

# 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](#).

<b>NQF #:</b> 0650 <b>NQF Project:</b> Cancer Project
(for Endorsement Maintenance Review) <b>Original Endorsement Date:</b> May 05, 2010 <b>Most Recent Endorsement Date:</b> May 05, 2010
<b>BRIEF MEASURE INFORMATION</b>
<b>De.1 Measure Title:</b> Melanoma Continuity of Care – Recall System
<b>Co.1.1 Measure Steward:</b> American Medical Association - Physician Consortium for Performance Improvement
<b>De.2 Brief Description of Measure:</b> Percentage of patients, regardless of age, with a current diagnosis of melanoma or a history of melanoma whose information was entered, at least once within a 12 month reporting period into a recall system that includes: <ul style="list-style-type: none"> <li>• A target date for the next complete physical skin exam , AND</li> <li>• A process to follow up with patients who either did not make an appointment within the specified timeframe or who missed a scheduled appointment</li> </ul>
<b>2a1.1 Numerator Statement:</b> Patients whose information is entered, at least once within a 12 month period, into a recall system* that includes: <ul style="list-style-type: none"> <li>• A target date for the next complete physical skin exam , AND</li> <li>• A process to follow up with patients who either did not make an appointment within the specified timeframe or who missed a scheduled appointment</li> </ul>
<b>2a1.4 Denominator Statement:</b> All patients, regardless of age, with a current diagnosis of melanoma or a history of melanoma.
<b>2a1.8 Denominator Exclusions:</b> Documentation of system reason(s) for not entering patients into a recall system (eg, melanoma being monitored by another physician provider)
<b>1.1 Measure Type:</b> Structure <b>2a1. 25-26 Data Source:</b> Administrative claims, Electronic Clinical Data, Electronic Clinical Data : Registry, Other, Paper Records <b>2a1.33 Level of Analysis:</b> Clinician : Group/Practice, Clinician : Individual, Clinician : Team
<b>1.2-1.4 Is this measure paired with another measure?</b> No
<b>De.3 If included in a composite, please identify the composite measure (title and NQF number if endorsed):</b> This measure is not included in a composite.

<b>STAFF NOTES</b> <i>(issues or questions regarding any criteria)</i>
<b>Comments on Conditions for Consideration:</b>
<b>Is the measure untested?</b> Yes <input type="checkbox"/> No <input type="checkbox"/> If untested, explain how it meets criteria for consideration for time-limited endorsement:
<b>1a. Specific national health goal/priority identified by DHHS or NPP addressed by the measure (check De.5):</b> <b>5. Similar/related <a href="#">endorsed</a> or submitted measures (check 5.1):</b> <b>Other Criteria:</b>
<b>Staff Reviewer Name(s):</b>

**1. IMPACT, OPPORTUNITY, 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 : Skin

**De.5 Cross Cutting Areas** (Check all the areas that apply): Care Coordination

**1a.1 Demonstrated High Impact Aspect of Healthcare:** Affects large numbers, Patient/societal consequences of poor quality

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

**1a.3 Summary of Evidence of High Impact** (Provide epidemiologic or resource use data):

In the year 2010, an estimated 68,130 new cases of melanoma were diagnosed and about 8,700 patients died of the disease in the United States. However, these figures for new cases may represent a substantial underestimation, because many superficial and in situ melanomas treated in the outpatient setting are not reported. The incidence of melanoma continues to increase dramatically. Melanoma is increasing in men more rapidly than any other malignancy and, in women more rapidly than any other malignancy except lung cancer. The lifetime risk of developing melanoma in the year 2005 for someone born in the United States may be as high as one in 55. The median age at diagnosis is 59 years. As such, melanoma ranks second to adult leukemia in terms of loss of years of potential life, per death.(1)

Melanoma is among the top 10 new cancer diagnoses for both American men and women. Nationally, melanoma incidence has increased 2.4% annually in the last decade.(2)

Estimates of the overall risk of recurrence in local, cutaneous melanoma vary between 15–35%.(3)

**1a.4 Citations for Evidence of High Impact cited in 1a.3:** 1. National Comprehensive Cancer Network (NCCN). Clinical Practice Guidelines in Oncology: Melanoma. 3.2012. Available at: [www.nccn.org](http://www.nccn.org)

2. Hu A, Parmet Y, Allen G, Parker DF, et al. Disparity in Melanoma. A Trend Analysis of Melanoma Incidence and Stage at Diagnosis Among Whites, Hispanics, and Blacks in Florida. Arch Dermatol. 2009;145(12):1369-1374.

3. Mooney MM, Mettlin C, Michalek AM, Petrelli NJ, Kraybill WG. Lifelong Screening of Patients with Intermediate-Thickness Cutaneous Melanoma for Asymptomatic Pulmonary Recurrences. A Cost Effective Analysis. Cancer 1997;80:1052-64.

**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:**

Follow-up for skin examination and surveillance is an important aspect in the management of patients with a current diagnosis or a history of melanoma. The presence of a recall system, whether it is electronic or paper based, enables providers to ensure that patients receive follow-up appointments and annual screening examinations in a timely manner.

Estimates of the overall risk of recurrence in local, cutaneous melanoma vary between 15–35%. Although this risk diminishes with time, melanoma patients have been known to recur even 10–35 years after the initial diagnosis. Physicians have recommended life-long screening because of this persistent risk and the risk of developing second primary tumors.(1)

The primary goal for follow-up of patients with a history of cutaneous melanoma is early detection of surgically resectable recurrent disease and additional primary melanoma.(2)

1. Mooney MM, Mettlin C, Michalek AM, Petrelli NJ, Kraybill WG. Lifelong Screening of Patients with Intermediate-Thickness

Cutaneous Melanoma for Asymptomatic Pulmonary Recurrences. A Cost Effective Analysis. Cancer 1997;80:1052-64.

