. (Consensus-based)

C. The following additional screening tests are either less-preferred options or not recommended for screening. However, an adult who has had one of these tests is considered screened. Follow-up screening using a preferred option is recommended.

- An annual standard gFOBT is a less-preferred option.*** (Consensus-based)
- ACBE is not recommended as a screening strategy for average-risk adults. (Evidence-based: I)
- Virtual colonoscopy is not recommended as a screening strategy for average-risk adults.* (Consensus-based)
- Fecal DNA is not recommended as a screening strategy for average-risk adults.****(Consensus-based)

Note: For fecal blood tests, inform patients of the potential risks associated with false-positive test and false-negative test results, as well as the need for prompt follow-up of a positive test result. For FSIG, inform patients that the test has a small risk of complications and is not a complete examination of the entire colon.

*There is insufficient evidence to choose one screening test over another.

**If a patient has had a normal colonoscopy within the last 10 years, there is insufficient evidence that supplemental FOBT adds any incremental benefit.

***Even though there is sufficient evidence in support of this screening modality, it is not a preferred option due to its low sensitivity and low compliance rates.

****Please note that fecal DNA testing and virtual colonoscopy are not listed as "appropriate screening tests" in 2008 HEDIS (Health Plan Employer Data and Information Set) specifications for colorectal cancer screening, and therefore regions may choose to screen members with other appropriate tests.

Recommendation: Frequency of Colorectal Cancer Screening

A. The following intervals for colorectal cancer screening in asymptomatic, average-risk adults are recommended*:

- FSIG: at least every 10 years. (Consensus-based)
- High-sensitivity guaiac or immunochemical FOBT (iFOBT/FIT): every 1-2 years. (Consensus-based)
- Colonoscopy: every 10 years. (Consensus-based)
- Combined FOBT and FSIG: every 1-2 years for FOBT, at least every 10 years for flexible sigmoidoscopy. (Consensus-based)

B. The following additional screening tests are either less-preferred options or not recommended for screening. However, if these tests are performed, then the recommended intervals are as indicated below. Follow-up screening using a preferred option is recommended.

- Standard gFOBT: every 1-2 years. (Consensus-based)
- ACBE:** every 5 years. (Consensus-based)
- Virtual colonoscopy:** every 10 years. (Consensus-based)
- Fecal DNA:** every 5 years. (Consensus-based)

* The GDT recognizes that these screening intervals differ from current HEDIS measures. Some regions may choose to offer screening at more frequent intervals. HEDIS intervals are as follows: FOBT (annual), flexible sigmoidoscopy (every 5 years), air contrast barium enema (every 5 years), colonoscopy (every 10 years).

**These modalities are not recommended for screening average-risk adults (see Recommendation #2 above).

Recommendation: Age to Begin and End Colorectal Cancer Screening

In the absence of sufficient evidence, the following ages at which to begin and end colorectal cancer screening in asymptomatic average-risk adults are recommended:

A. Initiation of screening is recommended at age 50. (Consensus-based)

B. Discontinuation of screening is generally recommended at age 75, provided that there is a history of routine screening. For those with no history of routine screening, discontinuation is recommended at age 80. The decision to discontinue screening should be based on physician judgment, patient preference, the increased risk of complications in older adults, and existing comorbidities. (Consensus-based)

USPSTF (2008)

Summary of Recommendations

- The USPSTF recommends screening for CRC using FOBT, sigmoidoscopy, or colonoscopy, in adults, beginning at age 50 years and continuing until age 75 years. Grade: A recommendation.
- The USPSTF recommends against routine screening for CRC cancer in adults age 76 to 85 years. There may be considerations that support colorectal cancer screening in an individual patient. Grade: C recommendation.
- The USPSTF recommends against screening for CRC in adults older than age 85 years. Grade: D recommendation.
- The USPSTF concludes that the evidence is insufficient to assess the benefits and harms of CTC and fecal DNA testing as screening modalities for colorectal cancer. Grade: I statement.

