Measure Number: IEP-005-10

Measure Title: Pulmonary CT Imaging for Patients at Low Risk for Pulmonary Embolism

Description: Percent of patients undergoing CT pulmonary angiogram for the evaluation of possible PE who are at low-risk for PE consistent with guidelines prior to CT imaging.

Numerator Statement: The number of denominator patients with either: a low clinical probability and any negative D-dimer, or an intermediate clinical probability and a negative high-sensitivity D-dimer, or no pretest probability documented.

Denominator statement: Number of patients who have a CT pulmonary angiogram (CTPA) for the evaluation of possible pulmonary embolism


Data Source: paper medical record/flowsheet, Electronic clinical data, electronic Health/Medical Record

Measure developer: Partners Healthcare System, Inc.

Type of Endorsement (full or time-limited): Time-limited endorsement

Attachments: N/A
This form contains the measure information submitted by stewards. Blank fields indicate no information was provided. Attachments also may have been submitted and are provided to reviewers. The sub-criteria and most of the footnotes from the evaluation criteria are provided in Word comments and will appear if your cursor is over the highlighted area (or in the margin if your Word program is set to show revisions in balloons). Hyperlinks to the evaluation criteria and ratings are provided in each section.

**TAP/Workgroup** (if utilized): Complete all yellow highlighted areas of the form. Evaluate the extent to which each sub-criterion is met. Based on your evaluation, summarize the strengths and weaknesses in each section.

**Note:** If there is no TAP or workgroup, the SC also evaluates the sub-criteria (yellow highlighted areas).

**Steering Committee:** Complete all pink highlighted areas of the form. Review the workgroup/TAP assessment of the sub-criterion, noting any areas of disagreement; then evaluate the extent to which each major criterion is met; and finally, indicate your recommendation for the endorsement. Provide the rationale for your ratings.

**Evaluation ratings of the extent to which the criteria are met**

- **C** = Completely (unquestionably demonstrated to meet the criterion)
- **P** = Partially (demonstrated to partially meet the criterion)
- **M** = Minimally (addressed BUT demonstrated to only minimally meet the criterion)
- **N** = Not at all (NOT addressed; OR incorrectly addressed; OR demonstrated to NOT meet the criterion)
- **NA** = Not applicable (only an option for a few sub-criteria as indicated)

---

**DE.1 Measure Title:** Pulmonary CT Imaging for Patients at Low Risk for Pulmonary Embolism

**DE.2 Brief description of measure:** Percent of patients undergoing CT pulmonary angiogram for the evaluation of possible PE who are at low-risk for PE consistent with guidelines(1) prior to CT imaging.


**1.1 Type of Measure:** efficiency/cost

**De.3 If included in a composite or paired with another measure, please identify composite or paired measure**

**De.4 National Priority Partners Priority Area:** Overuse

**De.5 IOM Quality Domain:** efficiency, safety

**De.6 Consumer Care Need:**
A.3 Measure Steward Agreement: agreement signed and submitted
A.4 Measure Steward Agreement attached:

B. The measure owner/steward verifies there is an identified responsible entity and process to maintain and update the measure on a schedule that is commensurate with the rate of clinical innovation, but at least every 3 years. Yes, information provided in contact section

C. The intended use of the measure includes both public reporting and quality improvement.
   ►Purpose: public reporting, quality improvement 0,0,0,

D. The requested measure submission information is complete. Generally, measures should be fully developed and tested so that all the evaluation criteria have been addressed and information needed to evaluate the measure is provided. Measures that have not been tested are only potentially eligible for a time-limited endorsement and in that case, measure owners must verify that testing will be completed within 12 months of endorsement.
   D.1 Testing: No, testing will be completed within 12 months
   D.2 Have NQF-endorsed measures been reviewed to identify if there are similar or related measures? Yes

(for NQF staff use) Have all conditions for consideration been met?
Staff Notes to Steward (if submission returned):

Staff Notes to Reviewers (issues or questions regarding any criteria):

Staff Reviewer Name(s):

TAP/Workgroup Reviewer Name:

Steering Committee Reviewer Name:

1. IMPORTANCE TO MEASURE AND REPORT

Extent to which the specific measure focus is important to making significant gains in health care quality (safety, timeliness, effectiveness, efficiency, equity, patient-centeredness) and improving health outcomes for a specific high impact aspect of healthcare where there is variation in or overall poor performance. Measures must be judged to be important to measure and report in order to be evaluated against the remaining criteria. (evaluation criteria)

1a. High Impact

(for NQF staff use) Specific NPP goal:

1a.1 Demonstrated High Impact Aspect of Healthcare: affects large numbers, patient/societal consequences of poor quality, frequently performed procedure, high resource use

1a.2

1a.3 Summary of Evidence of High Impact: The symptoms suggestive of pulmonary embolism (PE) and deep vein thrombosis (DVT) are very common. Over 10 million persons present each year in the US with chest pain or breathing difficulties, the main symptoms of PE (1). While the exact prevalence of PE in the Emergency Department (ED) setting is unknown given no reliable measurement of missed cases, it has been estimated that 1 in every 500 to 1 in every 1000 ED patients has a PE (2). In addition, A recent multicenter study of 12 US EDs and 1 New Zealand ED found that approximately 1.5% of all ED patients underwent a CTPA to evaluate for suspected pulmonary embolism. (3)

The past decade has also seen the implementation and validation of new technologies including d-dimer serological testing and CTPA for the evaluation and diagnosis of suspected venous thromboembolism. The adoption and application of these technologies has resulted in significant change in US practice with CTPA use becoming increasingly common and considered the definitive test for PE3.