2. Bichakjian CK, Halpern AC, Johnson TM, Foote Hood A, Grichnik JM, Swetter SM, Tsao H, Barbosa VH, Chuang TY, Duvic M, Ho VC, Sober AJ, Beutner KR, Bhushan R, Smith Begolka W. Guidelines of Care for the Management of Primary Cutaneous Melanoma. Journal of the American Academy of Dermatology 2011;65(5):1032-47.

**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.]**

This measure was included in the CMS Physician Quality Reporting Initiative/System (PQRI/S) in 2009 through 2011 in the claims and registry options for 2009 (in registry only option for 2010 and beyond) as PQRI/S #137 (Melanoma Continuity of Care – Recall System). The number of professionals reporting on this measure in 2009 was approximately 2,196.

The 2009 PQRI/S Performance Rate reveals that there is a gap in care as shown by the following data: 9.19% of patients reported on did not receive the optimal care.

**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]**  
Appendix B. 2009 Physician Quality Reporting System and eRx Experience Report Detailed Tables

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

At least two of the reviewed analyses in urban counties showed that the supply of primary care physicians is less closely related to the health of urban African Americans than it is for urban whites or for African Americans in rural areas. This is likely due to the poorer distribution of primary care physicians in more deprived urban areas, with the consequently greater need to seek care in such places as hospital outpatient units and emergency rooms.(1)

Research and public education efforts have focused on melanoma prevention in white populations because of their higher risk of developing melanoma. Improved secondary prevention measures with earlier detection of thin (early-stage) melanoma likely account for the improved survival among whites from 68% in the early 1970s to 92% in recent years. Such advances, however, have not occurred in other racial and ethnic groups in the United States. Emerging data call attention to disparity in melanoma diagnosis and survival in minorities such as Hispanics and blacks. Multiple reports found that US blacks have more advanced melanoma in association with worse survival rates; however, melanoma disparity among Hispanics is less recognized. The dearth of studies on melanoma among Hispanics partly reflects the small number of cases in many areas of the United States, as well as limitations of ethnicity information in cancer registries. In fact, the National Cancer Institute’s Surveillance, Epidemiology, and End Results (SEER) program and most other cancer registries did not begin classifying data for "Hispanic" until the late 1990s. As a result, few studies included data regarding Hispanics.

**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]**

1. Starfield B, Shi L, Macinko J. Contribution of Primary Care to Health Systems and Health. The Milbank Quarterly 2005;83(3):457-502.

2. Hu A, Parmet Y, Allen G, Parker DF, et al. Disparity in Melanoma. A Trend Analysis of Melanoma Incidence and Stage at Diagnosis Among Whites, Hispanics, and Blacks in Florida. Arch Dermatol. 2009;145(12):1369-1374.

**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 <input type="checkbox"/>

NQF #0650 Melanoma Continuity of Care - Recall System

L	M-H	M	Yes <input type="checkbox"/> IF additional research unlikely to change conclusion that benefits to patients outweigh harms: otherwise No <input type="checkbox"/>
M-H	L	M-H	Yes <input type="checkbox"/> IF potential benefits to patients clearly outweigh potential harms: otherwise No <input type="checkbox"/>
L-M-H	L-M-H	L	No <input type="checkbox"/>
Health outcome – rationale supports relationship to at least one healthcare structure, process, intervention, or service		Does the measure pass subcriterion 1c? Yes <input type="checkbox"/> IF rationale supports relationship	
<p><b>1c.1 Structure-Process-Outcome Relationship</b> (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):</p> <p>This is a structure measure, which focuses on entering melanoma patients into a recall system, at least once within a one year period. Having this structure measure in place, will lead to the process of melanoma patients being screened and examined at least once a year. Having the examinations on an annual basis will improve outcomes, as it will lead to early detection of any signs or symptoms of a relapse and/or systemic spread of melanoma, therefore, potentially reducing morbidity and mortality rates.</p> <p><b>1c.2-3 Type of Evidence</b> (Check all that apply): Clinical Practice Guideline</p> <p><b>1c.4 Directness of Evidence to the Specified Measure</b> (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 evidence clearly supports the measure, which facilitates screening melanoma patients at least once a year. There is clear guideline support for the measure topic and improved outcomes for melanoma patients. There are no differences in the target population as the measure includes all melanoma patients, regardless of age.</p> <p><b>1c.5 Quantity of Studies in the Body of Evidence</b> (Total number of studies, not articles): The description of the evidence review in the guideline did not address the overall quantity of studies in the body of evidence. However, 143 articles are cited in NCCN’s melanoma guideline’s reference section. 143 articles are also cited in AAD’s Guidelines of care for the management of primary cutaneous melanoma.</p> <p><b>1c.6 Quality of Body of Evidence</b> (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 NCCN guideline recommendation is summarized according to the NCCN categories of evidence and consensus as being based on "lower-level evidence."</p> <p>The quality of the body of evidence supporting the AAD guideline recommendation is summarized according to a 3-point scale based on the quality of methodology, and is categorized as "limited-quality patient-oriented evidence."</p> <p><b>1c.7 Consistency of Results across Studies</b> (Summarize the consistency of the magnitude and direction of the effect): Although there is no explicit statement regarding the overall consistency of results across studies in the NCCN or AAD guidelines supporting this measure, the NCCN recommendation received uniform NCCN consensus that the intervention is appropriate. The AAD recommendation was ranked based on inconsistent or limited-quality patient oriented evidence.</p> <p><b>1c.8 Net Benefit</b> (Provide estimates of effect for benefit/outcome; identify harms addressed and estimates of effect; and net benefit - benefit over harms): Melanoma patients being entered into the recall system will lead to annual screening for signs/symptoms of relapse and/or systemic spread. This will lead to early detection and may reduce morbidity and mortality rates of melanoma patients. No harms have been identified.</p> <p><b>1c.9 Grading of Strength/Quality of the Body of Evidence.</b> Has the body of evidence been graded? Yes</p>			

