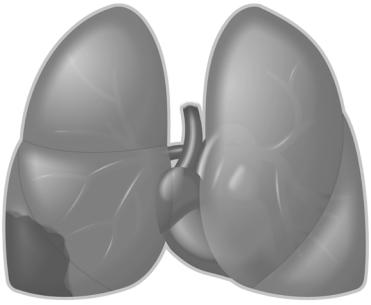
CHAPTER 4

Performance of 4 clinical decision rules in the diagnostic management of acute pulmonary embolism: a prospective cohort study.

I.C.M. Mos*, R.A. Douma*, P.M.G. Erkens, T.A.C. Nizet, M.F. Durian, M.M.C. Hovens, A.A. van Houten, H.M.A. Hofstee, F.A. Klok, H. ten Cate, E.F. Ullmann, H.R. Büller, P.W. Kamphuisen and M.V. Huisman - on behalf of the Prometheus Study Group

* Equally contributed

Ann Intern Med 2011; 154:709-718



ABSTRACT

Background

Several clinical decision rules (CDRs) are available to exclude acute pulmonary embolism (PE), but they have not been directly compared.

Objectives

To directly compare the performance of 4 CDRs (Wells rule, revised Geneva score, simplified Wells rule, and simplified revised Geneva score) in combination with D-dimer testing to exclude PE.

Design

Prospective cohort study.

Setting

Seven hospitals in the Netherlands.

Patients

807 consecutive patients with suspected acute PE.

Interventions

The clinical probability of PE was assessed by using a computer program that calculated all CDRs and indicated the next diagnostic step. Results of the CDRs and D-dimer tests guided clinical care.

Measurements

Results of the CDRs were compared with the prevalence of PE identified by computed tomography or venous thromboembolism at 3-month follow-up.

Results

Prevalence of PE was 23%. The proportion of patients categorized as PE-unlikely ranged from 62% (simplified Wells rule) to 72% (Wells rule). Combined with a normal D-dimer result, the CDRs excluded PE in 22% to 24% of patients. The total failure rates of the CDR and D-dimer combinations were similar (1 failure, 0.5% to 0.6% [upper-limit 95% CI, 2.9% to 3.1%]). Even though 30% of patients had discordant CDR outcomes, PE was not detected in any patient with discordant CDRs and a normal D-dimer result.

Limitation

Management was based on a combination of decision rules and d-dimer testing rather than only 1 CDR combined with D-dimer testing.

Primary Funding Source

This study was supported by unrestricted grants from the participating hospitals.

Conclusions

All 4 CDRs show similar performance for exclusion of acute PE in combination with a normal D-dimer result. This prospective validation indicates that the simplified scores may be used in clinical practice.

INTRODUCTION

The introduction of standardized clinical decision rules (CDRs) to determine the clinical probability of pulmonary embolism (PE) has improved the diagnostic workup of patients with suspected PE. A CDR indicated PE "unlikely" in combination with a normal D-dimer test result can exclude the diagnosis of PE in a large proportion of the patients who present for evaluation (20-40%), without the need for additional imaging with computed tomographic pulmonary angiography (CT) or ventilation-perfusion scintigraphy, which both involve radiation and intravenous contrast or radioisotopes. In these patients anticoagulants can be safely withheld.[1-4]

Several clinical decision rules, which incorporate information from medical history and physical examination, have been developed and validated. Next to six objective variables, the Wells rule contains one subjective variable: the physician should consider the possibility of an alternative diagnosis than PE for the patient's complaints (Table 1).[5] In contrast, the more recently introduced revised Geneva score is composed of eight objective clinical variables.[6] Both scores assign different weights to the variables, meaning that depending on the variable either 1, 1.5, 2, 3, 4, or 5 points need to be assigned (Table 1). Because miscalculations can occur, the scores have recently been simplified (Table 1).[7,8]

Until now, the simplified Wells rule and the simplified revised Geneva score have not been validated prospectively. Also, while some of the scores have retrospectively or prospectively been compared with each other,[9-12] the four scores have never been directly compared for the performance of excluding PE in combination with a normal D-dimer test result. Therefore, we performed a prospective multi-center clinical accuracy study to assess and directly compare the performance of these four different CDRs (Wells rule, revised Geneva score, simplified Wells rule and simplified revised Geneva score) in

excluding PE in combination with D-dimer testing, using a computer-based program to calculate the CDR scores.