1a.4 Citations for Evidence of High Impact: (1). McCaig LF, Burt CW. National Hospital Ambulatory

Comment [KP1]: 1a. The measure focus addresses:
   • a specific national health goal/priority identified by NQF’s National Priorities Partners; OR
   • a demonstrated high impact aspect of healthcare (e.g., affects large numbers, leading cause of morbidity/mortality, high resource use (current and/or future), severity of illness, and patient/societal consequences of poor quality).
**1b. Opportunity for Improvement**

**1b.1 Benefits (improvements in quality) envisioned by use of this measure:** This measure will improve efficiency by reducing the inappropriate ordering of CTPA for pulmonary embolism based on pre-test probability estimation. This measure does not require utilization of a structured clinical prediction rule such as the Wells Score or Geneva Score, however the measure aims to improve efficiency by guiding clinical practice towards use of initial d-dimer testing rather than immediate CTPA in low or intermediate probability patients as indicated.

In addition to imaging efficiency, the overuse of CTPA in ED patients with suspected pulmonary embolism has tangible implications for patient safety. Ionizing radiation from CTPA can increase the lifetime risk of cancer, particularly in young women due to the added vulnerability of breast tissue(1). Also, the use iodinated dye places patients at risk of contrast induced nephropathy, which a study by Mitchell and Kline estimated at approximately 8% of all patients undergoing CTPA in the ED(2),(3).


**1b.2 Summary of data demonstrating performance gap (variation or overall poor performance) across providers:**

Despite significant evidence supporting the use of structured clinical assessment in combination with d-dimer testing to develop an evaluation of patients with suspected PE, there remains poor application of these algorithms in the ED setting(1). There are numerous studies demonstrating poor application of clinical pre-test assessment to PE testing strategies including:

- Single-center study demonstrated suboptimal application of Wells criteria as 25% of patients with a normal or intermediate probability d-dimer assays subsequently had CTPA ordered to evaluate for PE, with only 2.7% (0.7% of cohort) subsequently having PE.(2)
- A large (5,344 patient) single center cohort study demonstrated that of 2,285 patients with negative d-dimer testing, 166 (7%) underwent CTPA, demonstrating inappropriate use of radiography outside established clinical algorithms.(3)
- Use of an ED protocol that combined structured clinical assessment with d-dimer testing doubles the rate of testing for PE, without increased imaging.(4)

**1b.3 Citations for data on performance gap:**

(1) Runyon MS, Richman PB, Kline JA. Emergency medicine practitioner knowledge and use of decision rules for the evaluation of patients with suspected pulmonary embolism: variations by practice setting and training level. Acad Emerg Med. 2007;14:53-57.

Comment [KP2]: 1b. Demonstration of quality problems and opportunity for improvement, i.e., data demonstrating considerable variation, or overall poor performance, in the quality of care across providers and/or population groups (disparities in care).

Comment [k3]: 1 Examples of data on opportunity for improvement include, but are not limited to: prior studies, epidemiologic data, measure data from pilot testing or implementation. If data are not available, the measure focus is systematically assessed (e.g., expert panel rating) and judged to be a quality problem.
1c. Outcome or Evidence to Support Measure Focus

1c.1 Relationship to Outcomes (For non-outcome measures, briefly describe the relationship to desired outcome. For outcomes, describe why it is relevant to the target population): The proposed measure is not an outcome measure. There will be no reporting of diagnostic outcomes or patient specific outcomes. Rather the measure is related to the overall use of CTPA technology for patients with suspected pulmonary embolism, and this measure will report total CTPA use as an outcome. The desired outcome is more appropriate use of CTPA technology, which this measure will report by demonstrating higher rates of guideline adherence.

1c.2-3. Type of Evidence: cohort study, observational study, evidence based guideline

1c.4 Summary of Evidence (as described in the criteria; for outcomes, summarize any evidence that health care services/care processes influence the outcome): There are single center studies that demonstrate that the use of structured clinical assessment, when combined with d-dimer testing, can result in more “efficient” use of CTPA as the number of patients tested for disease increased without increasing the use of CTPA(1). Furthermore there is significant multi-center data suggesting that the use of pre-defined clinical algorithms can improve the diagnostic rate in acute pulmonary embolism. The intention of this measure is not to create a new guideline or measure patient level outcomes, but rather use an imaging efficiency measure to improve guideline adherence.


1c.5 Rating of strength/quality of evidence (also provide narrative description of the rating and by whom): The strength of evidence has been rated and summarized by the Task Force responsible for the guidelines used for this measure. Specific to this measure, the table on page 2291 summarizes the strength of evidence used for this measure. This measure only draws on Level A evidence within the guideline and is therefore not subject to any Level C recommendations made by the group in the absence of strong evidence.

1c.6 Method for rating evidence: Evidence rating by the ESC was conducted by an expert panel including global experts in pulmonary embolism. A full description of the evidence review process is included on pages 2277-2278 of the ESC guidelines.