Patient Population Under Consideration

These recommendations apply to adults 50 years of age and older, excluding those with specific inherited syndromes (the Lynch syndrome or familial adenomatous polyposis) and those with inflammatory bowel disease. The recommendations do apply to those with first-degree relatives who have had colorectal adenomas or cancer, although for those with first-degree relatives who

developed cancer at a younger age or those with multiple affected first-degree relatives, an earlier start to screening may be reasonable. Data suggest that colorectal cancer has a higher mortality rate in African Americans. The reasons for this differential are not well known, and the recommendations are intended to apply to all ethnic and racial groups.
When the screening test results in the diagnosis of clinically significant colorectal adenomas or cancer, the patient will be followed by a surveillance regimen and recommendations for screening are no longer applicable. The USPSTF did not address evidence for the effectiveness of any particular surveillance regimen after diagnosis and/or removal of adenomatous polyps.
Screening Tests
The relative sensitivity and specificity of the different colorectal screening tests with adequate data to assess cancer detection — colonoscopy, FSIG, and fecal tests — can be depicted as follows:
Sensitivity: Hemoccult II < FIT = Hemoccult SENSA ~ FSIG < colonoscopy
Specificity: Hemoccult SENSA < FIT ~ Hemoccult II < FSIG = colonoscopy
For the operator-dependent tests—FSIG, CT colonography, and colonoscopy—better operator training and more experience have a high likelihood of improving sensitivity. Approaches related to certification, such as quality standards and possibly minimum volume requirements, could be used to achieve the goal of improving operator performance and therefore test sensitivity. Assurance of performance of high-quality endoscopy should be part of all screening programs.
Because several screening strategies have similar efficacy, efforts to reduce colon cancer deaths should focus on implementation of strategies that maximize the number of individuals who get screening of some type. The different options for CRC screening tests are variably acceptable to patients; eliciting patient preferences is one step in improving adherence. Ideally, shared decision making between clinicians and patients would incorporate information on local test availability and quality as well as patient preference.
Screening Intervals and Starting and Stopping Ages
Screening programs incorporating FOBT, FSIG, or colonoscopy will all be effective in reducing mortality. Modeling evidence suggests that population screening programs between the ages of 50 and 75 years using any of the following 3 regimens will be approximately equally effective in life-years gained, assuming 100% adherence to the same regimen for that period: 1) annual high-sensitivity FOBT, 2) sigmoidoscopy every 5 years combined with high-sensitivity FOBT every 3 years, and 3) screening colonoscopy at intervals of 10 years.
The strategies differ in the total number of colonoscopies that would be required to gain similar numbers of life-years. The first strategy, use of annual high-sensitivity FOBT (sensitivity for cancer > 70%) that has a false-positive rate less than 10% (that is, specificity > 90%), is estimated to require the fewest colonoscopies while achieving a gain in life-years similar to that seen with screening colonoscopy every 10 years. Currently available tests that meet both specifications include SENSA guaiac testing (Beckman Coulter, Fullerton, California) and FIT with characteristics similar to those of the Magstream quantitative test (Fujirebio, Tokyo, Japan).
Although use of an annual FOBT with a lower sensitivity has been demonstrated to reduce CRC mortality in randomized, controlled trials, modeling suggests that the number of life-years gained will be greater with the strategies using higher sensitivity tests.
For all screening modalities, the effectiveness decreases substantially as adherence to the regimen declines. At the individual level, adherence to a screening regimen will be more important in life-years gained than will the particular regimen selected. Current data are insufficient to predict adherence to any specific screening regimen at the population level.

1c.17 Clinical Practice Guideline Citation: 1. American Cancer Society/US Multisociety Task Force on Colorectal Cancer/American College of Radiology (ACS/USMSTF/ACR). 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 2008 May-Jun;58(3):130-60. [210 references]

2. Kaiser Permanente Care Management Institute (KPCMI). Colorectal cancer screening clinical practice guideline. Oakland (CA): Kaiser Permanente Care Management Institute; 2008 Dec. 190 p. [195 references]

3. US Preventive Services Task Force (USPSTF). Screening for colorectal cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med 2008 Nov 4;149(9):627-37.

1c.18 National Guideline Clearinghouse or other URL: http://www.uspreventiveservicestaskforce.org/uspstf/uspscolo.htm

1c.19 Grading of Strength of Guideline Recommendation. Has the recommendation been graded? No

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

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

1c.22 **If other, identify and describe the grading scale with definitions:**

1c.23 **Grade Assigned to the Recommendation:** 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. D.The USPSTF recommends against the service. There is moderate or high certainty that the service has no net benefit or that the harms outweigh the benefits.I. The USPSTF concludes that the current evidence is insufficient to assess the balance of benefits and harms of the service. Evidence is lacking, of poor quality, or conflicting, and the balance of benefits and harms cannot be determined.

1c.24 **Rationale for Using this Guideline Over Others:** It is NCQA policy to use guidelines which are evidence-based, applicable to physicians and other healthcare providers, and developed by a national specialty organization or government agency. NCQA convened an expert panel of diverse stakeholders to review the guidelines and evidence for this measure. The panel determined the measure was scientifically sound using the full body of evidence and guidelines for this measure concept.

