Evaluation of Pulmonary Embolism in the Emergency Department and Consistency With a National Quality Measure: Quantifying the opportunity for improvement

Arjun K. Venkatesh, MD, MBA, Jeffrey A. Kline, MD, D. Mark Courtney, MD, Carlos A. Camargo Jr, MD, DrPH, Michael C. Plewa, MD, Kristen E. Nordenholz, MD, Christopher L. Moore, MD, Peter B. Richman, MD, MBA, Howard A. Smithline, MD, Daren M. Beam, MD, MS, and Christopher Kabrhel, MD, MPH

Brigham and Women’s Hospital-Massachusetts General Hospital–Harvard Affiliated Emergency Medicine Residency, and Harvard Medical School, Boston (Drs Venkatesh, Camargo, and Kabrhel); Department of Emergency Medicine, Carolinas Medical Center, Charlotte, North Carolina (Drs Kline and Beam); Department of Emergency Medicine, Northwestern University, Chicago, Illinois (Dr Courtney); Department of Emergency Medicine, Mercy St. Vincent Medical Center, Toledo, Ohio (Dr Plewa); Department of Emergency Medicine, University of Colorado, Denver (Dr Nordenholz); Department of Emergency Medicine, Yale University Medical Center, New Haven, Connecticut (Dr Moore); Department of Emergency Medicine, Mayo Clinic Arizona, Scottsdale (Dr Richman); and Department of Emergency Medicine, Baystate Medical Center, Springfield, Massachusetts (Dr Smithline).

Abstract

Background—The National Quality Forum (NQF) has endorsed a performance measure designed to increase imaging efficiency for the evaluation of pulmonary embolism (PE) in the emergency department (ED). To our knowledge, no published data have examined the effect of patient-level predictors on performance.

Methods—To quantify the prevalence of avoidable imaging in ED patients with suspected PE, we performed a prospective, multicenter observational study of ED patients evaluated for PE from 2004 through 2007 at 11 US EDs. Adult patients tested for PE were enrolled, with data collected in a standardized instrument. The primary outcome was the proportion of imaging that was potentially avoidable according to the NQF measure. Avoidable imaging was defined as imaging in a patient with low pretest probability for PE, who either did not have a D-dimer test ordered or who had a negative D-dimer test result. We performed subanalyses testing alternative pretest
probability cutoffs and imaging definitions on measure performance as well as a secondary analysis to identify factors associated with inappropriate imaging. \(\chi^2\) Test was used for bivariate analysis of categorical variables and multivariable logistic regression for the secondary analysis.

**Results**—We enrolled 5940 patients, of whom 4113 (69%) had low pretest probability of PE. Imaging was performed in 2238 low-risk patients (38%), of whom 811 had no D-dimer testing, and 394 had negative D-dimer test results. Imaging was avoidable, according to the NQF measure, in 1205 patients (32%; 95% CI, 31%-34%). Avoidable imaging owing to not ordering a D-dimer test was associated with age (odds ratio [OR], 1.15 per decade; 95% CI, 1.10-1.21). Avoidable imaging owing to imaging after a negative D-dimer test result was associated with inactive malignant disease (OR, 1.66; 95% CI, 1.11-2.49).

**Conclusions**—One-third of imaging performed for suspected PE may be categorized as avoidable. Improving adherence to established diagnostic protocols is likely to result in significantly fewer patients receiving unnecessary irradiation and substantial savings.

Approximately 120 million patients present each year to US emergency departments (EDs),\(^1\) of whom 1.5% undergo computed tomography (CT) of the pulmonary arteries (CTPA) to evaluate for pulmonary embolism (PE).\(^2\) Despite evidence that structured diagnostic pathways can safely exclude PE without imaging, there has been poor application of clinical decision rules in the ED.\(^3\)-\(^6\) As a result, imaging for PE may be overused,\(^4,5,7-9\) with potentially negative cost and health consequences.\(^10-13\)

In 2011 the National Quality Forum (NQF) endorsed an imaging efficiency measure directed at the appropriateness of CTPA use in ED patients with low pretest probability (PTP) of PE.\(^14,15\) Based on retrospective data, the NQF estimated that 7% to 25% of imaging studies are avoidable.\(^14\) However, prospective data are required to assess imaging appropriateness and therefore, the potential for performance improvement. The goal of this study was to quantify the “performance gap” based on the NQF measure.

