NATIONAL QUALITY FORUM
Driving measurable health improvements together

Patient Safety Fall 2021 Cycle: Public and Member Comments

Contents

https://www.qualityforum.org

1099 14th Street NW, Suite 500 | Washington, DC 20005 | M 202.783.1300 F 202.783.3434

Post-Evaluation Measure-Specific Comments on Patient Safety Fall 2021 Submissions

NQF #3636 Quarterly Reporting of COVID-19 Vaccination Coverage among Healthcare Personnel (Recommended)

Lynne Batshon, SHEA; Submitted by Geeta Sood

Comment ID#: 7988 (Submitted: 04/29/2022)

Council / Public: HPR

Level of Support: N/A

Comment

Dear NQF Patient Safety Committee, The Society for Healthcare Epidemiology of America (SHEA) is committed to improving the quality of care in healthcare settings. We appreciate the thoughtful review of the NQF 3636 Quarterly Reporting of COVID-19 Vaccination Coverage among Healthcare Personnel in the Nursing Homes metric. COVID-19 has caused substantial morbidity and mortality in older adults. Vaccines for COVID-19 were shown to be effective in preventing severe disease and in reducing transmission and higher nursing home HCP vaccination rates are associated with better residents' outcome for COVID-19. SHEA is strongly supportive of healthcare personnel vaccinations to best protect the patients we serve. We appreciate that the metric minimizes the reporting burden by requiring quarterly reporting while maintaining meaningful measurement. SHEA supports endorsement of the NQF 3636 Quarterly Reporting of COVID-19 Vaccination Coverage among Healthcare Personnel in Nursing Homes. Sincerely, Sharon B. Wright, MD, MPH, FIDSA, FSHEA 2022 President, Society for Healthcare Epidemiology of America

Developer Response

N/A

NQF Response

Thank you for your comment. It has been shared with the Standing Committee and the measure developer.

NQF Committee Response

N/A

Stephanie Collingwood, UnityPoint Health; Submitted by Stephanie Collingwood

Comment ID#: 7968 (Submitted: 04/25/2022)

Council / Public: PRO

Level of Support: N/A

Comment

UnityPoint Health respectfully offers comments in opposition to measure 3636 as outlined below. UnityPoint Health is one of the nation's most integrated health care systems. Through more than

32,000 employees and our relationships with more than 480 physician clinics, 40 hospitals in urban and rural communities and 14 home health agencies throughout our 9 regions, UnityPoint Health provides care throughout Iowa, central Illinois, and southern Wisconsin. On an annual basis, UnityPoint Health hospitals, clinics and home health provide a full range of coordinated care to patients and families through more than 8.4 million patient visits. Today, UnityPoint Health reports vaccination information under the HHS COVID-19 reporting requirement as directed through the federal public health emergency (PHE) and thus, additional reporting of this measure becomes duplicative. In addition, hospitals typically keep employee health records outside of their electronic health record (EHR) due to health privacy concerns. With that said, attempting to identify and collect data on employee vaccine adherence is inherently difficult and burdensome. Additionally, as proposed, some of the measurement categories are difficult to capture, such as contract personnel. Due to the recommendation of the CDC, health care facilities are one of the only remaining locations to require masking and have longer exposure restrictions and testing requirements. This higher burden to health care, provides an exponentially more conservative work environment than general industry. The impact of this industry variation has been experienced within CMS's COVID-19 vaccine regulation. Due to the CMS regulation, UnityPoint Health is already ensuring all construction personnel are fully vaccinated or have an approved exemption to the COVID-19 vaccination. While collecting this information, UnityPoint Health has experienced many barriers, including vendors unwilling to share their employee vaccination records citing personnel information is confidential. We believe it may become more difficult to find partners for construction projects with this proposed measure. Furthermore, this type of immunization recordkeeping is not reported for other transmissible diseases. Scientifically speaking, tracking, and reporting of vaccination status is not an evidence-based intervention that results in improved outcomes. UnityPoint Health opposes measuring COVID-19 Vaccination Coverage among Healthcare Personnel (HCP) as a quality measure.

Developer Response

It is not entirely clear to which specific reporting the Commenter is referring. The Commenter may be referring to acute care facility reporting of healthcare personnel vaccination information as a component of the CMS public health emergency response (CMS-152-F) and as a component of CMS quality measurement programs (CMS-1762-F)(https://www.hhs.gov/sites/default/files/covid-19-faqs-hospitals-hospital-laboratory-acute-care-facility-data-reporting.pdf and https://www.federalregister.gov/documents/2021/08/13/2021-16519/medicare-program-hospital-inpatient-prospective-payment-systems-for-acute-care-hospitals-and-the). If so, this comment may reflect concern about duplicative reporting requirements of specific programs but does not appear to reflect opposition to NQF endorsement of measure #3636 itself. Healthcare worker COVID-19 vaccination is associated with reduced patient COVID-19 infections and deaths (N Engl J Med. 2022 Jan 27;386(4):397-398). Recording healthcare worker COVID-19 vaccination information may pose challenges and some associated burden, but healthcare worker COVID-19 vaccination is an important intermediate outcome directly relevant to patient safety. Reporting vaccination coverage for contractors (e.g., construction personnel) is not required for NQF #3636. There has been immunization record keeping and reporting of influenza vaccination coverage among healthcare personnel for many years across many healthcare facility types. CMS Quality reporting programs have required reporting influenza vaccination coverage among healthcare personnel by

acute care hospitals beginning in 2013, by inpatient rehabilitation facilities and long-term acute care hospitals beginning in 2014, and by Prospective payment system (PPS)-exempt cancer hospitals beginning in 2016. Evidence that tracking COVID-19 vaccination rates has directly and independently improved outcomes may not currently be available. However as noted above, there is evidence that reduced patient COVID-19 infections and deaths are associated with high healthcare personnel COVID-19 vaccination coverage, which provide supporting evidence for tracking vaccination rates. Tracking COVID-19 vaccination rates is feasible and continued monitoring of COVID-19 vaccination coverage is important as new personnel are hired, and additional doses of vaccine are recommended (Public Health Rep. Mar-Apr 2022;137(2):239-243).

NQF Response

Thank you for your comment. It has been shared with the Standing Committee and the measure developer.

NQF Committee Response

Proposed Response: Thank you for your comment. The Standing Committee found the specifications clear and does not anticipate the measure will add undue burden to measured entities.

NQF #3633e Excessive Radiation Dose or Inadequate Image Quality for Diagnostic Computed Tomography (CT) in Adults (Clinician Level) (Recommended)

J. Daniel Bourland, AAPM President, American Association of Physicists in Medicine; Submitted by Richard Martin

Comment ID#: 8009 (Submitted: 04/29/2022)

Council / Public: Member

Level of Support: N/A

Comment

The American Association of Physicists in Medicine (AAPM), is pleased to submit comments to the National Quality Forum (NQF) regarding its Patient Safety Standing Committee (PSSC) evaluation report of the following measures that the PSSC recommended for endorsement: NQF #: 3633e - Excessive Radiation Dose or Inadequate Image Quality for Diagnostic Computed Tomography (CT) in Adults (Clinician Level) NQF #: 3662e - Excessive Radiation Dose or Inadequate Image Quality for Diagnostic Computed Tomography (CT) in Adults (Clinician Group Level) NQF #: 3663e - Excessive Radiation Dose or Inadequate Image Quality for Diagnostic Computed Tomography (CT) in Adults (Facility Level) Background These electronic clinical quality measures (eCQM) are intended to monitor CT performance to discourage unnecessarily high radiation dose while maintaining adequate image quality. The proposed metrics require CT Category (i.e., the CT exam type), the size adjusted radiation dose [the patient's dose length product (DLP) adjusted by patient size], and the global noise (associated with the variance of the voxel values in CT images). The two reported measures are the percentage of eligible CT cases in a particular category deemed to be "out-of-range" compared to defined thresholds with respect to the size-adjusted radiation dose or the

global noise in a set time period. The measures are intended to advance quality assurance. In January 2022, prior to the Patient Safety Standing Committee's meeting to evaluate these proposed measures, AAPM provided comments on the measure application to the committee. AAPM attended the committee meeting and now responds to the committee's evaluation report. The AAPM and our leadership in medical physics AAPM, as the primary scientific and professional organization of physics in radiology and radiation oncology in the United States, is the foremost organization with expertise to speak to the topic under consideration. With 9717 members in 94 countries, AAPM supports the Medical Physics community with a focus on advancing patient care through education, improving safety and efficacy of medical imaging procedures through research, education and the maintenance of professional standards. Medical physicists contribute to the effectiveness of medical imaging by ensuring the safe and effective use of radiant energy (e.g., optical, ionizing, ultrasonic, or radiofrequency) to obtain detailed information about the form and function of the human body. Medical physicists continue to play a leading role in the development of novel imaging technologies, as well as in guiding the optimization of existing imaging modalities. General Comments AAPM commends NQF's efforts in advancing and evaluating quality assurance measures. The last 15 years of CT technology development has included new reconstruction algorithms and tube current modulation techniques resulting in substantial reductions in dose. AAPM supports efforts to enhance consistency of CT practice as evidenced by AAPM's proactive engagement in efforts to ensure diagnostic quality CT imaging, optimizing CT dose, and achieving consistency across facilities, considering differing technologies and practices. AAPM, together with other non-profit entities, including the American College of Radiology (ACR), and Image Wisely and Image Gently Alliances has spent decades working towards this goal and continues to do so through many initiatives. AAPM does not support the endorsement of NQF #3633e, #3662e, and #3663e. AAPM cautions that the measures recommended for endorsement by the PSSC have significant limitations that impact their scientific and practical value and overall likelihood of clinical acceptance. These limitations include improper representation of image quality, improper estimation of radiation risk, and substantial oversimplified representation of implementation in practice, including not addressing the challenges of implementation. We will address these concerns in the following paragraphs. Specific Comments PSSC failed to adequately review and consider expert opinion The PSSC failed to adequately review or consider AAPM's expert comments, as required. AAPM review of the proposed measures consisted of a detailed analysis by four prominent senior physicists from four separate institutions. AAPM's comments, however, were not considered as evidenced by the deliberations of the committee at its meeting and in the present report. AAPM's leadership in medical physics – national and international expertise and recognition AAPM's expertise in medical physics is widely recognized and valued by the Nuclear Regulatory Commission (NRC), Food and Drug Administration (FDA), National Institute of Biomedical Imaging and Bioengineering (NIBIB), National Council on Radiation Protection and Measurements (NCRP), other federal agencies and state radiation safety agencies. These agencies routinely engage AAPM on clinical practice, emerging technology and radiation safety issues and seek out AAPM members to serve on their advisory committees addressing the most cutting-edge issues in the radiation medicine field. Thus, AAPM's expert voice on this topic is of high scientific and practical relevance to provide consensus guidance on this important topic. Unscientific characterization of CT scan risk The measure developers include specific numbers estimating the number of cancers and deaths due to these cancers from the dose imparted from the CT scans. The authors describe these risks and the resulting estimates as based on models only. The applied

linear non-threshold model is currently HIGHLY disputed at diagnostic CT radiation dose levels. The resultant estimates of risk are known to involve large uncertainties. Moreover, the science of radiation risk estimation from CT examinations is based on calculation of dose to individual organs, age, and sex. The measures of risk proposed here, however, mention none of these factors or offer a strategy to incorporate it. The proposed measures are primarily based on radiation output of the CT system, not the risk to the patient. The benefit, if any, of minimizing patient dose cannot be scientifically statistically determined. AAPM is concerned that the stated risk of patient radiation dose and financial savings are hypothetical, exaggerated, and may contribute to fear of diagnostic medical exams that may in turn lead some patients to refuse safe and appropriate medical imaging, to the detriment of the patient. Diagnostic imaging doses are typically much lower than 100 mSv, and the anticipated benefits to the patient of medically appropriate imaging are highly likely to outweigh any small potential risks. Measures lack usability The usability of data resulting from these measures is not clear. In their pilot study, 30% of the CT cases for individual clinicians being out-of-range was the median value with half of the clinicians having between 16% and 43% of their cases out-of-range, as shown in Figure 1b-2 of the application. The measures do not provide the clinician with an analysis of or methodology for determining what improvements should be made to address a poor showing with these parameters. It may not be clear to practitioners what a poor score means or how to address it. Complexity of CT categorization The measures rely on the categorization of CT data into cohesive groups. There is, however, significant variability in the CT protocol lexicon across institutions that results in making assignment of a given protocol to one of these categories very challenging. The proposal does not address the magnitude of this challenge or present the means to overcome it, given that current standards lack uniform characterization of protocols. Inadequate measure of noise The proposed noise measure is not an adequate or sufficient parameter of overall image quality. Visually different texture patterns can have similar noise values, and each may be of more, or less, diagnostic value for the radiologist. As mentioned in the proposal, noise can be influenced by many different parameters, such as slice thickness, kV, and mAs. The effect on noise of these parameters is mostly predictable (particularly in a well-defined “subject”, such as a phantom). Noise is commonly determined in a standardized phantom. Noise measured in clinical images is another matter. There has been limited scientific work in that area and none is cited as having been performed by the authors. There is no information provided in the proposal about how the proposed global noise measure is calculated. In particular, the approach does not take into consideration the CT reconstruction settings that can have a dramatic impact on the appearance of the images, including noise, contrast (or CNR), and sharpness. Further, a “global noise” ignores the diversity within the CT series, especially within the (usually) limited locations that depict the abnormality of interest. Inadequate assessment of image quality Image noise alone is an insufficient descriptor of image quality. Noise in an image may also be justifiably varied to meet certain clinical needs (such as high resolution). Many other factors must be considered when attempting to define image quality. Spatial resolution, which includes visualizing small objects and image boundaries, and contrast resolution, of which noise is one component, are also critical aspects of image quality. Widely different noise values may be acceptable under different circumstances for similar protocols. Spatial resolution and contrast are as important as image noise. It is not all clear that improvements in global noise will in turn lead to improved clinical performance. Flawed assumption regarding clinical CT practice There is substantial variation in the radiation doses used in CT exams because the radiation delivered is protocol-specific. The implication in the proposed measures is that radiologists vary these