1c.7 Summary of Controversy/Contradictory Evidence: The wealth of evidence to date supports the
The notion that a combination of a low-probability clinical assessment and negative d-dimer testing is sufficient to exclude the diagnosis of acute pulmonary embolism. There remains some controversy, however, regarding the use of d-dimer testing in intermediate probability patients. This controversy is largely driven by variation in test characteristics between d-dimer assays and variation in the application of Wells criteria or modified Wells criteria in the literature.

The guidelines that are the basis for this measure make a distinction between high sensitivity and non-high sensitivity d-dimer assays. High sensitivity d-dimer assays (ELISA, and ELISA derived) have been shown to effectively exclude the diagnosis of PE in patients with low and intermediate probability of pulmonary embolism with a sensitivity greater than 95%. This is in contrast with lower sensitivity d-dimer assays (quantitative latex-derived assays and whole-blood agglutination), which have only been shown to exclude PE in patients with low-probability of pulmonary embolism in three level probability schemes or low probability in two level schemes (Wells score =4).


1c.9 Quote the Specific guideline recommendation (including guideline number and/or page number):
Guidelines are not numbered, however the actual guideline text is noted on page 2291 of the guidelines.

“Suspected non-high risk PE: Plasma D-dimer measurement is recommended in emergency department patients to reduce the need for unnecessary imaging and irradiation, preferably using a highly sensitive assay” (Class I Level A recommendation)

“Suspect non-high risk PE and low-clinical probability: Normal D-dimer level using either a highly or moderately sensitive assay excludes PE” (Class I Level A recommendation)

“Suspect non-high risk PE and intermediate-clinical probability: Normal D-dimer level using a highly sensitive assay excludes PE, ” (Class I Level A recommendation)

“Suspect non-high risk PE and intermediate-clinical probability: Further testing should be considered if D-dimer level is normal when using a less sensitive assay” (Class Ila Level B recommendation)


1c.11 National Guideline Clearinghouse or other URL: There is no pulmonary embolism testing guideline
in the National Guideline Clearinghouse, however a complimentary NGC guideline on the appropriateness of imaging from the American College of Radiology is posted. This guideline is outdated compared to the ESC guideline. Bettmann MA, Lyders EM, Yucel EK, Khan A, Haramati LB, Ho VB, Rozenstein A, Rybicki FJ, Schoepf UJ, Stanford W, Woodard PK, Jaff M, Expert Panel on Cardiac Imaging. Acute chest pain--suspected pulmonary embolism. [online publication]. Reston (VA): American College of Radiology (ACR); 2006. 5 p. [42 references]

1c.12 Rating of strength of recommendation (also provide narrative description of the rating and by whom):
The European Society of Cardiology utilizes a three tier evidence grading system (A, B, or C) and Level of Recommendation system (I, II, III) consistent with the American College of Cardiology and well accepted for guideline and performance measure development. The recommendations made by this measure are with regard to the use of CTPA in ED patients with suspected pulmonary embolism, and are all Class I Level of Evidence A recommendations in the guidelines.

1c.13 Method for rating strength of recommendation (If different from USPSTF system, also describe rating and how it relates to USPSTF):
Evidence grading by the European Society of Cardiology follows a three-tier scale with similar description to the USPSTF except that level of evidence is graded a letter scheme (A,B,C) rather than as good, fair and poor by the USPSTF. ESC definitions for evidence classification are:
• Level of evidence A Data derived from multiple randomized clinical trials or meta-analyses
• Level of evidence B Data derived from a single randomized clinical trials or large non-randomized studies
• Level of evidence C Consensus of opinion of the experts and/or small studies, retrospective studies, registries

Recommendations from the ESC follow a Class system mirrored by most professional societies:
• Class I Evidence and/or general agreement that a given treatment or procedure is beneficial, useful, and effective
• Class II Conflicting evidence and/or a divergence of opinion about the usefulness/efficacy of the given treatment or procedure
  o Class IIa Weight of evidence/opinion is in favor of usefulness/efficacy
  o Class IIb Usefulness/efficacy is less well established by evidence/opinion
• Class III Evidence or general agreement that the given treatment or procedure is not useful/ effective, and in some cases may be harmful.

1c.14 Rationale for using this guideline over others:
The European Society of Cardiology guideline was selected because it reflects the current state of evidence with respect to structured clinical assessment, d-dimer assays, and CTPA technology. While this guideline was developed by a European society, it incorporated predominantly North American studies as its literature basis, and an accompanying editorial with the guidelines by Dr. Goldhaber demonstrates its applicability to the United States.

All other guidelines for evaluation and testing strategies in pulmonary embolism are severly years old and therefore are based on studies utilizing less sensitive d-dimer assays, evidence from non Emergency Department ambulatory settings, or lower resolution CT scanners. The ESC guidelines are also consistent with major review papers written on acute pulmonary embolism over the past two years, which again speak to the consensus around these clinical algorithms.

TAP/Workgroup: What are the strengths and weaknesses in relation to the sub-criteria for Importance to Measure and Report? 1

Steering Committee: Was the threshold criterion, Importance to Measure and Report, met? Y

2. SCIENTIFIC ACCEPTABILITY OF MEASURE PROPERTIES

Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about
the quality of care when implemented. (evaluation criteria)

2a. MEASURE SPECIFICATIONS

| S.1 Do you have a web page where current detailed measure specifications can be obtained? |
| S.2 If yes, provide web page URL: |
| **Rating** |
| S.1 | S.2 |

2a. Precisely Specified

2a.1 Numerator Statement *(Brief, text description of the numerator - what is being measured about the target population, e.g. target condition, event, or outcome):*
The number of denominator patients with either: a low clinical probability and any negative D-dimer, or an intermediate clinical probability and a negative high-sensitivity D-dimer, or no pretest probability documented.