**Methods**

**Design and Setting**

We analyzed a prospective, multicenter, observational study of ED patients undergoing testing for suspected PE. Patients were enrolled at 12 US hospitals (10 academic and 2 community hospitals) from July 2003 through March 2007. The methods of this study have been described previously.\(^2\) The institutional review board of each institution approved the protocol.

**Population**

Eligible ED patients had an order for an objective diagnostic test (D-dimer, CTPA, ventilation/perfusion scan [V/Q], or pulmonary angiogram), written under supervision of a board-certified emergency physician to evaluate possible PE.\(^2\) Patients were enrolled consecutively or during randomly assigned shifts representative of all ED shifts.\(^2\) For this analysis, we included patients enrolled at hospitals using a high-sensitivity, quantitative D-dimer and with multidetector spiral CT available to the ED. Cutoffs used to define positive tests were institutionally determined, as previously described.\(^16\)

After a diagnostic test for PE was ordered, but before results were known, we prospectively collected 74 data points, including the objective elements of the Wells score by interviewing the patient and reviewing the medical record.\(^17\) We also asked the clinician ordering the initial trigger test to provide their most likely diagnosis and gestalt PTP. We classified gestalt PTP as low (<15%), medium (15%-40%), or high (>40%). This was a
noninterventional study, so all diagnostic decisions were made by the treating physician with data available in the ED.

We excluded patients in whom the treating physician had knowledge of a positive imaging study for PE within 7 days and patients being evaluated for deep vein thrombosis without PE.

Outcomes
The primary outcome was avoidable imaging, defined as either CTPA or V/Q in hemodynamically stable (systolic blood pressure ≥90 mm Hg) patients with low PTP (Wells score <2),\(^\text{18}\) and in whom D-dimer testing was either not done or the D-dimer result was negative. While the NQF measure is specific to CTPA use, we included both CTPA and V/Q in the primary outcome since both expose patients to ionizing radiation and incur health care costs.

Subanalyses
We measured the impact of loosening the PTP cutoff to define D-dimer testing as appropriate for patients classified as (1) unlikely to have PE by the modified Wells score (≤4 points)\(^\text{19}\) and (2) patients with low or intermediate PTP (<6 points), in accordance with clinical guidelines.\(^\text{6,20}\) We also quantified avoidable imaging using the physician's gestalt PTP. Finally, since the NQF-endorsed measure focuses on CTPA only, we performed a subanalysis to quantify avoidable CTPA (but not V/Q) imaging.

Secondary Analysis
To assess the association between patient-level predictors and potentially avoidable imaging, we analyzed a logistic regression model using NQF-defined avoidable imaging as the outcome variable.

Statistical Analysis
Baseline characteristics are presented as means, medians, and binomial proportions with 95% CIs. We performed logistic regression to identify variables associated with avoidable imaging. Our model included the D-dimer assay used, variables associated with avoidable imaging (P < .20) on univariate analysis, and variables previously associated with positive D-dimer test results.\(^\text{16}\) Statistical significance was defined by odds ratios (ORs) for which the 95% CI did not cross unity. Analysis was performed using SAS software, version 9.2 (SAS Institute Inc).

Result
We enrolled 6089 patients of whom 5940 were hemodynamically stable. Their mean (SD) age was 47 (17) years; 3966 (67%) were female; 3552 (60%) were white; 121 (31%), black; 23 (6%), Hispanic; and 9 (3%), other race; and 482 (8%) were uninsured. The Wells score classified 4113 patients (69%) as having low PTP (<2 points), 1634 (28%) as having intermediate PTP (2-6 points), and 193 (3%) as having high PTP (>6 points). A D-dimer test was performed in 4263 patients: Liatest (Diagnostica Stago) (1440 [33.8%]), VIDAS (BioMerieux) (1231 [28.9%]), MDA (BioMerieux) (704 [16.5%]), Advanced (Dade Behring) (452 [10.6%]), HemosIL (Instrumentation Laboratory) (329 [7.7%]), and Minutex (Trinity Biotech) (107 [2.5%]).

A total of 3849 patients (65% of enrolled) underwent imaging, of whom 139 were excluded from the measure for hemodynamic instability. Of the remaining 3710 patients, 2238 (54%) had low PTP (Wells score <2 points).
Of patients who had imaging, 1205 (32%) met the NQF definition of avoidable imaging (Figure): 811 (67%) because no D-dimer testing was performed and 394 (33%) because imaging was performed despite a negative D-dimer test result. Results of subanalyses are shown in Table 1. For alternative PTP cutoffs (Wells scores of <4 and <6) and when PTP was determined by clinical gestalt, about one-third of imaging studies were potentially avoidable. In two-thirds of cases, this was because no D-dimer testing was performed. Results were similar when we limited our analysis to CTPA use.