parameters indiscriminately. In most cases, however, these protocols are established by the institutions based on available equipment, patient population, expertise, scientific evidence, and the nature of cases presented at that institution. With the proposed measures, an optimum study is one that delivers the least radiation dose with an acceptable global noise level, but no evidence is provided that clinicians with high values for the proposed measures perform better or even adequately, only that they perform with less radiation dose. AAPM recommends using an alternative approach that would optimize the acquisition for the task at hand to deliver the least radiation dose necessary while still providing the diagnostic image quality necessary for the task. Dose reduction in and of itself is not enough to improve CT practice. There should also be no loss of clinical performance which is not guaranteed by these measurements. Global noise is not an adequate sufficient metric for image quality. Inadequate accuracy in patient size estimation Assessing a patient size can be challenging because of significant variability in the differences in the habitus of different patients, coupled with the existential challenge that there is no single metric capturing the size of a patient of varying diameter at different cross-sectional locations. The proposed measures rely on the calibration of the company's black-box size estimation to prior work of Cheng 2013 and Christianson 2012, both of which have been updated to newer versions to correct erroneous measures. The error in size measurements needs transparency and validated results. Limited expertise and track record of Alara Imaging Alara Imaging is a new (2020) company without a significant track record of having previously performed a project of such wide scope, scientifically or technically. While the measure developers have published on the topic of radiation dose, they have limited expertise or history with clinical CT, CT image quality, or CT technology. The company has no scientific track record on CT technology, size estimation, or image quality assessment to be considered steward of measures on which it lacks expertise, publication, or scientific history. Moreover, Alara Imaging has limited experience in IT development, with no demonstrated history of interfacing with complex EHR or Radiology Clinical Data systems. The software interface is problematic because it is a vector to a 3rd party product, which can expose healthcare organizations to ransomware attacks by malicious actors seeking valuable patient medical information. In summary, AAPM does not support the endorsement of NQF #3633e, #3662e, and #3663e. AAPM urges NQF to: • Address the concerns identified by AAPM experts; and • Reconsider its recommendation endorsing these measures as proposed. AAPM recognizes that this topic is complex, including scientific, technical and clinical components, and we would welcome the opportunity for greater in-depth discussion on meaningful measures of quality imaging practice. Thank you again for the opportunity to comment on the PSSC evaluation report. If you have any questions or require additional information, please contact Richard J. Martin, JD, Government Relations Project Manager, at 571-298-1227 or Richard@aapm.org.

Developer Response

UCSF thanks the American Association of Physicists in Medicine (AAPM) for their additional comments. UCSF would like to respond and address several inaccuracies and misunderstandings in how the measure is calculated and its intent. COMMENT 1: [Medical Physicists and the American Association of Physicists in Medicine have extensive content expertise that should be considered]…and "while the measure developers have published on the topic of radiation dose, they have limited expertise or history with clinical CT, CT image quality, or CT technology." RESPONSE 1: The measure developers agree that medical physicists have relevant and important

expertise, and they have involved medical physicists in all aspects of our work including both the measure development itself and all of the work in the preceding decade that laid the foundation for UCSF's development of this measure. FIRST: J. Anthony Seibert, PhD was included as a member of our Technical Expert Panel (TEP) to ensure consideration of the perspectives of medical physicists at every step of measure development. Dr. Seibert recently retired as Professor of Diagnostic Imaging Physics and Associate Chair of Radiology Informatics at UC Davis Health and is a past president of the AAPM (2011). In addition to serving on our TEP, Dr. Seibert led UC Davis as a measure testing site and wrote a letter of support for the measure confirming it was "highly feasible" to calculate the measure and noting his belief "that this quality measure can significantly reduce the use of excessive high radiation dose as well as inadequate, sub-optimal low dose used for clinical CT studies." SECOND, the developers also worked closely with another medical physicist, Tim Szczkutowicz, PhD on measure development. Dr. Szczkutowicz guided the work of automating the calculation of image noise, expanding on his earlier published work in this area (Malkus 2017). Dr. Szczkutowicz is an Associate Professor in the Department of Radiology at the University of Wisconsin Madison School of Medicine and Public Health with affiliations in the Department of Medical Physicists and Biomedical Imaging. Lastly, this measure development effort has been led by Rebecca Smith-Bindman, MD, a radiologist and epidemiologist whose primary area of research for the last 15 years has been in quantifying the radiation doses used for CT scanning and identifying ways to safely reduce excessive doses. Over the last 10 years, her research team created a CT radiation dose registry of more than 8 million exams from over 160 facilities, which has allowed the team to quantify the variation in dose, to understand the cause of the variation, and to develop and study interventions to help facilities appropriately lower doses without loss of image quality. The development of this quality measure was a natural extension of this work, and the registry has allowed for the testing of the adult measures. As part of this past work, Dr. Smith-Bindman led a randomized controlled trial of two interventions to optimize CT radiation doses across 100 hospitals and imaging facilities and found that providing feedback (similar to that proposed for these quality measures) along with education and opportunities for sharing best practices resulted in meaningful dose reductions (up to 40%) without any loss in image quality (Smith-Bindman 2020). In total, 13 medical physicists served as site-Principal Investigators for this NIH funded-trial (R01CA181191). Thus, medical physicists have contributed substantially to the body of work that led to the measure as well as measure development. In large part based on Dr. Smith-Bindman's 15-year track record in this area, involving medical physicists, CMS awarded UCSF a cooperative agreement to develop these CT quality measures under the Medicare Access and CHIP Reauthorization Act of 2015 ("MACRA"). COMMENT 2: [There is] unscientific characterization of CT scan risk … the AAPM is concerned that the stated risk of patient radiation dose .. may contribute to fear of diagnostic exams that may in turn lead some patients to refuse safe and appropriate imaging. RESPONSE 2: The measure is not focused on radiation risk and does not calculate nor report radiation risk. A review of the published epidemiological evidence summarizing radiation risk is provided in the application as background and context for the measure (see sections 1a.01-1b.01). This includes several systematic reviews, cohort studies, and an extensive and comprehensive review from the National Academies that conclude that exposure to CT (or radiation doses in the same range as CT) increases a person's risk of developing cancer. Most of these studies do not rely upon the linear no threshold model that the AAPM criticizes in its comment. Indeed, the systematic reviews are based on an observed elevated risk of cancer among patients exposed to medical imaging. The estimates used in the application are based on this

extensive literature review. HOWEVER, radiation risk is not part of the measure at all: it is not calculated nor is it reported. INSTEAD, the measure evaluates dose length product (DLP), and specifically whether size-adjusted DLP exceeds thresholds specific to CT category. DLP is the radiation dose measure most directly under the control of providers, determined by specific parameters that were chosen for the scan. Further, DLP is universally reported by CT scanner manufacturers (unlike other metrics of radiation dose or risk). THE TEP, including the American College of Radiology (ACR), several radiologists, and the medical physicist serving on the committee, unanimously recommended and supported the radiation dose measure used (DLP) and unanimously agreed that it is a relevant metric of quality for CT imaging, as noted in Validity Results, 2b.03. There is also considerable precedent for using DLP to evaluate radiation dose in CT. The American College of Radiology has used DLP to set benchmarks [Kanal 2017] and to measure CT radiation dose in their own NQF-endorsed quality measure #3621. There is no reason to believe that endorsing this measure, which seeks to standardize practice and reduce extreme radiation dose outliers based on DLP, would result in patients refusing appropriate imaging. COMMENT 3: The usability of data resulting from these measures is not clear. The measures do not provide the clinician with an analysis of or methodology for determining what improvements should be made to address a poor showing. It may not be clear to practitioners what a poor score means or how to address it. RESPONSE 3: Entities that report the measure using the measure steward's software will be provided information to both identify causes of performance gaps and make targeted changes to improve quality. There are only two conditions that would push a CT out of compliance - high radiation dose, and poor image quality. The reason for failure (high dose or low quality) will be available to sites on a scan-by-scan basis for those that report the measure using the measure steward's software Those scans where the radiation dose is too high, the dose should be lowered through usual means (technologist education, protocol changes). For those where image quality is too low, more radiation dose should be used through similar means. COMMENTS IN SUPPORT of the measure from many of the testing sites describe how useful the information provided was to allow them to understand and improve their practice. (Available here, beginning page 113: https://www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=96982) As described in our submission and noted above, UCSF found in a randomized controlled trial in 100 hospitals and outpatient radiology practices that providing detailed audit and feedback on radiation doses, similar to what will be provided as part of the feedback on this measure, resulted in significant reductions in radiation dose with no impact on satisfaction with image quality, described in Usability, 4b.01. (Smith-Bindman, 2020) The measure steward does not have control over how 3rd party vendors will report information back to reporting entities. COMMENT 4. The measures rely on categorization of CT data into cohesive groups… There is, however, significant variability in the CT protocol lexicon across institutions that results in making assignment of a given protocol to one of these categories challenging. [There is] substantial oversimplified representation of implementation in practice RESPONSE 4: For the reason the AAPM highlights and for another important reason describe below, the CT category assigned by the measure (reflecting the indication and appropriate radiation dose level for the scan) does not rely on the protocol name at all. As described in Specifications, sp-11, clinical indication for imaging is determined using an algorithm that combines procedure (CPT®) and diagnosis (ICD-10-CM) codes associated with the clinical visit when the test was ordered, information provided as part of the order, and information on the final bill. The codes are available in the radiology electronic systems and/or the EHR or billing systems. The goal in creating the CT categorization decision rules was to identify exams that

are exceptions to the routine dose category (i.e., either high or low dose), (Smith-Bindman, 2021.) Details of the approach for developing and validating the assignment of CT exams to categories in an automated fashion are provided in the measure submission (see Validity sections 2b.02 and 2b.03) and in detailed answers provided to initial comments made by the ACR and AAPM. This approach was first developed using records from over 4.5 million CT exams in the UCSF International CT Dose Registry (Smith-Bindman, 2021), and then turned into an algorithm that combines procedure (CPT®) and diagnosis codes (ICD-10-CM) associated with the clinical visit when the test was ordered, information provided as part of the order, and information on the final bill, provided in Specifications, sp-11. This algorithm was developed using detailed review of over 10,000 patient records from UCSF Health, and validated against "gold standard" chart review, as described in Validity sections 2b.02 and 2b.03. When the algorithm was deployed at our testing sites (including 16 hospitals and 13 outpatient imaging centers), the correct classification rate of the assignment of CT exams to CT category in field-testing was excellent (over 90% for all reporting levels: clinician, clinician group, and facility). Knowing that the algorithm was developed using data from a single health system, the developers performed detailed investigation of the categorization results at testing sites – comparing the assigned CT category against full radiology reports – for the purpose of improving the algorithm. One of the strengths of these measures is that they do not determine the CT category using the protocol name, as this would mask an important quality improvement opportunity (namely, the selection of which protocol to use to scan the patient). Two key process of care components determine radiation doses: (A) the choice of imaging protocol, for example, whether a patient with a suspected pulmonary embolism is imaged with a single- or double-phase CT exam (a decision usually made by the performing radiologist); and (B) the technical settings used for that type of CT exam, which are usually at the discretion of the technologist or medical physicist who oversee and operate the machines. As both of these components contribute to radiation dose, a comprehensive quality measure must encompass both of these decision-making processes. By determining the CT category independent of the protocol used, the measures is able to evaluate both components of quality. COMMENT 5 The noise measure is not an adequate or sufficient parameter of overall image quality. RESPONSE 5: Several comments focus on image quality and the concern that the measure does not offer a comprehensive assessment of image quality. The measure is not intended as a robust measurement of image quality. The primary focus of our measure is to assess radiation dose adjusted for body size. The image quality component was included to protect against the unlikely possibility of substantial degradation of image quality as an unintended consequence of dose reduction. Our measure of image quality uses thresholds developed based on radiologists' satisfaction with images, reflecting what in practice is regarded as adequate for diagnosis. Others might have an interest in more nuanced assessment of image quality for other purposes, but that was not our intent. If the measure is adopted and used, the Steward will closely monitor image noise and measure failure due to low image quality. The Steward will be sensitive to any signal that there is a problem and will revise the measure if changes are needed. COMMENT 6: With these measures, an optimum study is one that delivers the least radiation dose with an acceptable global noise level.. but no evidence is provided that clinicians with high values for the measures perform better or even adequately, only that they perform with less radiation dose. AAPM recommends using an alternative approach that would optimize the acquisition for the task at hand to deliver the least radiation dose necessary while still providing the diagnostic image quality necessary for the task. RESPONSE 6: This measure provides a standardized method for monitoring the

