

NATIONAL QUALITY FORUM

In-Person Meeting for Imaging Efficiency Meeting Summary February 23-24, 2010

A meeting of the Imaging Efficiency Steering Committee was held on February 23-24, 2010 in Washington, D.C. Topics discussed were the NQF Measure Evaluation Process, and the 17 Imaging Efficiency performance standard measures that were submitted by the American College of Radiology (ACR), Centers for Medicare and Medicaid Services (CMS), American College of Cardiology (ACC), and Brigham and Women's Hospital (BWH).

Imaging Efficiency Steering Committee members present: G.Scott Gazelle, MD, MPH (Co-chair); Eric D. Peterson, MD, MPH (Co-chair); Michael Backus, MBA; Jacqueline A. Bello, MD, FACR; Stephen V. Cantrill, MD, FACEP; Carl D'Orsi, MD; Troy Fiesinger, MD, FAAFP; Howard P. Forman, MD, MBA; Mary Gemignani, MD; Raymond Gibbons, MD; Richard Griffey, MD, MPH; Laszlo Mechtler, MD; Patti Raksin, MD; Gavin Setzen, MD, FACS, FAAOA; Rebecca Smith-Bindman, MD; Roger L. Snow, MD, MPH; Dr. Kirk Spencer, MD; Arthur Stillman, MD, PhD; Judy Zerzan, MD, MPH.

National Quality Forum (NQF) Staff Present: Helen Burstin, MD, MPH; Ian Corbridge, MPH, RN, BSN; Heidi Bossley, MSN, MBA; Sarah Fanta

WELCOME AND INTRODUCTIONS

Dr. Burstin welcomed the Imaging Efficiency Steering Committee and reviewed the agenda for the in-person meeting. Dr. Burstin clarified that the purpose of the in-person meeting was to discuss the 17 submitted Imaging Efficiency measures. Dr. Gazelle and Dr. Peterson asked Steering Committee members to disclose any conflicts of interest that had arisen since the February 19, 2010 conference call; there were no conflicts stated.

NATIONAL QUALITY FORUM

IMAGING EFFICIENCY MEASURE EVALUATIONS

Mammography-Related Measures (American College of Radiology)

#IEP-001-10 Cancer Detection Rate

#IEP-003-10 Diagnostic Mammography Positive Predictive Value 2 (PPV3-Biopsy Recommended)

#IEP-004-10 Abnormal Interpretation Rate of Screening Mammography Exams (Recall Rate)

The American College of Radiology (ACR) submitted a series of mammography-related measures to the Imaging Efficiency Project for consideration. The measure reviewers and Steering Committee members recognized that these measures address an important area in the imaging efficiency arena. The Committee acknowledged that the measures were not formally designed to be a composite measure but was unclear as to how the measures were intended to be interpreted individually. Although Committee members noted the importance of measuring this clinical area, they were concerned that any one individual measure may not provide a comprehensive view of mammography for public reporting. Furthermore, the Committee acknowledged that these measures may not provide actionable information at the facility level. Facilities must have enough breast cancer events to make the measures meaningful. The Committee further acknowledged that there may be value in stratifying the measures by patient age, because cancer prevalence is known to vary by age.

#IEP-001-10 Cancer Detection Rate (American College of Radiology)

Discussion Points:

- The measure offers very clear definitions and calculations along with a great deal of data to support the measure. Committee members expressed concern about the applicability of this measure to small facilities, that is, will there be a meaningful number of patients?

Response from Measure Developer:

This measure addresses a new measurement subject matter; ACR has only a year's worth of data. ACR indicated that, because of the nature of the request and the data available, it may not be able to meet the Committee's conditions for recommendation (see below).

#IEP-003-10 Diagnostic Mammography Positive Predictive Value 2 (PPV3-Biopsy Recommended) (American College of Radiology)

Discussion Points:

- The Committee agreed that the measure is feasible and that most centers would have the necessary data.
- The Committee noted that there was a discrepancy between the measure title and the actual intent of the measure. The measure determines the number of patients who were recommended to receive a biopsy—not the number of those who had an actual biopsy. The numerator is the diagnostic mammograms, and the denominator is the cancer. The

NATIONAL QUALITY FORUM

actual biopsy is what is important, but the recommendations mean that the facility would actually have to look at follow-up.

- The Committee noted the potential for unintended consequences. Without knowing the volume, detection rate, and other rates, it may be hard to interpret this measure.

#IEP-004-10 Abnormal Interpretation Rate of Screening Mammography Exams (Recall Rate) (American College of Radiology)

Discussion Points:

- The Committee noted that there may be issues in stratification when working with different demographic populations.
- The measure has the potential to provide a desirable rate when performance is not known, that is, is a high recall rate a good thing or not? A patient would not know which facility to use based on the reported results. The question is how useful is the measure if it's not interpretable.

In-person voting results #IEP-001-10, #IEP-003-10, and #IEP-004-10:

The Committee voted down the measures as individual stand-alone measures. The Committee then considered the measures in the form of a composite. The initial vote for a composite of the three measures was 9 recommending the measures for endorsement vs. 11 not recommending the measures for endorsement. Because of the closeness of the vote, the Committee decided to provide the measure developer with the agreed-upon conditions for potential recommendation and to re-vote after the measure developer responds to the conditions.

***The following contains additional information related to the project and measure evaluation review that took place after the in person meeting.**

Conditions for the Measure Developer for #IEP-001-10, #IEP-003-10, and #IEP-004-10:

The Committee made the following conditional recommendations:

- Stratify by the following proposed age categories:
 - 40-50, 50-64 and 65+ or
 - 40-59, 60-69 and 70-80
- Stratify by first and subsequent mammogram exams
- Provide a sample size for the measure or indicate the minimum size requirement of the reporting organizations
- Provide information on how many facilities would or would not have sufficient data
- The Steering Committee recognized that measures #IEP-001-10, #IEP-003-10, and #IEP-004-10 were not formally designed to be a composite measure, but it believed that there would be value in combining and presenting the measures as a composite

Detailed Response from Measure Developer:

- Based on data from facilities participating in registry collection for auditing of Cancer Detection Rate for two 12 month periods, the volume from most sites would not be adequate to provide meaningful statistics, either stratified by any age groupings or an overall rate. In 2008, from 58 facilities, the mean for frequency of exams by age groups

NATIONAL QUALITY FORUM

40-49, 50-59, 60-69 and 70+ was roughly even (23%, 30%, 24%, 20%) but across facilities there was substantial variance of the minimum/maximum distribution (12-37%, 17-38%, 17-37% and 9-34% respectively by age category). In 2007, from 72 facilities, the mean for frequency of exams by age groups 40-49, 50-59, 60-69 and 70+ was 24%, 30%, 23% and 20%, respectively by age category) and across facilities the variance of the distribution was 10-51%, 20-38%, 8-40%, 4-36% (respectively by age category). Additionally, from the same facility data in 2008, the mean number of exams by those age categories was around 1,000 (1398, 1698, 1256, 955 respectively by age category). However, the variance of the minimum/maximum distribution was significant (18 - 11,567; 25 - 10,410; 55-5918 and 47-3986, respectively). In 2007, the mean number of exams by those age categories was also around 1,000 (1329, 1567, 1102 and 888, respectively). The variance of the minimum/maximum distribution was again significant in 2007 (12 - 11,950; 23 - 10,792; 45-5787 and 28-4288, respectively). This stratification will not work per facility due to the lack of numbers per strata and the low probability of malignancy. Few facilities will even have the volume to report a meaningful overall facility rate without age stratification.