**1c.10 If body of evidence graded, identify the entity that graded the evidence including balance of representation and any disclosures regarding bias: The AAD Work Group:**

Work Group: Christopher K. Bichakjian, MD, Allan C. Halpern, MD (Co-chair), Timothy M. Johnson, MD (Co-Chair), Antoinette Foote Hood, MD, James M. Grichnik, MD, PhD, Susan M. Swetter, MD, Hensin Tsao, MD, PhD, Victoria Holloway Barbosa, MD, Tsu-Yi Chuang, MD, MPH,, Madeleine Duvic, MD, Vincent C. Ho, MD, Arthur J. Sober, MD, Karl R. Beutner, MD, PhD, Reva Bhushan, PhD, and Wendy Smith Begolka, MS

Ann Arbor, Michigan; New York, New York; Norfolk, Virginia; Miami, Florida; Palo Alto, Los Angeles, Palm Springs, San Francisco, and Fairfield, California; Boston, Massachusetts; Chicago and Schaumburg, Illinois; Houston, Texas; and Vancouver, British Columbia, Canada

**Disclosures:**

Allan C. Halpern, MD, served on the Advisory Board for DermTech and Roche receiving other financial benefits, was a consultant with Canfield Scientific receiving other financial benefits, and was an investigator with Lucid, Inc receiving no compensation. James M. Grichnik, MD, PhD, served as founder of Digital Derm Inc receiving stock and was consultant for Genentech, MELA Science, Inc and Spectral Image, In. receiving honoraria. Hensin Tsao, MD, PhD, served as consultant for Genentech, Quest Diagnostics, SciBASE, and Metamark receiving honoraria. Victoria Holloway Barbosa, MD, served as founder of Dermal Insights Inc receiving stocks, and as consultant for L'Oreal USA receiving other benefits, and served another role with Pierre Fabre receiving other benefits. Madeleine Duvic, MD, served as investigator and on the advisory board for Allos and BioCryst receiving grants and honoraria, and as investigator and consultant for Celgeene, Kyowa Hakko Kirin Pharma, and Merck receiving grants and honoraria, serving as consulting for Dermatech, Hoffman-La Roche, Millennium Pharmaceuticals, Vertex, and Upside Endeavors, LLC, receiving honoraria; serving as consultant, investigator, and speaker for Eisai receiving grants and honoraria, serving as investigator for Eli Lilly, Genmab, Hannah Biosciences, NAAF, Hobartis, OrthoBiotech MSK, Pfizer, Sloan Kettering, Spectrum, Therakos, Topotarget, and Yapon Therapeutics receiving grants and also as investigator for NIH receiving salary; served as a speaker for P4 Healthcare and Peer Direct receiving honoraria, and, lastly, served on advisory board for Quintiles Pharma and Seattle Genetics receiving honoraria. Vincent C. Ho, MD, served on the advisory board and as an investigator and speaker for Abbott, Janssen Ortho and Schering, receiving grants and honoraria, served on advisory board and as investigator for Amgen receiving grants and honoraria, served on the advisory board for Astellas and Basilea receiving honoraria, and served as investigator for Centocor, Novartis and Pfizer receiving grants. Arthur J. Sober, MD served as a consultant for MelaScience receiving other benefits. Karl R. Beutner, MD, PhD, Chair Clinical Research Committee, served as a consultant of Anacor receiving stock, stock options, and honoraria. Christopher K. Bichakjian, MD, Timothy M. Johnson, MD, Antoinette Foote Hood, MD, Susan M. Swetter, MD, Tsu-Yi Chuang, MD, MPH, Reva Bhushan, MA, PhD, and Wendy Smith Begolka, MS, had no relevant conflicts of interest to disclose.

**NCCN Work Group:**

NCCN Guidelines Version 3.2012 Panel Members Melanoma: Daniel G. Coit, MD/Chair (Surgery/Surgical oncology), Robert Andtbacka, MD (Surgery/Surgical oncology), Christopher J. Anker, MD (Radiotherapy/Radiation oncology), Christopher K. Bivhakjian, MD (Dermatology), William E. Carson, III, MD (Surgery/Surgical oncology), Adil Daud, MD (Medical oncology, Internal medicine), Raza A. Dilawari, MD (Surgery/Surgical oncology), Dominick DiMail, MD (Pathology), Valerie Guild (Patient Advocacy), Allan C. Halpern, MD (Dermatology, Internal medicine), F. Stephen Hodi, Jr., MD (Medical oncology), Mark C. Kelley, MD (Surgery/Surgical oncology), Nikhil I. Khushalani, MD (Medical oncology), Ragini R. Kudchadkar, MD (Medical oncology), Julie R. Lange, MD, ScM (Surgery/Surgical oncology), Anne Lind, MD (Pathology), Mary C. Martini, MD (Dermatology), Anthony J. Olszanski, MD (Medical oncology), Scott K. Pruitt, MD, PhD (Surgery/Surgical oncology), Merrick I. Ross, MD (Surgery/Surgical oncology), Susan M. Swetter, MD (Dermatology), Kenneth K. Tanabe, MD (Surgery/Surgical oncology), John A. Thompson, MD (Hematology/Hematology oncology), Vijay Trisal, MD (Surgery/Surgical oncology), Marshall M. Urist, MD (Surgery/Surgical oncology), Lauren Gallagher, RPh, PhD, Maria Ho, PhD, Nicole McMillian, MS. NCCN Panel Disclosures: All member disclosures can be accessed at: [www.nccn.org](http://www.nccn.org)

