Prospective Validation of Wells Criteria in the Evaluation of Patients With Suspected Pulmonary Embolism

Stephen J. Wolf, MD
Tracy R. McCubbin, MD
Kim M. Feldhaus, MD
Jeffrey P. Faragher, MD
Dorothy M. Adcock, MD

From the Department of Emergency Medicine, Denver Health Medical Center (Wolf, Feldhaus, Faragher); the Department of Emergency Medicine, Kaiser Permanente/Exempla St. Joseph Hospital (McCubbin); and Esoterix, Inc. (Adcock), Denver, CO.

Study objective: The literature suggests that the D-dimer is useful in patients suspected of having pulmonary embolism and who have a low pretest probability of disease. A previously defined clinical decision rule, the Wells Criteria, may provide a reliable and reproducible means of determining this pretest probability. We evaluate the interrater agreement and external validity of Wells Criteria in determining pretest probability in patients suspected of having pulmonary embolism.

Methods: This was a prospective observational study. Trained research assistants enrolled patients during 120 random 8-hour shifts. Patients who underwent imaging for pulmonary embolism after a medical history, physical examination, and chest radiograph were enrolled. Treating providers and research assistants determined pretest probability according to Wells Criteria in a blinded fashion. Two D-dimer assays were run. Three-month follow-up for the diagnosis of pulmonary embolism was performed. Interrater agreement tables were created. κ Values, sensitivities, and specificities were determined.

Results: Of the 153 eligible patients, 3 patients were missed, 16 patients declined, and 134 (88%) patients were enrolled. Sixteen (12%) patients were diagnosed with pulmonary embolism. The κ values for Wells Criteria were 0.54 and 0.72 for the trichotomized and dichotomized scorings, respectively. When Wells Criteria were trichotomized into low pretest probability (n=59, 44%), moderate pretest probability (n=61, 46%), or high pretest probability (n=14, 10%), the pulmonary embolism prevalence was 2%, 15%, and 43%, respectively. When Wells Criteria were dichotomized into pulmonary embolism–unlikely (n=88, 66%) or pulmonary embolism–likely (n=46, 34%), the prevalence was 3% and 28%, respectively. The immunoturbidimetric and rapid enzyme-linked immunosorbent assay D-dimer assays had similar sensitivities (94%) and specificities (45% versus 46%).

Conclusion: Wells Criteria have a moderate to substantial interrater agreement and reliably risk stratify pretest probability in patients with suspected pulmonary embolism.

[*Ann Emerg Med.* 2004;44:503-510.]

0196-0644/$30.00

doi:10.1016/j.annemergmed.2004.04.002

Editor's Capsule Summary

What is already known on this topic

The Wells Criteria have been proposed as a clinical prediction rule that can be used to determine the pretest probability of pulmonary embolism in emergency department patients.

What question this study addressed

This study was designed to evaluate the interrater agreement and externally validate the ability of the Wells Criteria to prospectively determine pretest probability of pulmonary embolism.

What this study adds to our knowledge

In 134 patients prospectively studied with 3-month follow-up, there was moderate to substantial interrater agreement in the evaluation of the criteria. Low prior probability patients by Wells Criteria had a 2% incidence of pulmonary embolism, moderate 15%, and high 43%. D-Dimer screening performed similarly to the Wells Criteria.

How this might change clinical practice

Clinicians can have increased confidence in the reproducibility and reliability of the Wells Criteria in predicting the pretest probability of pulmonary embolism in emergency patients.

INTRODUCTION

The role of the D-dimer assay in evaluating patients with suspected pulmonary embolism remains unclear to many physicians. Many clinicians are uncertain about the differences between the various types of D-dimer assays and the appropriate clinical setting in which to use them. Recent reviews[1-6] and published studies[7-13] have shown that the sensitivities and negative likelihood ratios of the rapid enzyme-linked immunosorbent assay and the immunoturbidimetric D-dimer assays are the most promising for the emergency department (ED) setting. However, use of these assays alone is insufficient to exclude the diagnosis of pulmonary embolism in all patients.[2,4-7] Several authors suggest that the D-dimer assay has the greatest utility in the patient with a low pretest probability.[2,3,5-8]

A renewed interest in pretest probability has developed in recent years, leading investigators to develop more reliable means of determining pretest probability using clinical decision rules. Several such rules have been recently published.[14-19] Unfortunately, many of these rules are complicated and difficult to apply, especially in a busy ED environment.[14,15,17,18] In recent years, Wells et al[14,19,20] published, refined, and then internally validated a clinical decision rule for pretest probability of pulmonary embolism.