Fifty patients (1.3% of those imaged) were diagnosed as having PE by imaging considered potentially avoidable by the NQF measure because no D-dimer testing was performed, whereas 8 (0.2%) were diagnosed as having PE by imaging considered potentially avoidable in patients with a negative D-dimer test result.

Our multivariable model identified few patient-level predictors of avoidable imaging (Table 2): older age (OR, 1.09; 95% CI, 1.04-1.13) and inactive cancer (OR, 1.48; 95% CI, 1.13-1.95). When we limited our analysis to patients whose imaging was avoidable because no D-dimer testing was performed, older age (OR, 1.15; 95%, CI, 1.10-1.21), sickle cell disease (OR, 4.84; 95% CI, 1.81-12.95), and pregnancy (OR, 1.96; 95% CI, 1.29-2.99) were associated with avoidable imaging. When we limited our analysis to patients whose imaging was avoidable because imaging was performed despite a negative D-dimer test result, only inactive cancer (OR, 1.66; 95% CI, 1.11-2.49) was associated with avoidable imaging.

Comment

Rising health care costs from advanced imaging and improved understanding of the risks associated with ionizing radiation and intravenous contrast are driving efforts to improve imaging efficiency. To our knowledge, our analysis is the first to quantify imaging appropriateness for PE, we found that 1 in 3 ED imaging studies (33%) was potentially avoidable. This “performance gap” persisted whether we assessed all imaging or CTPA alone and when PTP cutoffs were varied.

Our study is unique in that our prospective assessment of PTP, and our enrollment of patients regardless of imaging use enabled us to assess imaging appropriateness, rather than imaging utilization, as estimated by previous retrospective studies. Failure to perform D-dimer testing was responsible for nearly two-thirds of potentially avoidable imaging studies. This phenomenon may be explained by physician bias toward more “definitive” testing with CTPA, use of CTPA for evaluation of possible non-PE alternative diagnoses, overestimation of expected D-dimer testing false-positivity, or underestimation of D-dimer testing sensitivity. It could also be explained by physician overestimation of gestalt PTP. However, our subanalysis does not support this.

We found few patient-level predictors of avoidable imaging. This suggests that imaging inefficiency is likely to be more related to variation in physician-level risk tolerance, patient preference, or hospital characteristics not measured by our study. Future work in this area will be valuable. Age, pregnancy, and sickle cell disease, factors known to be associated with positive D-dimer test results, were predictive of avoidable imaging owing to no D-dimer test being performed. Clinicians may have bypassed D-dimer testing in these patients, anticipating positive results. Further research is needed to determine whether patients in whom D-dimer specificity is low may be suitable for measure exclusion. Imaging performed following a negative D-dimer test result in patients with inactive cancer may represent an opportunity for quality improvement, because clinicians may overestimate the risk of PE in these patients.
Our results demonstrate the validity of the NQF measure and refute the notion that high measure performance is associated with the unintended consequence of missed PE. Assuming 100% imaging specificity, measure adherence would have resulted in 11 “missed” PEs: 8 patients with a negative D-dimer test result and 3 patients who would have undergone D-dimer testing [93% sensitivity] testing according to the guideline.22

Several considerations in the interpretation of our results warrant mention. Consistent with the NQF guideline, we report potentially avoidable, as opposed to definitely avoidable, imaging for patients with no D-dimer testing performed. All patients were enrolled in hospitals participating in PE research. Such hospitals may be more likely to follow clinical guidelines, so our results may underestimate avoidable imaging. We also recognize that some CTPA may have been ordered to evaluate alternative diagnoses in addition to PE.

In summary, we found that one-third of ED imaging studies for suspected PE are potentially avoidable. The opportunity for improving the efficiency of imaging for suspected PE is large. Future work should focus on interventions to close this performance gap.

Acknowledgments

Funding/Support: This study was funded by the National Institutes of Health grant No. 2R42HL074415. Role of the Sponsor: The sponsor played no role in the design or conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript.