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

**1c.12 If other, identify and describe the grading scale with definitions:** NCCN Categories of Evidence and Consensus

Category 1: Based upon high-level evidence, there is uniform NCCN consensus that the intervention is appropriate.

Category 2A: Based upon lower-level evidence, there is uniform NCCN consensus that the intervention is appropriate.

Category 2B: Based upon lower-level evidence, there is NCCN consensus that the intervention is appropriate.

Category 3: Based upon any level of evidence, there is major NCCN disagreement that the intervention is appropriate.

AAD Guideline (3-point scale):

I. Good-quality patient oriented evidence (i.e. evidence measuring outcomes that matter to patients: morbidity, mortality, symptom improvement, cost reduction, and quality of life).

II. Limited-quality patient-oriented evidence

III. Other evidence including consensus guidelines, opinion, case studies, or disease-oriented evidence (i.e. evidence measuring intermediate, physiologic, or surrogate end points that may or may not reflect improvements in patient outcomes).

**1c.13 Grade Assigned to the Body of Evidence:** NCCN grade 2A, AAD grade II

**1c.14 Summary of Controversy/Contradictory Evidence:** In the absence of clear data, opinions vary widely regarding the appropriate follow-up of patients with melanoma. The follow-up schedule is influenced by risk of recurrence, previous primary melanoma, and family history of melanoma; other factors, such as the presence and extent of dysplastic nevi and patient anxiety will impact follow-up schedule as well. The optimal duration of follow-up remains controversial. Although most patients who are going to recur will do so in the first five years after treatment, late recurrence (more than ten years later) is well documented especially for patients initially presenting with early-stage melanoma. It is probably not cost effective to follow all patients intensively for metastatic disease beyond five to ten years (depending on relative risk for recurrence). However, because the lifetime risk of developing a second primary melanoma is 4-8% the panel felt that a recommendation for lifetime dermatologic surveillance for melanoma patients was justified.

National Comprehensive Cancer Network (NCCN). Clinical Practice Guidelines in Oncology: Melanoma. 3.2012. Available at: [www.nccn.org](http://www.nccn.org)

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

Not applicable

**1c.16 Quote verbatim, the specific guideline recommendation (Including guideline # and/or page #):**

Skin examination and surveillance at least once a year for life is recommended for all melanoma patients, including those with stage 0, in situ melanoma. Clinicians should educate all patients about post-treatment monthly self-exam of their skin and of their lymph nodes if they had stage 1A to IV melanoma. Specific signs or symptoms are indications for additional radiologic imaging.(1)

No clear data regarding follow-up interval exists, but at least annual history and physical examination with attention to the skin and lymph nodes is recommended.(2)

Regular clinical follow-up and interval patient self exam of skin and regional lymph nodes are the most important means of detecting recurrent disease or new primary melanoma; findings from history and physical exam should direct the need for further studies to detect local, regional, and distant metastasis.(2)

**1c.17 Clinical Practice Guideline Citation:** 1. National Comprehensive Cancer Network (NCCN). Clinical Practice Guidelines in Oncology: Melanoma. 3.2012. Available at: [www.nccn.org](http://www.nccn.org)

2. Bichakjian CK, Halpern AC, Johnson TM, Foote Hood A, Grichnik JM, Swetter SM, Tsao H, Barbosa VH, Chuang TY, Duvic M, Ho VC, Sober AJ, Beutner KR, Bhushan R, Smith Begolka W. Guidelines of Care for the Management of Primary Cutaneous Melanoma. Journal of the American Academy of Dermatology 2011;65(5):1032-47.

1c.18 National Guideline Clearinghouse or other URL: [www.nccn.org](http://www.nccn.org) and <http://www.aad.org/education-and-quality-care/clinical-guidelines/current-and-upcoming-guidelines>

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: Please see section 1c.10 for NCCN and AAD guideline Work Group information.

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

1c.22 If other, identify and describe the grading scale with definitions: 1. NCCN Categories of Evidence and Consensus

Category 1: Based upon high-level evidence, there is uniform NCCN consensus that the intervention is appropriate.

Category 2A: Based upon lower-level evidence, there is uniform NCCN consensus that the intervention is appropriate.

Category 2B: Based upon lower-level evidence, there is NCCN consensus that the intervention is appropriate.

Category 3: Based upon any level of evidence, there is major NCCN disagreement that the intervention is appropriate.