2a.2 Numerator Time Window *(The time period in which cases are eligible for inclusion in the numerator):*
This measure does not measure across time intervals as all numerator and denominator elements are available at the index visit.

2a.3 Numerator Details *(All information required to collect/calculate the numerator, including all codes, logic, and definitions):*
Number of hemodynamically stable patients who receive CT pulmonary angiograms for suspected pulmonary embolism who have of either†:
1. a low clinical probability* of PE and a negative D-Dimer
2. a low clinical probability* of PE and no D-Dimer performed
3. No documentation of a pre-test probability
†Documentation at the time of test ordering, timed prior to test initiation.
*clinical probability can be determined by a structured prediction tool (Wells, Revised Geneva) or implicit judgment.
Specific test cutoffs will be determined by each ED or institution a priori.

2a.4 Denominator Statement *(Brief, text description of the denominator - target population being measured):*
Number of patients who have a CT pulmonary angiogram (CTPA) for the evaluation of possible pulmonary embolism

2a.5 Target population gender: Female, Male
2a.6 Target population age range: This measure will be applied to all adult patients (Age=18).

2a.7 Denominator Time Window *(The time period in which cases are eligible for inclusion in the denominator):*
This measure does not measure across time intervals as all numerator and denominator elements are available at the index visit.

2a.8 Denominator Details *(All information required to collect/calculate the denominator - the target population being measured - including all codes, logic, and definitions):*
Denominator Inclusions:
Age =18
CTPA performed

2a.9 Denominator Exclusions *(Brief text description of exclusions from the target population):*
Hemodynamically unstable pulmonary embolism suspected by hypotension and/or shock.

Comment [KPB]: 2a. The measure is well defined and precisely specified so that it can be implemented consistently within and across organizations and allow for comparability. The required data elements are of high quality as defined by NQF’s Health Information Technology Expert Panel (HITEP).

Comment [k9]: 11 Risk factors that influence outcomes should not be specified as exclusions.
12 Patient preference is not a clinical exception to eligibility and can be influenced by provider interventions.
<table>
<thead>
<tr>
<th>Section</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>2a.10 Denominator Exclusion Details</td>
<td>(All information required to collect exclusions to the denominator, including all codes, logic, and definitions): Definition of Systemic Hypotension: systolic blood pressure &lt;90mm Hg or a reduction of at least 40mmHg for at least 15 min (1).</td>
</tr>
<tr>
<td>2a.11 Stratification Details/Variables</td>
<td>(All information required to stratify the measure including the stratification variables, all codes, logic, and definitions):</td>
</tr>
<tr>
<td>2a.12-13 Risk Adjustment Type</td>
<td>no risk adjustment necessary</td>
</tr>
<tr>
<td>2a.14 Risk Adjustment Methodology/Variables</td>
<td>(List risk adjustment variables and describe conceptual models, statistical models, or other aspects of model or method):</td>
</tr>
<tr>
<td>2a.15-17 Detailed risk model available Web page URL or attachment:</td>
<td></td>
</tr>
<tr>
<td>2a.18-19 Type of Score</td>
<td>rate/proportion</td>
</tr>
<tr>
<td>2a.20 Interpretation of Score</td>
<td>better quality = lower score</td>
</tr>
<tr>
<td>2a.21 Calculation Algorithm</td>
<td>(Describe the calculation of the measure as a flowchart or series of steps): See attached data sheet</td>
</tr>
<tr>
<td>1.</td>
<td>identify all e.g. patients undergoing CT PA using appropriate procedure codes</td>
</tr>
<tr>
<td>2.</td>
<td>review available data for evidence of pretest probability. This can include the medical record, and/or computerized or paper-based physician orders,</td>
</tr>
<tr>
<td>3.</td>
<td>divide number of patients with CT PA and low risk or no pretest probability BY the total number of patients with CT PA.</td>
</tr>
<tr>
<td>2a.22 Describe the method for discriminating performance (e.g., significance testing):</td>
<td>This measure does not require any significance testing. Rates of appropriate imaging use will be reported with the opportunity for classification by quintiles or other similar mechanisms based on initial reporting.</td>
</tr>
<tr>
<td>2a.23 Sampling (Survey) Methodology</td>
<td>If measure is based on a sample (or survey), provide instructions for obtaining the sample, conducting the survey and guidance on minimum sample size (response rate): Sampling is acceptable and will follow a standard sample size methodologies for process measures.</td>
</tr>
<tr>
<td>2a.24 Data Source</td>
<td>(Check the source(s) for which the measure is specified and tested) paper medical record/flowsheet, Electronic clinical data, electronic Health/Medical Record</td>
</tr>
<tr>
<td>2a.25 Data source/data collection instrument</td>
<td>(Identify the specific data source/data collection instrument, e.g. name of database, clinical registry, collection instrument, etc.): Data will be collected from the medical record, specifically from the provider’s order for a CTPA. No specific data collection instrument needs to be used since the determination of guideline adherence will be made solely on the criteria mentioned in the guideline. These can be easily recorded either electronically or on paper using institution-specific instruments.</td>
</tr>
<tr>
<td>2a.29-31 Data dictionary/code table web page URL or attachment:</td>
<td></td>
</tr>
<tr>
<td>2a.32-35 Level of Measurement/Analysis</td>
<td>(Check the level(s) for which the measure is specified and tested) Clinicians: Group, Facility/Agency, Population: national, Population: regional/network, Program: QIO</td>
</tr>
<tr>
<td>2a.36-37 Care Settings</td>
<td>(Check the setting(s) for which the measure is specified and tested)</td>
</tr>
</tbody>
</table>