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: High 1c.26 Quality: High1c.27 Consistency**:** High

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)***:**
One or more screenings for colorectal cancer. Appropriate screenings are defined by any one of the four criteria below:
•fecal occult blood test (FOBT) during the measurement year
•flexible sigmoidoscopy during the measurement year or the four years prior to the measurement year
•double contrast barium enema (DCBE) during the measurement year or the four years prior to the measurement year.
•Colonoscopy during the measurement year or the nine years prior to the measurement year

2a1.2 Numerator Time Window *(The time period in which the target process, condition, event, or outcome is eligible for inclusion)*:
- Fecal occult blood test (FOBT) during the measurement year.
- Flexible sigmoidoscopy during the measurement year or the four years prior to the measurement year
- Double contrast barium enema (DCBE) or air contrast barium enema during the measurement year or the four years prior to the measurement year
- Colonoscopy during the measurement year or the nine years prior to the measurement year

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*:
Appropriate screenings are defined by any one of the following criteria.
- Fecal occult blood test (FOBT) during the measurement year. Regardless of FOBT type, guaiac (gFOBT) or immunochemical (iFOBT), assume that the required number of samples was returned.
- Flexible sigmoidoscopy during the measurement year or the four years prior to the measurement year
- Double contrast barium enema (DCBE) or air contrast barium enema during the measurement year or the four years prior to the measurement year
- Colonoscopy during the measurement year or the nine years prior to the measurement year

There are two types of FOBT tests: guaiac (gFOBT) and immunochemical (iFOBT). Depending on the type of FOBT test, a certain number of samples are required for numerator compliance. Follow the instructions below to determine member compliance.
- If the medical record does not indicate the type of test and there is no indication as to how many samples were returned, assume the required number was returned. The member meets the screening criteria for inclusion in the numerator.
- If the medical record does not indicate the type of test and the number of returned samples is specified, the member would only meet the screening criteria if the number of samples specified is greater than or equal to three samples. If the number of samples is less than three, the member does not meet the screening criteria for inclusion in the numerator.
- iFOBT tests may require fewer than three samples. If the medical record indicates that an iFOBT was done, the member meets the screening criteria for inclusion in the numerator regardless of the number of returned samples.
- If the medical record indicates that a gFOBT was done, follow the scenarios below.
  - If the medical record does not indicate the number of returned samples, assume the required number was returned. The member meets the screening criteria for inclusion in the numerator.
  - If the medical record indicates that three or more samples were returned, the member meets the screening criteria for inclusion in the numerator.
  - If the medical record indicates that fewer than three samples were returned, the member does not meet the screening criteria.

FOBT: CPT codes (82270, 82274), HCPCS (G0328, G0394), ICD-9-CM Diagnosis (V76.51), LOINC (2335-8, 12503-9, 12504-7, 14563-1, 14564-9, 14565-6, 27396-1, 27401-9, 27925-7, 27926-5, 29771-3)

Flexible Signoidoscopy: CPT codes (45330-45335, 45337-45342, 45345), HCPCS codes (G0104), ICD-9-CM Procedure (45.24)

Colonoscopy: CPT codes (44388-44394, 44397, 45355, 45378-45387, 45391, 45392), HCPCS codes (G0105, G0121), ICD-9-CM Procedure (45.22, 45.23, 45.25, 45.42, 45.43)

2a1.4 Denominator Statement *(Brief, narrative description of the target population being measured)*:
Patients 51–75 years of age as of December 31 of the measurement year.

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)*:
Annually, from December 31 of the measurement year.

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)*:
Patients 51–75 years of age as of December 31 of the measurement year.

2a1.8 Denominator Exclusions *(Brief narrative description of exclusions from the target population)*:
Patients with a diagnosis of colorectal cancer or total colectomy. Look for evidence of colorectal cancer or total colectomy as far back as possible in the patient's history, through either administrative data or medical record review. Exclusionary evidence in the medical record must include a note indicating a diagnosis of colorectal cancer or total colectomy, which must have occurred by December 31 of the measurement year.

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)*:
Use the following codes or descriptions of the codes to identify allowable exclusions:

Colorectal Cancer HCPCS codes (G0213-G0215, G0231) ICD-9-CM codes (153., 154.0, 154.1, 197.5, V10.05)

Total colectomy CPT codes (44150-44153, 44155-44158, 44210-44212) ICD-9-CM codes (45.8)

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)*:
Measure is stratified by Commercial, Medicaid, and Medicare health plans.