Wells Criteria for pretest probability of pulmonary embolism consist of the 7 following weighted criteria: (1) clinical signs and symptoms of deep venous thrombosis (+3.0); (2) an alternative diagnosis that is less likely than pulmonary embolism (+3.0); (3) pulse rate greater than 100 beats/min (+1.5); (4) immobilization or surgery in the previous 4 weeks (+1.5); (5) previous deep venous thrombosis/pulmonary embolism (+1.5); (6) hemoptysis (+1.0); and (7) malignancy (on treatment, treated in the past 6 months, or palliative; +1.0). Summation of these point values can be trichotomized into low (<2), moderate (2 to 6), or high (>6) pretest probability with prevalences for pulmonary embolism of 2% to 4%, 19% to 21%, and 50% to 67%, respectively. Alternatively, the total score can be dichotomized into a pulmonary embolism–unlikely (≤4) or a pulmonary embolism–likely (>4) pretest probability with prevalences for pulmonary embolism of 5% to 8% and 39% to 41%, respectively.[14,19]

The studies by Wells et al[14,19,20] suggest that for a patient with a low pretest probability, as defined by Wells Criteria, and a negative whole-blood D-dimer assay, the diagnosis of pulmonary embolism can be reliably excluded. External, retrospective validation of Wells Criteria's ability to reliably risk stratify patients according to pulmonary embolism prevalence has also recently been published.[21] However, the interrater agreement and validity of Wells Criteria have not been reported independently and prospectively.

We therefore conducted a study with 2 objectives: (1) to evaluate the interrater agreement of the Wells Criteria; and (2) to externally validate the ability of the Wells Criteria to determine pretest probability prospectively.

MATERIALS AND METHODS

Study Design

This was a prospective, observational study of patients presenting to the ED with suspected pulmonary embolism and who underwent diagnostic imaging for pulmonary embolism. The study hospital's institutional review board approved the study protocol, and written informed consent was obtained from all study participants.

Setting

The study was conducted from August 2001 through June 2002 at an emergency medicine residency–affiliated, community-based ED, serving a predominantly managed-care patient population (Kaiser Permanente). This ED has an annual census of approximately 55,000 and performs an estimated 1,500 radiographic evaluations for pulmonary embolism annually, about 1 pulmonary embolism workup per 8-hour shift.

Data Collection

Before the data collection phase of our study, all departmental emergency care providers (emergency physicians, emergency medicine residents, and physician's assistants) participated in a training session about the proper application of Wells Criteria. In addition, all research assistants (emergency physicians, emergency medicine residents, medical students, and nurses) were trained in a 2-hour session in data collection and the application of Wells Criteria. Data were collected during 120 randomly generated 8-hour periods, representing all hours of the day and all days of the week, throughout a 6-month period. During these shifts, the ED was staffed with a research assistant to perform data collection. Any shift that could not be filled at the originally assigned time block was moved to the same time block and day of the week of the nearest following week.

All patients presenting during a study shift who were referred for imaging for pulmonary embolism, after the emergency care provider's history taking, physical examination, chest radiograph, and ECG, were considered eligible. Eligibility and subsequent enrollment were completed before the D-dimer assays or the diagnostic imaging was obtained. Patients were excluded if they were non–English speaking, recently (<6 months before) or currently pregnant, morbidly obese (>350 lb), diagnosed with a previous genetic clotting disorder, younger than 18 years or older than 85 years, critically ill or unable to consent, or known to have a recently elevated or normal D-dimer assay result. Exclusion criteria 2 to 5 were established because of potential limitations of the D-dimer assays[22-24] and because of the weight limitations of the table of the computed tomographic (CT) scanner used. The purpose of the last exclusion criterion was to exclude patients referred from outlying clinics for evaluation of a positive D-dimer. Eligible patients who consented for participation underwent an interview by the research assistant, who collected demographic data and historical information and performed a physical examination.