METHODS

The study was a prospective multi-center cohort study on clinical accuracy study of 4 CDRs in consecutive patients with a suspected first episode of acute PE. The study population consisted of consecutive outpatients and inpatients in whom a first acute PE was clinically suspected. Clinically suspected acute PE was defined as sudden onset of dyspnea, deterioration of existing dyspnea and/or sudden onset of pleuritic chest pain. Patients were included in seven participating academic or non-academic hospitals in the Netherlands.

Exclusion criteria were age below 18 years of age, life expectancy of less than 3 months, treatment with therapeutic-dose low molecular weight heparin or unfractionated heparin that was initiated 24 hours or more prior to eligibility assessment, treatment with vitamin K antagonists, previous PE, contraindication to helical CT scan because of allergy to intravenous iodinated contrast or renal insufficiency (creatinine clearance < 30 ml/min using the Cockroft-Gault formula), pregnancy and inability to return for follow-up. Institutional review boards of all participating hospitals approved the study protocol and written informed consent was obtained from all included patients.

Study flow

Patients included in the study underwent a sequential work-up of clinical probability assessment, D-dimer testing and CT scanning. In all patients, the items of four clinical decision rules were assessed by the treating physicians (Table 1). In addition, a high-sensitivity quantitative D-dimer test was performed (VIDAS D-dimer assay, Biomerieux, Marcy L'Etoile, France; Tinaquant assay, Roche Diagnostica, Mannheim, Germany; STA-Liatest D-Di, Diagnostica Stago, Asnieres, France; or Innovance D-dimer, Siemens, Marburg, Germany), in all included patients, irrespective of CDR results. The type of D-dimer assay that was used depended on local practice. Pulmonary embolism was considered "unlikely" in case of a Wells rule of 4 points or less, a simplified Wells rule of 1 point or less,[7,10] and a simplified revised Geneva score with a score of 2 points or less (Table 1).[8] The revised Geneva score, until now only available in a three-category scheme, was transformed to a two-category scheme similar to the other scores. This was done by a beforehand calculation of the optimal cut-off, using an existing cohort of patients with suspected PE[8] for whom the revised Geneva score variables were available for calculation of the score. The optimal cut-off point was determined by calculation of the area under the Receiver Operating Characteristic (ROC) curve, and the proportions of patients in

Table 1. Clinical decision rules.

Wells rule			**Revised Geneva score**		
Items	**Original**	**Simplified**	**Items**	**Original**	**Simplified**
Previous PE or DVT	1.5	1	Previous DVT or PE	3	1
Heart rate >100/min	1.5	1	Heart rate 75 – 94/min Heart rate ≥ 95/min	3 5	1 2
Surgery or immobilization within 4 weeks	1.5	1	Surgery or fracture within 1 month	2	1
Hemoptysis	1	1	Hemoptysis	2	1
Active malignancy	1	1	Active malignancy	2	1
Clinical signs of DVT	3	1	Unilateral lower limb pain	3	1
Alternative diagnosis less likely than PE	3	1	Pain on lower limb deep vein palpation and unilateral edema	4	1
			Age > 65 years	1	1
Clinical probability			**Clinical probability**		
PE unlikely	≤ 4	≤ 1	PE unlikely	≤ 5	≤ 2
PE likely	> 4	> 1	PE likely	> 5	> 2

PE: pulmonary embolism; DVT: deep vein thrombosis.

the "likely" and "unlikely" categories were calculated. Based on these calculations, PE was considered "unlikely" with a score of 5 points or less (Table 1). For any of the CDRs, a score above the respective cut-off indicated PE "likely".

Clinical care was guided by the results of the CDRs and D-dimer results (Figure 1). When PE was considered "unlikely" according to all four CDRs in combination with a normal D-dimer test result (cut-off < 500 µg/L), PE was excluded. In all remaining patients (i.e. a "likely" result according to at least one of the CDRs or an abnormal D-dimer test result), CT scanning was indicated to confirm or exclude the diagnosis. Patients with CT indicating PE were treated with anticoagulants and treatment was withheld from all patients in whom the diagnosis was excluded. These latter patients were followed for a 3-month period. The study flow is illustrated in Figure 1.