- A minimum sample size of 21,000 exams per age category (whether the age category is broken out by 40-50, 40-59 or 40 – 80 or total) per year is needed in order to estimate CDR to within 1 per 1000 (i.e., if the rate is actually 5 per 1000 as BCSC data shows, and we want to estimate it to be from 4-6 per 1000). Statistically, for a two-sided 95% Agresti-Coull confidence interval for a binomial proportion whose true value is 0.005 a sample size of 21,000 yields a half-width of at most 0.001 with a conditional probability of 0.80.
- We do not have data for facility exam volume overall (non-stratified) for the 58 (2008) facilities and 72 (2007) facilities described in our previous answer. We also do not have the data on how many facilities in the US would or would not have sufficient data. However, an average site with one mammography machine could be estimated to have approximately 7,500 exams per year (estimating 4 exams per hour x 8 hrs a day = 32 a day x 250d/yr = 7,500). This average number would need to be tripled to have sufficient data. We do not know how many facilities will be able to use the measure package for public reporting under that restriction.
- Applied to a multi-facility consortium or a large facility with established data collection this would be OK. Applied to typical facilities would give no useful information due to the lack of numbers combined with the low prior probability of malignancy.
- While the Steering Committee recognizes the value of presenting these measures as a package and the ACR again stresses the importance of the combined set, we feel it is premature to publicly report such data until sufficient evidence base guidance has been developed as to how these measures could best be presented in the composite. Until such time, we again recommend that public reporting be limited to facility attestation of participation in a systematic registry or database that collects data capable of calculating CDR, PPV and RR. In the case where insufficiency of data to determine publicly reported rates is an issue, but the exercise of collecting and self-auditing is a valuable quality improvement tool, such a structural measure can be used to inform the public in a

NATIONAL QUALITY FORUM

simplified manner of a facility's efforts to provide quality care. Tools such as the ACR National Mammography Database and the Breast Cancer Surveillance Consortium registries provide a participating facility with a method to track their results in a meaningful, comprehensive way with comparative national and regional benchmarked mammography data.

- We also do not have the data on how many facilities in the US would or would not have sufficient data. However, an average site with one mammography machine could be estimated to have approximately 7,500 exams per year (estimating 4 exams per hour x 8 hrs a day = 32 a day x 250d/yr = 7,500). This average number would need to be tripled to have sufficient data. We do not know how many facilities will be able to use the measure package for public reporting under that restriction.
- The ACR does agree that the measures should be presented as a package but the data must be sufficient to report an actual rate (e.g. a facility has 21,000 + exams per year). We are unclear what information the SC suggests to report in the package – rates, compliance within a range? There has been much discussion and lack of consensus on what exactly the measures mean. Describing what rates mean to the public will be difficult. For this reason, the ACR recommends either waiting until a validly constructed composite can be developed, or use of a structural measure describing a facility's participation in data collection for auditing its performance based on the PPV, CDR and RR constructs.

Steering Committee Recommendation for #IEP-001-10, #IEP-003-10, and #IEP-004-10:

The Committee reviewed the measure developer's responses to the conditional recommendations on Thursday, April 22, 2010, and subsequently voted on the measures electronically. The results of the vote were as follows:

- Recommend for Endorsement: 2
- Do not Recommend for Endorsement: 14
- Abstain: 3

The measures were not recommended for endorsement.

Steering Committee voting comments:

- This is an important area to measure, but concerns remain regarding the volume needed. Many rural areas and small facilities would not be able to meet this standard.
- These measures taken together are a better measure of performance than if they were left on their own. One, however, does not give individual feed back to an individual facility or reader. There are articles which demonstrate the ranges for these metrics based on over 2 million mammograms read. A better solution is to use membership in a national database that collects these and other metrics is a requisite for additional reimbursement .These metrics, since they are just becoming established by large studies, should not be public at this point. Thus membership in a national data base that collects this information will demonstrate that a facility and their individual readers are interested in quality performance and more importantly allow comparison metrics based on other

NATIONAL QUALITY FORUM

facilities and readers which will allow a more cogent profile of comparison for the facility and reader to begin to implement improvement where needed.

#IEP-009-10 Mammography Follow-up Rates (Centers for Medicare and Medicaid Services)

Committee members acknowledged that the measure addresses a critical topic area in the outpatient imaging realm, but they expressed concern about the measure's specifications (i.e., age stratification), value in public reporting, and usability. Although the measure assesses recall rates, as currently specified it does not assess cancer detection rates. The Committee was concerned about the unintended consequences that would result when a clinician or facility performs well on this measure even though it has low recall rates simultaneous with a substantial number of missed cancers, highlighting the importance of having both. The Steering Committee further questioned why the measure targets only outpatient hospital settings and not other outpatient ambulatory facilities. With hospitals increasingly purchasing outpatient imaging facilities, the measure as currently stipulated stands to miss a segment of the market.

Discussion Points:

- The Committee discussed what follow-up rates should be—a high rate is not good, but an extremely low rate is not good either. A recall rate of more than 20 percent could signal low quality and high costs. A recall rate of 10 percent to 14 percent could be considered ideal. A recall rate of less than 10 percent could mean that the facility is missing cancers. An ideal or excellent facility could have a low recall rate, but a high cancer detection rate. Committee members noted that recall rates in the UK and the Netherlands are 5 percent (closer to ideal) and 1.8 percent, respectively; however, it appears that these countries are missing some cancers because they have high cancer rates.

Response from the Measure Developer:

When looking at high recall rates one must consider if lowering recall rates would have negative unintended consequences. Currently the CMS numbers are in the 7 percent to 8 percent range. The measure as stipulated supports a rate of 10 percent, which applies a standard that half of your facility can achieve. The measure developer indicated that it was open for input from the Committee on where thresholds should be set.

In-person voting results:

The Committee voted to not recommend the measure at the time of the in-person meeting. The results of the vote were as follows: 9 recommending the measures for endorsement vs. 11 not recommending the measures for endorsement. Because of the closeness of the vote the Committee decided to provide the developer with the agreed-upon conditions for recommendation (see below) and to re-vote after the measure developer responds to the conditions.

***The following contains additional information related to the project and measure evaluation review that took place after the in person meeting.**

NATIONAL QUALITY FORUM

Conditions for Measure Developer for #IEP-009-10:

The Committee made the following conditional recommendations:

- Stratify according to the following proposed age categories:
 - 40-50, 50-64 and 65+ or
 - 40-59, 60-69 and 70-80
- Expand the scope of the measure to include both hospital outpatient and other outpatient facility settings
- Validate the accuracy of the CPT codes, that is, show the distribution of current screening/diagnostic codes

Detailed Response from Measure Developer:

Age Stratification for Medicare Population

- One of the questions raised by the NQF steering committee is whether the mammography follow-up rate measure should be stratified by age. As noted in the first of the suggested stratifications shown above, the Steering Committee was particularly concerned with differentiating between elderly and non-elderly populations. The Medicare population is predominantly elderly, with disabled female beneficiaries making up approximately 15% of the overall female Medicare population. An analysis of the Medicare Current Beneficiary Survey cost and use files for 2005 found that mammography follow-up rates decline by age, but follow-up rates do not vary to any meaningful degree across different elderly age cohorts. See Table 1 below.
- The measure developer and CMS **do not** think that it is necessary to stratify the mammography follow-up measure by age cohort for the Medicare population. The Medicare population is overwhelming age 65 and older and our analysis indicates that there is no need to break this population group into finer age cohorts. The under-65 population comprises only a small share of the Medicare population and as such, does not impact a facility's Medicare measure rate such that age stratification would be warranted.