All blood samples were collected by standard venipuncture into glass evacuated tubes containing 3.2% sodium citrate and centrifuged within 1 hour of collection. Plasma aliquots were frozen within 2 hours of collection at −70°C (–94°F) in polypropylene tubes after a second centrifugation at 1500*g* for 10 minutes. An automated immunoturbidimetric D-dimer assay (Liatest d-di; Diagnostica Stago, Parsippanny, NJ) was performed on the fresh plasma within 1 hour of centrifugation, according to the manufacturer's instructions. Thawed frozen samples were evaluated within 3 months using a rapid enzyme-linked immunosorbent assay D-dimer assay (VIDAS D-dimer; bioMérieux, Marcy L'etoile, France), following the manufacturer's instructions. The cutoff values to determine "positive" or "negative" D-dimer values were 400 ng/mL fibrinogen equivalent units for the immunoturbidimetric assay and 500 ng/mL for the automated enzyme-linked immunosorbent assay.[25-27]

The research assistant and the treating provider evaluated patients for Wells Criteria. They were blinded to each other's Wells Criteria results, the D-dimer values, and the results of any diagnostic radiographic evaluation. The patient then underwent imaging for pulmonary embolism according to a clinical diagnostic algorithm used at the study institution. Patients with signs of deep venous thrombosis underwent lower-extremity duplex ultrasonography first. Otherwise, patients underwent either a ventilation/perfusion lung scan (if their chest radiograph was interpreted as normal) or a helical CT angiogram of the chest (if the chest radiograph was interpreted as abnormal). An abnormal chest radiograph was defined as one demonstrating congestive heart failure, pleural effusion, parenchymal infiltrate, or frank bullae suggesting significant chronic obstructive pulmonary disease.

Because the immunoturbidimetric D-dimer assay was in current use by the study hospital, the results were disclosed to the treating provider after completion of Wells Criteria. The rapid enzyme-linked immunosorbent assay results were not available to the treating care provider at the initial ED evaluation.

All enrolled patients not diagnosed with pulmonary embolism on their initial evaluation were followed up with a 3-month telephone interview. Contacted patients were asked whether they had been diagnosed with a blood clot since their initial ED visit and, if so, where and when. A positive response was confirmed in the Kaiser Permanente medical records. Patients who could not be reached by telephone for the follow-up interview after 3 attempts had a review of their Kaiser Permanente and hospital medical record and the hospital's anticoagulation clinic enrollment for evidence of pulmonary embolism as a proxy for the interview. Kaiser Permanente maintains an electronic medical record that is comprehensive and includes clinic, emergency, and hospital visits. Patient demographics are continually updated. Any hospital, ED, or clinic visit to any local area hospital is logged into the electronic medical record system during authorization for the visit. If there was no entry after the initial ED encounter, the Kaiser Permanente database was queried to confirm that Kaiser Permanente insurance was still active.

A research assistant interviewed each enrolled patient, obtaining historical and physical data. Wells Criteria were determined and recorded by the research assistant and the treating provider. Wells Criteria were then dichotomized to yield a pulmonary embolism–likely or pulmonary embolism–unlikely pretest probability, as well as trichotomized to yield a low, moderate, or high pretest probability.

The diagnosis of pulmonary embolism was made when any of the following criteria were met: (1) high-probability ventilation/perfusion scan using modified Prospective Investigation of Pulmonary Embolism Diagnosis (PIOPED) study criteria; (2) contrast-enhanced CT scan of the chest diagnostic for pulmonary embolism (General Electric single detector helical scanner, General Electric CTI, General Electric, Fairfield, CT); (3) intermediate probability ventilation/perfusion scan with a high pretest clinical suspicion as determined by the treating provider; (4) pulmonary angiogram diagnostic for pulmonary embolism; (5) follow-up telephone interview; or (6) medical record review documenting diagnosis of pulmonary embolism or deep venous thrombosis. Patients not diagnosed with a pulmonary embolism were considered to be ruled out for the diagnosis of pulmonary embolism.

Primary Data Analysis

Statistical analysis was performed by an independent statistician using SPSS software (version 10.1, SPSS, Inc., Chicago, IL). Confidence intervals (CIs) were calculated using Confidence Interval Analysis (version 1.0, CIA, BMJ Publishing Group, London, England). Interrater agreement tables were created. κ Values were calculated for Wells Criteria as a measure of agreement and were interpreted according to a traditionally accepted scale, with 0.41 to 0.60, 0.61 to 0.80, and 0.81 to 1.0 being considered moderate, substantial, and excellent agreement, respectively.[28]