Standard contrast enhanced MDCT was performed using a 4-slice, 16-slice, or 64-slice MDCT scanner with acquisition of 0.5 or 1 mm sections (depending on the weight of the subject) of the entire chest for diagnosing or excluding PE. The rotation time is 0.4 s and the pitch factor 1.4; the tube current is 250-300 mA and the tube voltage 100 kV. Acquisitions are performed during a single breath-hold, lasting 10-12 seconds or less, depending on the type of scanner. 80-100 ml of contrast agent is injected in the antecubital vein with an injection rate of 4.0 ml/sec. The acquisition of the static pulmonary angiography scan is started after automated threshold enhancement detection in the pulmonary trunk. A threshold difference of 100 Hounsfield units is selected for starting the acquisition. CT images were read by skilled radiologists to determine if PE was present or could be excluded. The radiologists were aware of an indication for CT-

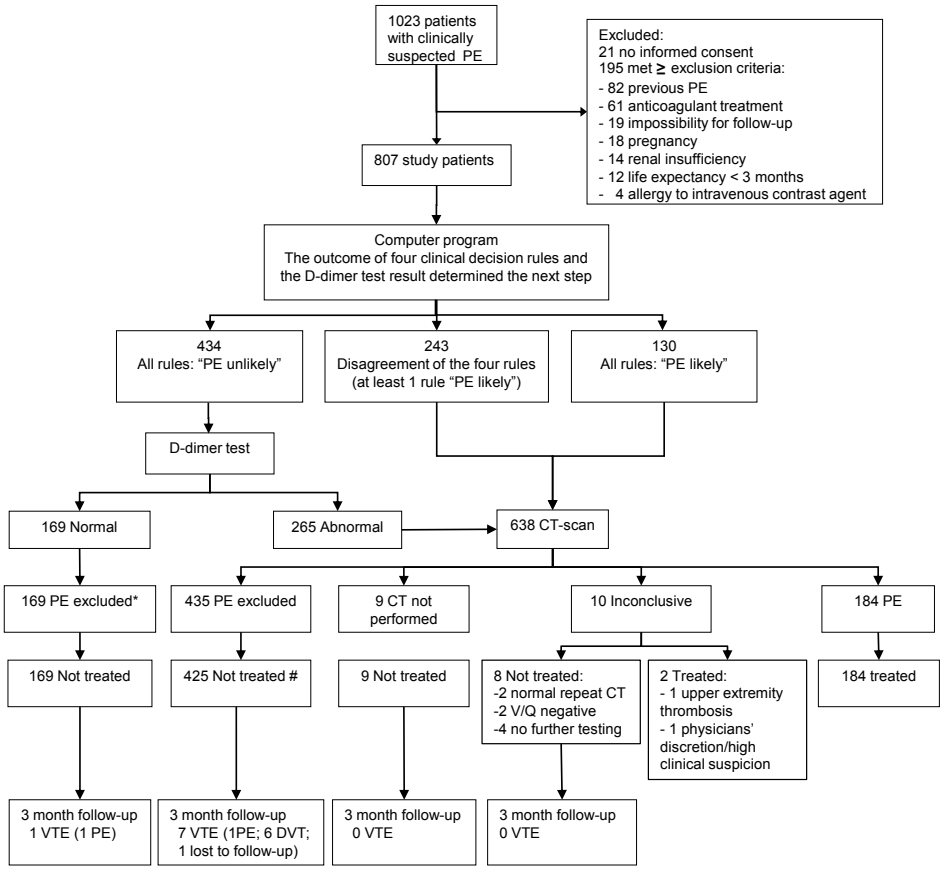


Figure 1. Flow-chart with results of the diagnostic strategy. PE: pulmonary embolism; DVT: deep vein thrombosis; VTE: venous thromboembolism; CT: computed tomography; V/Q: ventilation perfusion scintigraphy; *in seven patients a CT-scan was performed while not indicated and confirmed the diagnosis in one patient; "Treated or not treated" concerns treatment with anticoagulants; #ten patients in whom PE was excluded by CT-scan received anticoagulant treatment for reasons other than VTE.

scanning but not if this was based on a high CDR and/or an elevated D-dimer test result. The diagnosis of PE was confirmed by the presence of at least one filling defect in the pulmonary artery tree. Management of patients with an inconclusive CT result was left to the attending physician and could include repeat CT scanning, ventilation-perfusion scintigraphy or conventional pulmonary angiography.