Rating: C=Completely; P=Partially; M=Minimally; N=Not at all; NA=Not applicable
Ambulatory Care: Emergency Dept, Other (specify) This measure was developed for use in the ED, but the guideline upon which it is based is not specific for the ED. It would be reasonable to consider the measure for the following additional care settings: Office, Clinic, and Hospital Outpatient.

2a.38-41 Clinical Services (Healthcare services being measured, check all that apply) Clinicians: PA/NP/Advanced Practice Nurse, Clinicians: Physicians (MD/DO)

### 2b. Reliability testing

2b.1 Data/sample (description of data/sample and size): The guidelines used as the basis for the measure are drawn from large randomized controlled trials of diagnostic strategies for pulmonary embolism conducted in the United States and Europe. Numerous previous guidelines and review papers have cited these trials as sufficiently generalizable for guideline development and practice guidance.

In addition to the evidence base of these guidelines, we (Brigham and Women’s Hospital) are currently engaging in internal quality improvement initiatives to measure efficiency in CTPA use for ED patients with suspected pulmonary embolism.

2b.2 Analytic Method (type of reliability & rationale, method for testing):

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

### 2c. Validity testing

2c.1 Data/sample (description of data/sample and size): Validity testing is ongoing and will be completed within 24 months.

2c.2 Analytic Method (type of validity & rationale, method for testing):

2c.3 Testing Results (statistical results, assessment of adequacy in the context of norms for the test conducted):

### 2d. Exclusions Justified

2d.1 Summary of Evidence supporting exclusion(s):
The evidence surrounding the prognostic significance of shock and hypotension in acute pulmonary embolism is largely derived from two observational registries: ICOPER and Management and Prognosis in Pulmonary Embolism Trial (MAPPET)(1),(2),(3). Both registries demonstrate significant mortality associated with both conditions, and these patients are too unstable for structured assessment and serial testing strategies in the ED setting. These patients represent a very small percentage of ED patients tested for pulmonary embolism, and will therefore represent a small exclusion.

2d.2 Citations for Evidence:


10 Examples of evidence that an exclusion distorts measure results include, but are not limited to: frequency of occurrence, sensitivity analyses with and without the exclusion, and variability of exclusions across providers.
2d.3 Data/sample (description of data/sample and size):
2d.4 Analytic Method (type analysis & rationale):
2d.5 Testing Results (e.g., frequency, variability, sensitivity analyses):

2e. Risk Adjustment for Outcomes/ Resource Use Measures
2e.1 Data/sample (description of data/sample and size):
2e.2 Analytic Method (type of risk adjustment, analysis, & rationale):
2e.3 Testing Results (risk model performance metrics):
2e.4 If outcome or resource use measure is not risk adjusted, provide rationale:

2f. Identification of Meaningful Differences in Performance
2f.1 Data/sample from Testing or Current Use (description of data/sample and size):
2f.2 Methods to identify statistically significant and practicallymeaningfully differences in performance (type of analysis & rationale):
2f.3 Provide Measure Scores from Testing or Current Use (description of scores, e.g., distribution by quartile, mean, median, SD, etc.; identification of statistically significant and meaningfully differences in performance):

2g. Comparability of Multiple Data Sources/Methods
2g.1 Data/sample (description of data/sample and size):
2g.2 Analytic Method (type of analysis & rationale):
2g.3 Testing Results (e.g., correlation statistics, comparison of rankings):

2h. Disparities in Care
2h.1 If measure is stratified, provide stratified results (scores by stratified categories/cohorts):
2h.2 If disparities have been reported/identified, but measure is not specified to detect disparities, provide follow-up plans:

TAP/Workgroup: What are the strengths and weaknesses in relation to the sub-criteria for Scientific Acceptability of Measure Properties?

Steering Committee: Overall, to what extent was the criterion, Scientific Acceptability of Measure Properties, met?
Rationale:

3. USABILITY

Comment [KP16]: 2e. For outcome measures and other measures (e.g., resource use) when indicated:
• an evidence-based risk-adjustment strategy (e.g., risk models, risk stratification) is specified and is based on patient clinical factors that influence the measured outcome (but not disparities in care) and are present at start of care.
(Proposed BM: “or” can be replaced with a note that risk adjustment is not necessary).

Comment [KP17]: 2f. Risk models should not obscure disparities in care for populations by including factors that are associated with differences/inequalities in care such as race, socioeconomic status, gender (e.g., poorer treatment outcomes of African American men with prostate cancer, inequalities in treatment for cardiovascular risk factors between men and women). It is preferable to stratify measures by race and socioeconomic status rather than adjusting out differences.

Comment [KP18]: 2f. Data analysis demonstrates that methods for scoring and analysis of the specified measure allow for identification of statistically significant and pracitically/clinically meaningful differences in performance.