RESULTS

One hundred seventy-six patients who presented to the ED during the 120 study shifts met our enrollment criteria. Demographic data are listed in Table 1. Of these patients, 23 (23/176, 13%) were excluded for the following reasons: non–English speaking (2); pregnant at the time of the study (1); pregnant within the previous 6 months (3); morbidly obese (3); age older than 85 years (5); age younger than 18 years (2); unable to consent or critically ill (3); and known to have a recently elevated or normal D-dimer assay result (6). Two excluded patients met more than 1 exclusion criterion. Of the 153 (153/176, 87%) patients who met the inclusion criteria, 16 (16/153, 10%) declined to participate and 3 (3/153, 2%) were missed because of high volume in the ED, leaving 134 of the eligible patients to be enrolled (134/153, 88%). Despite being enrolled in the study, 17 patients (17/134, 13%) were discharged without an imaging procedure for pulmonary embolism after the treating provider received negative results of an immunoturbidimetric D-dimer. The remaining 117 (117/134, 87%) patients underwent imaging for pulmonary embolism on their initial ED visits. One patient had severely hemolyzed blood samples, and because of concern that this might compromise results, neither D-dimer assay was performed. Although this patient's data were not used for statistical calculations involving the D-dimer assays, they were used in the determination of interrater agreement and external validation of Wells Criteria. Three-month follow-up data were obtained on 112 (112/120, 93%) enrolled patients not diagnosed with pulmonary embolism during their initial visit. Of these, 14 (14/112, 13%) patients had the proxy follow-up because they could not be reached on the telephone. Eight (8/120, 7%) patients were lost to follow-up. The Figure depicts patient flow and follow-up in the study.

Fourteen patients (14/134, 10%) were diagnosed with pulmonary embolism (pulmonary embolism–positive) on their initial ED evaluation. Two patients (2/112, 2%) were found to have venous thromboembolism on 3-month follow-up (1 deep venous thrombosis, 1 pulmonary embolism). The overall venous thromboembolism prevalence for the study was 12% (16/134). All of the 17 patients who received no initial imaging study and had a negative D-dimer result had no venous thromboembolism on direct telephone follow-up. The 8 patients who were lost to follow-up but had negative imaging on initial presentation were considered pulmonary

Table 1.
Demographic characteristics of enrolled patients.

Demographic Characteristic	No. (n=134)	%
Male sex	61	46
Median age, y	58	
Interquartile range	(43–72)	
Race/ethnicity		
White	99	74
Hispanic	19	14
Black	11	8
Asian/Pacific Islander	2	2
Other	3	2

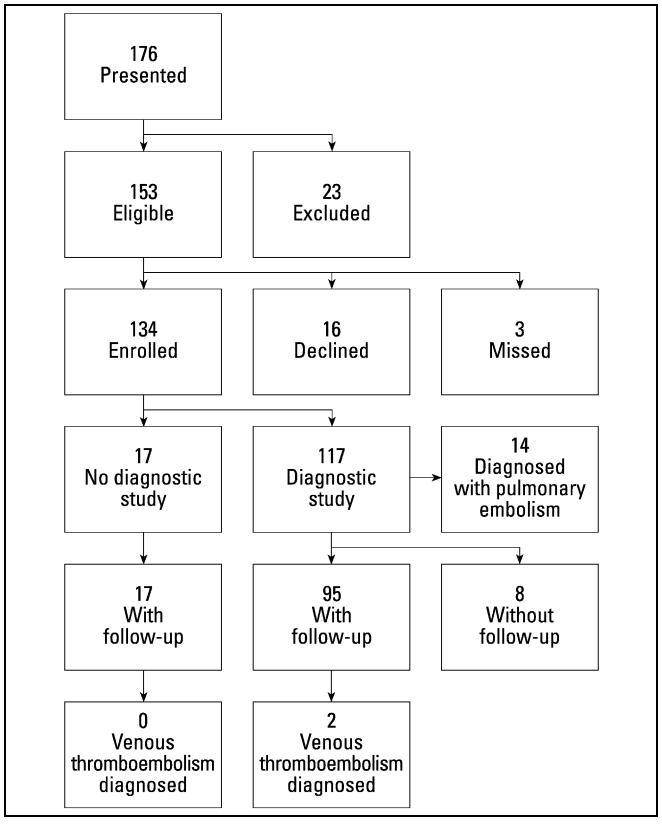
embolism–negative. Patient classification according to pulmonary embolism diagnosis is listed in Table 2.

The patient diagnosed with a pulmonary embolism on follow-up had a Wells Criteria score of 7.5, an immunoturbidimetric D-dimer value of 2,610, and a low-probability ventilation/perfusion scan with emphysema on initial ED evaluation and was discharged; pulmonary embolism was diagnosed by CT scan within 72 hours of the initial visit. The patient diagnosed with a deep venous thrombosis had a Wells Criteria score of 1.0 (for cancer), an immunoturbidimetric D-dimer of 1,330, and a ventilation/perfusion scan read as very low probability on his initial ED visit. The patient was discharged and re-presented 33 days later, when he was diagnosed with the deep venous thrombosis on lower-extremity ultrasonography.