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: Rate/proportion

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.)*:

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

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):

2a1.25 Data Source *(Check all the sources for which the measure is specified and tested*). If other, please describe:
Administrative claims, Electronic Clinical Data : Imaging/Diagnostic Study, Electronic Clinical Data : Laboratory, 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.*)**:** HEDIS data collection tool

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:

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, Health Plan

2a1.34-35 Care Setting *(Check all the settings for which the measure is specified and tested)***:** Ambulatory Care : Clinician Office

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)***:**
HEDIS Health Plan performance data 2010

2a2.2 Analytic Method *(Describe method of reliability testing & rationale)***:**
Reliability was estimated by using the beta-binomial model. Beta-binomial is a better fit when estimating the reliability of simple pass/fail rate measures as is the case with most HEDIS® health plan measures. The beta-binomial model assumes the plan score is a binomial random variable conditional on the plan´s true value that comes from the beta distribution. The beta distribution is usually defined by two parameters, alpha and beta. Alpha and beta can be thought of as intermediate calculations to get to the needed variance estimates. The beta distribution can be symmetric, skewed or even U-shaped.

Reliability used here is the ratio of signal to noise. The signal in this case is the proportion of the variability in measured performance that can be explained by real differences in performance. A reliability of zero implies that all the variability in a measure is attributable to measurement error. A reliability of one implies that all the variability is attributable to real differences in performance. The higher the reliability score, the greater is the confidence with which one can distinguish the performance of one plan from another. A reliability score greater than or equal to 0.7 is considered very good.

2a2.3 Testing Results *(Reliability statistics, assessment of adequacy in the context of norms for the test conducted)***:**
Commercial Plans 2010: reliability 0.994468

Medicaid 2010: Not available

Medicare 2010: reliability 0.993543

2b. VALIDITY. Validity, Testing, <u>including all Threats to Validity</u>: 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:**
Consistent, not differences noted.

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)***:**

2b2.2 **Analytic Method** *(Describe method of validity testing and rationale; if face validity, describe systematic assessment)*:
NCQA tested the measure for face validity using a panel of stakeholders with specific expertise in measurement. This panel included representatives from key stake holder groups, including oncologists, family practitioners, health plans, state Medicaid agencies and researchers. Experts reviewed the results of the field test and assessed whether the results were consistent with expectation, whether the measure represented quality care, and whether we were measuring the most important aspects of care in this area.

In the pilot test, we explored periodicities associated with colorectal cancer screening, as long periodicities in light of average lengths of enrollment in MCOs can be a threat to validity. We examined whether the rates of screening would differ depending on the length of time an individual had been enrolled in the plan and found little effect as shown in Table 2. Although the rates increase a small amount each year in each plan, the relative rates of screening remain about the same. The sample sizes decline significantly with increased lengths of continuous enrollment; at 10 years, only two MCOs had enough data to estimate the rate.

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)*:

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)*:

2b3.2 **Analytic Method** (*Describe type of analysis and rationale for examining exclusions, including exclusion related to patient preference)*:

2b3.3 **Results** *(Provide statistical results for analysis of exclusions, e.g., frequency, variability, sensitivity analyses):*

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)*:
No risk adjustment

2b4.2 **Analytic Method (***Describe methods and rationale for development and testing of risk model or risk stratification including selection of factors/variables)*:

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)*:

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):
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.

2b5.2 **Analytic Method** *(Describe methods and rationale to identify statistically significant and practically/meaningfully differences in performance)*:
Comparison of means and percentiles; analysis of variance against established benchmarks; if sample size is >400, we would use an analysis of variance

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)*:

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)*:
Field test data

2b6.2 **Analytic Method** *(Describe methods and rationale for testing comparability of scores produced by the different data sources specified in the measure)*:
During field testing, performance rates are calculated from administrative claims and compared to rates calculated from medical record review.

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)*:
The measure was deemed valid by NCQA's expert panels.

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)*: The measure is not stratified to detect disparities.

