



Published in final edited form as:

*J Am Coll Radiol.* 2019 August ; 16(8): 1064–1072. doi:10.1016/j.jacr.2018.12.015.

## An evaluation of guideline discordant ordering behavior for CT pulmonary angiography in the emergency department

Emma Simon<sup>1,2</sup>, Isomi M. Miake-Lye, PhD<sup>2,3</sup>, Silas W. Smith, MD<sup>4,5</sup>, Jordan L. Swartz, MD, MA<sup>4</sup>, Leora I. Horwitz, MD, MHS<sup>1,2,6</sup>, Danil V. Makarov, MD, MHS<sup>1,7,8</sup>, Soterios Gyftopoulos, MD, MS<sup>9,10</sup>

<sup>1</sup>NYU School of Medicine, Department of Population Health, New York, NY

<sup>2</sup>NYU School of Medicine, Center for Healthcare Innovation and Delivery Science, New York, NY

<sup>3</sup>VA Greater Los Angeles Healthcare System, Los Angeles, CA

<sup>4</sup>NYU School of Medicine, Ronald O. Perelman Department of Emergency Medicine, New York, NY

<sup>5</sup>NYU Langone Health, Institute for Innovations in Medical Education, New York, NY

<sup>6</sup>NYU School of Medicine, Department of Medicine, New York, NY

<sup>7</sup>NYU School of Medicine, Department of Urology, New York, NY

<sup>8</sup>VA New York Harbor Healthcare System, New York University School of Medicine, New York, NY

<sup>9</sup>NYU School of Medicine, Department of Radiology, New York, NY

<sup>10</sup>NYU School of Medicine, Department of Orthopedic Surgery, New York, NY

### Abstract

**Purpose:** To determine rates of and possible reasons for guideline discordant ordering of CT pulmonary angiography (CTPA) for the evaluation of suspected pulmonary embolism (PE) in the emergency department (ED).

**Methods:** We performed a retrospective review of 212 consecutive encounters (01/06/2016–02/25/2016) with 208 unique patients in the emergency department that resulted in a CTPA order. For each encounter, we calculated the revised Geneva score and two versions of the Wells criteria. We then classified each encounter using a two-tiered risk stratification method (PE unlikely vs. PE likely). Finally, we assessed the rate of and possible explanations for guideline discordant ordering via in-depth chart review.

**Results:** Frequency of guideline discordant studies ranged from 53 (25%) to 79 (37%), depending on the scoring system used; a total of 46 (22%) of which were guideline discordant under all three scoring systems. Of these, 18 (39%) had at least one patient-specific factor

---

**Corresponding Author:** Emma Simon, 227 E 30<sup>th</sup> Street, 6<sup>th</sup> Floor, New York, NY 10016, *Phone:* 646-501-2917, emma.simon@nyumc.org.

**Author Contributions:** LIH, DVM, and SG conceived the study and obtained research funding. LIH, DVM, and SG supervised the conduct of the study. ES, SS, and JS collected and managed the data, including quality control. ES and IML analyzed the data. ES drafted the manuscript, and all authors contributed substantially to its revision. ES takes responsibility for the paper as a whole.

associated with increased risk for PE, but not included in the risk stratification scores (e.g. travel, thrombophilia).

**Conclusions:** Many of the guideline discordant orders were placed for patients who presented with evidence-based risk factors for PE that are not included in the risk stratification scores. Therefore, guideline discordant ordering may indicate that, in the presence of these factors, the assessment of risk made by current scoring systems may not align with clinical suspicion.

### Summary Sentence

Through analysis of CTPA ordering in our institution's emergency department, we found that many guideline discordant orders were placed for patients who presented with clinically-reasonable, evidence-based risk factors for PE that are not included in the Wells and revised Geneva risk stratification scores.

### Keywords

Pulmonary embolism; guideline discordance; guideline concordance; computed tomography pulmonary angiography; evidence-based medicine

---

### Introduction

Pulmonary embolism (PE) is an emergent condition that affects 1 in 1000 people in the United States each year, causing as many as 200,000 deaths annually [1, 2]. Its morbidity and treatability make PE a “can’t miss” diagnosis, often driving emergency department (ED) providers to include PE in their initial differential diagnosis.

Imaging is the most accurate tool for identifying PE. The current gold standard imaging examination used to diagnose PE is CT pulmonary angiography (CTPA) [2, 3]. In recent years, the use of CTPAs has increased dramatically, leading multiple specialty societies, including the American Thoracic Society, American College of Chest Physicians, American College of Emergency Physicians, and the American College of Radiology have selected the reduction of unnecessary CTPAs as a priority [4, 5]. Though CTPA use provides the benefit of clear diagnosis, overuse exposes patients unnecessarily to risks associated with imaging (i.e. radiation, IV contrast administration, incidental findings, costs, and anxiety) without any potential benefit [6, 7].