Comment [KP19]: 2f. Data analysis demonstrates that methods for scoring and analysis of the specified measure allow for identification of statistically significant and practically/clinically meaningful differences in performance.

Comment [KP20]: 2g. If multiple data sources/methods are allowed, there is demonstration they produce comparable results.

Comment [KP21]: 2h. If disparities in care have been identified, measure specifications, scoring, and analysis allow for identification of disparities through stratification of results (e.g., by race, ethnicity, socioeconomic status, gender); OR rationale/data justifies why stratification is not necessary or not feasible.
Extent to which intended audiences (e.g., consumers, purchasers, providers, policy makers) can understand the results of the measure and are likely to find them useful for decision making. (evaluation criteria)

3a.1. Current Use: testing not yet completed

3a.2 Use in a public reporting initiative (disclosure of performance results to the public at large) (If used in a public reporting initiative, provide name of initiative(s), locations, Web page URL(s). If not publicly reported, state the plans to achieve public reporting within 3 years):

We intend this measure to be suitable for public reporting in the future. We plan to continue our internal Quality Improvement study to demonstrate the efficiencies in imaging, which can be result from use of the measure.

3a.3. If used in other programs/initiatives (If used in quality improvement or other programs/initiatives, name of initiative(s), locations, Web page URL(s). If not used for QI, state the plans to achieve use for QI within 3 years):

This measure is currently being used in a quality improvement program at Brigham and Women’s hospital.

Testing of Interpretabillity (Testing that demonstrates the results are understood by the potential users for public reporting and quality improvement)

3a.4 Data/sample (description of data/sample and size):

3a.5 Methods (e.g., focus group, survey, QI project):

3a.6 Results (qualitative and/or quantitative results and conclusions):

3b/3c. Relation to other NQF-endorsed measures

3b.1 NQF # and Title of similar or related measures:

There are currently no directly related measures, however there are some associated NQF measures with which this measure is complimentary. These include: 1. Institute for Clinical Systems Improvement (ICSI). Venous thromboembolism diagnosis and treatment. Bloomington (MN): Institute for Clinical Systems Improvement (ICSI); 2009 Feb. 79 p. [220 references] This measure is not specific to ED patients and is written with respect to treatment, evaluation and diagnosis, as such there is no discrepancy between these two complimentary measures. 2. Venous thromboembolism (VTE) diagnosis and treatment: percentage of adult patients suspected of deep vein thrombosis (DVT) who have leg duplex ultrasound with compression performed despite a low clinical pretest probability and a negative D-dimer test. Institute for Clinical Systems Improvement (ICSI). Venous thromboembolism diagnosis and treatment. Bloomington (MN): Institute for Clinical Systems Improvement (ICSI); 2009 Feb. 79 p. [220 references] This measure is similar to our measure with the goal of improving the appropriateness of duplex ultrasound use rather than CTPA use. Since this measure addresses a different technology in the same patient population, our measure would be complimentary to this measure.

(for NQF staff use) Notes on similar/related endorsed or submitted measures:

3b. Harmonization

If this measure is related to measure(s) already endorsed by NQF (e.g., same topic, but different target population/settings/data source or different topic but same target population):

3b.2 Are the measure specifications harmonized? If not, why?

As above, while this measure is related to other NQF measures there is no actual discrepancy or overlap requiring additional harmonization.

3c. Distinctive or Additive Value

3c.1 Describe the distinctive, improved, or additive value this measure provides to existing NQF-endorsed measures:

Rating: C=Completely; P=Partially; M=Minimally; N=Not at all; NA=Not applicable
5.1 Competing Measures  If this measure is similar to measure(s) already endorsed by NQF (i.e., on the same topic and the same target population), describe why it is a more valid or efficient way to measure quality:

TAP/Workgroup: What are the strengths and weaknesses in relation to the sub-criteria for Usability?

Steering Committee: Overall, to what extent was the criterion, Usability, met?

Rationale:

4. FEASIBILITY

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

4a. Data Generated as a Byproduct of Care Processes

4a.1-2 How are the data elements that are needed to compute measure scores generated? data generated as byproduct of care processes during delivery.

4b. Electronic Sources

4b.1 Are all the data elements available electronically? (elements that are needed to compute measure scores are in defined, computer-readable fields, e.g., electronic health record, electronic claims)

No

4b.2 If not, specify the near-term path to achieve electronic capture by most providers. All data elements are not likely to be available electronically to most providers currently. Although many electronic health records include computerized physician order entry (CPOE) for radiologic tests, most are not currently programmed to have guideline-based decision support. At Brigham and Women's Hospital, the Center for Evidence Based Imaging has developed a CPOE interface that can collect specific clinical information at the time of ordering and offer interactive decision support. This measure is one of several for which there is ongoing quality improvement work utilizing this interface. Although most electronic health records do not currently have the exact specifications for this measure in their CPOE, it is technically feasible for them to be reprogrammed to include such data. The measure specifications provided include all information needed for any EHR to be reprogrammed to collect the needed data elements.

Providers who do not have CPOE could implement a templated paper order entry form that included all data fields. Alternatively they could conduct chart review to identify if the data fields were present at the time of test ordering, but this would likely have a low yield as most clinical charts do not have time to data entry and many are completed at the end of the patient visit. If approved by the NQF, we would produce a model templated paper order entry form for this measure. Ultimately, this and other measures will be significantly aided by the transition to electronic health records.