performance of diagnostic CT to discourage unnecessarily high radiation doses, a risk factor for cancer, while preserving image quality. It is expressed as a percentage of eligible CT exams that are out-of-range based on having either excessive radiation dose or inadequate image quality, relative to evidence-based thresholds based on the clinical indication for the exam. THE HIGHER THE SCORE, the higher the proportion of out-of-range exams and the worse the performance. The measure is NOT INTENDED to improve diagnostic accuracy. The purpose of the measure is to establish a radiation dose ceiling to avoid excessive radiation exposure, and an image quality floor to safeguard against unintended deterioration of image quality. ADDITIONALLY, the entire framework for the measure is to ensure the radiation dose and image quality are acceptable for the specific clinical indication, aligned with what the AAPM recommends. The CT categories were created based on radiation dose and image quality requirements specific to the clinical indications for imaging (Smith-Bindman 2021). Using radiologists' satisfaction with image quality, an image quality floor for each category was established, below which an exam is considered to have inadequate quality, and a radiation dose ceiling, beyond which doses are considered unnecessarily high. The purpose is to allow detailed assessment of each CT exam to ensure the dose is optimal based on the clinical indication for imaging. In our testing data, far more CT exams exceeded the radiation dose ceiling (average = 30%) than failed to meet the image quality requirement (average << 1%) (see section 1b.02). The measure encourages entities to reduce the proportion of exams that may "be overdosed for their exact need and condition" while preserving the minimum image quality. COMMENT 7: [There is] Inadequate accuracy in patient size estimation… Assessing a patient size can be challenging because of significant variability in differences in the habitus of different patients, coupled with the existential challenge that there is no single metric capturing the size of a patient of varying diameter at different cross-sectional locations. RESPONSE 7: The developers agree that measuring patient size is important and provided a detailed response to the ACR and AAPM in their earlier comments. FIRST, our approach for using mid-scan diameter is highly correlated with patient weight. In separate, NIH-funded research on CT use in children up to age 21 (Kwan 2022), UCSF has shown that diameter in 4,239 children as measured on mid-scan axial images is highly predictive of patient weight, correlation = 0.904.SECOND, for this measure, patient size is measured using CT image pixel data, either on the mid-scan axial image or the coronal scout image when the mid-scan axial image was not available. This approach has been validated using data from UCSF Health, the UCSF Registry, as well as the data assembled for measure testing from 16 hospitals and 13 outpatient imaging centers. While there may be different ways to measure patient size, and different reasons for measuring patient size, the developers are adjusting for patient size primarily to ensure that entities that see larger patients are not penalized for doing so. using data from the UCSF Registry for abdomen CT we asse the relationship between radiation dose (in DLP) and patient diameter. Abdomen CT was selected as this is the category most influenced by patient size, meaning that patient mix could impact an entity's out-of-range rate. The raw correlation between patient diameter and unadjusted DLP is 0.50, and the marginal R-squared of the log-linear model used for adjustment is 0.15. After size-adjustment, the relationship is nearly removed: the raw correlation is far lower -(0.09), and the modeled marginal R-squared post-adjustment is 0. THIS DEMONSTRATES ADEQUACY OF THE APPROACH FOR PATIENT SIZE ADJUSTMENT TO REMOVE BIAS CAUSED BY CASE MIX. THIRD, the adequacy of size adjustment was shown empirically using data assembled from the testing sites. Out-of-range proportions for routine abdomen exams at 16 hospitals in our testing data based on unadjusted DLP, by decile in patient size are strongly associated by decile in size. Among patients in the highest size decile the

out-of-range proportions across the 16 hospitals ranged from 93-100%. ON THE OTHER HAND The out-of-range rates based on adjusted DLP are not higher among the larger patients. Among patients in the highest size decile, out-of-range rates ranged from 11-53%. THIS ALSO DEMONSTRATES ADEQUACY OF THE APPROACH FOR ADJUSTMENT OF PATIENT SIZE. COMMENT 8 Limited expertise… Alara Imaging Alara is a new company without a significant track record of having previously performed a project of such wide scope, scientifically or technically … Alara has limited IT development experience with no demonstrated history of interfacing with complex EHR or Radiology Clinical Data systems. ..The software interface is problematic because it is a vector to a 3rd party product, which can expose healthcare organizations to ransomware attacks by malicious actors seeking valuable patient medical information. RESPONSE 8: These assertions are incorrect. Measure stewardship is in collaboration with the University of California San Francisco (UCSF). UCSF was responsible for all measure development, scientific research, and measure validation work. UCSF also has a significant track record of successfully performing projects of this scope. About mid-way into the cooperative agreement, CMS asked UCSF to develop and report these measures as eCQMs. UCSF sought guidance from the measures' Technical Expert Panel and asked if members of the TEP would step forward to develop the software and steward the measure. When no group presented itself, it became clear that developing the eCQM and managing nationwide implementation and reporting of this measure was beyond the scope of the UCSF academic team and other TEP member organizations. Accordingly, Dr. Smith-Bindman worked with UCSF to create a company, Alara Imaging, that would help serve as measure steward. UCSF created Alara Imaging to develop the eCQM software and support measure stewardship. Alara is comprised of a team specifically assembled for the creation of this measure software, with deep radiology informatics and technical expertise. The Alara team has over 50 combined years of experience deploying software in hospital environments. The company is new, but the team is well versed in secure implementation. Alara's software is secure. The software is both HIPAA certified and SOC II certified with an independent third-party audit. The software protects against cyberattacks. Tools and information are provided to protect site data and support product security. Alara Imaging's software was used to perform the extensive testing described in the measure application, including correctly calculating measure results for 35,729 CT exams assembled from 7 hospital systems and 1 ambulatory imaging network. Software to calculate the measure will be made available to sites without charge. Measure specifications are in the public domain. If practices do not want to work with Alara, they may work with other vendors to report on the measure. Burden was found to be no more or less onerous than the effort required by participation in other eCQMs or national registries, such as the ACR Dose Index Registry (Feasibility, 3.06). References: KANAL KM et al. U.S. Diagnostic Reference Levels and Achievable Doses for 10 Adult CT Examinations. Radiology. 2017;284(1):120-133. MALKUS A, SZCZYKUTOWICZ TP. A method to extract image noise level from patient images in CT. Med Phys. 2017 Jun;44(6):2173-2184. SMITH-BINDMAN R, Yu S, Wang Y, et al. An Image Quality-informed Framework for CT Characterization. Radiology. 2021 Nov 9:210591. SMITH-BINDMAN R et al. Comparison of the Effectiveness of Single-Component and Multicomponent Interventions for Reducing Radiation Doses in Patients Undergoing Computed Tomography: A Randomized Clinical Trial. JAMA Intern Med. 2020 May 1;180(5):666-675. KWAN M et al. Smith-Bindman senior Author. Quantifying cancer risk from exposures to medical imaging in the Risk of Pediatric and Adolescent Cancer Associated with Medical Imaging (RIC) Study: Research Methods and Cohort Profile Marilyn Kwan et al. Cancer Causes Control 2022 May;33(5):711-726. doi: 10.1007/s10552-022-01556-z. Epub 2022 Feb 2.

NQF Response

Thank you for your comment. It has been shared with the Standing Committee and the measure developer.

NQF Committee Response

Proposed Response: The Patient Safety Standing Committee thanks American Association of Physicists in Medicine for their comment. The Standing Committee does take public comments into account when discussing and recommending measures for endorsement. The Standing Committee made the decision to endorse the measures after reviewing and considering the original comment and the measure developer's response. The Standing Committee stands by their decision to recommend to endorse the measures.

Stephanie Collingwood, UnityPoint Health; Submitted by Stephanie Collingwood

Comment ID#: 7966 (Submitted: 04/25/2022)

Council / Public: PRO

Level of Support: N/A

Comment

UnityPoint Health respectfully offers comments in support of measures 3633, 3662e and 3663e with additional considerations outlined below. UnityPoint Health is one of the nation's most integrated health care systems. Through more than 32,000 employees and our relationships with more than 480 physician clinics, 40 hospitals in urban and rural communities and 14 home health agencies throughout our 9 regions, UnityPoint Health provides care throughout Iowa, central Illinois, and southern Wisconsin. On an annual basis, UnityPoint Health hospitals, clinics and home health provide a full range of coordinated care to patients and families through more than 8.4 million patient visits. Multiple clinically relevant details come into play when determining the appropriate safe dose of radiation for a patient versus obtaining a clear image. Implementation of electronic health record tools requiring this level of documentation within a charting system would be required, along with tools to determine the point system applied. While UnityPoint Health fully understands the value of appropriate CT imaging, operational concerns exist regarding the capability of detailed tracking required to determine excessive CT use on a collective patient population. Additionally, reporting challenges exist today as multiple vendors are used within a health care system. UnityPoint Health supports the concept of this measure but would recommend developing exclusion criteria for overuse.

Developer Response

We thank UnityPoint Health for their comments. Given their large size, the large number of providers and clinics they work with, and the large number of patients they care for, we are grateful they appreciate the value of appropriate CT Imaging. We want to address their misunderstanding in how the measure works. FIRST, The measure is an electronic Clinical Quality Measure (eCQM) and relies on existing electronic data stored in the EHR, billing claims and

radiology information systems to calculate the measure. There is no charting nor new documentation required for measure calculation nor a requirement from sites to assign a point system to CT scans. All data elements used to calculate the measure come from existing structured variables listed in the feasibility scorecards and in Specifications, Table sp-2: CPT® and ICD-10-CM codes; dose length product stored in the DICOM data; and patient diameter and image noise calculated on imaging data. The measure would not have met the requirements of an eCQM had it relied on unstructured or newly created variables. The measure was tested across diverse EHR systems and diverse Radiology Information Systems, including those used by 7 hospital systems and 1 outpatient ambulatory practice group. Data were found to be widely available. SECOND, we strongly agree with UnityPoint Health that relevant clinical details (e.g. the clinical indication for scanning) are required to determine the appropriate radiation dose for each CT scan; e.g. the radiation dose and image quality required for a chest CT performed for lung cancer screening is not the same as required for the surveillance of known lung cancer. The approach of assigning CT examinations to the different CT categories (reflecting the clinical indications and required radiation dose and image quality) as specified in the measure was first developed using records from over 4.5 million CT exams in the UCSF International CT Dose Registry (Smith-Bindman, 2021). We then developed an approach for determining the clinical indication for imaging using an algorithm that combines procedure (CPT®) and diagnosis codes (ICD-10-CM) associated with the clinical visit when the test was ordered, information provided as part of the order, and information on the final bill. These are provided in Specifications, sp-11. This algorithm was developed using detailed review of over 10,000 patient records from UCSF Health. We validated the CT category assignment using the algorithm against “gold standard” chart review, as described in Validity sections 2b.02 and 2b.03. When the algorithm was deployed at our testing sites (including 16 hospitals and 13 outpatient imaging centers) the correct classification rate of the assignment of CT exams to CT category in field-testing was highly accurate across clinicians, clinician groups and hospitals. The goal in creating the CT categorization decision rules was to identify exams that are exceptions to the routine dose category (i.e. either high or low dose); most scans fall in the routine dose categories. IN SUMMARY, the calculation of the measure score does not require any new charting, does not require change in clinical practice, and does not require new documentation. Reference: Smith-Bindman R, Yu S, Wang Y, et al. An Image Quality-informed Framework for CT Characterization. Radiology. 2021 Nov 9:210591.

NQF Response

Thank you for your comment. It has been shared with the Standing Committee and the measure developer.