Wells Criteria were applied to all 134 enrolled patients. Tables 3 and 4 depict the interrater agreement tables for each Wells Criteria criterion, Wells Criteria trichotomized scoring system, and Wells Criteria dichotomized scoring system. Both tables show data for all care providers to all research assistants (n=134) and only physician care providers to only physician research assistants (n=79). κ Values are included as a means of measurement for interrater agreement. Table 5 shows the patients' risk stratification as derived by Wells Criteria.

Plasma samples for the immunoturbidimetric and rapid enzyme-linked immunosorbent assay D-dimer samples were collected from all 134 enrolled patients; however, 1 patient's samples were severely hemolyzed, and thus only 133 patient results were considered for data analysis. The sensitivity and specificity were 94% (95% CI 70% to 100%) and 45% (95% CI 36% to 54%) for the immunoturbidimetric assay and 94% (95% CI 70% to 100%) and 46% (95% CI 37% to 55%) for the rapid enzyme-linked immunosorbent assay.

Figure.
Patient flow through follow-up for the study.



LIMITATIONS

Our study has several important limitations. First, our study is limited in that 17 patients did not undergo an initial imaging study for pulmonary embolism because of negative D-dimer results but were considered to be ruled out for a pulmonary embolism according to a negative telephone follow-up. In all of these 17 follow-up instances, direct contact was made with the patients, and a proxy follow-up was unnecessary. One patient with hemolyzed D-dimers and a negative workup and follow-up was also considered

Table 2.
Ultimate diagnoses in 134 enrolled patients.

Diagnostic Modality	No.
Pulmonary embolism–negative (n=118)*	
Normal/very low probability ventilation/perfusion scan	43
Low pretest probability and low probability ventilation/perfusion scan	13
Non-low pretest probability, low probability ventilation/perfusion scan and no pulmonary embolism/deep venous thrombosis on follow-up	11
CTA negative for pulmonary embolism	34
Normal PA	0
No radiographic study in ED and no pulmonary embolism/deep venous thrombosis on follow-up	17
Pulmonary embolism–positive (n=16)	
High probability ventilation/perfusion scan	6
High pretest probability/intermediate probability ventilation/perfusion scan	1
Lower-extremity duplex, ultrasonography diagnostic for deep venous thrombosis with a suspicion for pulmonary embolism	0
CTA diagnostic for pulmonary embolism	7
PA diagnostic for pulmonary embolism	0
Pulmonary embolism/deep venous thrombosis on follow-up	2

CTA, Computed tomography angiogram of the chest; **PA**, pulmonary angiogram.
*Includes 8 patients who were pulmonary embolism–negative on initial ED evaluation but had no follow-up.

to be ruled out for pulmonary embolism. We believed that excluding these patients, for whom there was a possibility of venous thromboembolism on follow-up, would have had a greater potential to distort our findings than including them, with the possibility of a false negative follow-up.

In addition, there were 8 patients who did not receive follow-up. These patients could not be contacted by telephone, and they did not have reliable medical records in the hospital and managed-care system. They were categorized as pulmonary embolism–negative according to their initial ED visit with a negative result for pulmonary embolism. Finally, 3 patients were missed and no data were available on them.

A second limitation is related to our decision to consider pulmonary embolism excluded by a negative CT angiogram of the chest, which continues to be an area of debate in the literature and in EDs across the country. Several studies do suggest that CT angiography can be used as a primary modality of diagnosis of pulmonary embolism and that the false negative rate can be low.[29-33] However, our study was not designed to evaluate the CT angiogram's performance. Of note, both of the pulmonary embolisms diagnosed on follow-up in our study were in patients who underwent an initial ventilation/perfusion scan.

A final limitation is the potential selection bias introduced by including only patients who underwent an

Table 3A.
Interrater agreement table comparing all care providers (n=134).

	Rater 2													
	Criterion 1		Criterion 2		Criterion 3		Criterion 4		Criterion 5		Criterion 6		Criterion 7	
Rater 1	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Yes	17	4	54	17	34	7	15	4	24	1	6	0	9	0
No	5	108	11	52	6	87	1	112	0	109	0	128	3	122
κ (95% CI)	0.76 (0.60–0.92)		0.58 (0.44–0.72)		0.78 (0.66–0.90)		0.79 (0.63–0.95)		0.97 (0.92–1.00)		1.00		0.83 (0.64–1.00)	

Table 3B.
Interrater agreement table comparing physician care providers to physician research assistants (n=79).