Computerized program

Clinical evaluations and the collection of data were performed by the treating physicians at baseline. In each participating center, a study coordinator was available for advice regarding the study. This coordinator also checked the completeness and correctness of the data. Demographic data and additional relevant information

(e.g. recent trauma or surgery, cancer, use of anticoagulants, duration of time since symptom onset and D-dimer test result) were collected on a Case Report Form (CRF), available in paper and digital format. The computerized design forced the physician to start the diagnostic process with clinical evaluation of the patient and to enter all variables necessary to calculate the four CDRs and the D-dimer test result into the computer. The computer program calculated the four individual CDR scores and, after combining these scores with the D-dimer result indicated to the physician the next recommended step in the diagnostic process according to the predefined study flow: either exclusion of PE based on the CDR and D-dimer level or performing a CT scan (Figure 1).

Follow-up

Patients in whom PE was excluded, either based on the CDR/D-dimer combination (for all four CDRs) or based on a normal CT, were followed up for 3 months. All patients were instructed to return to the hospital should complaints of venous thromboembolism (PE or DVT) or bleeding occur. Objective diagnostic tests were performed if a suspicion of VTE was raised e.g. CT-scanning, V/Q-scanning and/or compression ultrasonography. Patients were interviewed by telephone by one of the study coordinators at the end of a 3-month follow-up period and were questioned on health-related events during the past three months, especially for symptoms suggestive of PE or DVT, interval initiation of anticoagulants and possible haemorrhagic complications. If relevant, the patient's general practitioner was contacted for additional information. If a patient had died, the case of death was obtained from hospital records, autopsy reports or from information of the general practitioner. Deaths were classified as due to pulmonary embolism in case of confirmation by autopsy, an objective diagnostic test positive for PE prior to death, or if the cause of death could not completely be explained by reasons other than VTE. All outcomes were adjudicated by a panel of three experts.

Statistical analysis

We calculated that, based on a β of 10% (power 90%) and an alpha of 0.05, 128 positive CT-scans would be needed to detect a difference of more than 5% (55% vs. 50%) in sensitivity among the two primary CDRs (i.e. Wells and revised Geneva score). Based on a prevalence of PE of 20%,[2] a sample size of 753 participants with suspected pulmonary embolism was required. All additional sample size calculations on other outcomes needed a smaller sample size. Because of possible dropout, we aimed for a total sample size of 800 patients.

In this study, the four CDRs were directly compared for their performance in identifying patients as having PE or not. This included four primary analyses: 1) the ability of each CDR to correctly categorize patients with suspected PE as "unlikely" or "likely";

2) the proportion of patients in whom the diagnosis was excluded based on an "unlikely" CDR combined with a normal D-dimer test at the time of the acute evaluation; 3) the safety of clinical management based upon each CDR-D-dimer combination to exclude the diagnosis, i.e. the true negative results (proportion of patients safely managed without CT scan) and false negative results. The latter was defined as the VTE rate during the 3-month follow-up in patients in whom PE was considered ruled out by the initial diagnostic work-up and who did not receive anticoagulants during follow-up; 4) The distribution of patients in the probability categories according to the four CDRs was studied using sensitivity, specificity and receiving operating characteristic (ROC) analysis. Also, the discordant cases (patients classified as "unlikely" by one CDR but "likely" by another) were described. The reference standard in patients, in whom CT scanning was not indicated, was the recurrent VTE rate during 3-month follow-up. For patients who had to undergo CT scanning, the reference standard was CT scanning and 3-month follow-up.

Performance of the four CDRs and the combination of the CDRs and D-dimer testing were examined using sensitivity, specificity, ROC analysis, event rates and predictive values. To assess differences between the four CDRs in sensitivity, specificity, predictive values and to compare the categorization of patients into the probability groups (paired data), multiple testing was performed using McNemar's test. Each CDR was compared with the other CDRs individually. Furthermore, stratification by type of hospital (academic and non-academic hospitals) was performed to give insight into possible type of hospital-associated differences using stratified Mantel-Haenszel test (CDR "likely"/"unlikely" versus outcome of PE stratified for academic versus non-academic hospitals). Exact 95% confidence intervals (CI) were calculated around the observed incidences using Confidence Interval Analysis.[13] Descriptive parameters were calculated using SPSS software, version 16.0 (SPSS Inc, Chicago, Ill). Mean values and frequencies such as the clinical characteristics of subgroups were compared using Students t-test and χ^2-test respectively. Statistical significance was set at $p<0.05$.