2. AAD guideline recommendations were developed on the best available evidence tabled in the guideline. These are ranked as follows:

A. Recommendation based on consistent and good-quality patient-oriented evidence.

B. Recommendation based on inconsistent or limited-quality patient-oriented evidence

C. Recommendation based on consensus, opinion, case studies, or disease-oriented evidence.

1c.23 Grade Assigned to the Recommendation: NCCN grade 2A, AAD grade B

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: Moderate 1c.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? [Yes](#)

**S.2 If yes, provide web page URL:** [www.physicianconsortium.org](http://www.physicianconsortium.org)

**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):

Patients whose information is entered, at least once within a 12 month period, into a recall system\* that includes:

- A target date for the next complete physical skin exam , AND
- A process to follow up with patients who either did not make an appointment within the specified timeframe or who missed a scheduled appointment

**2a1.2 Numerator Time Window** (The time period in which the target process, condition, event, or outcome is eligible for inclusion):  
At least once 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):

Numerator Instructions:

To satisfy this measure, the recall system must be linked to a process to notify patients when their next physical exam is due and to follow up with patients who either did not make an appointment within the specified timeframe or who missed a scheduled appointment and must include the following elements at a minimum; patient identifier, patient contact information, cancer diagnosis(es), date(s) of initial cancer diagnosis (if known), and the target date for the next complete physical exam.

For Claims/Administrative:

Report CPT Category II code:

7010F -- Patient information entered into a recall system with the target date for the next complete physical skin exam specified

For EHR:

This measure does not lend itself to a “traditional specification” for EHR reporting. This is a structural measure; each facility may have a different process or software system for tracking and transmitting recalls as well as different appointment tracking systems.

**2a1.4 Denominator Statement** (Brief, narrative description of the target population being measured):

All patients, regardless of age, with a current diagnosis of melanoma or a history of melanoma.

**2a1.5 Target Population Category** (Check all the populations for which the measure is specified and tested if any): [Adult/Elderly Care](#), [Children's Health](#)

**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:

This measure does not lend itself to a “traditional specification” for EHR reporting. This is a structural measure; each facility may have a different process or software system for tracking and transmitting recalls as well as different appointment tracking systems.

For Claims/Administrative:

ICD-9-CM diagnosis codes: 172.0, 172.1, 172.2, 172.3, 172.4, 172.5, 172.6, 172.7, 172.8, 172.9, V10.82



ICD-10-CM diagnosis codes: C41.10, C41.11, C41.12, C43.0, C43.20, C43.21, C43.22, C43.30, C43.31, C43.39, C43.4, C43.51, C43.52, C43.59, C43.60, C43.61, C43.62, C43.70, C43.71, C43.72, C43.8, C43.9, D03.0, D03.10, D03.11, D03.12, D03.20, D03.21, D03.22, D03.30, D03.39, D03.4, D03.51, D03.52, D03.59, D03.60, D03.61, D03.62, D03.70, D03.71, D03.72, D03.8, D03.9, Z85.820

AND

CPT codes: 99201, 99202, 99203, 99204, 99205, 99212, 99213, 99214, 99215, 99241, 99242, 99243, 99244, 99245

**2a1.8 Denominator Exclusions** *(Brief narrative description of exclusions from the target population):*

Documentation of system reason(s) for not entering patients into a recall system (eg, melanoma being monitored by another physician provider)

**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 system reason(s) for not entering patients into a recall system (eg, melanoma being monitored by another physician provider). 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:

This measure does not lend itself to a "traditional specification" for EHR reporting. This is a structural measure; each facility may have a different process or software system for tracking and transmitting recalls as well as different appointment tracking systems.

For Claims/Administrative:

Documentation of system reason exception

- Append modifier to CPT Category II code: 7010F-3P

**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.):*

Not applicable

**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.):

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: system reason(s) (eg, melanoma being monitored by another physician provider)]. 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.

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

[Attachment](#)

[AMA-PCPI\\_Measure Calculation-Standard Measures650.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:

[Administrative claims, Electronic Clinical Data, Electronic Clinical Data : Registry, Other, 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.): [Not Applicable](#)

**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](#)

**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):

AMA-PCPI Testing Project

3 dermatology practice sites representing various types, locations and sizes were identified to participate in testing the melanoma measures.

- Site A was a dermatology practice with 3 physicians that cared for mainly geriatric patients.
- Site B was a general dermatology practice with 9 physicians that cared for a pediatric to geriatric patient population.
- Site C was a medical dermatology practice with 3 physicians that cared for a pediatric to geriatric patient population.
- Site A was paper chart-based. Sites B and C utilized EHRs.
- Sites were located in the northeastern and southern regions of the United States.
- Patient visit volume was 125 visits per day at site A and 115-130 visits per day at site C. For site B, physician 1 had 40 patient visits per day, physician 2 had 38 patient visits per day and patient 3 had 57 patient visits per day.
- All sites were instructed to select the number of charts that they had entered into the American Academy of Dermatology (AAD) 2011 PQRS Registry. For this measure, abstractors reviewed 397 patient records.
- The difference in the number of patient records (N=397) and the number of melanoma diagnoses (N=466) is due to the fact that it is possible for an individual patient to have several different types of melanomas.
- The measurement period (data collected from patients seen) was January 2011 through July 2011.
- Chart auditing was performed mid-July 2011 through mid-August 2011 (1 month total).

**2a2.2 Analytic Method** (Describe method of reliability testing & rationale):

PCPI Testing Project

Data abstracted from patient records were used to calculate inter-rater reliability for the measure.

Data analysis at the data element level included:

- Percent agreement; and
- Kappa statistic to adjust for chance agreement.

Reliability was measured as 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 physician performance. Reliability at the level of the specific physician is given by:

Reliability = Variance (physician-to-physician) / [Variance (physician-to-physician) + Variance (physician-specific-error)]

Reliability is the ratio of the physician-to-physician variance divided by the sum of the physician-to-physician variance plus the error variance specific to a physician. 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 physician performance.