NQF Committee Response

N/A

NQF #3662e Excessive Radiation Dose or Inadequate Image Quality for Diagnostic Computed Tomography (CT) in Adults (Clinician Group Level) (Recommended)

J. Daniel Bourland, AAPM President, American Association of Physicists in Medicine; Submitted by Richard Martin

Comment ID#: 8010 (Submitted: 04/29/2022)

Council / Public: Public

Level of Support: N/A

Comment

The American Association of Physicists in Medicine (AAPM), is pleased to submit comments to the National Quality Forum (NQF) regarding its Patient Safety Standing Committee (PSSC) evaluation report of the following measures that the PSSC recommended for endorsement: NQF #: 3633e - Excessive Radiation Dose or Inadequate Image Quality for Diagnostic Computed Tomography (CT) in Adults (Clinician Level) NQF #: 3662e - Excessive Radiation Dose or Inadequate Image Quality for Diagnostic Computed Tomography (CT) in Adults (Clinician Group Level) NQF #: 3663e - Excessive Radiation Dose or Inadequate Image Quality for Diagnostic Computed Tomography (CT) in Adults (Facility Level) Background These electronic clinical quality measures (eCQM) are intended to monitor CT performance to discourage unnecessarily high radiation dose while maintaining adequate image quality. The proposed metrics require CT Category (i.e., the CT exam type), the size adjusted radiation dose [the patient's dose length product (DLP) adjusted by patient size], and the global noise (associated with the variance of the voxel values in CT images). The two reported measures are the percentage of eligible CT cases in a particular category deemed to be "out-of-range" compared to defined thresholds with respect to the size-adjusted radiation dose or the global noise in a set time period. The measures are intended to advance quality assurance. In January 2022, prior to the Patient Safety Standing Committee's meeting to evaluate these proposed measures, AAPM provided comments on the measure application to the committee. AAPM attended the committee meeting and now responds to the committee's evaluation report. The AAPM and our leadership in medical physics AAPM, as the primary scientific and professional organization of physics in radiology and radiation oncology in the United States, is the foremost organization with expertise to speak to the topic under consideration. With 9717 members in 94 countries, AAPM supports the Medical Physics community with a focus on advancing patient care through education, improving safety and efficacy of medical imaging procedures through research, education and the maintenance of professional standards. Medical physicists contribute to the effectiveness of medical imaging by ensuring the safe and effective use of radiant energy (e.g., optical, ionizing, ultrasonic, or radiofrequency) to obtain detailed information about the form and function of the human body. Medical physicists continue to play a leading role in the development of novel imaging technologies, as well as in guiding the optimization of existing imaging modalities. General Comments AAPM commends NQF's efforts in advancing and evaluating quality assurance measures. The last 15 years of CT technology development has included new reconstruction algorithms and tube current modulation techniques resulting in substantial reductions in dose. AAPM supports efforts to enhance consistency of CT practice as evidenced by AAPM's proactive engagement in efforts to ensure diagnostic quality CT imaging, optimizing CT dose, and achieving consistency across facilities, considering differing technologies and practices. AAPM, together with other non-profit entities, including the American College of Radiology (ACR), and Image Wisely and

Image Gently Alliances has spent decades working towards this goal and continues to do so through many initiatives. AAPM does not support the endorsement of NQF #3633e, #3662e, and #3663e. AAPM cautions that the measures recommended for endorsement by the PSSC have significant limitations that impact their scientific and practical value and overall likelihood of clinical acceptance. These limitations include improper representation of image quality, improper estimation of radiation risk, and substantial oversimplified representation of implementation in practice, including not addressing the challenges of implementation. We will address these concerns in the following paragraphs. Specific Comments PSSC failed to adequately review and consider expert opinion The PSSC failed to adequately review or consider AAPM's expert comments, as required. AAPM review of the proposed measures consisted of a detailed analysis by four prominent senior physicists from four separate institutions. AAPM's comments, however, were not considered as evidenced by the deliberations of the committee at its meeting and in the present report. AAPM's leadership in medical physics – national and international expertise and recognition AAPM's expertise in medical physics is widely recognized and valued by the Nuclear Regulatory Commission (NRC), Food and Drug Administration (FDA), National Institute of Biomedical Imaging and Bioengineering (NIBIB), National Council on Radiation Protection and Measurements (NCRP), other federal agencies and state radiation safety agencies. These agencies routinely engage AAPM on clinical practice, emerging technology and radiation safety issues and seek out AAPM members to serve on their advisory committees addressing the most cutting-edge issues in the radiation medicine field. Thus, AAPM's expert voice on this topic is of high scientific and practical relevance to provide consensus guidance on this important topic. Unscientific characterization of CT scan risk The measure developers include specific numbers estimating the number of cancers and deaths due to these cancers from the dose imparted from the CT scans. The authors describe these risks and the resulting estimates as based on models only. The applied linear non-threshold model is currently HIGHLY disputed at diagnostic CT radiation dose levels. The resultant estimates of risk are known to involve large uncertainties. Moreover, the science of radiation risk estimation from CT examinations is based on calculation of dose to individual organs, age, and sex. The measures of risk proposed here, however, mention none of these factors or offer a strategy to incorporate it. The proposed measures are primarily based on radiation output of the CT system, not the risk to the patient. The benefit, if any, of minimizing patient dose cannot be scientifically statistically determined. AAPM is concerned that the stated risk of patient radiation dose and financial savings are hypothetical, exaggerated, and may contribute to fear of diagnostic medical exams that may in turn lead some patients to refuse safe and appropriate medical imaging, to the detriment of the patient. Diagnostic imaging doses are typically much lower than 100 mSv, and the anticipated benefits to the patient of medically appropriate imaging are highly likely to outweigh any small potential risks. Measures lack usability The usability of data resulting from these measures is not clear. In their pilot study, 30% of the CT cases for individual clinicians being out-of-range was the median value with half of the clinicians having between 16% and 43% of their cases out-of-range, as shown in Figure 1b-2 of the application. The measures do not provide the clinician with an analysis of or methodology for determining what improvements should be made to address a poor showing with these parameters. It may not be clear to practitioners what a poor score means or how to address it. Complexity of CT categorization The measures rely on the categorization of CT data into cohesive groups. There is, however, significant variability in the CT protocol lexicon across institutions that results in making assignment of a given protocol to one of these categories very challenging. The proposal does not address the magnitude of this challenge

or present the means to overcome it, given that current standards lack uniform characterization of protocols. Inadequate measure of noise The proposed noise measure is not an adequate or sufficient parameter of overall image quality. Visually different texture patterns can have similar noise values, and each may be of more, or less, diagnostic value for the radiologist. As mentioned in the proposal, noise can be influenced by many different parameters, such as slice thickness, kV, and mAs. The effect on noise of these parameters is mostly predictable (particularly in a well-defined "subject", such as a phantom). Noise is commonly determined in a standardized phantom. Noise measured in clinical images is another matter. There has been limited scientific work in that area and none is cited as having been performed by the authors. There is no information provided in the proposal about how the proposed global noise measure is calculated. In particular, the approach does not take into consideration the CT reconstruction settings that can have a dramatic impact on the appearance of the images, including noise, contrast (or CNR), and sharpness. Further, a "global noise" ignores the diversity within the CT series, especially within the (usually) limited locations that depict the abnormality of interest. Inadequate assessment of image quality Image noise alone is an insufficient descriptor of image quality. Noise in an image may also be justifiably varied to meet certain clinical needs (such as high resolution). Many other factors must be considered when attempting to define image quality. Spatial resolution, which includes visualizing small objects and image boundaries, and contrast resolution, of which noise is one component, are also critical aspects of image quality. Widely different noise values may be acceptable under different circumstances for similar protocols. Spatial resolution and contrast are as important as image noise. It is not all clear that improvements in global noise will in turn lead to improved clinical performance. Flawed assumption regarding clinical CT practice There is substantial variation in the radiation doses used in CT exams because the radiation delivered is protocol-specific. The implication in the proposed measures is that radiologists vary these parameters indiscriminately. In most cases, however, these protocols are established by the institutions based on available equipment, patient population, expertise, scientific evidence, and the nature of cases presented at that institution. With the proposed measures, an optimum study is one that delivers the least radiation dose with an acceptable global noise level, but no evidence is provided that clinicians with high values for the proposed measures perform better or even adequately, only that they perform with less radiation dose. AAPM recommends using an alternative approach that would optimize the acquisition for the task at hand to deliver the least radiation dose necessary while still providing the diagnostic image quality necessary for the task. Dose reduction in and of itself is not enough to improve CT practice. There should also be no loss of clinical performance which is not guaranteed by these measurements. Global noise is not an adequate sufficient metric for image quality. Inadequate accuracy in patient size estimation Assessing a patient size can be challenging because of significant variability in the differences in the habitus of different patients, coupled with the existential challenge that there is no single metric capturing the size of a patient of varying diameter at different cross-sectional locations. The proposed measures rely on the calibration of the company's black-box size estimation to prior work of Cheng 2013 and Christianson 2012, both of which have been updated to newer versions to correct erroneous measures. The error in size measurements needs transparency and validated results. Limited expertise and track record of Alara Imaging Alara Imaging is a new (2020) company without a significant track record of having previously performed a project of such wide scope, scientifically or technically. While the measure developers have published on the topic of radiation dose, they have limited expertise or history with clinical CT, CT image quality, or CT technology. The

company has no scientific track record on CT technology, size estimation, or image quality assessment to be considered steward of measures on which it lacks expertise, publication, or scientific history. Moreover, Alara Imaging has limited experience in IT development, with no demonstrated history of interfacing with complex EHR or Radiology Clinical Data systems. The software interface is problematic because it is a vector to a 3rd party product, which can expose healthcare organizations to ransomware attacks by malicious actors seeking valuable patient medical information. In summary, AAPM does not support the endorsement of NQF #3633e, #3662e, and #3663e. AAPM urges NQF to: • Address the concerns identified by AAPM experts; and • Reconsider its recommendation endorsing these measures as proposed. AAPM recognizes that this topic is complex, including scientific, technical and clinical components, and we would welcome the opportunity for greater in-depth discussion on meaningful measures of quality imaging practice. Thank you again for the opportunity to comment on the PSSC evaluation report. If you have any questions or require additional information, please contact Richard J. Martin, JD, Government Relations Project Manager, at 571-298-1227 or Richard@aapm.org.

Developer Response

UCSF thanks the American Association of Physicists in Medicine (AAPM) for their additional comments. UCSF would like to respond and address several inaccuracies and misunderstandings in how the measure is calculated and its intent. COMMENT 1: [Medical Physicists and the American Association of Physicists in Medicine have extensive content expertise that should be considered]…and "while the measure developers have published on the topic of radiation dose, they have limited expertise or history with clinical CT, CT image quality, or CT technology." RESPONSE 1: The measure developers agree that medical physicists have relevant and important expertise, and they have involved medical physicists in all aspects of our work including both the measure development itself and all of the work in the preceding decade that laid the foundation for UCSF's development of this measure. FIRST: J. Anthony Seibert, PhD was included as a member of our Technical Expert Panel (TEP) to ensure consideration of the perspectives of medical physicists at every step of measure development. Dr. Seibert recently retired as Professor of Diagnostic Imaging Physics and Associate Chair of Radiology Informatics at UC Davis Health and is a past president of the AAPM (2011). In addition to serving on our TEP, Dr. Seibert led UC Davis as a measure testing site and wrote a letter of support for the measure confirming it was "highly feasible" to calculate the measure and noting his belief "that this quality measure can significantly reduce the use of excessive high radiation dose as well as inadequate, sub-optimal low dose used for clinical CT studies." SECOND, the developers also worked closely with another medical physicist, Tim Szczkutowicz, PhD on measure development. Dr. Szczkutowicz guided the work of automating the calculation of image noise, expanding on his earlier published work in this area (Malkus 2017). Dr. Szczkutowicz is an Associate Professor in the Department of Radiology at the University of Wisconsin Madison School of Medicine and Public Health with affiliations in the Department of Medical Physicists and Biomedical Imaging. Lastly, this measure development effort has been led by Rebecca Smith-Bindman, MD, a radiologist and epidemiologist whose primary area of research for the last 15 years has been in quantifying the radiation doses used for CT scanning and identifying ways to safely reduce excessive doses. Over the last 10 years, her research team created a CT radiation dose registry of more than 8 million exams from over 160 facilities, which has allowed the team to quantify the variation in dose, to understand the cause of the variation, and to