A number of tools have been introduced to aid providers in diagnosing PE. Validated tools such as the Wells criteria and revised Geneva score stratify patients by level of risk [8–11]. In addition, when used in conjunction with these scores, the D-dimer test can be employed to avoid CT scanning in low and some intermediate risk patients [8, 10]. The utility of these tools has led to their widespread adoption in clinical guidelines for PE management [2]. Some debate still exists, however, about the performance of objective risk stratification tools versus provider gestalt (an unstructured assessment of pretest probability based on clinical experience and suspicion), with some studies reporting superior diagnostic outcomes using the latter [2, 12].

In previous work, several groups have retrospectively assessed the rate of guideline discordant CTPA ordering in their institutions' emergency departments [13, 14]. However, while these studies reported discordant ordering rates, they did not attempt to determine why providers were not following established guidelines. In prior work, we conducted qualitative interviews to determine provider opinions on CTPA utilization and guideline discordant ordering behavior [15]. In this study, utilizing insights from these interviews, we sought to not only determine our institution's rate of guideline discordance, but also to identify potential reasons for discordant imaging. To do so, we performed in-depth case reviews of the discordant cases to pinpoint factors that may have affected the provider's confidence in the established guidelines. Such analyses are important to create a more accurate picture of the current state of CTPA ordering and to inform efforts by providers and policy makers to decrease rates of guideline discordance going forward.

## Methods

### Study Design and Setting

We performed a retrospective review of ED encounters to determine rates of yield and guideline discordant ordering of CTPA. All patients were seen in the ED of an academic medical center in New York City, which has 726 inpatient beds, 29 ED beds and treats approximately 70,000 patients annually. CTPA orders were placed by 41 resident physicians and 35 physician assistants who were supervised in decision making by 63 attending physicians. This study was approved by the hospital's institutional review board. All medical records were reviewed under a waiver of consent and a waiver of HIPAA authorization.

### Selection of Participants

We analyzed all ED encounters with CTPA studies ordered on adult patients (≥ 18 years of age) for the purpose of evaluating a suspected PE over a 7 week period (01/06/2016–02/25/2016). We excluded encounters in which the CTPA was performed after the patient had been admitted to an inpatient unit or for which charts could not be accessed post-hoc (only paper charts were used due to unexpected EHR downtime).

### Methods and Measurements

The chart review was performed by one member of the study team (ES). A data dictionary and standard method of abstraction were compiled prior to review. Data were then abstracted for variables as specified in this data dictionary (Appendix A.1.). For any questions that arose during the abstraction process, a board-certified, emergency medicine physician (JLS) was consulted. To determine the reliability of the data abstracted, 20 charts (approximately 10% of the sample) from the validation cohort underwent a blinded, secondary review by a board-certified, emergency medicine physician (SWS).

Data were abstracted solely from the clinical notes and results affiliated with each ED encounter from the EHR (Epic Systems Corporation, Verona, WI). We abstracted for age, sex, race, the Wells and revised Geneva score risk factors (Table 1), travel history, thrombophilia status (e.g. Factor V Leiden, prothrombin gene mutation), relevant laboratory and radiological results (D-dimer assay, CTPA and ventilation-perfusion scan), mention of

previous visits (to the ED or to an outside MD) for similar symptoms in the ED note, and post-ED disposition (admitted as an inpatient, placed in observation status, or discharged). In cases of contradictory documentation, the variable was documented as positive. In cases of missing data, the variable was documented as negative. We then used these data to calculate the Wells criteria and revised Geneva score for each patient.

Due to the partially subjective nature of the Wells criteria, which includes points for the category “pulmonary embolism as likely as or more likely than an alternative diagnosis,” we chose to calculate the score by two different methods [16]. The first method, “automatic Wells,” assumed that pulmonary embolism was as likely as or more likely than an alternative diagnosis for all patients, thereby giving every patient at least 3 points regardless of patient presentation and past medical history. This method effectively afforded the provider the “benefit of the doubt” in their ordering practices and maximized the impact of provider gestalt [17]. In the second method, “non-automatic Wells,” we awarded points discerningly depending upon the provider’s differential diagnosis noted in the chart. For example, if the provider mentioned PE in the differential diagnosis, we awarded the three points associated with this category. However, if the provider did not mention PE in the differential diagnosis, we did not award the three points. The full criteria used for this discerning scoring are shown in Appendix A.2. We awarded points associated with all other categories according to definitions established by the original Wells criteria validation study [8, 9].