Reliability testing was performed by using a beta-binomial model. The beta-binomial model assumes the physician performance score is a binomial random variable conditional on the physician'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.

Reliability is estimated at two different points, at the minimum number of quality reporting events for the measure and at the mean number of quality reporting events per physician. For this measure, the minimum number required to be included is 10 events. Given the structure of the PQRS program, a physician may choose to submit or not submit to PQRS on any given claim. Since these data contain results on a large number of physicians, limiting the reliability analysis to only those physicians who are participating in the program will eliminate the bias introduced by the inclusion of from physicians who are in the data but are not submitting claims to PQRS.

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

Reliability: N, % Agreement, Kappa ( 95% Confidence Interval)

Melanoma diagnosis

Overall: 466, 95.92%, 0.8603 (0.7988-0.9218)

Denominator: 466, 97.00%, 0.4012 (0.0923-0.7101)

Numerator: 466, 95.92%, 0.8603 (0.7988-0.9218)

Exceptions: 466, 100%, kappa statistic is noncalculable\*

This measure demonstrates moderate reliability, as shown in results from the above analysis.

Follow-up scheduled in recall system

Overall: 465, 93.55%, 0.6606 (0.5431-0.7781)

Denominator: 465, 100%, kappa statistic is noncalculable\*

Numerator: 465, 93.55%, 0.6606 (0.5431-0.7781)

Exceptions: 465, 100%, kappa statistic is noncalculable\*

This measure demonstrates almost perfect reliability, as shown in results from the above analysis.

Follow-up information documented in recall system

Overall: 402, 89.05%, kappa statistic is noncalculable\*

Denominator: 402, 100%, kappa statistic is noncalculable\*

Numerator: 402, 89.05%, kappa statistic is noncalculable\*

Exceptions: 402, 100%, kappa statistic is noncalculable\*

This measure demonstrates almost perfect reliability, as shown in results from the above analysis.

History of melanoma

Overall: 466, 97.00%, 0.8622 (0.7911-0.9333)

Denominator: 466, 100%, kappa statistic is noncalculable\*

Numerator: 466, 97.00%, 0.8622 (0.7911-0.9333)

Exceptions: 466, 100%, kappa statistic is noncalculable\*

This measure demonstrates almost perfect reliability, as shown in results from the above analysis.

A complete skin exam performed and changing pigmented lesions noted at visit

Overall: 288, 81.60%, 0.3778 (0.2265-0.5291)

Denominator: 288, 100%, kappa statistic is noncalculable\*

Numerator: 288, 81.60%, 0.3778 (0.2265-0.5291)

Exceptions: 288, 100%, kappa statistic is noncalculable\*

This measure demonstrates almost perfect reliability, as shown in results from the above analysis.

During visit, patient was asked specifically if acquired new moles

Overall: 465, 92.90%, 0.6515 (0.5369-0.7661)

Denominator: 465, 100%, kappa statistic is noncalculable\*

Numerator: 465, 92.90%, 0.6515 (0.5369-0.7661)

Exceptions: 465, 100%, kappa statistic is noncalculable\*

This measure demonstrates almost perfect reliability, as shown in results from the above analysis.

Patient evaluated for new episode at visit

Overall: 465, 97.42%, 0.9137 (0.8655-0.9619)

Denominator: 465, 100%, kappa statistic is noncalculable\*

Numerator: 465, 97.42%, 0.9137 (0.8655-0.9619)

Exceptions: 465, 100%, kappa statistic is noncalculable\*

This measure demonstrates almost perfect reliability, as shown in results from the above analysis.

Patient received a physical exam at visit

Overall: 288, 92.90%, 0.6428 (0.4935-0.7921)

Denominator: 288, 100%, kappa statistic is noncalculable\*

Numerator: 288, 92.90%, 0.6428 (0.4935-0.7921)

Exceptions: 288, 100%, kappa statistic is noncalculable\*

This measure demonstrates almost perfect reliability, as shown in results from the above analysis.

\*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.

The total number of physicians reporting on this measure is 535. Of those, 405 met the minimum number of quality reporting events for inclusion in the reliability analysis. For this measure 75.70 percent of physicians are included in the analysis, and the average number of quality reporting events for physicians included is 50.53 for a total of 20,465 events. The average number of quality reporting events for the remaining 24.30 percent of physicians who aren't included is 4.18 for a total of 543 events.

For this measure, the reliability at the minimum level of quality reporting events (ie 10 quality events) was .8727 . The reliability at the average number of quality reporting events was .9719

This measure has high reliability when evaluated at both the minimum level of quality reporting events and at the average number of quality events.

Data analyses were conducted by using SAS/STAT software, version 8.2 (SAS Institute, Cary, North Carolina).

**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:**

The guidelines focus on the importance and frequency of follow up for melanoma patients. The measure specifications are consistent with the evidence cited in support of the measure focus, as the measure captures visits of all patients, regardless of age, with a current diagnosis or history of melanoma being entered into a recall system, at least once within a 12 month period.

**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 16 members, with representation from the following specialties:

List of Work Group Members

Raj Behal, MD, MPH (Co-Chair)(methodology)

Dirk Elston, MD (Co-Chair)(dermatology)

Stephen Bines, MD (general surgery)

Peter C. Dandalides, MD (health plan)

Evan Farmer, MD (dermatology)

Rutledge Forney, MD (dermatology)

Andrea Gelzer, MD, MS FACP (health plan)

Robert Gilson, MD (dermatology)

Stephen Helms, MD (dermatology)

Abrar Qureshi, MD (dermatology)

Todd Schlessinger, MD (dermatology)

John Schneider, MD, PhD (family medicine)  
 Janet (Jessie) Sullivan, MD(dermatology)  
 Arthur Sober, MD (dermatology)  
 Steven Strode, MD, Med, MPH (family medicine)  
 William Wooden, MD (plastic surgery)

**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 16 members, with representation from the following specialties: general surgery, dermatology, family medicine, plastic surgery, and health plan representatives.