	MD 2													
	Criterion 1		Criterion 2		Criterion 3		Criterion 4		Criterion 5		Criterion 6		Criterion 7	
MD 1	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Yes	13	2	35	6	19	2	11	3	14	0	4	0	4	0
No	3	61	8	30	4	54	1	64	0	65	0	75	1	74
κ (95% CI)	0.78 (0.60–0.96)		0.64 (0.46–0.82)		0.81 (0.67–0.96)		0.82 (0.64–0.99)		1.00		1.00		0.85 (0.57–1.00)	

MD, Physician.

Table 4A.
*Trichotomized scoring table comparing all care providers with all care providers.**

	Rater 2		
Rater 1	Low	Moderate	High
Low	40	8	0
Moderate	18	52	5
High	0	2	9

*κ=0.54 (95% CI 0.40–0.68)

Table 4B.
*Trichotomized scoring table comparing physician care providers with physician research assistants.**

	MD 2		
MD 1	Low	Mod	High
Low	23	7	0
Moderate	7	34	2
High	0	0	66

*κ=0.62 (95% CI 0.44–0.80)

imaging study for pulmonary embolism. Enrolling all patients with chest pain, shortness of breath, or syncope would have made this study prohibitively costly, and would also have resulted in a lower apparent prevalence of pulmonary embolism in our population, diluted by patients for whom the suspicion or pulmonary embolism was extremely low to begin with. This inclusion would have severely limited the generalizability of this study. In any case, our goal was to study the application of Wells Criteria to patients who were going to undergo imaging for pulmonary embolism.

DISCUSSION

To our knowledge, this is the first study to examine the interrater agreement and prospective external validity of Wells Criteria. This clinical decision rule appears to reproducibly and reliably risk stratify patients with suspected pulmonary embolism into either a low, moderate, and high classification or a pulmonary embolism–unlikely and pulmonary embolism–likely classification. Furthermore, the 2 D-dimer assays used in our study appeared to perform with equal sensitivities and specificities. Our results do not seem to support the use of either assay alone to rule out the diagnosis of pulmonary embolism, although the CIs are large.

As the first study looking at the interrater agreement of the Wells Criteria, our findings are encouraging. We found that the dichotomized scoring system had substantial interrater agreement, whereas the trichotomized system had moderate interrater agreement. Furthermore, our study found that the interrater agreement for both scoring systems was essentially independent of the level of training of the emergency care provider (0.72 versus 0.70 and 0.54 versus 0.62), which is interesting, given the subjective nature of the second criterion, which asks whether an alternative diagnosis is considered as likely or more likely than pulmonary embolism. This answer undoubtedly would vary with education and experience, as the traditional "gestalt" means of determining pretest probability does.[34] In fact, the second criterion generated the lowest κ value of all the Wells Criteria. We speculate that when the clinical decision rule is considered as a whole, the other 6 criteria may act to moderate the results of the second, causing experience and education to influence the final score less.

Wells et al[14,19] previously published studies reporting the incidence of pulmonary embolism to be 2% to 4% for patients with a low pretest probability, 19% to 21% for patients with a moderate pretest probability, and 50% to 67% for patients with a high pretest probability. Additionally, they found that the pulmonary embolism–unlikely and pulmonary embolism–likely groups had incidences of 5% to 8% and 39% to 41%, respectively.[19] Our findings are consistent with these numbers, providing independent external validity to Wells Criteria for both the dichotomized and trichotomized scoring systems. It is interesting to note that none of the patients with a negative D-dimer result in the low pretest probability or pulmonary embolism–unlikely groupings were diagnosed with a pulmonary embolism.

Table 4C.
*Dichotomized scoring table comparing all care providers with all care providers.**

	Rater 2	
Rater 1	**PE-UL**	**PE-L**
PE-UL	77	6
PE-L	12	39

PE-L, Pulmonary embolism–likely; **PE-UL,** pulmonary embolism–unlikely.
*κ=0.72 (95% CI 0.60–0.84)

Table 4D.
*Dichotomized scoring table comparing physician care providers with physician research assistants.**

	MD 2	
MD 1	**PE-UL**	**PE-L**
PE-UL	44	5
PE-L	6	24

*κ=0.70 (95% CI 0.53–0.87)

Table 5.
Results of Wells Criteria.