References

Figure.
Application of the National Quality Forum pulmonary embolism (PE) imaging efficiency measure. Pretest probability is based on the Wells score: low pretest probability, less than 2 points; intermediate pretest probability, 2 to 6 points; high pretest probability, more than 6 points. All percentages were calculated using the number of hemodynamically stable (systolic blood pressure ≥90 mm Hg) patients who underwent workup for PE with imaging as the denominator.
Table 1

Primary, Subanalyses, and Sensitivity Analyses of NQF Measure Performance\(^a\)

<table>
<thead>
<tr>
<th>Pretest Probability Assessment</th>
<th>All Imaging</th>
<th>CTPA Imaging Only</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Potentially AI</td>
<td>AI Owing to No D-Dimer Testing Done</td>
</tr>
<tr>
<td>NQF measure: (Wells score &lt;2)</td>
<td>1205 (32)</td>
<td>811 (22)</td>
</tr>
<tr>
<td>95% CI</td>
<td>31-34</td>
<td>21-23</td>
</tr>
<tr>
<td>Modified Wells score: (Wells score &lt;4)</td>
<td>1589 (43)</td>
<td>1098 (30)</td>
</tr>
<tr>
<td>95% CI</td>
<td>41-44</td>
<td>28-31</td>
</tr>
<tr>
<td>Clinical guideline (Wells score &lt;6)</td>
<td>1977 (33)</td>
<td>1410 (24)</td>
</tr>
<tr>
<td>95% CI</td>
<td>32-35</td>
<td>23-25</td>
</tr>
<tr>
<td>Clinical gestalt: low(^b)</td>
<td>1090 (29)</td>
<td>740 (20)</td>
</tr>
<tr>
<td>95% CI</td>
<td>28-31</td>
<td>19-21</td>
</tr>
</tbody>
</table>

Abbreviations: AI, avoidable imaging; CTPA, computed tomography of the pulmonary arteries; NQF, National Quality Forum.

\(^a\)Data are given as number (percentage). Percentages apply to the number of all hemodynamically patients undergoing imaging, in accordance with the NQF measure.

\(^b\)Does not include 5 patients for whom gestalt pretest probability assessment was not available.
### Table 2
Multivariate Analysis of Predictors of Avoidable Imaging (AI) for Pulmonary Embolism in Patients With Low Pretest Probability