**Table 1: Medicare Female Mammography Follow-Up Rates
by Age Cohort**

Age Range	Rate
≤ 64	11.37%
65 to 69	8.96%
70 to 74	8.09%
75 to 79	7.98%
80 to 84	7.26%
≥ 85	6.22%
Total	8.43%

Source: Medicare Current Beneficiary Survey 2005

NATIONAL QUALITY FORUM

The Committee’s concern was whether stratification would be needed if a facility served a Medicare population with a different age distribution than the national average. The age distribution for Medicare females receiving screening mammograms is shown below in Table 2.

Table 2: National Medicare Population Age Distribution for Women Receiving Screening Mammograms

Age Range	Distribution, National (%)
≤ 64	11.0%
65 to 69	25.7%
70 to 74	25.2%
75 to 79	18.7%
80 to 84	12.9%
≥ 85	6.6%
Source: MCBS 2005 (Cost and Use Files)	

The measure developer also conducted a sensitivity analysis using the 2007 case volumes from actual facilities to determine the effect on the follow-up rate of various age distributions. Using an extreme assumption that skews the distribution to having 70 percent of the elderly population in the 65 to 69-age cohort, the impact on facility follow-up rates was found to be less than one percentage point, and this difference was not statistically significant. See Table 3 below.

Table 3: Modeled Facility Measures, Statistical Indicators

Facility	Denom Total	Num Facility	Measure, Facility	Standard Error	Numerator, Adjusted	Measure Adjusted	Standard Error	Z Statistic	P-Value
Example 1	501	43	8.58%	0.0125	46 (45.48)	9.08%	0.0129	0.3331	0.3695
Example 2	46	6	13.0%	0.0497	7 (6.34)	13.78%	0.0530	0.2993	0.3824
Example 3	136	13	9.56%	0.0252	14(13.75)	10.11%	0.0261	0.2028	0.4197
Example 4	2,050	177	8.63%	0.0062	188 (187.20)	9.13%	0.0064	0.6032	0.2732

P-values greater than 0.05 are assumed to be not significant. Under a second scenario which used the Example 3 facility, 20 percent of the patient population was assigned to <65 cohort and 65 percent of the population was assigned to the 65 to 69 cohort. Even with this extreme

NATIONAL QUALITY FORUM

skewing of the population, the resulting change to the facility follow-up rate is still less than one percentage point.

Thus stratifying the mammography follow-up measure by age cohort does not appear to be necessary for the Medicare population.

Age Stratification for Non-Elderly

If the measure is used in a predominantly non-elderly patient population, stratification by age cohort may be appropriate. In our review of the literature concerning screening and cancer detection, we found several different age category groupings. Studies of screening and/or diagnostic mammography performance in cancer detection generally seem to use 10 year age groupings beginning with age 40 (e.g., 40-49, 50-59, 60-69, and greater than/equal to age 70).

Studies of mammography in the context of breast cancer incidence and mortality seem to use smaller age groupings. For example, a 2009 study of the absolute benefit of screening mammography by Keen and Keen uses an age range of 40 to 65, with five year age groupings. Predictive modeling of mammography screening seems to use smaller age categories. For example, the SPECTRUM model covers a broad age range of 26 to over 85. (e.g., 26-39, 40-44, 45-49, ... 80-84, greater than or equal to 85).

The 2009 update of clinical guidelines from the U.S. Preventive Services Task Force uses ten year age categories, ranging from 39-74 (e.g., 39-49, 50-59, 60-69, 70-74).

For the purpose of implementing the Mammography Follow-up Studies Measure, hospital outpatient facilities and payers will have varying amounts of data. For example, small facilities may not have a sufficient number of observations to group patients in less than ten year categories. Facilities and payers predominately serving particular age groups will be limited to the data they have, and the age groupings will need to be determined by the data.

Response clarification submitted to NQF on April 8, 2010:

The measure developer and CMS do think that stratification by age cohort may be appropriate if the measure is being used for commercial and other non-elderly patient populations. As noted in our submitted response, the literature uses both 5 and 10 year age breaks in the non-elderly population. However, patient population sizes contained in various data sources may preclude the use of 5 year age breaks due to issues related to adequate sample sizes. Therefore an approach similar to the first approach that the Steering Committee proposed may be appropriate, but without overlapping years, i.e., we suggest 40-49, 50-64 and 65+ may be the most consistent with our finding that age stratification is not needed for the Medicare population, but that age stratification may be useful for predominantly non-elderly patient populations using larger age band cohorts.

Expand Measure to Other Settings

CMS as measure steward is supportive of the expansion of the use of the measure to settings beyond outpatient hospital departments.

Validate the Accuracy of CPT Codes

The Steering Committee requested information on the overall distribution of diagnostic versus screening mammography CPT codes in the Medicare claims data. See Table 4 for the distribution for both 2007 and 2008 data.

NATIONAL QUALITY FORUM

**Table 4: Relative Distribution of Screening vs. Diagnostic Mammograms
Analysis of Medicare Claims Data, 2007-2008**

Service	Services 2007	2007(%)	Services 2008	2008(%)
Screening	6,159,934	77.52%	6,133,612	77.83%
Diagnostic	1,786,621	22.48%	1,746,790	22.17%
Total	7,946,555	100%	7,880,402	100%

Per the suggestion of one of the Steering Committee members, the measure developer contacted the Breast Cancer Surveillance Consortium (BCBS) at the National Cancer Institute (NCI) in order to obtain a point of comparison concerning the distribution of screening versus diagnostic exams found in the registry data. The distribution found in the Medicare claims data is fairly similar to the distribution found in the BCBS registry data. Data from BCSC for women 65 and older from 2007-2008, excluding HMOs (for consistency because HMOs are not in Medicare claims), indicate that 85% of mammograms were coded as screening and 15% were diagnostic. The Statistical Coordinating Center indicated that it is currently in the process of updating an earlier study by Dr. Rebecca Smith-Bindman (“Can Medicare Billing Claims Data Be Used to Assess Mammography Utilization Among Women Ages 65 and Older?,” *Medical Care*, May 2006) in which Medicare claims data classification of mammograms as screening or diagnostic was compared to the patient level registry data against the specific patient level claims data.

Steering Committee Recommendation for #IEP-009-10:

The Steering Committee reviewed the measure developer’s responses to the conditional recommendations on Thursday, April 22, 2010, and subsequently voted on the measure electronically. The results of the vote were as follows:

- Recommend for Endorsement: 6
- Do not Recommend for Endorsement: 10
- Abstain: 3

The measure was not recommended for endorsement.

#IEP-015-10 Cardiac stress imaging not meeting appropriate use criteria: Routine testing after percutaneous coronary intervention (PCI) (American College of Cardiology)

The Committee acknowledged that the measure targets a critical imaging area with significant opportunities to improve efficiency in an expanding and evolving field. Overall, the Committee was generally agreeable to the measure as written and noted that the goal should be zero. As a public reporting tool this measure is precise and interpretable by members of the public. The Committee’s primary concern was specific to the measure’s denominator exclusion criteria, “patients without sufficient patient selection criteria recorded,” which would create a perverse incentive for individuals to not record criteria to improve their performance. Because the Committee was concerned with the perceived narrow scope of the measure, it requested that ACC expand the measure scope to include stress MRI and CTA, even though such an addition may not have a large affect on the measure.

NATIONAL QUALITY FORUM

Although the measure's reliance on chart review presents feasibility issues, the Committee did not consider it reason enough to stop the advancement of the measure because it targets a large problem area.