Wells Criteria Score	No.	%	Pulmonary Embolism, No.	Pulmonary Embolism, %	95% CI
Low (<2)	59	44	1	2	0–9
Moderate (2–6)	61	46	9	15	7–26
High (>6)	14	10	6	43	18–71
PE-UL(≤4)	88	66	3	3	0–9
PE-L (>4)	46	34	13	28	18–71

The sensitivities we found for both of the D-dimer assays were not as promising as those that have previously been reported. However, our sample size is relatively small, and it may be that with a greater sample size our numbers would converge with those previously published. Still, keeping in mind the very large CIs, these numbers are consistent with recent literature that states that a negative D-dimer assay alone should not be used to rule out pulmonary embolism.[2,4-7]

In summary, we conclude that Wells Criteria appears to have a moderate to substantial interrater reliability, affording emergency medicine care providers with a reproducible means of determining pretest probability among ED patients for whom the diagnosis of pulmonary embolism is being considered.

Author contributions: TRM and DMA conceived the study and obtained funding. SJW, TRM, and KMF supervised the conduct of the study. All authors were involved in data collection, and SJW managed the data. SJW, TRM, and KMF were responsible for quality control. Statistical analysis was performed by an independent statistician and in part by SJW. SJW drafted the manuscript, and SJW, TRM, RMF, JPF, and DMA contributed substantially to its revision. SJW takes responsibility for the paper as a whole.

Received for publication May 18, 2003. Revisions received October 13, 2003; December 15, 2003; March 8, 2004; and April 5, 2004. Accepted for publication April 7, 2004. Available online August 25, 2004.

Presented at the Society for Academic Emergency Medicine annual meeting, St. Louis, MO, May 2002.

Supported in part by a Medicare grant (Medicare 2001 Kaiser Permanente/Exempla Saint Joseph Hospital Research Fund).

Reprints not available from the authors.

Address for correspondence: Stephen J. Wolf, MD, Department of Emergency Medicine, Denver Health Medical Center, Mail Code 0108, 777 Bannock Street, Denver, CO 80204; 303-436-8842, fax 303-436-7541; E-mail Stephen.Wolf@dhha.org.