develop and study interventions to help facilities appropriately lower doses without loss of image quality. The development of this quality measure was a natural extension of this work, and the registry has allowed for the testing of the adult measures. As part of this past work, Dr. Smith-Bindman led a randomized controlled trial of two interventions to optimize CT radiation doses across 100 hospitals and imaging facilities and found that providing feedback (similar to that proposed for these quality measures) along with education and opportunities for sharing best practices resulted in meaningful dose reductions (up to 40%) without any loss in image quality (Smith-Bindman 2020). In total, 13 medical physicists served as site-Principal Investigators for this NIH funded-trial (R01CA181191). Thus, medical physicists have contributed substantially to the body of work that led to the measure as well as measure development. In large part based on Dr. Smith-Bindman's 15-year track record in this area, involving medical physicists, CMS awarded UCSF a cooperative agreement to develop these CT quality measures under the Medicare Access and CHIP Reauthorization Act of 2015 ("MACRA"). COMMENT 2: [There is] unscientific characterization of CT scan risk … the AAPM is concerned that the stated risk of patient radiation dose .. may contribute to fear of diagnostic exams that may in turn lead some patients to refuse safe and appropriate imaging. RESPONSE 2: The measure is not focused on radiation risk and does not calculate nor report radiation risk. A review of the published epidemiological evidence summarizing radiation risk is provided in the application as background and context for the measure (see sections 1a.01-1b.01). This includes several systematic reviews, cohort studies, and an extensive and comprehensive review from the National Academies that conclude that exposure to CT (or radiation doses in the same range as CT) increases a person's risk of developing cancer. Most of these studies do not rely upon the linear no threshold model that the AAPM criticizes in its comment. Indeed, the systematic reviews are based on an observed elevated risk of cancer among patients exposed to medical imaging. The estimates used in the application are based on this extensive literature review. HOWEVER, radiation risk is not part of the measure at all: it is not calculated nor is it reported. INSTEAD, the measure evaluates dose length product (DLP), and specifically whether size-adjusted DLP exceeds thresholds specific to CT category. DLP is the radiation dose measure most directly under the control of providers, determined by specific parameters that were chosen for the scan. Further, DLP is universally reported by CT scanner manufacturers (unlike other metrics of radiation dose or risk). THE TEP, including the American College of Radiology (ACR), several radiologists, and the medical physicist serving on the committee, unanimously recommended and supported the radiation dose measure used (DLP) and unanimously agreed that it is a relevant metric of quality for CT imaging, as noted in Validity Results, 2b.03. There is also considerable precedent for using DLP to evaluate radiation dose in CT. The American College of Radiology has used DLP to set benchmarks [Kanal 2017] and to measure CT radiation dose in their own NQF-endorsed quality measure #3621. There is no reason to believe that endorsing this measure, which seeks to standardize practice and reduce extreme radiation dose outliers based on DLP, would result in patients refusing appropriate imaging. COMMENT 3: The usability of data resulting from these measures is not clear. The measures do not provide the clinician with an analysis of or methodology for determining what improvements should be made to address a poor showing. It may not be clear to practitioners what a poor score means or how to address it. RESPONSE 3: Entities that report the measure using the measure steward's software will be provided information to both identify causes of performance gaps and make targeted changes to improve quality. There are only two conditions that would push a CT out of compliance - high radiation dose, and poor image quality. The reason for failure (high dose or low quality) will be

available to sites on a scan-by-scan basis for those that report the measure using the measure steward's software Those scans where the radiation dose is too high, the dose should be lowered through usual means (technologist education, protocol changes). For those where image quality is too low, more radiation dose should be used through similar means. COMMENTS IN SUPPORT of the measure from many of the testing sites describe how useful the information provided was to allow them to understand and improve their practice. (Available here, beginning page 113: https://www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=96982) As described in our submission and noted above, UCSF found in a randomized controlled trial in 100 hospitals and outpatient radiology practices that providing detailed audit and feedback on radiation doses, similar to what will be provided as part of the feedback on this measure, resulted in significant reductions in radiation dose with no impact on satisfaction with image quality, described in Usability, 4b.01. (Smith-Bindman, 2020) The measure steward does not have control over how 3rd party vendors will report information back to reporting entities. COMMENT 4. The measures rely on categorization of CT data into cohesive groups… There is, however, significant variability in the CT protocol lexicon across institutions that results in making assignment of a given protocol to one of these categories challenging. [There is] substantial oversimplified representation of implementation in practice RESPONSE 4: For the reason the AAPM highlights and for another important reason describe below, the CT category assigned by the measure (reflecting the indication and appropriate radiation dose level for the scan) does not rely on the protocol name at all. As described in Specifications, sp-11, clinical indication for imaging is determined using an algorithm that combines procedure (CPT®) and diagnosis (ICD-10-CM) codes associated with the clinical visit when the test was ordered, information provided as part of the order, and information on the final bill. The codes are available in the radiology electronic systems and/or the EHR or billing systems. The goal in creating the CT categorization decision rules was to identify exams that are exceptions to the routine dose category (i.e., either high or low dose), (Smith-Bindman, 2021.) Details of the approach for developing and validating the assignment of CT exams to categories in an automated fashion are provided in the measure submission (see Validity sections 2b.02 and 2b.03) and in detailed answers provided to initial comments made by the ACR and AAPM. This approach was first developed using records from over 4.5 million CT exams in the UCSF International CT Dose Registry (Smith-Bindman, 2021), and then turned into an algorithm that combines procedure (CPT®) and diagnosis codes (ICD-10-CM) associated with the clinical visit when the test was ordered, information provided as part of the order, and information on the final bill, provided in Specifications, sp-11. This algorithm was developed using detailed review of over 10,000 patient records from UCSF Health, and validated against "gold standard" chart review, as described in Validity sections 2b.02 and 2b.03. When the algorithm was deployed at our testing sites (including 16 hospitals and 13 outpatient imaging centers), the correct classification rate of the assignment of CT exams to CT category in field-testing was excellent (over 90% for all reporting levels: clinician, clinician group, and facility). Knowing that the algorithm was developed using data from a single health system, the developers performed detailed investigation of the categorization results at testing sites – comparing the assigned CT category against full radiology reports – for the purpose of improving the algorithm. One of the strengths of these measures is that they do not determine the CT category using the protocol name, as this would mask an important quality improvement opportunity (namely, the selection of which protocol to use to scan the patient). Two key process of care components determine radiation doses: (A) the choice of imaging protocol, for example, whether a patient with a suspected pulmonary embolism is imaged with a single- or

double-phase CT exam (a decision usually made by the performing radiologist); and (B) the technical settings used for that type of CT exam, which are usually at the discretion of the technologist or medical physicist who oversee and operate the machines. As both of these components contribute to radiation dose, a comprehensive quality measure must encompass both of these decision-making processes. By determining the CT category independent of the protocol used, the measures is able to evaluate both components of quality. COMMENT 5 The noise measure is not an adequate or sufficient parameter of overall image quality. RESPONSE 5: Several comments focus on image quality and the concern that the measure does not offer a comprehensive assessment of image quality. The measure is not intended as a robust measurement of image quality. The primary focus of our measure is to assess radiation dose adjusted for body size. The image quality component was included to protect against the unlikely possibility of substantial degradation of image quality as an unintended consequence of dose reduction. Our measure of image quality uses thresholds developed based on radiologists' satisfaction with images, reflecting what in practice is regarded as adequate for diagnosis. Others might have an interest in more nuanced assessment of image quality for other purposes, but that was not our intent. If the measure is adopted and used, the Steward will closely monitor image noise and measure failure due to low image quality. The Steward will be sensitive to any signal that there is a problem and will revise the measure if changes are needed. COMMENT 6: With these measures, an optimum study is one that delivers the least radiation dose with an acceptable global noise level.. but no evidence is provided that clinicians with high values for the measures perform better or even adequately, only that they perform with less radiation dose. AAPM recommends using an alternative approach that would optimize the acquisition for the task at hand to deliver the least radiation dose necessary while still providing the diagnostic image quality necessary for the task. RESPONSE 6: This measure provides a standardized method for monitoring the performance of diagnostic CT to discourage unnecessarily high radiation doses, a risk factor for cancer, while preserving image quality. It is expressed as a percentage of eligible CT exams that are out-of-range based on having either excessive radiation dose or inadequate image quality, relative to evidence-based thresholds based on the clinical indication for the exam. THE HIGHER THE SCORE, the higher the proportion of out-of-range exams and the worse the performance. The measure is NOT INTENDED to improve diagnostic accuracy. The purpose of the measure is to establish a radiation dose ceiling to avoid excessive radiation exposure, and an image quality floor to safeguard against unintended deterioration of image quality. ADDITIONALLY, the entire framework for the measure is to ensure the radiation dose and image quality are acceptable for the specific clinical indication, aligned with what the AAPM recommends. The CT categories were created based on radiation dose and image quality requirements specific to the clinical indications for imaging (Smith-Bindman 2021). Using radiologists' satisfaction with image quality, an image quality floor for each category was established, below which an exam is considered to have inadequate quality, and a radiation dose ceiling, beyond which doses are considered unnecessarily high. The purpose is to allow detailed assessment of each CT exam to ensure the dose is optimal based on the clinical indication for imaging. In our testing data, far more CT exams exceeded the radiation dose ceiling (average = 30%) than failed to meet the image quality requirement (average << 1%) (see section 1b.02). The measure encourages entities to reduce the proportion of exams that may "be overdosed for their exact need and condition" while preserving the minimum image quality. COMMENT 7: [There is] Inadequate accuracy in patient size estimation… Assessing a patient size can be challenging because of significant variability in differences in the habitus of

different patients, coupled with the existential challenge that there is no single metric capturing the size of a patient of varying diameter at different cross-sectional locations. RESPONSE 7: The developers agree that measuring patient size is important and provided a detailed response to the ACR and AAPM in their earlier comments. FIRST, our approach for using mid-scan diameter is highly correlated with patient weight. In separate, NIH-funded research on CT use in children up to age 21 (Kwan 2022), UCSF has shown that diameter in 4,239 children as measured on mid-scan axial images is highly predictive of patient weight, correlation = 0.904.SECOND, for this measure, patient size is measured using CT image pixel data, either on the mid-scan axial image or the coronal scout image when the mid-scan axial image was not available. This approach has been validated using data from UCSF Health, the UCSF Registry, as well as the data assembled for measure testing from 16 hospitals and 13 outpatient imaging centers. While there may be different ways to measure patient size, and different reasons for measuring patient size, the developers are adjusting for patient size primarily to ensure that entities that see larger patients are not penalized for doing so. using data from the UCSF Registry for abdomen CT we asse the relationship between radiation dose (in DLP) and patient diameter. Abdomen CT was selected as this is the category most influenced by patient size, meaning that patient mix could impact an entity's out-of-range rate. The raw correlation between patient diameter and unadjusted DLP is 0.50, and the marginal R-squared of the log-linear model used for adjustment is 0.15. After size-adjustment, the relationship is nearly removed: the raw correlation is far lower -(0.09), and the modeled marginal R-squared post-adjustment is 0. THIS DEMONSTRATES ADEQUACY OF THE APPROACH FOR PATIENT SIZE ADJUSTMENT TO REMOVE BIAS CAUSED BY CASE MIX. THIRD, the adequacy of size adjustment was shown empirically using data assembled from the testing sites. Out-of-range proportions for routine abdomen exams at 16 hospitals in our testing data based on unadjusted DLP, by decile in patient size are strongly associated by decile in size. Among patients in the highest size decile the out-of-range proportions across the 16 hospitals ranged from 93-100%. ON THE OTHER HAND The out-of-range rates based on adjusted DLP are not higher among the larger patients. Among patients in the highest size decile, out-of-range rates ranged from 11-53%. THIS ALSO DEMONSTRATES ADEQUACY OF THE APPROACH FOR ADJUSTMENT OF PATIENT SIZE. COMMENT 8 Limited expertise… Alara Imaging Alara is a new company without a significant track record of having previously performed a project of such wide scope, scientifically or technically … Alara has limited IT development experience with no demonstrated history of interfacing with complex EHR or Radiology Clinical Data systems. ..The software interface is problematic because it is a vector to a 3rd party product, which can expose healthcare organizations to ransomware attacks by malicious actors seeking valuable patient medical information. RESPONSE 8: These assertions are incorrect. Measure stewardship is in collaboration with the University of California San Francisco (UCSF). UCSF was responsible for all measure development, scientific research, and measure validation work. UCSF also has a significant track record of successfully performing projects of this scope. About mid-way into the cooperative agreement, CMS asked UCSF to develop and report these measures as eCQMs. UCSF sought guidance from the measures' Technical Expert Panel and asked if members of the TEP would step forward to develop the software and steward the measure. When no group presented itself, it became clear that developing the eCQM and managing nationwide implementation and reporting of this measure was beyond the scope of the UCSF academic team and other TEP member organizations. Accordingly, Dr. Smith-Bindman worked with UCSF to create a company, Alara Imaging, that would help serve as measure steward. UCSF created Alara Imaging to develop the eCQM software and support measure stewardship. Alara is comprised of a team

specifically assembled for the creation of this measure software, with deep radiology informatics and technical expertise. The Alara team has over 50 combined years of experience deploying software in hospital environments. The company is new, but the team is well versed in secure implementation. Alara's software is secure. The software is both HIPAA certified and SOC II certified with an independent third-party audit. The software protects against cyberattacks. Tools and information are provided to protect site data and support product security. Alara Imaging's software was used to perform the extensive testing described in the measure application, including correctly calculating measure results for 35,729 CT exams assembled from 7 hospital systems and 1 ambulatory imaging network. Software to calculate the measure will be made available to sites without charge. Measure specifications are in the public domain. If practices do not want to work with Alara, they may work with other vendors to report on the measure. Burden was found to be no more or less onerous than the effort required by participation in other eCQMs or national registries, such as the ACR Dose Index Registry (Feasibility, 3.06). References: KANAL KM et al. U.S. Diagnostic Reference Levels and Achievable Doses for 10 Adult CT Examinations. Radiology. 2017;284(1):120-133. MALKUS A, SZCZYKUTOWICZ TP. A method to extract image noise level from patient images in CT. Med Phys. 2017 Jun;44(6):2173-2184. SMITH-BINDMAN R, Yu S, Wang Y, et al. An Image Quality-informed Framework for CT Characterization. Radiology. 2021 Nov 9:210591. SMITH-BINDMAN R et al. Comparison of the Effectiveness of Single-Component and Multicomponent Interventions for Reducing Radiation Doses in Patients Undergoing Computed Tomography: A Randomized Clinical Trial. JAMA Intern Med. 2020 May 1;180(5):666-675. KWAN M et al. Smith-Bindman senior Author. Quantifying cancer risk from exposures to medical imaging in the Risk of Pediatric and Adolescent Cancer Associated with Medical Imaging (RIC) Study: Research Methods and Cohort Profile Marilyn Kwan et al. Cancer Causes Control 2022 May;33(5):711-726. doi: 10.1007/s10552-022-01556-z. Epub 2022 Feb 2.