Unlike the Wells criteria, the revised Geneva score employs entirely objective data to evaluate risk level [11]. Therefore, we calculated only one version of the revised Geneva score according to the definitions established by the revised Geneva score validation study [10].

Though both risk stratification scores can be stratified by a two-tiered (PE unlikely or PE likely) and three-tiered model (low, intermediate, or high), we chose to stratify patients by the two-tiered model because of the consensus within the ED community regarding the D-dimer’s ability to sufficiently exclude PE unlikely patients (NPV = 99%) [2]. Therefore, we grouped patients into PE unlikely and PE likely categories based on their calculated point value [18, 19].

Our institution uses the HemosIL D-dimer HS (Instrumentation Laboratory, Orangeburg, NY). The units of measurement are ng/mL D-dimer units (DDU) with a cut-off of 230 ng/mL DDU. D-dimer levels below this cut-off are deemed “normal” and those above are deemed “elevated.” CTPAs ordered from the ED at our institution are imaged using the 384 (2 × 192) slice Somatom Force, a Dual Source CT Scanner (Siemens Healthcare, Malvern, PA), and the ISOVUE-300 contrast agent (60 cc Iopromide and 300 mg Iodine/mL solution injected; Bracco Imaging, Milan, Italy). The mean CT dose index per scan was 7.81 mGy and the mean dose length product was 209.36 mGy cm.

## Outcomes

We assessed three main outcomes: yield, rate of guideline discordant ordering, and explanations for guideline discordant ordering [2]. We defined yield as the percentage of completed imaging examinations that demonstrated a PE of any size. According to accepted

recommendations, we defined guideline concordant ordering as a CTPA performed on a patient who was either PE likely or was PE unlikely with an elevated D-dimer result [2, 18, 19]. Conversely, guideline discordant ordering was defined as a CTPA performed on a PE unlikely patient who had either a normal D-dimer result or had not received a D-dimer before proceeding to scan [2, 18, 19]. This definition, however, contains two classes of guideline discordant orders that are qualitatively different. We therefore chose to categorize these two classes separately. We defined a CTPA performed on a PE unlikely patient with a normal D-dimer result as “avoidable” because of confidence in the medical literature that PE can be reliably ruled out in such cases [2]. In contrast, we defined a CTPA performed on a PE unlikely patient with no D-dimer result, without which we lack a component needed to determine if the patient was truly at low risk for PE, as “potentially avoidable.”

Finally, for those studies determined to be potentially avoidable by all three scoring systems, we sought to determine potential reasons why a provider might have chosen to forego the D-dimer. We used Virchow’s triad (hypercoagulability, hemodynamic changes, and endothelial injury), prior emergency medicine provider qualitative interviews and prior PE literature to identify factors that may be associated with increased risk for PE, but are not included in the Wells and revised Geneva scores [20–22]. Despite being absent from the risk stratification scores, the presence of these factors at patient presentation could elevate provider gestalt, resulting in a divergent risk assessment that could explain guideline discordant ordering behavior. For example, we reviewed for the presence of thrombophilia, leading to a hypercoagulable state, and travel > 4 hours, leading to a hemodynamic change [21]. We additionally took note of any factors within the patient chart specifically denoted by the provider to have elevated their gestalt. A full list of these factors can be found in Appendix A.3.

## Statistical Analysis

We reported cohort characteristics and distribution of risk scores using standard descriptive statistics. We conducted blinded double-abstractions of 20 randomly sampled encounters (approximately 10% of the sample) to calculate the inter-rater reliability for our review. This was done by comparing the concordance of the blinded abstractions in their determination of the two-tier risk classification (PE unlikely vs. PE likely) for each of the three scoring methods.

## Results

### Characteristics of Study Subjects

During the study period, a total of 254 ED encounters resulted in a completed CTPA. We excluded 42 of these encounters from in-depth review: 40 were excluded in which the CTPA was performed after the patient had been admitted to an inpatient unit and 2 were excluded due to EHR downtime which led to use of paper charts that could not be accessed post-hoc (n=2). In total, we included 212 encounters (208 unique patients) in the study.