Response from Measure Developer:

- The majority of facilities would have enough volume (more than 20) for the measure to be applicable.
- ACC uses paper and electronic forms to collect data. However, many facilities do not have electronic systems.

In-person voting results:

- Recommend for endorsement with conditions: 14
- Do not Recommend for Endorsement: 3

The measure was recommended for endorsement with conditions.

***The following contains additional information related to the project and measure evaluation review that took place after the in person meeting.**

Conditions for the Measure Developer for #IEP-015-10:

The Committee made the following conditional recommendations:

- Potentially eligible for time-limited endorsement, would need to affirm a 12 month testing strategy
- Expand the measure to include MRI and CTA or explain why MRI and CTA were not part of or needed in the initial measure
- Remove the denominator exclusion criteria, "patients without sufficient patient selection criteria recorded"
- Expand the denominator population to include CABG
- Consider changing the title of the measure to eliminate potentially negative connotations

Detailed Response from Measure Developer:

- ACC has data available from a recently completed pilot project to demonstrate the validity and reliability of this measure. Given the extensive testing in the pilot project, ACC would argue that time limited endorsement and a further testing strategy is not needed. ACC needs further clarification about what additional testing must be completed. In addition, the ACC has a current quality improvement project with an online data collection tool for national measurement.
- ACC is willing to add MRI and CTA to the measure. However, ACC does not anticipate a large number of cases to be documented for these imaging modalities for follow-up in this patient population. CTA is limited technically in stents of small diameter currently, and MRI is not widely used in this patient population.
- ACC is willing to remove the denominator exclusion criteria, "patients without sufficient patient selection criteria recorded." However, this change will inflate the denominator of

NATIONAL QUALITY FORUM

the measure for imaging laboratories that are unable to locate the information necessary to determine all components of the numerator. As such, the removal could create an incentive not to obtain enough data to clearly indicate a patient qualifies for this measure.

- ACC has decided not to include CABG in the numerator of the measure. The denominator is all stress imaging so it is not appropriate for the denominator. CABG has different timeframes for follow-up testing, generally is done in more complex patients, and may be reasonable in some patients. By focusing on post PCI, the measure maintains a focus on a patient indication that did not meet appropriate use criteria.
- ACC does not think a change of the title to be less negative is warranted. The College revised its initial draft measure title of “Inappropriate cardiac stress imaging” to the current title in an attempt to make the title more neutral. The measure is designed to examine imaging that is not reasonable. Development of a more positive title would not reflect the focus of this measure which is overuse.

Steering Committee Recommendation for #IEP-015-10:

The Steering Committee reviewed the measure developer’s responses to the conditional recommendations on Thursday, April 22, 2010, and subsequently voted on the measure electronically. The results of the vote were as follows:

- Recommend for Endorsement: 19
- Do Not Recommend for Endorsement: 1
- Abstain: 1

The measure was recommended for endorsement with conditions.

Steering Committee voting comments:

- Some members of the Steering Committee would like to see CABG included in this measure.

#IEP-011-10 Use of Stress Echocardiography, SPECT MPI, and Cardiac Stress MRI Post CABG (Centers for Medicare and Medicaid Services)

The Committee expressed significant concerns with the measure as submitted. The Committee’s primary concerns related to the measure’s numerator exclusions, which were considered difficult to justify and narrow in scope and introduce potential unintended consequences for small facilities. The Committee noted that in facilities with a smaller volume the events may be rare enough to compromise the precision of the measure. The Committee expects measures to have the broadest applicability. The measure as currently specified does not capture a critical setting of care (e.g., free-standing outpatient facilities) where there are potentially even larger opportunities for improvement.

Response from Measure Developer:

- There seems to be a number of acquisitions of practices by hospitals; about 80 percent of these procedures are performed in outpatient clinics. Endorsing this measure may serve

NATIONAL QUALITY FORUM

as a baseline for future measures, moving imaging back to the hospital setting. The measure developer noted that this evolution is worth monitoring.

- The measure developer indicated that the measure's scope is at the outpatient hospital level, but it does not have data to look at other free-standing outpatient facilities.

In-person voting results:

The Committee did not vote on the measure at the time of the in-person meeting. The Committee decided to vote on the measure after the measure developer responds to the conditions for recommendation.

***The following contains additional information related to the project and measure evaluation review that took place after the in person meeting.**

Conditions for the Measure Developer for #IEP-011-10:

The Committee made the following conditional recommendations:

- Consider removing the six-month exclusion criteria from the numerator statement or provide justification for the use of the six-month exclusion criteria
- Expand the measure to include PCI, but report CABG and PCI separately
- Expand the measure to include both hospital outpatient and free-standing imaging facilities
- Expand the sample size or provide justification for how the feasibility and validity concerns are addressed for smaller or rural hospitals with small patient populations

Detailed Response from Measure Developer:

The request to remove the exclusion criteria for “patients with catheterization, percutaneous coronary intervention (PCI) or CABG procedure in 6 months following the imaging study” requires additional consideration by the measure developer. The exclusions for this and all submitted measures were developed by a technical expert panel (TEP), subjected to sensitivity analysis using several alternatives for exclusion length, and agreed upon by both the developer and the TEP. Modification of the exclusions would require a similar level of effort and is not possible to complete within the timeframe provided. Further, although CMS is very interested in expanding the measure to include PCI, but report PCI and CABG separately, completing this request is also not feasible within the timeframe given that the request requires additional data acquisition and substantial analysis. As such, although CMS is amenable to expanding the scope of the measure beyond the hospital outpatient setting, it is unable to address the other two aforementioned requests at this time.

With regards to the request to expand the sample size or provide justification on how the feasibility and validity of the measure is addressed for smaller or rural hospitals with small patient populations, the measure developer understands the concern that the committee is raising and believes that adjustment to increase sample size likely may be needed. Given the committee's comments for significant changes to the measure, consideration of adjustment to sample size needs to be addressed in the context of the additional data analyses involved in the modifications recommended by the committee. Thus the previously mentioned time constraints apply to the sample size question as well.

NATIONAL QUALITY FORUM

Steering Committee Recommendation for #IEP-011-10:

The Steering Committee reviewed the measure developer's responses to the conditional recommendations on Thursday, April 22, 2010, and subsequently voted on the measure electronically. The results of the vote were as follows:

- Recommend for Endorsement: 4
- Do not Recommend for Endorsement: 14
- Abstain: 1

The measure was not recommended for endorsement.

Steering Committee voting comments:

- The Steering Committee believes this is an interesting area, but would like CMS to do the extra analysis requested. It takes quite some time for the conditions requested to be met and the Committee understand that. This measure should be applicable for use in rural and small patient population facilities.
- Would like Medicare patients are included (not just hospital OP facilities).
- Steering Committee felt the responses from CMS inadequate as the response after the in-person meeting still did not justify the 6 month exclusion.

#IEP-016-10 Cardiac stress imaging not meeting appropriate use criteria: Testing in asymptomatic, low risk patients (American College of Cardiology)

Given the prevalence of cardiovascular disease and the rise in cardiac imaging expenditures, the Committee deemed this measure's focus to be highly important to improving care and encouraging efficiency in the market. Although the Committee supported the concept of the measure it raised some concerns with the measure's denominator exclusion criteria, lack of risk-adjustment, and narrow scope. The Committee was concerned that the denominator exclusion would create potentially negative unintended consequences. The Committee further noted that the measure may penalize smaller facilities, driving up costs and ultimately distorting the measure.