4c. Exclusions

4c.1 Do the specified exclusions require additional data sources beyond what is required for the numerator and denominator specifications?

Yes

4c.2 If yes, provide justification. The specified exclusions require additional data sources only if an electronic order entry system is not programmed to capture them. In this case, clinical records, either electronic or paper would be needed to indentify exclusions. An EHR can be programmed to collect all data on exclusions at the time of order entry.

4d. Susceptibility to Inaccuracies, Errors, or Unintended Consequences

Comment [K26]: 5. Demonstration that the measure is superior to competing measures - new submissions and/or endorsed measures (e.g., is a more valid or efficient way to measure).

Comment [KP27]: 4a. For clinical measures, required data elements are routinely generated concurrent with and as a byproduct of care processes during care delivery. (e.g., BP recorded in the electronic record, not abstracted from the record later by other personnel; patient self-assessment tools, e.g., depression scale; lab values, meds, etc.)

Comment [KP28]: 4b. The required data elements are available in electronic sources. If the required data are not in existing electronic sources, a credible, near-term path to electronic collection by most providers is specified and clinical data elements are specified for transition to the electronic health record.

Comment [KP29]: 4c. Exclusions should not require additional data sources beyond what is required for scoring the measure (e.g., numerator and denominator) unless justified as supporting measure validity.

Comment [KP30]: 4d. Susceptibility to inaccuracies, errors, or unintended consequences and the ability to audit the data items to detect such problems are identified.
4d.1 Identify susceptibility to inaccuracies, errors, or unintended consequences of the measure and describe how these potential problems could be audited. If audited, provide results.

As with most measures based on guideline recommendations, the major source of inaccuracy or error will be incomplete medical records. This measure is based on a set of specific clinical criteria outlined by the guideline and will require physicians to document the presence or absence of these criteria in patients undergoing CT imaging. As all CTPAs require an order from a clinician, it is completely feasible for the required information to be integrated into the ordering process, either on a paper or electronic ordering template.

The main unintended consequence of this measure is that CT images ordered by emergency physicians at the request of consultants may be attributed to the emergency physicians themselves. However, by analyzing this measure at the Group or Facility level, organizations can develop measure-specific policies that will apply to all physicians, including emergency physicians and consultants.

4e. Data Collection Strategy/Implementation

4e.1 Describe what you have learned/modified as a result of testing and/or operational use of the measure regarding data collection, availability of data/missing data, timing/frequency of data collection, patient confidentiality, time/cost of data collection, other feasibility/implementation issues:

Successful data collection using an electronic order entry system is dependent on designing an explicit order form with a method of categorizing indications for CT imaging. If these indications are categorized correctly, the inclusion and exclusion criteria can effectively sort the CT images obtained into those to which the guideline should apply and those to which it should not.

4e.2 Costs to implement the measure (costs of data collection, fees associated with proprietary measures):

The cost to implement this measure will depend on the method used to collect data. An electronic order entry system, after it is programmed, will be able to determine guideline-appropriateness for little or no cost other than that associated with the programming. Personnel time will be needed if paper medical records are to be reviewed in order to determine the appropriateness of individual CTs.

4e.3 Evidence for costs:

4e.4 Business case documentation:
### Additional Information

<table>
<thead>
<tr>
<th>Co.1 <strong>Organization</strong></th>
<th>Partners HealthCare System, Inc.</th>
<th>Prudential Tower, 800 Boylston Street, Suite 1150</th>
<th>Boston</th>
<th>Massachusetts</th>
<th>02199-8001</th>
</tr>
</thead>
<tbody>
<tr>
<td>Co.2 <strong>Point of Contact</strong></td>
<td>Sheridan</td>
<td>Kassirer, Vice President, Quality Management and Clinical Programs</td>
<td><a href="mailto:eesheppard@partners.org">eesheppard@partners.org</a></td>
<td>617-278-1036</td>
<td></td>
</tr>
</tbody>
</table>

#### Measure Developer if different from Measure Steward

<table>
<thead>
<tr>
<th>Co.3 <strong>Organization</strong></th>
<th>Partners HealthCare System, Inc.</th>
<th>Prudential Tower, 800 Boylston Street, Suite 1150</th>
<th>Boston</th>
<th>Massachusetts</th>
<th>02199-8001</th>
</tr>
</thead>
<tbody>
<tr>
<td>Co.4 <strong>Point of Contact</strong></td>
<td>Jeremiah</td>
<td>Schuur, MD, MHS</td>
<td><a href="mailto:jschuur@partners.org">jschuur@partners.org</a></td>
<td>617-525-8872</td>
<td></td>
</tr>
</tbody>
</table>

| Co.5 **Submitter if different from Measure Steward POC** | Jeremiah | Schuur, MD, MHS | jschuur@partners.org | 617-525-8872 | Partners HealthCare System, Inc. |

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

#### Additional Information

**Workgroup/Expert Panel involved in measure development**

Ad.1 Provide a list of sponsoring organizations and workgroup/panel members’ names and organizations. Describe the members’ role in measure development.

#### Measure Developer/Steward Updates and Ongoing Maintenance

Ad.6 Year the measure was first released: 2009

Ad.7 Month and Year of most recent revision: 2009-12

Ad.8 What is your frequency for review/update of this measure? Every 2 years.