The mean age of the patients was 58 years (range: 18 to 96). Of the total, 137 (65%) patients were female. Patients self-identified as Asian (n=9, 4.2%), black (n=31, 14.6%), white

(n=132, 62.3%), and other (n=37, 17.5%), and 3 (1.4%) did not specify. Additionally, for ethnicity, patients self-identified as of Spanish/Hispanic origin (n=12, 5.7%) and not of Spanish/Hispanic origin (n=71, 33.5%), and 129 (60.8%) did not specify. The most common risk factors observed were elevated heart rate, age, active malignancy, and previous history of deep vein thrombosis (DVT) or PE. The demographics of the patient encounters are shown in Table 2.

## Main Results

Interrater reliability tests resulted in an overall kappa value of .94. Individual kappa values were also calculated for each of the scores: non-automatic Wells had a kappa value of .80 and both automatic Wells and revised Geneva had kappa values of 1.

A total of 18 scans were positive for PE (8.5%; CI: 4.7–12.3). Other relevant diagnoses included pulmonary nodules (n=47, 22.2%), pleural effusion (n=30, 14.2%), pneumonia (n=27, 12.7%), COPD/emphysema (n=19, 9.0%), lung cancer (n=15, 7.1%), interstitial lung disease (n=2, 0.9%), and congestive heart failure (n=1, 0.5%). For patients categorized as PE unlikely by the automatic Wells criteria, non-automatic Wells criteria, and revised Geneva score 3/89 (3.4%), 6/138 (4.3%), and 7/119 (5.9%) were positive for PE, respectively. In contrast, for patients categorized as PE likely by the automatic Wells criteria, non-automatic Wells criteria, and revised Geneva score 15/123 (12.2%), 12/74 (16.2%), and 11/93 (11.8%) were positive for PE, respectively.

Overall rates of guideline discordant studies (including both avoidable and potentially avoidable orders) ranged from 25.0% to 37.3%. Using the automatic Wells criteria, 53 (25.0%; CI: 19.3–31.1) studies were guideline discordant. Of these studies, 4 were in PE unlikely patients with a normal D-dimer result and therefore classified as avoidable and the remaining 49 studies were ordered prior to a resulted D-dimer and were therefore classified as potentially avoidable (Figure 1). Using the non-automatic Wells criteria, 79 (37.3%; CI: 30.7–43.9) studies were guideline discordant (74 were potentially avoidable and 5 were avoidable). Using the revised Geneva score, 73 (34.4%; CI: 28.3–41.0) studies were guideline discordant (67 were potentially avoidable and 6 were avoidable).

A total of 46 (21.7%) studies were guideline discordant under all of the three scoring systems. Of these, 4 (8.7%) studies were deemed avoidable; none were positive for PE. Case summaries for these 4 studies can be found in Appendix B. The other 42 (91.3%) were found to be potentially avoidable; 2 (4.8%) were positive for PE. Of the 42, 18 (42.9%) had at least one factor that could elevate provider gestalt but is not included in the risk stratification scores. These risk factors fell into three categories: hypercoagulability (i.e. pregnancy or postpartum state, exogenous estrogen, cocaine use, dialysis-dependent chronic kidney disease, rheumatologic diseases, malignancy with no treatment within 1 year, thrombophilia, and family history of PE), hemodynamic changes (travel > 4 hours) or provider-acknowledged gestalt-elevating factors (i.e. patient history of non-VTE thrombosis and patent foramen ovale) (Table 3) [20–22].

## Discussion

In this study, we found that between 25–37% of CTPA studies were not ordered in accordance with specialty society guidelines. Of these, 46 (22%) were determined to be guideline discordant under all three scoring systems used. A total of 4 (9%) of these cases were performed in PE unlikely patients with a normal D-dimer result, which we considered to be clearly avoidable. As expected, none of these patients was diagnosed with a PE. Most guideline discordance, however, was due to providers proceeding directly to CTPA without prior D-dimer in PE unlikely patients (42 of the 46 cases). Had a D-dimer been obtained prior to the CTPA order in those patients, it is probable that some would have been elevated, making the CTPA order guideline concordant. Accordingly, we characterized these as *potentially* avoidable cases and analyzed provider rationale behind these orders via in-depth chart review. Our analysis of these potentially avoidable studies showed that 18 (43%) had at least one characteristic associated with increased risk for PE, but not included in the risk stratification scores, which could have elevated provider gestalt. Furthermore, 2 (5%) of these studies were positive for PE. Therefore, in a subset of such cases, it might in fact have been appropriate to proceed directly to CTPA without D-dimer testing.