REFERENCES

1. Goldhaber SZ. Pulmonary embolism. *N Engl J Med.* 1998;339:93-104.
2. Brown MD, Rowe BH, Reeves MJ, et al. The accuracy of the enzyme-linked immunosorbent assay D-dimer test in the diagnosis of pulmonary embolism: a meta-analysis. *Ann Emerg Med.* 2002;40:133-144.
3. Kline JA, Johns KL, Colucciello SA, et al. New diagnostic tests for pulmonary embolism. *Ann Emerg Med.* 2000;35:168-172.
4. Wolfe TR, Hartsell SC. Pulmonary embolism: making sense of the diagnostic evaluation. *Ann Emerg Med.* 2001;37:504-514.
5. Sadosty AT, Goyal DG, Boie ET, et al. Emergency department D-dimer testing. *J Emerg Med.* 2001;21:423-429.
6. Kline JA, Wells PS. Methodology for a rapid protocol to rule out pulmonary embolism in the emergency department. *Ann Emerg Med.* 2003;42:266-275.
7. Bounameaux H, de Moerloose P, Perrier A, et al. Plasma measurement of d-dimer as diagnostic aid in suspected venous thromboembolism: an overview. *Thromb Haemost.* 1994;71:1-6.
8. Goldhaber SZ, Simons GR, Elliott CG, et al. Quantitative plasma D-dimer levels among patients undergoing pulmonary angiography for suspected pulmonary embolism. *JAMA.* 1993;270:2819-2822.
9. Ginsberg JS, Brill-Edwards PA, Demers C, et al. D-Dimer in patients with clinically suspected pulmonary embolism. *Chest.* 1993;104:1679-1684.
10. De Moerloose P, Desmarais S, Bounameaux H, et al. Contribution of a new, rapid, individual and quantitative automated D-dimer ELISA to exclude pulmonary embolism. *Thromb Haemost.* 1996;75:11-13.
11. Perrier A, Desmarais S, Miron MJ, et al. Non-invasive diagnosis of venous thromboembolism in outpatients. *Lancet.* 1999;353:190-195.
12. Knecht MF, Heinrich F. Clinical evaluation of an immunoturbidimetric D-dimer assay in the diagnostic procedure of deep vein thrombosis and pulmonary embolism. *Thromb Res.* 1997;88:413-417.
13. Duet M, Benelhdj S, Kedra W, et al. A new quantitative D-dimer assay appropriate in emergency: reliability of the assay for pulmonary embolism exclusion diagnosis. *Thromb Res.* 1998;91:1-5.
14. Wells PS, Ginsberg JS, Anderson DR, et al. Use of a clinical model for safe management of patients with suspected pulmonary embolism. *Ann Intern Med.* 1998; 129:997-1005.
15. Wicki J, Perneger TV, Junod AF, et al. Assessing clinical probability of pulmonary embolism in the emergency ward: a simple score. *Arch Intern Med.* 2001;161:92-97.
16. Kline JA, Nelson RD, Jackson RE, et al. Criteria for the safe use of D-dimer testing in emergency department patients with suspected pulmonary embolism: a multicenter US study. *Ann Emerg Med.* 2002;39:144-152.
17. British Thoracic Society Standards of Care Committee Pulmonary Embolism Guideline Development Group. British Thoracic Society guidelines for the management of suspected acute pulmonary embolism. *Thorax.* 2003;58:470-484.
18. Miniati M, Monti S, Bottai M. A structured clinical model for predicting the probability of pulmonary embolism. *Am J Med.* 2003;114:173-179.
19. Wells PS, Anderson DR, Rodger M, et al. Derivation of a simple clinical model to categorize patients' probability of pulmonary embolism: increasing the model's utility with the SimpliRED D-dimer. *Thromb Haemost.* 2000;83:416-420.
20. 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. *Ann Intern Med.* 2001;135:98-107.
21. Chagnon I, Bounameaux H, Aujesky D, et al. Comparison of two clinical prediction rules and implicit assessment among patients with suspected pulmonary embolism. *Am J Med.* 2002;113:269-275.
22. Chabloz P, Reber G, Boehlen F, et al. TAFI antigen and D-dimer levels during normal pregnancy and at delivery. *Br J Haematol.* 2001;115:150-152.
23. Righini M, de Moerloose P, Reber G, et al. Should the D-dimer cut-off value be increased in elderly patients suspected of pulmonary embolism? [letter] *Thromb Haemost.* 2001;85:744.
24. De Lorenzo F, Mukherjee M, Kadziola Z, et al. Association of overall adiposity rather than body mass index with lipids and procoagulant factors. *Thromb Haemost.* 1998;80:603-606.
25. Perrier A, Desmarais S, Miron MJ, et al. Non-invasive diagnosis of venous thromboembolism in outpatients. *Lancet.* 1999;353:190-195.
26. Freyburger G, Trillaud H, Labrouche S, et al. D-Dimer strategy in thrombosis exclusion. *Thromb Haemost.* 1998;79:32-37.
27. Heit JA, Meyers BJ, Plumhoff EA, et al. Operating characteristics of automated latex immunoassay fibrin D-dimer in the diagnosis of angiographically-defined acute pulmonary embolism. *Thromb Haemost.* 2000;83:970.
28. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33:159-174.
29. Teigen CL, Maus TP, Sheedy PF II, et al. Pulmonary embolism: diagnosis with contrast-enhanced electron-beam CT and comparison with pulmonary angiography. *Radiology.* 1995;194:313-319.
30. Mayo JR, Rémy-Jardin M, Muller NL, et al. Pulmonary embolism: prospective comparison of spiral CT with ventilation-perfusion scintigraphy. *Radiology.* 1997;205: 447-452.
31. Garg K, Welsh CH, Feyerabend AJ, et al. Pulmonary embolism: diagnosis with spiral CT and ventilation-perfusion scanning: correlation with pulmonary angiographic results or clinical outcome. *Radiology.* 1998;208:201-208.
32. Goodman LR, Lipchik RJ, Kuzo RS, et al. Subsequent pulmonary embolism: risk after a negative helical CT pulmonary angiogram: prospective comparison with scintigraphy. *Radiology.* 2000;215:535-542.
33. Swensen SJ, Sheedy PF II, Ryu JH, et al. Outcomes after withholding anticoagulation from patients with suspected acute pulmonary embolism and negative computed tomographic findings: a cohort study. *Mayo Clin Proc.* 2002;77:130-138.
34. Stein PD, Terrin ML, Hales CA, et al. Clinical, laboratory, roentgenographic, and electrocardiographic findings in patients with acute pulmonary embolism and no pre-existing cardiac or pulmonary disease. *Chest.* 1991;100:598-603.