Ad.9 When is the next scheduled review/update for this measure? 2011-12

Ad.10 Copyright statement/disclaimers:

Ad.11-13 Additional Information web page URL or attachment:

**Date of Submission (MM/DD/YY):**
1c. The measure focus is:

- an outcome (e.g., morbidity, mortality, function, health-related quality of life) that is relevant to, or associated with, a national health goal/priority, the condition, population, and/or care being addressed;

OR

- if an intermediate outcome, process, structure, etc., there is evidence that supports the specific measure focus as follows:
  - Intermediate outcome - evidence that the measured intermediate outcome (e.g., blood pressure, Hba1c) leads to improved health/avoidance of harm or cost/benefit.
  - Process - evidence that the measured clinical or administrative process leads to improved health/avoidance of harm and if the measure focus is on one step in a multi-step care process, it measures the step that has the greatest effect on improving the specified desired outcome(s).
  - Structure - evidence that the measured structure supports the consistent delivery of effective processes or access that lead to improved health/avoidance of harm or cost/benefit.
  - Patient experience - evidence that an association exists between the measure of patient experience of health care and the outcomes, values and preferences of individuals/ the public.
  - Access - evidence that an association exists between access to a health service and the outcomes of, or experience with, care.
  - Efficiency - demonstration of an association between the measured resource use and level of performance with respect to one or more of the other five IOM aims of quality.

2d. Clinically necessary measure exclusions are identified and must be:

- supported by evidence of sufficient frequency of occurrence so that results are distorted without the exclusion; AND
- a clinically appropriate exception (e.g., contraindication) to eligibility for the measure focus; AND
- precisely defined and specified:
  - if there is substantial variability in exclusions across providers, the measure is specified so that exclusions are computable and the effect on the measure is transparent (i.e., impact clearly delineated, such as number of cases excluded, exclusion rates by type of exclusion);
  - if patient preference (e.g., informed decision-making) is a basis for exclusion, there must be evidence that it strongly impacts performance on the measure and the measure must be specified so that the information about patient preference and the effect on the measure is transparent (e.g., numerator category computed separately, denominator exclusion category computed separately).
### Measure #/Title/Steward

**#IEP-005-10 Appropriate Pulmonary CT Imaging for Pulmonary Embolism/BWH**

### Description

Percent of patients undergoing CT pulmonary angiogram for the evaluation of possible PE who have a documented indication consistent with guidelines prior to CT imaging.

### Initial In-person Vote

- Recommend for endorsement with conditions – **16**
- Not recommend for endorsement - **3**

### Steering Committee Questions/Conditions for Measure Developer:

<table>
<thead>
<tr>
<th>Question/Condition</th>
<th>Abbreviated Response from Measure Developer</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Potentially eligible for time-limited endorsement, would need to affirm a 12 month testing strategy.</td>
<td>• BWH affirmed that a 12-month testing strategy for this measure has begun.</td>
</tr>
</tbody>
</table>
| • Change measure to become an “Overuse” measure – the denominator would remain unchanged, the numerator would read the low clinical probability and either no high sensitivity D-dimer or a negative high sensitivity D-dimer | • The BWH supports this change – the numerator now reads: “The number of denominator patients with either: a low clinical probability and any negative D-dimer, or an intermediate clinical probability and a negative high-sensitivity D-dimer, or no pretest probability documented.”
  • Developer states that it is important to require a pretest probability score as part of the pre-test assessment, otherwise clinicians who do not assess pretest risk will not be measured. |
| • Provide additional information how the numerator data is captured, specifically at centers other than BWH, where electronic integration of order entry and EMR may not be sufficient to allow automated data collection | • Attached "Sample Data Collection Form for Measure #IEP-005-10.doc" data collection form.
  • See Attachment A
  • Prepared “CT order forms” that are paper based and that EDs without computerized physician order entry (CPOE) for radiology could use to both order the scan and record a pretest probability.
  • See Attachment B1 (high sensitivity) and B2 (low sensitivity) |
| • Consider changing the title of the measure, removing potentially negative connotations | • Suggested new title: "Pulmonary CT Imaging for Patients at low risk for Pulmonary Embolism" |

### Detailed Response from Measure Developer:

- We affirm that we have begun a 12-month testing strategy for this measure.
- The numerator now reads: “The number of denominator patients with either: a low clinical probability and any negative D-dimer, or an intermediate clinical probability and a negative high-sensitivity D-dimer, or no pretest probability documented.” We believe that it is important to require a pretest probability score as part of the pre-test assessment, otherwise clinicians who do not assess pretest risk will
not be measured.

- Additionally, we recommend that the NQF request a tracking measure for this measure that is “Percent of CT scans performed by pre-test risk category.” The measure would calculate the percent of patients getting a CT for PE that were pre-test probability low, medium and high. This is important as the current measure is open to gaming. Clinicians can use an unstructured pretest assessment to determine pretest probability (as recommended by the current evidence based guidelines) but this also means clinicians can choose to assign a pre-test risk of high to all patients that they decide to scan.

- Tracking the proportion of pretest probabilities is one way to monitor such decisions. Please see the attached "Sample Data Collection Form for Measure #IEP-005-10.doc" This is a data collection form. We will also prepare a “CT order form” that is paper based and that EDs without computerized physician order entry (CPOE) for radiology could use to both order the scan and record a pretest probability. Suggested new title: "Pulmonary CT Imaging for Patients at low risk for Pulmonary Embolism".