Prior work has reported rates of guideline discordant imaging at other institutions to be similar to that found in this study [13, 14]. However, our study extends this work by distinguishing between different types of guideline discordance. In addition, our study addresses the challenge of the subjective component of the Wells criteria, “PE as likely as or more likely than an alternative diagnosis” [8]. Most retrospective studies have used the Wells criteria to determine rates of guideline discordance without addressing this subjective component [13, 14]. In a retrospective review, this component presents a particular challenge because only the provider notes can be used to determine the rationale upon which the points awarded for this category depend. To provide a rate unaffected by the potential bias added by subjective points, we calculated the revised Geneva score, a fully objective risk stratification tool. Additionally, we utilized a scoring system that decreased bias in our evaluation of the subjective category itself. One prior study employed a “conservative” and “liberal” approach in calculating the Wells criteria—with the liberal approach more easily awarding points to the provider in their assessment of risk [16]. However, both the conservative and liberal approaches required discretion on the part of the reviewer; consequently, points for the subjective category were only awarded part of the time in both their conservative and liberal approaches [16]. We employed a similar two-pronged approach for calculating the Wells criteria. We differ from this study in that only one of our methods required the discretion of the reviewer to award subjective points while the other awarded them regardless of clinical context. Through these methods, we improved inter-rater reliability, gave maximal benefit of the doubt to providers, and aimed to report a more accurate range of the potential boundaries of guideline discordance.

Understanding the rationale for provider ordering is crucial for effective design of future interventions aimed at reducing unnecessary CTPA testing. By heeding factors that may elevate provider gestalt but are left out of the decision scores, we acknowledge the well-founded reliance on gestalt in clinical decision making [12, 15, 23]. By identifying these additional risk factors through a two-pronged approach of qualitative interviews conducted

by our own study team and a literature search of the work of others, we aimed to incorporate a level of nuance that assigning binary categorizations of risk lacks [20, 22]. These stratification scores provide an excellent approximation of risk, but they cannot and are not meant to capture the full clinical picture—particularly when utilized in retrospective review [9, 11]. For example, our results suggest that simply requiring D-dimer testing prior to CTPA in patients deemed PE unlikely according to standard scoring systems, as some have proposed, would not match the provider’s mental model of the patient’s risk nearly half of the time. In fact, previous studies have found that, while some EM providers are generally aware of guidelines for PE workup, they deliberately elect not to follow them [15, 23]. Furthermore, gestalt has been shown to perform better than both Wells and revised Geneva in a clinical setting—with Wells, the partially subjective score, performing better than revised Geneva, which is completely objective [12]. Therefore, the more that provider opinion is taken into account, the more accurate risk assessment becomes. In such cases of discord between assessments made by the standard scoring systems and a provider’s gestalt, providers are likely to ignore guidance, limiting the value of decision support or education and potentially biasing providers against attending to guidance even for clearly lower risk patients. Many clinical decision support tools have failed because of a similar mismatch between computer-generated and provider-generated risk assessments [13, 14].

Consequently, despite attempts at intervention, utilization remains high, yield has not significantly improved and guideline discordance is consistently problematic at both our institution and across the country [13, 14]. Though our list of additional gestalt-elevating risk factors is far from complete, considering the complexity of this clinical pathway and the decisions made within it provides a more accurate picture of clinical practice. Looking forward, we hope to apply this nuanced approach from our analysis to the development of quality improvement interventions such as audit and feedback and clinical decision support. For example, a more effective decision tool to reduce CTPA testing might calculate the patient’s risk and then query the provider for agreement with risk level before providing a recommendation. By focusing interventions more on the difficulties involved in these decisions, rather than further simplification of an inherently complex issue, it is possible that we will see an improvement in outcomes.

Through analysis of CTPA ordering in our institution’s emergency department, we found that many guideline discordant orders were placed on patients who presented with gestalt-elevating, evidence-based risk factors for PE that are not included in the risk stratification scores. Therefore, guideline discordant ordering may indicate that, in the presence of these factors, the assessment of risk made by current scoring systems and clinical suspicion can differ. In the future, we hope to incorporate this approach into the creation of quality improvement interventions to decrease guideline discordance at our institution.

## Limitations

There are several limitations to this study. This study was conducted retrospectively using only the information provided in the EHR. Retrospective calculation of these scores may be inaccurate because of a lack of clinical context. We tried to minimize potential bias in several ways. Firstly, we calculated multiple scores: one with a subjective component (Wells



criteria) and one without (revised Geneva score). Secondly, we calculated the subjective component of the Wells criteria in two different ways: one that automatically awarded the associated points, thereby giving the providers the benefit of the doubt, and another that discerningly awarded the associated points, thereby possibly providing a more accurate picture of rates of potentially avoidable imaging. Lastly, we calculated interrater reliability to ensure the accuracy of the review. Another limitation to our study is that we examined only patients who underwent CTPA. Patients in whom PE might have been suspected but who, following guidelines, did not receive a CTPA were not included. This is due to the fact that a complete cohort of patients with suspected PE but who did not receive a CTPA cannot be readily identified in the EHR. Consequently, the true incidence of guideline discordance is almost certainly lower than the numbers we report. Finally, ours is a single institution study, which may limit external validity, although the prevalence of PE and rates of guideline discordant ordering are within the range seen in previous studies [13, 14].