2c.2 **If disparities have been reported/identified (e.g., in 1b), but measure is not specified to detect disparities, please explain:**
NCQA has participated with IOM and others in attempting to include information on disparities in measure data collection. However, at the present time, this data, at all levels (claims data, paper chart review, and electronic records), is not coded in a standard manner, and is incompletely captured. There are no consistent standards for what entity (physician, group, plan, employer) should capture and report this data. While "requiring" reporting of the data could push the field forward, it has been our position that doing so would create substantial burden with inability to use the data because of its inconsistency. At the present time, we agree with the IOM report that disparities are best considered by the use of zip code analysis which has limited applicability in most reporting situations. At the health plan level, for HEDIS health plan data collection, NCQA does have extensive data related to our use of stratification by insurance status (Medicare, Medicaid and private-commercial) and would strongly recommend this process where the data base supporting the measurement includes this information. However, we believe that the measure specifications should NOT require this since the measure is still useful where the data needed to determine disparities cannot be ascertained from the data available.

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)*: Payment Program, Professional Certification or Recognition Program, Public Reporting, Quality Improvement (Internal to the specific organization), Quality Improvement with Benchmarking (external benchmarking to multiple organizations), Regulatory and Accreditation Programs

3.1 **Current Use** *(Check all that apply; for any that are checked, provide the specific program information in the following questions)*: Public Reporting, Payment Program, Public Health/ Disease Surveillance, Regulatory and Accreditation Programs, Quality Improvement with Benchmarking (external benchmarking to multiple organizations), Quality Improvement (Internal to the specific organization)

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.***]**
This measure is used in public reporting for plans only through Healthcare Effectiveness Data and Information Set (HEDIS) and is reported through venues such as the annual State of Healthcare Quality report, Quality Compass, America's Best Health Plans.

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**:**
Longstanding HEDIS measure widely reported in annual State of Health Care Quality and in Quality Compass Database, and other mediums.

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)**:**

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***]**.
This measure is a measure in the Healthcare Effectiveness Data and Information Set (HEDIS) and is used in NCQA's Health Plan Accreditation program.

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:
Widely used by health plans and providers for QI. Benchmarks and targets are set using HEDIS results.

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**:**
Coded by someone other than person obtaining original information (e.g., DRG, ICD-9 codes on claims)

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 are in a combination of electronic sources

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:
Susceptibility to inaccuracies, errors, or unintended consequences and the ability to audit the data items to detect such problems are identified. All measures that are used in NCQA programs are audited.

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

A.2 Please check if either of the following apply (*regarding proprietary measures*)**:** Proprietary measure
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)***:**

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): National Committee for Quality Assurance, 1100 13th Street NW, Suite 1000, Washington, District Of Columbia, 20005

Co.2 Point of Contact: Bob, Rehm, Assistant Vice President, Performance Measurement, Rehm@ncqa.org, 202-955-1728-

Co.3 Measure Developer if different from Measure Steward: National Committee for Quality Assurance, 1100 13th Street NW, Suite 1000, Washington, District Of Columbia, 20005

Co.4 Point of Contact: Bob, Rehm, Assistant Vice President, Performance Measurement, Rehm@ncqa.org, 202-955-1728-

Co.5 Submitter: Dawn, Alayon, MPH, CPH, Senior Health Care Analyst, alayon@ncqa.org, 202-955-3533-, National Committee for Quality Assurance

Co.6 Additional organizations that sponsored/participated in measure development:

Co.7 Public Contact: Bob, Rehm, Assistant Vice President, Performance Measurement, Rehm@ncqa.org, 202-955-1728-, 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.

Panel members:
Joel V. Brill, Predictive Health, LLC
Durado Brooks, American Cancer Society
Robert Fletcher, Harvard Medical School
William Lawrence, AHRQ Center for Outcomes and Effectiveness
T.R. Levin, Kaiser Permanente
Michael Pignone, UNC Hospital
Wvelyn Whitlock

The NCQA Colorectal Cancer Measurement Advisory Panel advised NCQA during measure development. They evaluated the way staff specified measures, assessed the content validity of measures, and reviewed field test results. As you can see from the list, the MAP consisted of a balanced group of experts, including representatives from medical research and education, cancer prevention and treatment associations, and internal medicine practitioners. Note that, in addition to the MAP, we also vetted these measures with a host of other stakeholders, as is our process. Thus, our measures are the result of consensus from a broad and diverse group of stakeholders, in addition to the MAP.

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: 2004
Ad.4 Month and Year of most recent revision: 2009
Ad.5 What is your frequency for review/update of this measure? Approximately every 3 years.

Ad.6 **When is the next scheduled review/update for this measure?**
Ad.7 **Copyright statement:** © June 29, 2011 by the National Committee for Quality Assurance 1100 13th Street, NW, Suite 1000 Washington, DC 20005
Ad.8 **Disclaimers:**
Ad.9 **Additional Information/Comments:**
Date of Submission (*MM/DD/YY*)**:** 07/12/2011