Response from Measure Developer:

Measure developer agreed to modify of the denominator (see below).

In-person voting results:

The Committee did not vote on the measure at the time of the in-person meeting. The Committee decided to vote on the measure after the measure developer responds to the conditions for recommendation.

***The following contains additional information related to the project and measure evaluation review that took place after the in person meeting.**

Conditions for the Measure Developer for #IEP-016-10:

The Committee made the following conditional recommendations:

- Address the need for risk-adjustment or provide a rationale for not adjusting for risk

NATIONAL QUALITY FORUM

- Expand the measure to include MRI or provide a justification for not including MRI in the scope
- Remove the denominator exclusion criteria, “patients without sufficient patient selection criteria recorded”
- Consider changing the title of the measure to eliminate potentially negative connotations

Detailed Response from Measure Developer:

- Risk is explicitly considered in the measures as it takes into account two clinical characteristics of the patient – symptom status and global risk for congestive heart disease (CHD). The latter includes numerous factors including age, gender, smoking status, blood pressure, lipid profile etc. Exclusions for a known history of CHD, perip evaluation, and prior testing also are included to ensure that patients who are not being seen for initial evaluation of CHD are excluded. Additional risk adjustment is not required since patient risk is already core to the definition of this measure.
- ACC is willing to add CTA and MRI to the measure. However, ACC does not anticipate a large number of cases to be documented for these imaging modalities in this patient population.
- ACC is willing to remove the denominator exclusion criteria, “patients without sufficient patient selection criteria recorded.” However, this change will inflate the denominator of the measure for imaging laboratories that are unable to locate the information necessary to determine all components of the numerator. As such, the removal could create an incentive not to obtain enough data to clearly indicate if a patient qualifies for this measure.
- ACC does not think a change of the title to be less negative is warranted. The College revised its initial draft measure title of “Inappropriate cardiac stress imaging” to the current title in an attempt to make the title more neutral. The measure is designed to examine imaging that is not reasonable. Development of a more positive title would not reflect the focus of this measure which is overuse.

Steering Committee Recommendation for #IEP-016-10:

The Steering Committee reviewed the measure developer’s responses to the conditional recommendations on Thursday, April 22, 2010, and subsequently voted on the measure electronically. The results of the vote were as follows:

- Recommend for Endorsement: 18
- Do not Recommend for Endorsement: 1
- Abstain: 1

The measure was recommended for endorsement.

Steering Committee voting comments:

- There are no comments at this time.

**Preoperative Evaluation for Low-Risk Non-Cardiac Surgery Risk Assessment Measures
(Centers for Medicare and Medicaid and American College of Cardiology)**

NATIONAL QUALITY FORUM

IEP-010-10 Preoperative Evaluation for Low-Risk Non-Cardiac Surgery Risk Assessment (Centers for Medicare and Medicaid (CMS))

IEP-014-10 Cardiac stress imaging not meeting appropriate use criteria: Preoperative evaluation in low risk surgery patients (American College of Cardiology (ACC))

This project's Call for Measures resulted in the submission of two similar measures focused on cardiac imaging for non-cardiac low-risk surgery patients from two different developers. The Committee reviewed both measures separately against the NQF evaluation criteria at the in-person meeting. Due to both the complexities and similarities of the measures, the Committee requested that an expanded discussion take place after the in-person meeting.

In-person voting results:

The Committee elected to not vote on the measures at the time of the in-person meeting. The Committee decided to vote on the measures after the expanded discussion between the Committee and the measure developers.

***The following contains additional information related to the project and measure evaluation review that took place after the in person meeting.**

NQF staff facilitated the expanded discussion via a conference call between the cardiac specialists on the Committee who were the primary reviewers for the measure and the measure developers from CMS and ACC.

Conference participants: Dr. Rucker and Dr. Fiesinger – NQF Imaging Efficiency Steering Committee, Sharman Stephens, Charlie, & Laura Peterson – Lewin Group, Susan Arday – CMS, Joseph Allen & Helen Smith – ACC, Heidi Bossley & Ian Corbridge – NQF staff.

IEP-010-10 Preoperative Evaluation for Low-Risk Non-Cardiac Surgery Risk Assessment (Centers for Medicare and Medicaid)

Expanded discussion following in-person meeting:

Members of the Committee noted that this measure has strong feasibility and is easy to report at no additional costs to facilities. The Committee noted that from a payer perspective this measure addresses a major problem area with significantly high rates of inappropriate testing and therefore would be very useful. The Committee reviewed the CMS measure against the NQF evaluation criteria and then in relation to the ACC measure. Committee members expressed concern with the 1 percent rate of stress MRI testing for cardiac risk presented in the measure. They believed that such a low rate signals the need to focus measurement efforts elsewhere. The measure developer explained that the study looked at only the surgeries recommended by its Technical Expert Panel. The Committee noted that the measure's primary limitation is that it does not include free-standing cardiac centers. The Committee requested that the measure developer expand the measure to include free-standing cardiac centers.

NATIONAL QUALITY FORUM

IEP-014-10 Cardiac stress imaging not meeting appropriate use criteria: Preoperative evaluation in low risk surgery patients (American College of Cardiology)

Expanded discussion following in-person meeting:

Committee members noted that from a payer perspective this measure addresses a major problem area with high rates of inappropriate testing; however, they also expressed concerns regarding the measure's sampling timeframe and exclusion criteria. Specifically, the measure sampling period of 60 days may be too short for some facilities to report a meaningful measure. The Committee requested that the measure developer explore options to expand the sampling timeframe. Regarding the denominator exclusion criteria, "patients without sufficient patient selection criteria recorded," individuals could manipulate the measure by simply not recording patients' selection criteria. The Committee requested that the measure developer remove this exclusion criterion. Further discussions focused on the need to include stress MRI and CTA to the measures. Although some Committee members did not consider it necessary, the Committee ultimately requested that the measure developer add stress MRI and CTA to the measure.

Combined Discussion on Measures IEP-010-10 and IEP-014-10:

IEP-010-10 Preoperative Evaluation for Low-Risk Non-Cardiac Surgery Risk Assessment (Centers for Medicare and Medicaid)

IEP-014-10 Cardiac stress imaging not meeting appropriate use criteria: Preoperative evaluation in low risk surgery patients (American College of Cardiology)

The Committee determined that although measures IEP-010-10 and IEP-014-10 have similar constructs, they are distinct with regard to target populations and data generation. The Committee acknowledged the value in having both measures, but it noted that the measures' lists of low-risk surgeries must be aligned to ensure clarity in public reporting. The Committee requested that the measure developers collaborate to develop an aligned list of low-risk surgeries.

Response from CMS:

- The measure developer indicated that it would explore expanding the measure to other settings to include free-standing cardiac centers.
- The measure developer indicated that it would align its list of low-risk surgeries with ACC's list.

Conditions for the Measure Developer for IEP-010-10:

The Committee made the following conditional recommendations:

- CMS and ACC must align the measures' lists of low-risk surgeries.
- Add free-standing cardiac centers to the measure.

Detailed Response from Measure Developer (CMS):

CMS and ACC discussed the list and believe the detailed CMS list (of low-risk surgeries) is now in alignment with the ACC broader categories listing included in ACC guidelines.

Steering Committee Recommendation for Measure IEP-010-10:

NATIONAL QUALITY FORUM

The Committee noted that the measure has the potential to be a strong indicator of quality and efficiency if the measure developer is able to meet the Committee's conditional recommendations (see conditions for measure developer). The Committee reviewed the measure developer's responses to the conditional recommendation on Thursday, April 22, 2010, and subsequently voted on the measure electronically. The results of the vote were as follows:

- Recommended for Endorsement: 15
- Not Recommended for Endorsement: 1
- Abstain: 3

The measure was recommended for endorsement.