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Funding:

This work was funded by the Agency for Healthcare Research and Quality (P30HS024376). DVM is a VA HSR&D Career Development awardee at the Manhattan VA Hospital (CDA11-257 & CDP 11-254). SWS is supported by grant funding from the U.S. Army Medical Research and Materiel Command (Project number DM160044) to New York University School of Medicine. No funding source had any role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the article for publication. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Agency for Healthcare Research and Quality, the Manhattan VA Hospital, or the US Army Medical Research and Materiel Command.

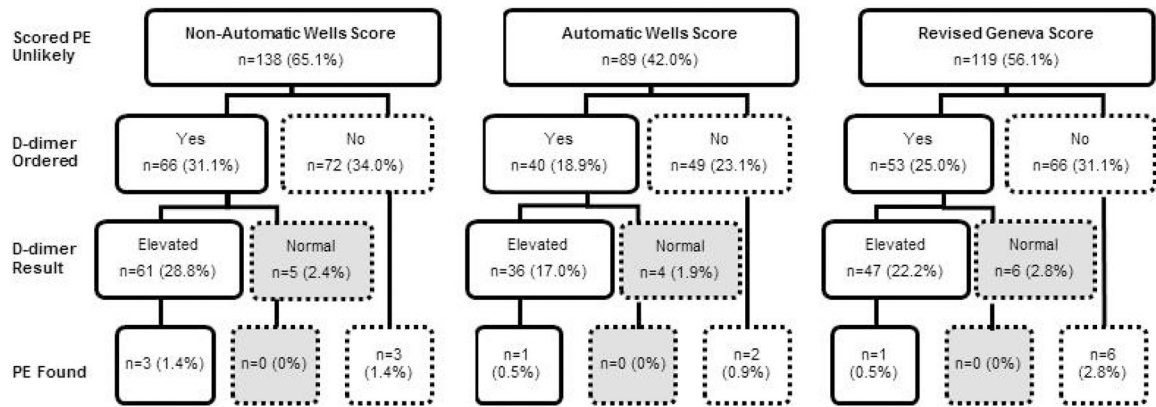
## References

1. Horlander KT, Mannino DM, Leeper KV. Pulmonary embolism mortality in the United States, 1979–1998: an analysis using multiple-cause mortality data. *Archives of internal medicine*. 2003;163(14):1711–7. [PubMed: 12885687]
2. Fesmire FM, Brown MD, Espinosa JA, Shih RD, Silvers SM, Wolf SJ, et al. Critical issues in the evaluation and management of adult patients presenting to the emergency department with suspected pulmonary embolism. *Annals of emergency medicine*. 2011;57(6):628–52. e75. [PubMed: 21621092]
3. Anderson DR, Kahn SR, Rodger MA, et al. Computed tomographic pulmonary angiography vs ventilation-perfusion lung scanning in patients with suspected pulmonary embolism: A randomized controlled trial. *JAMA*. 2007;298(23):2743–53. [PubMed: 18165667]
4. Wiener RS, Ouellette DR, Diamond E, Fan VS, Maurer JR, Mularski RA, et al. An official American Thoracic Society/American College of Chest Physicians policy statement: the Choosing Wisely top five list in adult pulmonary medicine. *CHEST Journal*. 2014;145(6):1383–91.
5. Physicians ACoE. ACEP - CT Pulmonary Angiography in ED Patients: Choosing Wisely; 2014 [Available from: <http://www.choosingwisely.org/clinician-lists/acep-ct-pulmonary-angiography-in-ed-patients/>].
6. Brenner DJ, Hall EJ. Computed Tomography — An Increasing Source of Radiation Exposure. *New England Journal of Medicine*. 2007;357(22):2277–84. [PubMed: 18046031]
7. Cochran ST, Bomyea K, Sayre JW. Trends in adverse events after IV administration of contrast media. *American Journal of Roentgenology*. 2001;176(6):1385–8. [PubMed: 11373197]