RESULTS

Study patients

Between July 2008 and November 2009, a total of 1023 consecutive patients with clinically suspected pulmonary embolism were screened, of whom 195 (19%) were excluded because of one or more of the predefined exclusion criteria (Figure 1). In addition, 21 patients refused to give informed consent. The final study population of 807 participants included 644 (80%) outpatients and 163 (20%) inpatients. The baseline demographic and clinical characteristics of the 807 study participants are shown in Table 2.

Table 2. Clinical characteristics of the 807 patients with suspected pulmonary embolism.

Characteristic	Value
Age, mean (SD), y	53 (17.7)
Female, n (%)	487 (60.3)
Outpatient, n (%)	644 (79.8)
Duration of complaints, median (IQR), d	2 (1-7)
Body mass index, mean (SD), kg/m2	26.3 (5.5)
Risk factors	
Immobilization or recent surgery, n (%)	176 (21.8)
Previous VTE, n (%)	39 (4.8)
COPD with treatment, n (%)	75 (9.3)
Heart failure with treatment, n (%)	47 (5.8)
Active malignancy, n (%)	114 (14.1)
Estrogen use, women, n (%)	97 (19.9)
Body mass index ≥ 30 kg/m2, n (%)	152 (4.6)
Symptoms and clinical presentation	
Clinical symptoms of deep vein thrombosis, n (%)	47 (5.8)
Heart rate, mean (SD), bpm	88 (18.8)
Hemoptysis, n (%)	40 (5.0)

Bpm: beats per minute; COPD: chronic obstructive pulmonary disease; IQR: interquartile range; SD: standard deviation; VTE: venous thromboembolism.

Result of Diagnostic Algorithm

Patients were managed according to the results of CDRs combined with the D-dimer test result as illustrated in Figure 1. Discordant CDR results were observed in 243 patients (29%), while in 564 patients the CDR results were concordant. In total, PE was ruled out by a combination of an "unlikely" CDR result according to all four CDRs and a normal D-dimer test result in 169 patients (21%). In 638 patients (79%) CT was indicated; either due to an abnormal D-dimer test result (265 patients) or due to having at least one of CDRs indicating "PE likely" in 373 patients.

D-dimer testing was not performed in 19 patients (protocol violations). This happened in one patient with "PE unlikely" according to all four CDRs; this patient was regarded as having a positive D-dimer test and a CTPA was performed (this patient is one of the 265 patients with an 'abnormal' D-dimer). In 18 other patients, the CDR results were discordant. The missing D-dimer result had no impact on the next step in the strategy, for a CTPA was to be performed based on the discordant CDRs.

Protocol violations regarding CTPA occurred in 16 patients: in nine of them CT scanning was indicated but not performed, these patients were all followed for three months; in seven patients, CT scanning was performed while it was not indicated, and showed PE in one of these patients.

In total, CT confirmed the diagnosis of PE in 185 patients, 184 with at least one "likely PE" CDR and a positive D-dimer result and 1 for whom the CT scan was not indicated based on the study criteria, but was done based on clinical judgment (Table 3, patient

Table 3. Characteristics of patients in whom venous thromboembolism was found during the 3-month follow-up, despite initial exclusion of the diagnosis.