NQF Response

Thank you for your comment. It has been shared with the Standing Committee and the measure developer.

NQF Committee Response

Proposed Response: The Patient Safety Standing Committee thanks American Association of Physicists in Medicine for their comment. The Standing Committee does take public comments into account when discussing and recommending measures for endorsement. The Standing Committee made the decision to endorse the measures after reviewing and considering the original comment and the measure developer's response. The Standing Committee stands by their decision to recommend to endorse the measures.

Stephanie Collingwood, UnityPoint Health; Submitted by Stephanie Collingwood

Comment ID#: 7965 (Submitted: 04/25/2022)

Council / Public: PRO

Level of Support: N/A

Comment

UnityPoint Health respectfully offers comments in support of measures 3633, 3662e and 3663e with additional considerations outlined below. UnityPoint Health is one of the nation's most integrated health care systems. Through more than 32,000 employees and our relationships with more than 480 physician clinics, 40 hospitals in urban and rural communities and 14 home health agencies throughout our 9 regions, UnityPoint Health provides care throughout Iowa, central Illinois, and southern Wisconsin. On an annual basis, UnityPoint Health hospitals, clinics and home health provide a full range of coordinated care to patients and families through more than 8.4 million patient visits. Multiple clinically relevant details come into play when determining the appropriate safe dose of radiation for a patient versus obtaining a clear image. Implementation of electronic health record tools requiring this level of documentation within a charting system would be required, along with tools to determine the point system applied. While UnityPoint Health fully understands the value of appropriate CT imaging, operational concerns exist regarding the capability of detailed tracking required to determine excessive CT use on a collective patient population. Additionally, reporting challenges exist today as multiple vendors are used within a health care system. UnityPoint Health supports the concept of this measure but would recommend developing exclusion criteria for overuse.

Developer Response

We thank UnityPoint Health for their comments. Given their large size, the large number of providers and clinics they work with, and the large number of patients they care for, we are grateful they appreciate the value of appropriate CT Imaging. We want to address their misunderstanding in how the measure works. FIRST, The measure is an electronic Clinical Quality Measure (eCQM) and relies on existing electronic data stored in the EHR, billing claims and radiology information systems to calculate the measure. There is no charting nor new documentation required for measure calculation nor a requirement from sites to assign a point system to CT scans. All data elements used to calculate the measure come from existing structured variables listed in the feasibility scorecards and in Specifications, Table sp-2: CPT® and ICD-10-CM codes; dose length product stored in the DICOM data; and patient diameter and image noise calculated on imaging data. The measure would not have met the requirements of an eCQM had it relied on unstructured or newly created variables. The measure was tested across diverse EHR systems and diverse Radiology Information Systems, including those used by 7 hospital systems and 1 outpatient ambulatory practice group. Data were found to be widely available. SECOND, we strongly agree with UnityPoint Health that relevant clinical details (e.g. the clinical indication for scanning) are required to determine the appropriate radiation dose for each CT scan; e.g. the radiation dose and image quality required for a chest CT performed for lung cancer screening is not the same as required for the surveillance of known lung cancer. The approach of assigning CT examinations to the different CT categories (reflecting the clinical indications and required radiation dose and image quality) as specified in the measure was first developed using records from over 4.5 million CT exams in the UCSF International CT Dose Registry (Smith-Bindman, 2021). We then developed an approach for determining the clinical indication for imaging using an algorithm that combines procedure (CPT®) and diagnosis codes (ICD-10-CM) associated with the clinical visit when the test was ordered, information provided as part of the order, and information on the final bill. These are provided in Specifications, sp-11. This algorithm was developed using

detailed review of over 10,000 patient records from UCSF Health. We validated the CT category assignment using the algorithm against "gold standard" chart review, as described in Validity sections 2b.02 and 2b.03. When the algorithm was deployed at our testing sites (including 16 hospitals and 13 outpatient imaging centers) the correct classification rate of the assignment of CT exams to CT category in field-testing was highly accurate across clinicians, clinician groups and hospitals. The goal in creating the CT categorization decision rules was to identify exams that are exceptions to the routine dose category (i.e. either high or low dose); most scans fall in the routine dose categories. IN SUMMARY, the calculation of the measure score does not require any new charting, does not require change in clinical practice, and does not require new documentation. Reference: Smith-Bindman R, Yu S, Wang Y, et al. An Image Quality-informed Framework for CT Characterization. Radiology. 2021 Nov 9:210591.

NQF Response

Thank you for your comment. It has been shared with the Standing Committee and the measure developer.

NQF Committee Response

N/A

NQF #3663e Excessive Radiation Dose or Inadequate Image Quality for Diagnostic Computed Tomography (CT) in Adults (Facility Level) (Recommended)

J. Daniel Bourland, AAPM President, American Association of Physicists in Medicine; Submitted by Richard Martin

Comment ID#: 8011 (Submitted: 04/29/2022)

Council / Public: Public

Level of Support: N/A

Comment

The American Association of Physicists in Medicine (AAPM), is pleased to submit comments to the National Quality Forum (NQF) regarding its Patient Safety Standing Committee (PSSC) evaluation report of the following measures that the PSSC recommended for endorsement: NQF #: 3633e - Excessive Radiation Dose or Inadequate Image Quality for Diagnostic Computed Tomography (CT) in Adults (Clinician Level) NQF #: 3662e - Excessive Radiation Dose or Inadequate Image Quality for Diagnostic Computed Tomography (CT) in Adults (Clinician Group Level) NQF #: 3663e - Excessive Radiation Dose or Inadequate Image Quality for Diagnostic Computed Tomography (CT) in Adults (Facility Level) Background These electronic clinical quality measures (eCQM) are intended to monitor CT performance to discourage unnecessarily high radiation dose while maintaining adequate image quality. The proposed metrics require CT Category (i.e., the CT exam type), the size adjusted radiation dose [the patient's dose length product (DLP) adjusted by patient size], and the global noise (associated with the variance of the voxel values in CT images). The two reported measures are the percentage of eligible CT cases in a particular category deemed to be "out-of-range" compared to defined thresholds with respect to the size-adjusted radiation dose or the

global noise in a set time period. The measures are intended to advance quality assurance. In January 2022, prior to the Patient Safety Standing Committee's meeting to evaluate these proposed measures, AAPM provided comments on the measure application to the committee. AAPM attended the committee meeting and now responds to the committee's evaluation report. The AAPM and our leadership in medical physics AAPM, as the primary scientific and professional organization of physics in radiology and radiation oncology in the United States, is the foremost organization with expertise to speak to the topic under consideration. With 9717 members in 94 countries, AAPM supports the Medical Physics community with a focus on advancing patient care through education, improving safety and efficacy of medical imaging procedures through research, education and the maintenance of professional standards. Medical physicists contribute to the effectiveness of medical imaging by ensuring the safe and effective use of radiant energy (e.g., optical, ionizing, ultrasonic, or radiofrequency) to obtain detailed information about the form and function of the human body. Medical physicists continue to play a leading role in the development of novel imaging technologies, as well as in guiding the optimization of existing imaging modalities. General Comments AAPM commends NQF's efforts in advancing and evaluating quality assurance measures. The last 15 years of CT technology development has included new reconstruction algorithms and tube current modulation techniques resulting in substantial reductions in dose. AAPM supports efforts to enhance consistency of CT practice as evidenced by AAPM's proactive engagement in efforts to ensure diagnostic quality CT imaging, optimizing CT dose, and achieving consistency across facilities, considering differing technologies and practices. AAPM, together with other non-profit entities, including the American College of Radiology (ACR), and Image Wisely and Image Gently Alliances has spent decades working towards this goal and continues to do so through many initiatives. AAPM does not support the endorsement of NQF #3633e, #3662e, and #3663e. AAPM cautions that the measures recommended for endorsement by the PSSC have significant limitations that impact their scientific and practical value and overall likelihood of clinical acceptance. These limitations include improper representation of image quality, improper estimation of radiation risk, and substantial oversimplified representation of implementation in practice, including not addressing the challenges of implementation. We will address these concerns in the following paragraphs. Specific Comments PSSC failed to adequately review and consider expert opinion The PSSC failed to adequately review or consider AAPM's expert comments, as required. AAPM review of the proposed measures consisted of a detailed analysis by four prominent senior physicists from four separate institutions. AAPM's comments, however, were not considered as evidenced by the deliberations of the committee at its meeting and in the present report. AAPM's leadership in medical physics – national and international expertise and recognition AAPM's expertise in medical physics is widely recognized and valued by the Nuclear Regulatory Commission (NRC), Food and Drug Administration (FDA), National Institute of Biomedical Imaging and Bioengineering (NIBIB), National Council on Radiation Protection and Measurements (NCRP), other federal agencies and state radiation safety agencies. These agencies routinely engage AAPM on clinical practice, emerging technology and radiation safety issues and seek out AAPM members to serve on their advisory committees addressing the most cutting-edge issues in the radiation medicine field. Thus, AAPM's expert voice on this topic is of high scientific and practical relevance to provide consensus guidance on this important topic. Unscientific characterization of CT scan risk The measure developers include specific numbers estimating the number of cancers and deaths due to these cancers from the dose imparted from the CT scans. The authors describe these risks and the resulting estimates as based on models only. The applied

linear non-threshold model is currently HIGHLY disputed at diagnostic CT radiation dose levels. The resultant estimates of risk are known to involve large uncertainties. Moreover, the science of radiation risk estimation from CT examinations is based on calculation of dose to individual organs, age, and sex. The measures of risk proposed here, however, mention none of these factors or offer a strategy to incorporate it. The proposed measures are primarily based on radiation output of the CT system, not the risk to the patient. The benefit, if any, of minimizing patient dose cannot be scientifically statistically determined. AAPM is concerned that the stated risk of patient radiation dose and financial savings are hypothetical, exaggerated, and may contribute to fear of diagnostic medical exams that may in turn lead some patients to refuse safe and appropriate medical imaging, to the detriment of the patient. Diagnostic imaging doses are typically much lower than 100 mSv, and the anticipated benefits to the patient of medically appropriate imaging are highly likely to outweigh any small potential risks. Measures lack usability The usability of data resulting from these measures is not clear. In their pilot study, 30% of the CT cases for individual clinicians being out-of-range was the median value with half of the clinicians having between 16% and 43% of their cases out-of-range, as shown in Figure 1b-2 of the application. The measures do not provide the clinician with an analysis of or methodology for determining what improvements should be made to address a poor showing with these parameters. It may not be clear to practitioners what a poor score means or how to address it. Complexity of CT categorization The measures rely on the categorization of CT data into cohesive groups. There is, however, significant variability in the CT protocol lexicon across institutions that results in making assignment of a given protocol to one of these categories very challenging. The proposal does not address the magnitude of this challenge or present the means to overcome it, given that current standards lack uniform characterization of protocols. Inadequate measure of noise The proposed noise measure is not an adequate or sufficient parameter of overall image quality. Visually different texture patterns can have similar noise values, and each may be of more, or less, diagnostic value for the radiologist. As mentioned in the proposal, noise can be influenced by many different parameters, such as slice thickness, kV, and mAs. The effect on noise of these parameters is mostly predictable (particularly in a well-defined “subject”, such as a phantom). Noise is commonly determined in a standardized phantom. Noise measured in clinical images is another matter. There has been limited scientific work in that area and none is cited as having been performed by the authors. There is no information provided in the proposal about how the proposed global noise measure is calculated. In particular, the approach does not take into consideration the CT reconstruction settings that can have a dramatic impact on the appearance of the images, including noise, contrast (or CNR), and sharpness. Further, a “global noise” ignores the diversity within the CT series, especially within the (usually) limited locations that depict the abnormality of interest. Inadequate assessment of image quality Image noise alone is an insufficient descriptor of image quality. Noise in an image may also be justifiably varied to meet certain clinical needs (such as high resolution). Many other factors must be considered when attempting to define image quality. Spatial resolution, which includes visualizing small objects and image boundaries, and contrast resolution, of which noise is one component, are also critical aspects of image quality. Widely different noise values may be acceptable under different circumstances for similar protocols. Spatial resolution and contrast are as important as image noise. It is not all clear that improvements in global noise will in turn lead to improved clinical performance. Flawed assumption regarding clinical CT practice There is substantial variation in the radiation doses used in CT exams because the radiation delivered is protocol-specific. The implication in the proposed measures is that radiologists vary these