8. Wells PS, Anderson DR, Rodger M, et al. Excluding pulmonary embolism at the bedside without diagnostic imaging: Management of patients with suspected pulmonary embolism presenting to the emergency department by using a simple clinical model and d-dimer. *Annals of Internal Medicine*. 2001;135(2):98–107. [PubMed: 11453709]
9. Wells PS, Anderson DR, Rodger M, Ginsberg JS, Kearon C, Gent M, et al. Derivation of a simple clinical model to categorize patients probability of pulmonary embolism-increasing the models utility with the SimpliRED D-dimer. *THROMBOSIS AND HAEMOSTASIS-STUTTART*. 2000;83(3):416–20.
10. Righini M, Le Gal G, Aujesky D, Roy P-M, Sanchez O, Verschuren F, et al. Diagnosis of pulmonary embolism by multidetector CT alone or combined with venous ultrasonography of the leg: a randomised non-inferiority trial. *The Lancet*. 2008;371(9621):1343–52.
11. Le Gal G, Righini M, Roy PM, Sanchez O, Aujesky D, Bounameaux H, et al. Prediction of pulmonary embolism in the emergency department: The revised Geneva score. *Annals of Internal Medicine*. 2006;144(3):165–71. [PubMed: 16461960]
12. Penalzoza A, Verschuren F, Meyer G, Quentin-Georget S, Soulie C, Thys F, et al. Comparison of the unstructured clinician gestalt, the wells score, and the revised Geneva score to estimate pretest probability for suspected pulmonary embolism. *Annals of emergency medicine*. 2013;62(2):117–24. e2. [PubMed: 23433653]
13. Goldzweig CL, Orshansky G, Paige NM, Miake-Lye IM, Beroes JM, Ewing BA, et al. Electronic health record-based interventions for improving appropriate diagnostic imaging: a systematic review and meta-analysis. *Ann Intern Med*. 2015;162(8):557–65. [PubMed: 25894025]
14. Wang RC, Bent S, Weber E, Neilson J, Smith-Bindman R, Fahimi J. The Impact of Clinical Decision Rules on Computed Tomography Use and Yield for Pulmonary Embolism: A Systematic Review and Meta-analysis. *Ann Emerg Med*. 2016;67(6):693–701.e3. [PubMed: 26747217]
15. Gyftopoulos S, Smith SW, Simon E, Kuznetsova M, Horwitz LI, Makarov DV. Qualitative Study to Understand Ordering of CT Angiography to Diagnose Pulmonary Embolism in the Emergency Room Setting. *Journal of the American College of Radiology*.
16. Tarr GP, Modahl L, Jones P. Wells score, D-dimer testing and computer tomographic pulmonary angiography appropriateness in the Auckland Hospital Adult Emergency Department. *N Z Med J*. 2015;128(1413):81–3.
17. Booker MT, Johnson JO. Optimizing CT Pulmonary Angiogram Utilization in a Community Emergency Department: A Pre- and Postintervention Study. *J Am Coll Radiol*. 2017;14(1):65–71. [PubMed: 27717576]
18. van Belle A, Buller HR, Huisman MV, Huisman P, Kaasjager K, Kamphuisen P, et al. Effectiveness of managing suspected pulmonary embolism using an algorithm combining clinical probability, D-dimer testing, and computed tomography. *Jama*. 2006;295(2):172–9. [PubMed: 16403929]
19. Douma RA, Mos IM, Erkens PG, et al. Performance of 4 clinical decision rules in the diagnostic management of acute pulmonary embolism: A prospective cohort study. *Annals of Internal Medicine*. 2011;154(11):709–18. [PubMed: 21646554]
20. Gyftopoulos S, Smith SW, Simon E, Kuznetsova M, Horwitz LI, Makarov DV. Qualitative Study to Understand Ordering of CT Angiography to Diagnose Pulmonary Embolism in the Emergency Room Setting. *Journal of the American College of Radiology*. 2017.
21. Tapson VF. Acute Pulmonary Embolism. *New England Journal of Medicine*. 2008;358(10):1037–52. [PubMed: 18322285]
22. Kabrhel C, Mark Courtney D, Camargo CA, Plewa MC, Nordenholz KE, Moore CL, et al. Factors Associated With Positive D-dimer Results in Patients Evaluated for Pulmonary Embolism. *Academic Emergency Medicine*. 2010;17(6):589–97. [PubMed: 20624138]
23. Lin MP, Nguyen T, Probst MA, Richardson LD, Schuur JD. Emergency Physician Knowledge, Attitudes, and Behavior Regarding ACEP’s Choosing Wisely Recommendations: A Survey Study. *Academic Emergency Medicine*. 2017.