Response from ACC:

The measure developer indicated that it would be willing to align the measure's list of low-risk surgeries with the CMS list and to add stress MRI and CTA to the measure.

Conditions for the Measure Developer for #IEP-014-10:

The Committee made the following conditional recommendations:

- Expand sampling period from 60 days (2 months) to a one year (12 months)
- Remove "patients without sufficient patient selection criteria recorded" from the denominator exclusions
- Add stress MRI and CTA to the measure or provide a detailed rationale for why they were excluded from the measure
- Potentially eligible for time-limited endorsement, would need to affirm a 12-month testing strategy and harmonize the low-risk procedure lists between the ACC and CMS measures
- Provide an overview/summary of the low-risk list

Detailed Response from Measure Developer (ACC):

- ACC is willing to align our example list of procedures with the CMS proposed measure. However, ACC also will maintain the broad definition for included surgeries as being less than 1% mortality (based on the guideline definitions) and maintain the general category list of endoscopic procedures, superficial procedures, cataract surgery, and breast surgery. There is no way to be exhaustive in the list. ACC wants to maintain these general categories to allow clinicians to document other low risk surgeries that cannot be included in the example list.
- ACC is willing to add CTA and MRI to the measure. However, ACC does not anticipate a large number of cases to be documented for these imaging modalities in this patient population.
- ACC is willing to remove the denominator exclusion criteria, "patients without sufficient patient selection criteria recorded." However, this change will inflate the denominator of the measure for imaging laboratories that are unable to locate the information necessary to determine all components of the numerator. As such, the removal could create an incentive not to obtain enough data to clearly indicate a patient qualifies for this measure.
- ACC is willing to consider a rolling sampling period that would change annually. The rolling sampling period would capture any changes in ordering patterns received by an

NATIONAL QUALITY FORUM

imaging laboratory that occur during the course of a year. The rolling sampling period would state that practices that collected during a particular 60 period (e.g. January – early March) the first year, conduct a different 60 day period the following year (early March – April). Addresses a potential concern about practices choosing a more “favorable” sampling period to report. The rolling sample period would ensure that performance was being maintained in the same way during different times of the year by measuring a different time period each measure year. However, our testing data have shown that the measure stability does not change if the data collection period is extended from 2 to 12 months. The rolling sampling period would state that practices that collected during a particular 60 period (e.g. January – early March) the first year, conduct a different 60 day period the following year (early March – April). This measure is based on clinical data collection and thus, ACC does not support extending the data collection period and creating additional collection burden without a measurable impact on the results. In addition, this measure is collected by the imaging laboratory and would require sharing its results with ordering physicians to impact a change. The 2 month sampling period encourages rapid sharing of the data and continuous quality improvement. A longer sampling period would discourage this improvement by putting more time between the data collection and results reporting. Since orders come to the imaging laboratory from multiple referring providers who won’t know the sampling period, the ability of any referring physician to change behavior during the sample period compared to the rest of the year is minimized. ACC does not support a data collection requirement that requires continuous 12 month data collection and supports only a sampling methodology for collection.

Steering Committee Recommendation for IEP-014-10:

The Committee noted that the measure has the potential to be a strong indicator of quality and efficiency if the measure developer is able to meet the Committee’s conditional recommendations (see conditions for measure developer). The Steering Committee reviewed the measure developer’s responses to the conditional recommendation on Thursday, April 22, 2010, and subsequently voted on the measure electronically. The results of the vote were as follows:

- Recommended for Endorsement: 16
- Not Recommended for Endorsement: 3
- Abstain: 0

The measure was recommended for endorsement.

Steering committee voting comments:

- The Steering Committee would like to ensure that there is harmonization of “low risk surgery definition” included in the measure.

IEP-005-10 Appropriate Pulmonary CT Imaging for Pulmonary Embolism (Brigham and Women’s Hospital):

The Steering Committee acknowledged the value of the measure and believed it was best suited as an “overuse” measure rather than strictly as an “efficiency” measure. In changing the measure to an overuse measure the Committee requested that the measure developer amend the numerator specifications, specifically relating to the D-dimer. The Committee noted feasibility issues,

NATIONAL QUALITY FORUM

because it is based on a proprietary electronic data collection tool used at the Brigham and Women's Hospital (BWH). The measure developer presented a paper data collection tool and indicated it would be available to the public for those institutions without an electronic data collection system. Because the paper data collection tool as specified has not been tested, the Committee recommended the measure for time-limited endorsement with conditions (see below).

In-person voting results:

- Recommend for endorsement with conditions: 16
- Do not recommend for endorsement: 3

The measure was recommended for time-limited endorsement with conditions.

***The following contains additional information related to the project and measure evaluation review that took place after the in person meeting.**

Conditions for the Measure Developer for #IEP-005-10:

The Committee made the following conditional recommendations:

- Potentially eligible for time-limited endorsement, would need to affirm a 12 month testing strategy
- Change the measure to become an “overuse” measure. The denominator would remain unchanged, the numerator would read the low clinical probability and either no high-sensitivity D-dimer or a negative high-sensitivity D-dimer
- Provide additional information about how the numerator data is to be captured, specifically at centers other than BWH, where electronic integration of order entry and electronic medical records may not be sufficient to allow for automated data collection
- Consider changing the title of the measure to eliminate potentially negative connotations

Detailed Response from Measure Developer:

- We affirm that we have begun a 12-month testing strategy for this measure.
- The numerator now reads: “The number of denominator patients with either: a low clinical probability and any negative D-dimer, or an intermediate clinical probability and a negative high-sensitivity D-dimer, or no pretest probability documented.” We believe that it is important to require a pretest probability score as part of the pre-test assessment, otherwise clinicians who do not assess pretest risk will not be measured.
- Additionally, we recommend that the NQF request a tracking measure for this measure that is “Percent of CT scans performed by pre-test risk category.” The measure would calculate the percent of patients getting a CT for PE that were pre-test probability low, medium and high. This is important as the current measure is open to gaming. Clinicians can use an unstructured pretest assessment to determine pretest probability (as recommended by the current evidence based guidelines) but this also means clinicians can choose to assign a pre-test risk of high to all patients that they decide to scan.
- Tracking the proportion of pretest probabilities is one way to monitor such decisions. Please see the attached "Sample Data Collection Form for Measure #IEP-005-10.doc" This is a data collection form. We will also prepare a “CT order form” that is paper based

NATIONAL QUALITY FORUM

and that EDs without computerized physician order entry (CPOE) for radiology could use to both order the scan and record a pretest probability. Suggested new title: "Pulmonary CT Imaging for Patients at low risk for Pulmonary Embolism".

Steering Committee Recommendation for IEP-005-10:

The Steering Committee reviewed the measure developer's responses to the conditional recommendations on Thursday, April 22, 2010, and subsequently voted on the measure electronically. The results of the vote were as follows:

- Recommend for Endorsement: 15
- Do Not Recommend for Endorsement: 3
- Abstain: 1

The measure was recommended for time-limited endorsement.

Steering Committee voting comments:

- This measure should be reconsidered for time-limited endorsement, in order for the paper version to be tested.