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 = 11; Mean rating = 4.6.

Percentage in the top two categories (4 and 5): 100%

Frequency Distribution of Ratings

1 – 0  
 2 – 0  
 3 – 0  
 4 – 4  
 5 – 7

**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):*

PCPI Testing Project

- o All sites were instructed to select the number of charts that they had entered into the American Academy of Dermatology (AAD) 2011 PQRS Registry. For this measure, abstractors reviewed 397 patient records.
- o The measurement period (data collected from patients seen) was January 2011 through July 2011.
- o Chart auditing was performed mid-July 2011 through mid-August 2011 (1 month total).

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

Exceptions were analyzed for frequency and variability across providers.

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

PCPI Testing Project

Reliability: N, % Agreement, Kappa ( 95% Confidence Interval)

Melanoma diagnosis

Exceptions: 466, 100%, kappa statistic is noncalculable\*

This measure demonstrates perfect reliability, as shown in results from the above analysis.

Follow-up scheduled in recall system

Exceptions: 465, 100%, kappa statistic is noncalculable\*

This measure demonstrates perfect reliability, as shown in results from the above analysis.

Follow-up information documented in recall system

Exceptions: 402, 100%, kappa statistic is noncalculable\*

This measure demonstrates perfect reliability, as shown in results from the above analysis.

History of melanoma

Exceptions: 466, 100%, kappa statistic is noncalculable\*

This measure demonstrates perfect reliability, as shown in results from the above analysis.

A complete skin exam performed and changing pigmented lesions noted at visit

Exceptions: 288, 100%, kappa statistic is noncalculable\*

This measure demonstrates perfect reliability, as shown in results from the above analysis.

During visit, patient was asked specifically if acquired new moles

Exceptions: 465, 100%, kappa statistic is noncalculable\*

This measure demonstrates perfect reliability, as shown in results from the above analysis.

Patient evaluated for new episode at visit

Exceptions: 465, 100%, kappa statistic is noncalculable\*

This measure demonstrates perfect reliability, as shown in results from the above analysis.

Patient received a physical exam at visit

Exceptions: 288, 100%, kappa statistic is noncalculable\*

This measure demonstrates perfect reliability, as shown in results from the above analysis.

\* 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.

**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:** As a structure measure, no risk adjustment is necessary.

**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):

PCPI Testing Project

- o All sites were instructed to select the number of charts that they had entered into the American Academy of Dermatology (AAD) 2011 PQRS Registry. For this measure, abstractors reviewed 397 patient records.
- o The measurement period (data collected from patients seen) was January 2011 through July 2011.
- o Chart auditing was performed mid-July 2011 through mid-August 2011 (1 month total).

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

AAD Testing Project

The variability in physician performance rates across the measure were calculated by decile

**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):

AAD Testing Project

N is the number of physicians with 10 or more quality events

Measure 137

N = 405

Performance Rate	Decile
0.9737	0.1
1.0000	0.2
1.0000	0.3
1.0000	0.4
1.0000	0.5
1.0000	0.6
1.0000	0.7
1.0000	0.8
1.0000	0.9
1.0000	1

**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):

PCPI Testing Project

- o All sites were instructed to select the number of charts that they had entered into the American Academy of Dermatology (AAD) 2011 PQRS Registry. For this measure, abstractors reviewed 397 patient records.
- o The measurement period (data collected from patients seen) was January 2011 through July 2011
- o Chart auditing was performed mid-July 2011 through mid-August 2011 (1 month total)

**2b6.2 Analytic Method** (Describe methods and rationale for testing comparability of scores produced by the different data sources specified in the measure):

Reliability was measured as 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 physician performance. Reliability at the level of the specific physician is given by:



Reliability = Variance (physician-to-physician) / [Variance (physician-to-physician ) + Variance (physician-specific-error)]

Reliability is the ratio of the physician-to-physician variance divided by the sum of the physician-to-physician variance plus the error variance specific to a physician. 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 physician performance.

Reliability testing was performed by using a beta-binomial model. The beta-binomial model assumes the physician performance score is a binomial random variable conditional on the physician'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.

Reliability is estimated at two different points, at the minimum number of quality reporting events for the measure and at the mean number of quality reporting events per physician. For this measure, the minimum number required to be included is 10 events. Given the structure of the PQRS program, a physician may choose to submit or not submit to PQRS on any given claim. Since these data contain results on a large number of physicians, limiting the reliability analysis to only those physicians who are participating in the program will eliminate the bias introduced by the inclusion of from physicians who are in the data but are not submitting claims to PQRS.

**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 total number of physicians reporting on this measure is 535. Of those, 405 met the minimum number of quality reporting events for inclusion in the reliability analysis. For this measure 75.70 percent of physicians are included in the analysis, and the average number of quality reporting events for physicians included is 50.53 for a total of 20,465 events. The average number of quality reporting events for the remaining 24.30 percent of physicians who aren't included is 4.18 for a total of 543 events.

For this measure, the reliability at the minimum level of quality reporting events (ie 10 quality events) was .8727 . The reliability at the average number of quality reporting events was .9719

This measure has high reliability when evaluated at both the minimum level of quality reporting events and at the average number of quality events.