Patient			Outcome of diagnostic tests at inclusion						Follow-up		
Pt.	**Sex**	**Age**	**Wells**	**SW**	**RGS**	**sRGS**	**DD**	**CT at presentation**	**VTE**	**Day (d)**	**Brief description**
1	Male	65	↓	↓	↓	↓	482	Indicating PE	PE	0	CT was performed although not indicated (all CDR unlikely and a normal D-dimer), and positive for PE: multiple subsegmental emboli were found, as well as signs suggesting pulmonary infarction
2	Female	63	↑	↑	↑	↑	1535	Normal	DVT	19	Also suspected for DVT at presentation, CUS was negative for DVT at presentation
3	Male	63	↓	↑	↑	↑	1100	Normal	PE	22	PE found by coincidence on CT-scan made for other reasons
4	Female	39	↓	↓	↓	↓	1100	Normal	DVT	27	DVT was found on CUS at day 27 (during hospitalisation). Despite this finding, anticoagulant treatment was delayed till day 51 after another CUS positive for DVT
5	Female	58	↓	↓	↓	↓	2600	Normal	DVT	62	DVT of jugular and subclavian vein, patient had Takayasu arteritis
6	Female	43	↓	↑	↓	↓	2100	Normal	DVT	21	DVT of jugular vein found by coincidence on staging CT-scan after chemo-radiotherapy
7	Male	87	↓	↓	↑	↑	3420	Normal	DVT	0	DVT found on CUS made directly after CT negative for PE, patient also had complaints of the leg
8	Female	62	↑	↑	↓	↓	-*	Normal	DVT	7	DVT after surgery and immobilisation

Wells: original Wells rule; SW: simplified Wells rule; RGS: revised Geneva score; sRGS: simplified revised Geneva score; DD: D-dimer test; CT: computed tomography; VTE: venous thromboembolism; ↑: clinical decision rule indicating "likely"; ↓: clinical decision rule indicating "unlikely"; PE: pulmonary embolism; DVT: deep vein thrombosis; CDR: clinical decision rule; CUS: compression ultrasonography

*D-dimer was not performed.

number 1; Figure 1). The diagnosis was excluded in 435 patients; in 164 of them an alternative diagnosis for the complaints was found. CT was inconclusive in 10 patients: repeat CT scanning excluded the diagnosis in two of these patients, and ventilation/perfusion scintigraphy excluded the diagnosis in two other patients. In one patient, anticoagulant treatment was started based on an inconclusive CT scan combined with high clinical suspicion of PE and in another, thrombosis of the subclavian vein was found with the same scan; these patients were treated accordingly. In the remaining 4 patients with inconclusive CT scans, the diagnosis was considered to be excluded without further testing and as a result these patients were not treated with anticoagulation medication. A final diagnosis could be established within an hour in the majority of the patients or at maximum within 24 hours after presentation. The overall prevalence of PE in this total study population was therefore 185/807 (23%, 95% CI: 20% to 26%).

Follow Up

In seven patients of the 169 patients in the all-unlikely group who had a normal D-dimer, the protocol was violated and a CT scan was performed while not indicated; in one of these patients, PE was diagnosed as already mentioned above. This was regarded as a diagnostic failure in the CDR-D-dimer strategy (1/169; 0.6%, 95% CI: 0.02% to 3.3%) (Figure 2; patient number 1 in Table 3). None of the remaining 168 patients in this group were treated with anticoagulants during follow-up and all of these patients had an uneventful follow-up. The nine patients, in whom a CT scan was indicated but not performed, were also left untreated and had an uneventful follow-up. Of the 435 patients in whom PE was excluded with CT scanning and the eight untreated patients with inconclusive results, 10 patients (2.3%) were treated with anticoagulants during follow-up for reasons other than VTE. Seven of the 433 patients with a normal or inconclusive CT scan result without anticoagulant treatment for other reasons returned with symptomatic and objectively confirmed VTE events during the 3-month follow up (Figure 1; patient numbers 2 to 8 in Table 3). Eighteen patients died during follow-up. In one of these patients, a DVT had already been diagnosed during follow-up; in another patient PE was excluded by autopsy as cause of death; while in the remaining 16 patients, the cause of death was adjudicated to be unrelated to a possible VTE. Therefore, the failure rate of a normal or inconclusive CT in this study was 7 in 433 (1.6%, 95% CI: 0.7% to 3.3%). One patient (1/807, 0.1%) was lost to follow-up. In a "worst case" scenario, in which this patient would have developed VTE, the failure rate after CT-scanning that excluded PE would have been 8 in 433 (1.9%, 95% CI: 0.8% to 3.6%). Allergy to intravenous iodinated contrast or contrast induced nephropathy was not recognized in the included patients during the study period.