parameters indiscriminately. In most cases, however, these protocols are established by the institutions based on available equipment, patient population, expertise, scientific evidence, and the nature of cases presented at that institution. With the proposed measures, an optimum study is one that delivers the least radiation dose with an acceptable global noise level, but no evidence is provided that clinicians with high values for the proposed measures perform better or even adequately, only that they perform with less radiation dose. AAPM recommends using an alternative approach that would optimize the acquisition for the task at hand to deliver the least radiation dose necessary while still providing the diagnostic image quality necessary for the task. Dose reduction in and of itself is not enough to improve CT practice. There should also be no loss of clinical performance which is not guaranteed by these measurements. Global noise is not an adequate sufficient metric for image quality. Inadequate accuracy in patient size estimation Assessing a patient size can be challenging because of significant variability in the differences in the habitus of different patients, coupled with the existential challenge that there is no single metric capturing the size of a patient of varying diameter at different cross-sectional locations. The proposed measures rely on the calibration of the company's black-box size estimation to prior work of Cheng 2013 and Christianson 2012, both of which have been updated to newer versions to correct erroneous measures. The error in size measurements needs transparency and validated results. Limited expertise and track record of Alara Imaging Alara Imaging is a new (2020) company without a significant track record of having previously performed a project of such wide scope, scientifically or technically. While the measure developers have published on the topic of radiation dose, they have limited expertise or history with clinical CT, CT image quality, or CT technology. The company has no scientific track record on CT technology, size estimation, or image quality assessment to be considered steward of measures on which it lacks expertise, publication, or scientific history. Moreover, Alara Imaging has limited experience in IT development, with no demonstrated history of interfacing with complex EHR or Radiology Clinical Data systems. The software interface is problematic because it is a vector to a 3rd party product, which can expose healthcare organizations to ransomware attacks by malicious actors seeking valuable patient medical information. In summary, AAPM does not support the endorsement of NQF #3633e, #3662e, and #3663e. AAPM urges NQF to: • Address the concerns identified by AAPM experts; and • Reconsider its recommendation endorsing these measures as proposed. AAPM recognizes that this topic is complex, including scientific, technical and clinical components, and we would welcome the opportunity for greater in-depth discussion on meaningful measures of quality imaging practice. Thank you again for the opportunity to comment on the PSSC evaluation report. If you have any questions or require additional information, please contact Richard J. Martin, JD, Government Relations Project Manager, at 571-298-1227 or Richard@aapm.org.

Developer Response

UCSF thanks the American Association of Physicists in Medicine (AAPM) for their additional comments. UCSF would like to respond and address several inaccuracies and misunderstandings in how the measure is calculated and its intent. COMMENT 1: [Medical Physicists and the American Association of Physicists in Medicine have extensive content expertise that should be considered]…and "while the measure developers have published on the topic of radiation dose, they have limited expertise or history with clinical CT, CT image quality, or CT technology." RESPONSE 1: The measure developers agree that medical physicists have relevant and important

expertise, and they have involved medical physicists in all aspects of our work including both the measure development itself and all of the work in the preceding decade that laid the foundation for UCSF's development of this measure. FIRST: J. Anthony Seibert, PhD was included as a member of our Technical Expert Panel (TEP) to ensure consideration of the perspectives of medical physicists at every step of measure development. Dr. Seibert recently retired as Professor of Diagnostic Imaging Physics and Associate Chair of Radiology Informatics at UC Davis Health and is a past president of the AAPM (2011). In addition to serving on our TEP, Dr. Seibert led UC Davis as a measure testing site and wrote a letter of support for the measure confirming it was "highly feasible" to calculate the measure and noting his belief "that this quality measure can significantly reduce the use of excessive high radiation dose as well as inadequate, sub-optimal low dose used for clinical CT studies." SECOND, the developers also worked closely with another medical physicist, Tim Szczkutowicz, PhD on measure development. Dr. Szczkutowicz guided the work of automating the calculation of image noise, expanding on his earlier published work in this area (Malkus 2017). Dr. Szczkutowicz is an Associate Professor in the Department of Radiology at the University of Wisconsin Madison School of Medicine and Public Health with affiliations in the Department of Medical Physicists and Biomedical Imaging. Lastly, this measure development effort has been led by Rebecca Smith-Bindman, MD, a radiologist and epidemiologist whose primary area of research for the last 15 years has been in quantifying the radiation doses used for CT scanning and identifying ways to safely reduce excessive doses. Over the last 10 years, her research team created a CT radiation dose registry of more than 8 million exams from over 160 facilities, which has allowed the team to quantify the variation in dose, to understand the cause of the variation, and to develop and study interventions to help facilities appropriately lower doses without loss of image quality. The development of this quality measure was a natural extension of this work, and the registry has allowed for the testing of the adult measures. As part of this past work, Dr. Smith-Bindman led a randomized controlled trial of two interventions to optimize CT radiation doses across 100 hospitals and imaging facilities and found that providing feedback (similar to that proposed for these quality measures) along with education and opportunities for sharing best practices resulted in meaningful dose reductions (up to 40%) without any loss in image quality (Smith-Bindman 2020). In total, 13 medical physicists served as site-Principal Investigators for this NIH funded-trial (R01CA181191). Thus, medical physicists have contributed substantially to the body of work that led to the measure as well as measure development. In large part based on Dr. Smith-Bindman's 15-year track record in this area, involving medical physicists, CMS awarded UCSF a cooperative agreement to develop these CT quality measures under the Medicare Access and CHIP Reauthorization Act of 2015 ("MACRA"). COMMENT 2: [There is] unscientific characterization of CT scan risk … the AAPM is concerned that the stated risk of patient radiation dose .. may contribute to fear of diagnostic exams that may in turn lead some patients to refuse safe and appropriate imaging. RESPONSE 2: The measure is not focused on radiation risk and does not calculate nor report radiation risk. A review of the published epidemiological evidence summarizing radiation risk is provided in the application as background and context for the measure (see sections 1a.01-1b.01). This includes several systematic reviews, cohort studies, and an extensive and comprehensive review from the National Academies that conclude that exposure to CT (or radiation doses in the same range as CT) increases a person's risk of developing cancer. Most of these studies do not rely upon the linear no threshold model that the AAPM criticizes in its comment. Indeed, the systematic reviews are based on an observed elevated risk of cancer among patients exposed to medical imaging. The estimates used in the application are based on this

extensive literature review. HOWEVER, radiation risk is not part of the measure at all: it is not calculated nor is it reported. INSTEAD, the measure evaluates dose length product (DLP), and specifically whether size-adjusted DLP exceeds thresholds specific to CT category. DLP is the radiation dose measure most directly under the control of providers, determined by specific parameters that were chosen for the scan. Further, DLP is universally reported by CT scanner manufacturers (unlike other metrics of radiation dose or risk). THE TEP, including the American College of Radiology (ACR), several radiologists, and the medical physicist serving on the committee, unanimously recommended and supported the radiation dose measure used (DLP) and unanimously agreed that it is a relevant metric of quality for CT imaging, as noted in Validity Results, 2b.03. There is also considerable precedent for using DLP to evaluate radiation dose in CT. The American College of Radiology has used DLP to set benchmarks [Kanal 2017] and to measure CT radiation dose in their own NQF-endorsed quality measure #3621. There is no reason to believe that endorsing this measure, which seeks to standardize practice and reduce extreme radiation dose outliers based on DLP, would result in patients refusing appropriate imaging. COMMENT 3: The usability of data resulting from these measures is not clear. The measures do not provide the clinician with an analysis of or methodology for determining what improvements should be made to address a poor showing. It may not be clear to practitioners what a poor score means or how to address it. RESPONSE 3: Entities that report the measure using the measure steward's software will be provided information to both identify causes of performance gaps and make targeted changes to improve quality. There are only two conditions that would push a CT out of compliance - high radiation dose, and poor image quality. The reason for failure (high dose or low quality) will be available to sites on a scan-by-scan basis for those that report the measure using the measure steward's software Those scans where the radiation dose is too high, the dose should be lowered through usual means (technologist education, protocol changes). For those where image quality is too low, more radiation dose should be used through similar means. COMMENTS IN SUPPORT of the measure from many of the testing sites describe how useful the information provided was to allow them to understand and improve their practice. (Available here, beginning page 113: https://www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=96982) As described in our submission and noted above, UCSF found in a randomized controlled trial in 100 hospitals and outpatient radiology practices that providing detailed audit and feedback on radiation doses, similar to what will be provided as part of the feedback on this measure, resulted in significant reductions in radiation dose with no impact on satisfaction with image quality, described in Usability, 4b.01. (Smith-Bindman, 2020) The measure steward does not have control over how 3rd party vendors will report information back to reporting entities. COMMENT 4. The measures rely on categorization of CT data into cohesive groups… There is, however, significant variability in the CT protocol lexicon across institutions that results in making assignment of a given protocol to one of these categories challenging. [There is] substantial oversimplified representation of implementation in practice RESPONSE 4: For the reason the AAPM highlights and for another important reason describe below, the CT category assigned by the measure (reflecting the indication and appropriate radiation dose level for the scan) does not rely on the protocol name at all. As described in Specifications, sp-11, clinical indication for imaging is determined using an algorithm that combines procedure (CPT®) and diagnosis (ICD-10-CM) codes associated with the clinical visit when the test was ordered, information provided as part of the order, and information on the final bill. The codes are available in the radiology electronic systems and/or the EHR or billing systems. The goal in creating the CT categorization decision rules was to identify exams that

are exceptions to the routine dose category (i.e., either high or low dose), (Smith-Bindman, 2021.) Details of the approach for developing and validating the assignment of CT exams to categories in an automated fashion are provided in the measure submission (see Validity sections 2b.02 and 2b.03) and in detailed answers provided to initial comments made by the ACR and AAPM. This approach was first developed using records from over 4.5 million CT exams in the UCSF International CT Dose Registry (Smith-Bindman, 2021), and then turned into an algorithm that combines procedure (CPT®) and diagnosis codes (ICD-10-CM) associated with the clinical visit when the test was ordered, information provided as part of the order, and information on the final bill, provided in Specifications, sp-11. This algorithm was developed using detailed review of over 10,000 patient records from UCSF Health, and validated against "gold standard" chart review, as described in Validity sections 2b.02 and 2b.03. When the algorithm was deployed at our testing sites (including 16 hospitals and 13 outpatient imaging centers), the correct classification rate of the assignment of CT exams to CT category in field-testing was excellent (over 90% for all reporting levels: clinician, clinician group, and facility). Knowing that the algorithm was developed using data from a single health system, the developers performed detailed investigation of the categorization results at testing sites – comparing the assigned CT category against full radiology reports – for the purpose of improving the algorithm. One of the strengths of these measures is that they do not determine the CT category using the protocol name, as this would mask an important quality improvement opportunity (namely, the selection of which protocol to use to scan the patient). Two key process of care components determine radiation doses: (A) the choice of imaging protocol, for example, whether a patient with a suspected pulmonary embolism is imaged with a single- or double-phase CT exam (a decision usually made by the performing radiologist); and (B) the technical settings used for that type of CT exam, which are usually at the discretion of the technologist or medical physicist who oversee and operate the machines. As both of these components contribute to radiation dose, a comprehensive quality measure must encompass both of these decision-making processes. By determining the CT category independent of the protocol used, the measures is able to evaluate both components of quality. COMMENT 5 The noise measure is not an adequate or sufficient parameter of overall image quality. RESPONSE 5: Several comments focus on image quality and the concern that the measure does not offer a comprehensive assessment of image quality. The measure is not intended as a robust measurement of image quality. The primary focus of our measure is to assess radiation dose adjusted for body size. The image quality component was included to protect against the unlikely possibility of substantial degradation of image quality as an unintended consequence of dose reduction. Our measure of image quality uses thresholds developed based on radiologists' satisfaction with images, reflecting what in practice is regarded as adequate for diagnosis. Others might have an interest in more nuanced assessment of image quality for other purposes, but that was not our intent. If the measure is adopted and used, the Steward will closely monitor image noise and measure failure due to low image quality. The Steward will be sensitive to any signal that there is a problem and will revise the measure if changes are needed. COMMENT 6: With these measures, an optimum study is one that delivers the least radiation dose with an acceptable global noise level.. but no evidence is provided that clinicians with high values for the measures perform better or even adequately, only that they perform with less radiation dose. AAPM recommends using an alternative approach that would optimize the acquisition for the task at hand to deliver the least radiation dose necessary while still providing the diagnostic image quality necessary for the task. RESPONSE 6: This measure provides a standardized method for monitoring the