**Take-Home Points**

- Overall frequency of guideline discordant CTPAs ordered for adult patients in the ED ranged from 52 (25%) to 79 (37%). A total of 46 (22%) studies were guideline discordant under all three scoring systems.
- Of these 46 studies, 18 (39%) were placed for patients who presented with evidence-based risk factors for PE that are not included in the risk stratification scores.
- Therefore, guideline discordant ordering may indicate that the assessment of risk made by current scoring systems may not align with clinical suspicion.



**Figure 1.** Guideline discordance of CTPA orders by risk stratification score. Boxes with a dashed outline represent guideline discordant studies. Boxes shaded in grey represent avoidable studies.

**Table 1.**

Risk stratification specifications for the Wells and Revised Geneva Scores

Wells Criteria[9]		Revised Geneva Score[11]	
<i>Risk Factors</i>	<i>Points</i>	<i>Risk Factors</i>	<i>Points</i>
Clinical signs and symptoms of deep venous thrombosis (objectively measured leg swelling and pain with palpation in the deep-vein region)	3.0	Unilateral lower limb pain	3
		Pain on leg deep-vein palpation and unilateral edema	4
Heart rate greater than 100 beats/min	1.5	Heart rate between 75–94 beats/min OR Heart rate ≥95 beats/min	3 OR 5
Immobilization (bedrest, except access to the bathroom, for ≥3 days) or surgery in the previous 4 weeks	1.5	Surgery (under general anesthesia) or fracture (of the lower limbs) within 1 month	2
Previous objectively diagnosed deep venous thrombosis or pulmonary embolism	1.5	Previous deep-venous thrombosis or pulmonary embolism	3
Hemoptysis	1.0	Hemoptysis	2
Malignancy (patients with cancer who were receiving treatment, those in whom treatment had been stopped within the past 6 months, or those who were receiving palliative care)	1.0	Active malignancy (solid or hematologic malignant condition, currently active or considered cured <1 year)	2
Pulmonary embolism as likely as or more likely than an alternative diagnosis*	3.0	--	--
--	--	Age >65 years	1
<b>Clinical Probability</b>			
<i>PE unlikely</i>	4	<i>PE unlikely</i>	5
<i>PE likely</i>	>4	<i>PE likely</i>	>5

\* This is the category subject to change in the automatic vs. non-automatic scoring versions. Points are automatically awarded for this category under automatic Wells and discerningly awarded under non-automatic Wells

**Table 2.**

Study participant demographics (n=212 encounters)

Characteristic	Number of Encounters (%)
<b>General Demographics</b>	
Average age in years; range	58;18–96
<b>Sex</b>	
Male	75 (35.0)
Female	137 (65.0)
<b>Race</b>	
Asian	9 (4.2)
Black	31 (14.6)
Other	37 (17.5)
White	132 (62.3)
Unspecified	3 (1.4)
<b>Ethnicity</b>	
Not of Spanish/Hispanic Origin	71 (33.5)
Spanish/Hispanic Origin	12 (5.7)
Unspecified	129 (60.8)
<b>Wells Score Risk Factors</b>	
Clinical signs and symptoms of DVT	9 (4.2)
Heart rate > 100 beats/min	75 (35.4)
Immobilization or surgery in the previous 4 weeks	19 (9.0)
History of DVT or PE	36 (17.0)
Hemoptysis	6 (2.8)
Active malignancy	43 (20.3)
<b>Revised Geneva Score Risk Factors</b>	
Unilateral lower limb pain	12 (5.7)
Pain on deep vein palpation and unilateral edema	9 (4.2)
Heart rate between 75–94 beats/min	75 (35.4)
Heart rate ≥ 95 beats/min	100 (47.2)
Surgery or lower limb fracture within 1 month	17 (8.0)
History of DVT or PE	36 (17.0)
Hemoptysis	6 (2.8)
Active malignancy	43 (20.3)
Age >65 years	83 (39.2)

**Table 3.**

Gestalt-elevating risk factors documented in potentially avoidable CTPA encounters

Factor Type	Reason Seen	Count of Times Factor was Present*
Hypercoagulability	Malignancy (last treatment >1 year prior)	6
	Thrombophilia	3
	Rheumatologic disease	2
	Pregnancy or postpartum state	1
	Exogenous estrogen	1
	Cocaine use	1
	Dialysis-dependent chronic kidney disease	1
	Family history of PE	1
Hemodynamic Changes	Travel 4 hours	7
Provider-acknowledged Gestalt-elevating Factors	Patient history of non-VTE thrombosis	2
	Patent foramen ovale	1

\* These categories were not mutually exclusive; there are 18 encounters represented in the 26 factors listed above due to patients presenting with two or more factors at once