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>All AI</th>
<th>AI Owing to No D-Dimer Test Done</th>
<th>AI Owing to Negative D-Dimer Test Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients, No.</td>
<td>1205</td>
<td>811</td>
<td>394</td>
</tr>
<tr>
<td>Age, by decade, y</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;30</td>
<td>1 [Reference]</td>
<td>1 [Reference]</td>
<td>1 [Reference]</td>
</tr>
<tr>
<td>30-40</td>
<td>1.15 (0.90-1.48)</td>
<td>1.16 (0.85-1.58)</td>
<td>1.09 (0.74-1.61)</td>
</tr>
<tr>
<td>41-50</td>
<td>1.60 (1.25-2.05)</td>
<td>1.51 (1.11-2.05) &lt;sup&gt;b&lt;/sup&gt;</td>
<td>1.54 (1.06-2.24)</td>
</tr>
<tr>
<td>51-60</td>
<td>1.60 (1.24-2.08) &lt;sup&gt;b&lt;/sup&gt;</td>
<td>1.56 (1.14-2.16) &lt;sup&gt;b&lt;/sup&gt;</td>
<td>1.48 (0.99-2.18)</td>
</tr>
<tr>
<td>61-70</td>
<td>1.54 (1.16-2.04) &lt;sup&gt;b&lt;/sup&gt;</td>
<td>1.69 (1.20-2.37) &lt;sup&gt;b&lt;/sup&gt;</td>
<td>1.20 (0.77-1.87)</td>
</tr>
<tr>
<td>71-80</td>
<td>1.77 (1.31-2.42) &lt;sup&gt;b&lt;/sup&gt;</td>
<td>2.36 (1.65-3.37) &lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.77 (0.45-1.35)</td>
</tr>
<tr>
<td>&gt;80</td>
<td>1.54 (1.09-2.17) &lt;sup&gt;b&lt;/sup&gt;</td>
<td>2.27 (1.54-3.36) &lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.45 (0.21-0.95)</td>
</tr>
<tr>
<td>Race/ethnicity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>1 [Reference]</td>
<td>1 [Reference]</td>
<td>1 [Reference]</td>
</tr>
<tr>
<td>Black</td>
<td>0.83 (0.70-0.97)</td>
<td>0.83 (0.69-1.01)</td>
<td>0.84 (0.66-1.07)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.88 (0.67-1.17)</td>
<td>0.86 (0.62-1.19)</td>
<td>0.99 (0.63-1.55)</td>
</tr>
<tr>
<td>Other</td>
<td>0.69 (0.45-1.06)</td>
<td>0.73 (0.44-1.21)</td>
<td>0.68 (0.34-1.36)</td>
</tr>
<tr>
<td>Immobility</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No immobility</td>
<td>1 [Reference]</td>
<td>1 [Reference]</td>
<td>1 [Reference]</td>
</tr>
<tr>
<td>Generalized</td>
<td>0.52 (0.37-0.72)</td>
<td>0.54 (0.37-0.79)</td>
<td>0.56 (0.31-1.02)</td>
</tr>
<tr>
<td>Limb</td>
<td>0.45 (0.24-0.83)</td>
<td>0.62 (0.32-1.21)</td>
<td>0.15 (0.02-1.06)</td>
</tr>
<tr>
<td>Neurological</td>
<td>0.75 (0.31-1.83)</td>
<td>1.41 (0.56-3.58)</td>
<td>NP</td>
</tr>
<tr>
<td>Cancer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No cancer history</td>
<td>1 [Reference]</td>
<td>1 [Reference]</td>
<td>1 [Reference]</td>
</tr>
<tr>
<td>Active cancer&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.67 (0.51-0.89)</td>
<td>0.86 (0.64-1.16)</td>
<td>0.33 (0.17-0.65)</td>
</tr>
<tr>
<td>Inactive cancer</td>
<td>1.48 (1.13-1.95) &lt;sup&gt;b&lt;/sup&gt;</td>
<td>1.31 (0.95-1.80)</td>
<td>1.66 (1.11-2.49) &lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Surgery</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No surgery</td>
<td>1 [Reference]</td>
<td>1 [Reference]</td>
<td>1 [Reference]</td>
</tr>
<tr>
<td>Abdominal</td>
<td>0.91 (0.59-1.39)</td>
<td>1.15 (0.73-1.82)</td>
<td>0.44 (0.16-1.20)</td>
</tr>
<tr>
<td>Chest</td>
<td>0.85 (0.46-1.56)</td>
<td>0.95 (0.49-1.84)</td>
<td>0.57 (0.14-2.35)</td>
</tr>
<tr>
<td>Orthopedic</td>
<td>1.08 (0.62-1.88)</td>
<td>1.50 (0.82-2.72)</td>
<td>0.21 (0.03-1.57)</td>
</tr>
<tr>
<td>Other</td>
<td>1.03 (0.29-3.56)</td>
<td>0.68 (0.08-5.44)</td>
<td>1.44 (0.32-6.51)</td>
</tr>
<tr>
<td>Warfarin use</td>
<td>0.81 (0.59-1.17)</td>
<td>0.79 (0.55-1.14)</td>
<td>0.89 (0.52-1.54)</td>
</tr>
<tr>
<td>Pregnancy</td>
<td>1.36 (0.94-1.98)</td>
<td>1.96 (1.29-2.99) &lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.54 (0.24-1.19)</td>
</tr>
<tr>
<td>Sickle cell disease</td>
<td>2.07 (0.84-5.12)</td>
<td>4.84 (1.81-12.95) &lt;sup&gt;b&lt;/sup&gt;</td>
<td>NP</td>
</tr>
<tr>
<td>Connective tissue disease</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No CTD</td>
<td>1 [Reference]</td>
<td>1 [Reference]</td>
<td>1 [Reference]</td>
</tr>
<tr>
<td>Lupus</td>
<td>1.44 (0.86-2.40)</td>
<td>1.61 (0.91-2.85)</td>
<td>1.06 (0.45-2.47)</td>
</tr>
</tbody>
</table>

<sup>a</sup> OR and 95% CI are reported. <sup>b</sup> P < 0.05.
<table>
<thead>
<tr>
<th>Characteristic</th>
<th>All AI</th>
<th>AI Owing to No D-Dimer Test Done</th>
<th>AI Owing to Negative D-Dimer Test Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rheumatoid</td>
<td>0.87 (0.53-1.43)</td>
<td>0.91 (0.51-1.62)</td>
<td>0.88 (0.38-2.04)</td>
</tr>
<tr>
<td>Other CTD</td>
<td>1.40 (0.62-3.16)</td>
<td>1.26 (0.47-3.39)</td>
<td>1.33 (0.40-4.48)</td>
</tr>
</tbody>
</table>

Abbreviations: CTD, connective tissue disease; NP, no patients in these categories; OR, odds ratio.

a. Low pretest probability is consistent with the definition used in the NQF-endorsed measure (Wells score <2)
b. Statistically significant based on 95% CIs not crossing unity.
c. Active cancer includes malignant disease being actively treated or palliated.