#IEP-007-10 Appropriate Head CT Imaging in Adults with Mild Traumatic Brain Injury (Brigham and Women's Hospital)

The Steering Committee agreed that the measure is based on strong evidence-based guidelines and targets a critical imaging practice in the emergency department (ED) setting. The Committee discussed the inclusion criteria of the Glasgow Coma Scale (GCS) greater than 13 (as specified) and as a result requested that the measure developer change the inclusion criteria to GCS greater than or equal to 13. As with other measures submitted by the Brigham and Women's Hospital, the Committee had concerns about the measure's feasibility, because it is based on a proprietary electronic system. The measure developer supplied a paper format of the data collection tool to be used at facilities without the proprietary electronic system. Although the paper format presents some challenges to feasibility, the Committee noted that the measure is of great value and will help improve the appropriateness of head CT imaging. Because the paper data collection tool has not been tested, the Steering Committee recommended the measure for time-limited endorsement.

Response from Measure Developer:

- Most of the studies used follow-up by telephone or medical records; the Canadian study has had a 95 percent follow-up. The New Orleans study uses CT and findings on radiology, while the Canadian study uses any finding that requires neurosurgical intervention. The Canadian study requires less of a scan, head to head study of the two. Most U.S. facilities use New Orleans because of a fear of legal actions, even if it doesn't require and specific treatment.
- Researchers have found that at this point, facilities are looking at 60-80% compliance with one of these reviews. They are able to click on a box they wouldn't have otherwise, which has resulted in very high compliance rates. In the Canadian head CT, in chart reviews of current practice there is a large gap with evidence based indication.

NATIONAL QUALITY FORUM

- If a hospital has an EMR imaging ordering platform, then it can develop a paper form and the exclusions and inclusions can be indicated using check boxes.

In-person voting results:

- Recommend for endorsement with conditions – 16
- Do not recommend for endorsement – 3

The measure was recommended for time-limited endorsement with conditions.

***The following contains additional information related to the project and measure evaluation review that took place after the in person meeting.**

Conditions for the Measure Developer for #IEP-007-10:

The Committee made the following conditional recommendations:

- Affirm a 12-month testing strategy using only the paper form of the data collection tool
- Consider changing the inclusion criteria to read a GCS greater than or equal to 13 or provide a rationale as to why the GCS should be greater than 13

Detailed Response from Measure Developer:

- While we have already begun a testing strategy it is based upon computerized physician order entry (CPOE). We cannot commit to conducting a paper testing strategy at this date, but are actively investigating this. We are looking into the possibility of doing this at another Partners HealthCare site or another ED and will be in touch with the committee as soon as we confirm. Most Partners HealthCare sites have CPOE for radiology, although all do not have the active decision support system that is at BWI.
- There have been a number of GCS criteria used in the various studies on which this measure is based. While the Canadian CT Head Rule uses an initial GCS of 13-15 (allowing two hours for normalization of the GCS to 15), the New Orleans Criteria and the majority of later studies have used either a GCS of 15 or a GCS of 14-15 as inclusion criteria. For this reason, the authors of the ACEP Clinical Policy based their recommendation on a GCS inclusion criterion of 14-15. We have created the performance measure to follow this most recent evidence based guideline.
- **Revised/Clarified MD response submitted to NQF on Thursday April 8, 2010:** PHS is committed to conducting a paper testing strategy within the next 12 months but have not yet started it at this time

Steering Committee Recommendation for #IEP-007-10:

The Steering Committee reviewed the measure developer's responses to the conditional recommendations on Thursday, April 22, 2010, and subsequently voted on the measure electronically. The results of the vote were as follows:

- Recommend for Endorsement: 14
- Do not Recommend for Endorsement: 4
- Abstain: 1

NATIONAL QUALITY FORUM

The measure was recommended for time-limited endorsement.

Steering Committee voting comments:

- This should be recommended for time-limited endorsement in order to allow time for paper testing; this will also have an impact on payers collecting this information.

#IEP-008-10 Appropriate Cervical Spine CT Imaging in Trauma (Brigham and Women's Hospital)

The Steering Committee agreed that the measure targets an important imaging modality with obvious potential for improvement. NQF has endorsed a cervical imaging measure related to the use of cervical spine radiographs (#0512); therefore, the Committee suggested that the submitted measure be combined with the endorsed measure. The Committee would then review the combined measure. In recommending the combination, the Committee is not supporting one imaging modality over the other.

Response from Measure Developer:

The measure developer indicated that it would work with the measure steward of the endorsed measure to explore ways to combine the two measures.

Initial in-person voting:

This measure was taken off line to allow for discussion about the potential combination.

#IEP-013-10 Use of Computed Tomography (CT) in the Emergency Department (ED) for Atraumatic Headache (Centers for Medicare & Medicaid Services)

This project's Call for Standards resulted in the submission of two similar measures focused on computed tomography in the ED setting. The Committee reviewed the measures individually and in comparison. The Committee acknowledged that the measure targets a clinical area with high levels of overuse and has the potential for quality improvement. Although important to the Medicare population, the measure would be more appropriately applied to a younger population. The Committee further noted that the measure is highly feasible because it relies on administrative data.

The Committee identified potential limitations in the measure's denominator exclusions and age range. Because this is a retrospective measure, the rates could range from zero to 80 percent, which is a large range and suggests the need for more studies. The Committee noted however that the rates should be zero or 80 percent. During its evidence review, the Committee noted that some of the evidence was not as strong as would be expected and was based on expert opinion.

The Committee acknowledged that the measure addresses an important clinical problem, but it had reservations as to whether the evidence is strong enough to validate the measure. Although there was a general degree of support for the measure, some members of the Committee expressed displeasure with the measure and cautioned against recommending it for endorsement.

NATIONAL QUALITY FORUM

Despite the measure's limitations, the Committee asked the measure developer to specify a more detailed set of implementation instructions to help improve the public reporting of the measure.

In-person voting results:

- Recommend for endorsement with conditions – 15
- Do not recommend for endorsement – 4

The measure was recommended for endorsement with conditions.

***The following contains additional information related to the project and measure evaluation review that took place after the in person meeting.**

Conditions for the Measure Developer for #IEP-013-10:

The Committee made the following conditional recommendations:

- Develop instructions on how to implement the measure

Detailed Response from Measure Developer:

For the purposes of this measure calculation, we assume that a visit is equal to a day. Therefore, if a patient had multiple Brain CTs in the ED the same day, the patient would only be included in the denominator once and in the numerator once. However, if the multiple Brain CTs were conducted on different days the patient would be included in the denominator and numerator more than once.

For visits to be counted in the measure, headache must be the primary diagnosis on the ED claim.

Given these parameters, to calculate this measure:

1. Identify patients seen in the ED for headache using Medicare hospital outpatient claims data on a specific day, including only those claims where headache is the primary diagnosis on the ED claim.
 - a. This is the denominator prior to exclusions.
2. Apply the measure exclusions to the denominator.
 - a. Exclusions include codes for lumbar puncture, dizziness, paresthesia, lack of coordination, subarachnoid hemorrhage, complicated or thunderclap headache, focal neurologic deficit, pregnancy, trauma, HIV, tumor/mass. These exclusion codes must be included on the ED claim.
 - b. Further, exclude all patients admitted to the hospital.
 - c. This is the final denominator.
3. Of the patients remaining in the denominator, determine which patients also received a Brain CT on the same day using Medicare hospital outpatient claims data.
 - a. This is the numerator.
4. Calculate the measure ratio of the numerator to the denominator.

The measure developer believes confusion over this measure stems from the way the denominator and exclusions are written. As such, we propose to rewrite the denominator

NATIONAL QUALITY FORUM

statement to make sure it is clear that the exclusions are the indications for use of Brain Computed Tomography (CT).