performance of diagnostic CT to discourage unnecessarily high radiation doses, a risk factor for cancer, while preserving image quality. It is expressed as a percentage of eligible CT exams that are out-of-range based on having either excessive radiation dose or inadequate image quality, relative to evidence-based thresholds based on the clinical indication for the exam. THE HIGHER THE SCORE, the higher the proportion of out-of-range exams and the worse the performance. The measure is NOT INTENDED to improve diagnostic accuracy. The purpose of the measure is to establish a radiation dose ceiling to avoid excessive radiation exposure, and an image quality floor to safeguard against unintended deterioration of image quality. ADDITIONALLY, the entire framework for the measure is to ensure the radiation dose and image quality are acceptable for the specific clinical indication, aligned with what the AAPM recommends. The CT categories were created based on radiation dose and image quality requirements specific to the clinical indications for imaging (Smith-Bindman 2021). Using radiologists' satisfaction with image quality, an image quality floor for each category was established, below which an exam is considered to have inadequate quality, and a radiation dose ceiling, beyond which doses are considered unnecessarily high. The purpose is to allow detailed assessment of each CT exam to ensure the dose is optimal based on the clinical indication for imaging. In our testing data, far more CT exams exceeded the radiation dose ceiling (average = 30%) than failed to meet the image quality requirement (average << 1%) (see section 1b.02). The measure encourages entities to reduce the proportion of exams that may "be overdosed for their exact need and condition" while preserving the minimum image quality. COMMENT 7: [There is] Inadequate accuracy in patient size estimation… Assessing a patient size can be challenging because of significant variability in differences in the habitus of different patients, coupled with the existential challenge that there is no single metric capturing the size of a patient of varying diameter at different cross-sectional locations. RESPONSE 7: The developers agree that measuring patient size is important and provided a detailed response to the ACR and AAPM in their earlier comments. FIRST, our approach for using mid-scan diameter is highly correlated with patient weight. In separate, NIH-funded research on CT use in children up to age 21 (Kwan 2022), UCSF has shown that diameter in 4,239 children as measured on mid-scan axial images is highly predictive of patient weight, correlation = 0.904.SECOND, for this measure, patient size is measured using CT image pixel data, either on the mid-scan axial image or the coronal scout image when the mid-scan axial image was not available. This approach has been validated using data from UCSF Health, the UCSF Registry, as well as the data assembled for measure testing from 16 hospitals and 13 outpatient imaging centers. While there may be different ways to measure patient size, and different reasons for measuring patient size, the developers are adjusting for patient size primarily to ensure that entities that see larger patients are not penalized for doing so. using data from the UCSF Registry for abdomen CT we asse the relationship between radiation dose (in DLP) and patient diameter. Abdomen CT was selected as this is the category most influenced by patient size, meaning that patient mix could impact an entity's out-of-range rate. The raw correlation between patient diameter and unadjusted DLP is 0.50, and the marginal R-squared of the log-linear model used for adjustment is 0.15. After size-adjustment, the relationship is nearly removed: the raw correlation is far lower -(0.09), and the modeled marginal R-squared post-adjustment is 0. THIS DEMONSTRATES ADEQUACY OF THE APPROACH FOR PATIENT SIZE ADJUSTMENT TO REMOVE BIAS CAUSED BY CASE MIX. THIRD, the adequacy of size adjustment was shown empirically using data assembled from the testing sites. Out-of-range proportions for routine abdomen exams at 16 hospitals in our testing data based on unadjusted DLP, by decile in patient size are strongly associated by decile in size. Among patients in the highest size decile the

out-of-range proportions across the 16 hospitals ranged from 93-100%. ON THE OTHER HAND The out-of-range rates based on adjusted DLP are not higher among the larger patients. Among patients in the highest size decile, out-of-range rates ranged from 11-53%. THIS ALSO DEMONSTRATES ADEQUACY OF THE APPROACH FOR ADJUSTMENT OF PATIENT SIZE. COMMENT 8 Limited expertise… Alara Imaging Alara is a new company without a significant track record of having previously performed a project of such wide scope, scientifically or technically … Alara has limited IT development experience with no demonstrated history of interfacing with complex EHR or Radiology Clinical Data systems. ..The software interface is problematic because it is a vector to a 3rd party product, which can expose healthcare organizations to ransomware attacks by malicious actors seeking valuable patient medical information. RESPONSE 8: These assertions are incorrect. Measure stewardship is in collaboration with the University of California San Francisco (UCSF). UCSF was responsible for all measure development, scientific research, and measure validation work. UCSF also has a significant track record of successfully performing projects of this scope. About mid-way into the cooperative agreement, CMS asked UCSF to develop and report these measures as eCQMs. UCSF sought guidance from the measures' Technical Expert Panel and asked if members of the TEP would step forward to develop the software and steward the measure. When no group presented itself, it became clear that developing the eCQM and managing nationwide implementation and reporting of this measure was beyond the scope of the UCSF academic team and other TEP member organizations. Accordingly, Dr. Smith-Bindman worked with UCSF to create a company, Alara Imaging, that would help serve as measure steward. UCSF created Alara Imaging to develop the eCQM software and support measure stewardship. Alara is comprised of a team specifically assembled for the creation of this measure software, with deep radiology informatics and technical expertise. The Alara team has over 50 combined years of experience deploying software in hospital environments. The company is new, but the team is well versed in secure implementation. Alara's software is secure. The software is both HIPAA certified and SOC II certified with an independent third-party audit. The software protects against cyberattacks. Tools and information are provided to protect site data and support product security. Alara Imaging's software was used to perform the extensive testing described in the measure application, including correctly calculating measure results for 35,729 CT exams assembled from 7 hospital systems and 1 ambulatory imaging network. Software to calculate the measure will be made available to sites without charge. Measure specifications are in the public domain. If practices do not want to work with Alara, they may work with other vendors to report on the measure. Burden was found to be no more or less onerous than the effort required by participation in other eCQMs or national registries, such as the ACR Dose Index Registry (Feasibility, 3.06). References: KANAL KM et al. U.S. Diagnostic Reference Levels and Achievable Doses for 10 Adult CT Examinations. Radiology. 2017;284(1):120-133. MALKUS A, SZCZYKUTOWICZ TP. A method to extract image noise level from patient images in CT. Med Phys. 2017 Jun;44(6):2173-2184. SMITH-BINDMAN R, Yu S, Wang Y, et al. An Image Quality-informed Framework for CT Characterization. Radiology. 2021 Nov 9:210591. SMITH-BINDMAN R et al. Comparison of the Effectiveness of Single-Component and Multicomponent Interventions for Reducing Radiation Doses in Patients Undergoing Computed Tomography: A Randomized Clinical Trial. JAMA Intern Med. 2020 May 1;180(5):666-675. KWAN M et al. Smith-Bindman senior Author. Quantifying cancer risk from exposures to medical imaging in the Risk of Pediatric and Adolescent Cancer Associated with Medical Imaging (RIC) Study: Research Methods and Cohort Profile Marilyn Kwan et al. Cancer Causes Control 2022 May;33(5):711-726. doi: 10.1007/s10552-022-01556-z. Epub 2022 Feb 2.

NQF Response

Thank you for your comment. It has been shared with the Standing Committee and the measure developer.

NQF Committee Response

Proposed Response: The Patient Safety Standing Committee thanks American Association of Physicists in Medicine for their comment. The Standing Committee does take public comments into account when discussing and recommending measures for endorsement. The Standing Committee made the decision to endorse the measures after reviewing and considering the original comment and the measure developer's response. The Standing Committee stands by their decision to recommend to endorse the measures.

Stephanie Collingwood, UnityPoint Health; Submitted by Stephanie Collingwood

Comment ID#: 7967 (Submitted: 04/25/2022)

Council / Public: PRO

Level of Support: N/A

Comment

UnityPoint Health respectfully offers comments in support of measures 3633, 3662e and 3663e with additional considerations outlined below. UnityPoint Health is one of the nation's most integrated health care systems. Through more than 32,000 employees and our relationships with more than 480 physician clinics, 40 hospitals in urban and rural communities and 14 home health agencies throughout our 9 regions, UnityPoint Health provides care throughout Iowa, central Illinois, and southern Wisconsin. On an annual basis, UnityPoint Health hospitals, clinics and home health provide a full range of coordinated care to patients and families through more than 8.4 million patient visits. Multiple clinically relevant details come into play when determining the appropriate safe dose of radiation for a patient versus obtaining a clear image. Implementation of electronic health record tools requiring this level of documentation within a charting system would be required, along with tools to determine the point system applied. While UnityPoint Health fully understands the value of appropriate CT imaging, operational concerns exist regarding the capability of detailed tracking required to determine excessive CT use on a collective patient population. Additionally, reporting challenges exist today as multiple vendors are used within a health care system. UnityPoint Health supports the concept of this measure but would recommend developing exclusion criteria for overuse.

Developer Response

We thank UnityPoint Health for their comments. Given their large size, the large number of providers and clinics they work with, and the large number of patients they care for, we are grateful they appreciate the value of appropriate CT Imaging. We want to address their misunderstanding in how the measure works. FIRST, The measure is an electronic Clinical Quality Measure (eCQM) and relies on existing electronic data stored in the EHR, billing claims and radiology information systems to calculate the measure. There is no charting nor new documentation required for measure calculation nor a requirement from sites to assign a point

system to CT scans. All data elements used to calculate the measure come from existing structured variables listed in the feasibility scorecards and in Specifications, Table sp-2: CPT® and ICD-10-CM codes; dose length product stored in the DICOM data; and patient diameter and image noise calculated on imaging data. The measure would not have met the requirements of an eCQM had it relied on unstructured or newly created variables. The measure was tested across diverse EHR systems and diverse Radiology Information Systems, including those used by 7 hospital systems and 1 outpatient ambulatory practice group. Data were found to be widely available. SECOND, we strongly agree with UnityPoint Health that relevant clinical details (e.g. the clinical indication for scanning) are required to determine the appropriate radiation dose for each CT scan; e.g. the radiation dose and image quality required for a chest CT performed for lung cancer screening is not the same as required for the surveillance of known lung cancer. The approach of assigning CT examinations to the different CT categories (reflecting the clinical indications and required radiation dose and image quality) as specified in the measure was first developed using records from over 4.5 million CT exams in the UCSF International CT Dose Registry (Smith-Bindman, 2021). We then developed an approach for determining the clinical indication for imaging using an algorithm that combines procedure (CPT®) and diagnosis codes (ICD-10-CM) associated with the clinical visit when the test was ordered, information provided as part of the order, and information on the final bill. These are provided in Specifications, sp-11. This algorithm was developed using detailed review of over 10,000 patient records from UCSF Health. We validated the CT category assignment using the algorithm against "gold standard" chart review, as described in Validity sections 2b.02 and 2b.03. When the algorithm was deployed at our testing sites (including 16 hospitals and 13 outpatient imaging centers) the correct classification rate of the assignment of CT exams to CT category in field-testing was highly accurate across clinicians, clinician groups and hospitals. The goal in creating the CT categorization decision rules was to identify exams that are exceptions to the routine dose category (i.e. either high or low dose); most scans fall in the routine dose categories. IN SUMMARY, the calculation of the measure score does not require any new charting, does not require change in clinical practice, and does not require new documentation. Reference: Smith-Bindman R, Yu S, Wang Y, et al. An Image Quality-informed Framework for CT Characterization. Radiology. 2021 Nov 9:210591.

NQF Response

Thank you for your comment. It has been shared with the Standing Committee and the measure developer.

NQF Committee Response

N/A