Data analyses were conducted by using SAS/STAT software, version 8.2 (SAS Institute, Cary, North Carolina).

**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): [Professional Certification or Recognition Program](#), [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): [Public Reporting](#), [Professional Certification or Recognition Program](#), [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 was used in the Physician Quality Reporting System from 2009 through 2011. This measure is currently included in PQRS 2012.

<http://www.cms.gov/pqrs>

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.

**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 is 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):**  
[No competing measures have been identified.](#)

## CONTACT INFORMATION

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

**Co.2 Point of Contact:** [Mark S., Antman, DDS, MBA, 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 St., Chicago, Illinois, 60654](#)

**Co.4 Point of Contact:** [Diedra, Joseph, MPH, diedra.joseph@ama-assn.com, 312-545-2813-](#)

**Co.5 Submitter:** [Diedra, Joseph, MPH, diedra.joseph@ama-assn.com, 312-545-2813-](#), [American Medical Association - Physician Consortium for Performance Improvement](#)

**Co.6 Additional organizations that sponsored/participated in measure development:**  
[American Academy of Dermatology and National Committee for Quality Assurance](#)

**Co.7 Public Contact:** [Diedra, Joseph, MPH, diedra.joseph@ama-assn.com, 312-545-2813-](#), [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.**

[Steven D. Bines, MD \(general surgery\)](#)

[Peter Dandalides, MD \(health plan\)](#)

[Evan R. Farmer, MD \(dermatology\)](#)

Rutledge Fourney, MD (dermatology)  
 Andrea Gelzer, MD MS FACP (health plan)  
 Robert T. Gilson, MD (dermatology)  
 Stephen E. Helms, MD (dermatology)  
 Abrar Qureshi, MD (dermatology)  
 Todd Schlessinger, MD (dermatology)  
 John Schneider, MD, PhD (family medicine)  
 Arthur Joel Sober, MD (dermatology)  
 Steven W. Strobe, MD, MEd, MPH (family medicine)  
 Janet (Jessie) Sullivan, MD (dermatology)  
 William Wooden, MD (plastic surgery)

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?** Please see Additional Information/Comments

**Ad.6 When is the next scheduled review/update for this measure?** 09, 2012

**Ad.7 Copyright statement:** Physician Performance Measures (Measures) and related data specifications, developed by the American Medical Association

(AMA) in collaboration with the Physician Consortium for Performance Improvement (the Consortium) and the National Committee for Quality Assurance (NCQA) pursuant to government sponsorship under subcontract 6205-05-054 with Mathematica Policy Research, Inc. under contract 500-00-0033 with Centers for Medicare & Medicaid Services.

These performance Measures are not clinical guidelines and do not establish a standard of medical care, and have not been tested for all potential applications.

The Measures, while copyrighted, can be reproduced and distributed, without modification, for noncommercial purposes, e.g., use by health care providers in connection with their practices. Commercial use is defined as the sale, license, or distribution of the Measures for commercial gain, or incorporation of the Measures into a product or service that is sold, licensed or distributed for commercial gain. Commercial uses of the Measures require a license agreement between the user and the AMA, (on behalf of the Consortium) or NCQA. Neither the AMA, NCQA, Consortium nor its members shall be responsible for any use of the Measures.

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

© 2004-6 American Medical Association and National Committee for Quality Assurance. All Rights Reserved.

Limited proprietary coding is contained in the Measure specifications for convenience. Users of the proprietary code sets should obtain all necessary licenses from the owners of these code sets. The AMA, NCQA, the Consortium and its members disclaim all liability for use or accuracy of any Current Procedural Terminology (CPT®) or other coding contained in the specifications.

CPT® contained in the Measures specifications is copyright 2005 American Medical Association.

**Ad.8 Disclaimers:**

**Ad.9 Additional Information/Comments:** Coding/Specifications updates occur annually. The PCPI has a formal measurement review process that stipulates regular (usually on a three-year cycle, when feasible) review of the measures. The process can also be activated if there is a major change in scientific evidence, results from testing or other issues are noted that materially affect the integrity of the measure.

**Date of Submission (MM/DD/YY):** 01/13/2012

NQF #0650 Melanoma Continuity of Care - Recall System

**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 counted when calculating the exception rate**

<p><b>Initial Patient Population (IPP)</b></p> <p><b>Definition:</b> The initial patient population identifies the general group of patients that the performance measure is 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 CAD who has at least 2 Visits during the measurement period.</p>	<p><b>Denominator (D)</b></p> <p><b>Definition:</b> 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 identical to the initial patient population.</p>	<p><b>Numerator (N)</b></p> <p><b>Definition:</b> The numerator defines the group of patients in the denominator for whom a process or outcome of care occurs (e.g., flu vaccine received).</p>	<p><b>Denominator Exceptions (E)</b></p> <p><b>Definition:</b> 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 there is a valid Denominator Exception.</p>
<p>Find the patients who meet the Initial Patient Population criteria (IPP)</p>	<p>Find the patients who qualify for the denominator (D):</p> <ul style="list-style-type: none"> <li>○ From the patients within the Patient Population criteria (IPP) select those people who meet Denominator selection criteria.</li> </ul> <p>(In some cases the IPP and D are identical).</p>	<p>Find the patients who qualify for the Numerator (N):</p> <ul style="list-style-type: none"> <li>○ From the patients within the Denominator (D) criteria, select those people who meet Numerator selection criteria.</li> <li>○ Validate that the number of patients in the numerator is less than or equal to the number of patients in the denominator</li> </ul>	<p>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.</p>