Original Denominator: ED patient visits with a primary diagnosis code of headache

Original Exclusions: Claims with secondary diagnosis codes related to:

- lumbar puncture,
- dizziness, paresthesia,
- lack of coordination,
- subarachnoid hemorrhage,
- complicated or thunderclap headache
- focal neurologic deficit
- pregnancy
- trauma
- HIV
- tumor/mass

Imaging studies for ED patients admitted to the hospital.

Proposed Revised Denominator: ED patient visits with a primary diagnosis code of headache who are not admitted to the hospital and with no secondary diagnosis codes related to:

- lumbar puncture,
- dizziness, paresthesia,
- lack of coordination,
- subarachnoid hemorrhage,
- complicated or thunderclap headache
- focal neurologic deficit
- pregnancy
- trauma
- HIV
- tumor/mass

For patients visiting the ED with a primary diagnosis of headache and one of the above secondary diagnoses, the presence of these secondary diagnoses potentially indicates that a CT brain imaging study may be indicated, depending upon the individual physician assessment of the particular patient.

Steering Committee Recommendation for #IEP-013-10:

The Steering Committee reviewed the measure developer's response to the conditional recommendation on Thursday, April 22, 2010, and subsequently voted on the measure electronically. The results of the vote were as follows:

- Recommend for Endorsement: 17
- Do not Recommend for Endorsement: 1
- Abstain: 1

The measure was recommended for endorsement.

Steering Committee voting comments:

- This is a poorly conceived measure which will receive lots of pushback.

NATIONAL QUALITY FORUM

#IEP-002-10 Screening Mammography Positive Predictive Value 2 (PPV 2- Biopsy Recommended) (American College of Radiology)

The Committee noted that this measure addresses a very important area but has challenges in its constructs and is minimally related to outcomes. The Committee's primary concern related to the measure's name: the name should read "positive predictive value 1" instead "positive predictive value 2." Although the Committee acknowledged the potential value of the measure, it decided that it cannot be used in isolation. Given the concerns with the measure's lack of actionable information at the facility level, the Committee did not recommend the measure for endorsement.

Steering Committee Recommendation for #IEP-002-10:

The Steering Committee unanimously did not recommend this measure for endorsement.

#IEP-017-10 Adequacy of data to assess appropriate use of cardiac stress imaging (American College of Cardiology)

The Committee ultimately reached the consensus that the measure does not sufficiently meet NQF's evaluation criterion of importance. Specifically, the specified numerator and denominator are identical, limiting or eliminating the meaningfulness of the measure. Furthermore, the measure is not a measure of *efficiency*; rather it is a measure that determines if a patient's chart indicates why a test was performed. The Committee noted further problems pertaining to the measure's data specifications. Given all of these concerns, the Committee did not recommend the measure for endorsement.

Response from Measure Developer:

The measure developer indicated that the measure was constructed to avoid one aspect of gaming and that the easiest way to overcome gaming is to force clinicians to document actions in the chart.

Steering Committee Recommendation for #IEP-017-10

The Steering Committee unanimously did not recommend this measure for endorsement.

#IEP-012-10 Simultaneous Use of Brain Computed Tomography (CT) and Sinus Computed Tomography (CT) (Centers for Medicaid and Medicare)

The Committee had immediate concerns about the measure's importance, because it does not target an imaging practice with a substantial or large magnitude of overuse. When the Committee inquired about the rate of overuse, the measure developer indicated that approximately 5 percent of patients who receive a brain CT also receive a sinus CT on the same day, thus reaffirming the Committee's low rating on importance. Furthermore, the Committee determined that a substantial number of facilities would not be able to report the measure because their sample size would be too small. The Committee noted that the measure, which excludes a substantial number

NATIONAL QUALITY FORUM

of facilities from quality measurement, is not ready for public reporting. Given the measure's low rating for importance and the substantial issues regarding its usability, the Committee did not recommend the measure for endorsement.

Steering Committee Recommendation for #IEP-012-10:

The Steering Committee unanimously did not recommend this measure for endorsement.

#IEP-006-10 Appropriate Head CT Imaging in Adults with Acute Atraumatic Headache (Brigham and Women's Hospital)

This project's Call for Standards resulted in the submission of two similar measures focused on appropriate head CT imaging in adults with acute atraumatic headache. This measure uses different specifications and data sources than the other head CT imaging measure (#IEP-012-10) and is based on the American College of Emergency Physicians' Clinical Policy. Although there is value in measuring this clinical practice, the Committee was concerned about the measure guidelines, which include both level B and level C recommendations. Level C recommendations include "panel consensus" in addition to recommendations based on lower quality studies. Because high-level evidence to support the efficient use of CT imaging in adults with acute atraumatic headache is lacking, the Committee was concerned about recommending the measure for endorsement.

***The following contains additional information related to the project and measure evaluation review that took place after the in person meeting.**

Response from Measure Developer:

The measure developer asked the Committee to revisit the measure based on the fact that the Committee had approved a similar measure that addresses the identical clinical area, but with an even older level of evidence. The Committee reviewed the additional documentation provided by the measure developer.

Excerpt from Appeal Letter from Dr. Jeremiah Schuur, MD, MHS sent on April 4, 2010:

We do not agree that the guideline upon which our measure is based is a "consensus statement." The ACEP Clinical Policy is a recent, multidisciplinary expert review committee that follows a systematic process to review the published literature and develop graded recommendations. The evidence base addressing Head CT for acute headache is of poor quality, and includes no Level I evidence. Of the 28 studies used in the 2000 US Headache Consortium Guidelines, all were considered to be Level IV evidence at best. To our knowledge, there are no studies evaluating Head CT for acute headache using blinding, comparison against a gold standard, or randomized or consecutive patient samples. In our review of the literature, there are no Level IA (or Class I) studies that address the issue of CT in acute headache, let alone the appropriateness of CT in acute headache. Furthermore, the guidelines used for current measure development are based on retrospective studies of acute headache patients and either exclude or do not address important sub-populations such as patients with a cancer history, neurosurgical interventions, or other chronic neurological conditions such as normal pressure hydrocephalus which may indicate the need for CT but not meet the stringent neurological exam based criteria used in these retrospective studies. This is why when faced with the question of "Which patients with

NATIONAL QUALITY FORUM

headache require neuroimaging in the ED?” the ACEP guidelines committee recommendations were graded Level B and Level C. The weakness of the evidence applies equally to measure #IEP-013-10.

In-person voting results:

The measure was not recommended for endorsement because it does not meet the threshold criterion of importance criteria. Three Committee members voted that the measure’s importance is moderate, and 15 voted that the importance is minimal.

Steering Committee Recommendation for #IEP-006-10:

The Committee reviewed the measure developer’s appeal and additional information on Thursday, April 22, 2010, and subsequently voted on the measure electronically. The Committee members again reiterated their professional judgment that the measure did not meet the importance criterion.

The measure was not recommended for endorsement.

Steering Committee voting comments:

- This measure represents an important area in overuse. I still find that the original measure lacked information in regards to anticoagulation. This measure has untested paper forms. The Steering Committee supports the idea of having a measure for headache pts that come to the ED as this is a problem where over-imaging is occurs the most. From a payer perspective is most expensive; however this measure is not ready for prime time.
- The data is not out there to evaluate but the significance of determining this metric. The numerator range seems broad enough to allow very reasonable criteria for use of CT. This measure should not be made public and perhaps should be time- limited to allow for paper form to be tested.