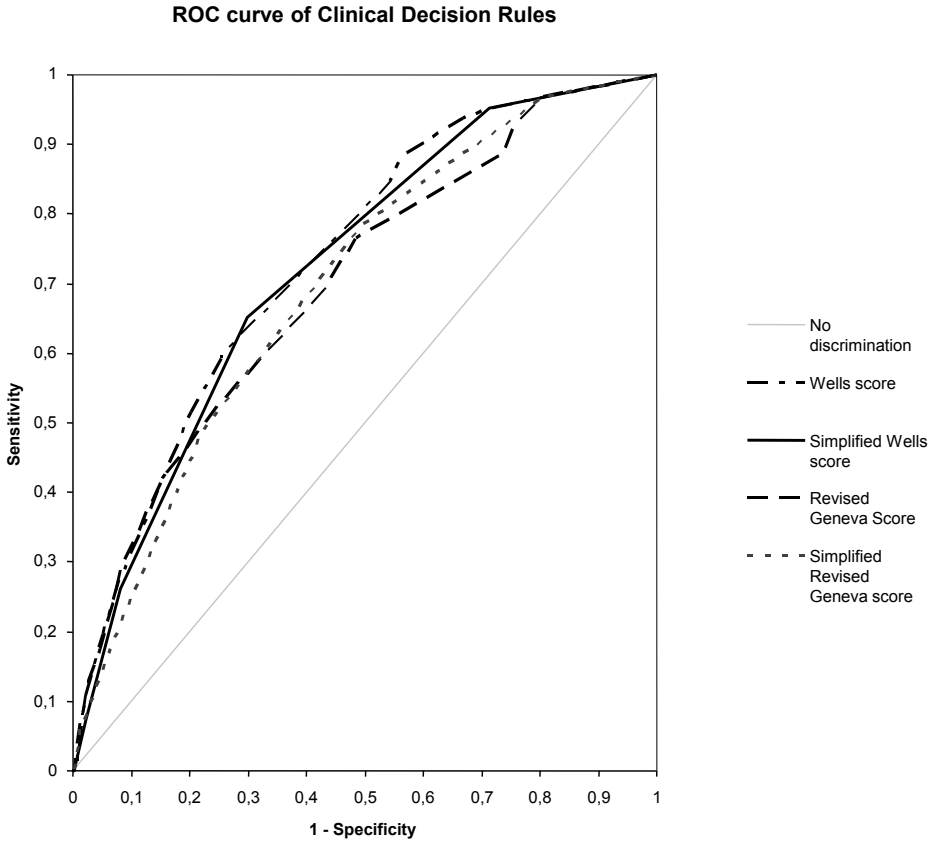


Figure 2. Receiver operating characteristics curves of the four clinical decision rules. Continuous clinical Decision Rules. Area under the curves: 0.73 (95% CI 0.69-0.77) for the Wells rule, 0.72 (95% CI 0.68-0.76) for the simplified Wells rule, 0.70 (95% CI 0.65-0.74) for the revised Geneva score and 0.69 (95% CI 0.65-0.74) for the simplified revised Geneva score, respectively.

Categorization of patients in probability groups with the four CDRs

Table 4 describes how the patients were categorized by the two probability categories of the four CDRs without taking the D-dimer test results into account. The proportion of patients classified as "unlikely PE" was similar for the four CDRs. Also, the prevalence of PE in the "unlikely" categories was comparable. Overall, the proportion of patients classified as "likely" was largest using the simplified Wells rule: 38% versus 28% to 32% with the other three CDRs (Table 4). The sensitivity and specificity of each CDR alone (without D-dimer results) ranged from 49% to 65% (sensitivity) and from 70% to 80% (specificity), respectively (Table 5a).

The Receiver Operator Characteristics-curves for the four CDRs were comparable and showed areas under the curve ranging from 0.69 to 0.73 (Figure 2).

Table 4. Distribution of patients in unlikely and likely clinical probability based on four clinical decision rules and the combination of the CDR and D-dimer test (n=807).

	Original Wells rule	Simplified Wells rule	Original Revised Geneva rule	Simplified revised Geneva rule
CDR unlikely, n (%, 95% CI)	584 (72, 69-76)	499 (62, 59-65)	553 (69, 65-72)	576 (71, 68-75)
Prevalence of PE in patients with CDR 'unlikely', n (%, 95% CI)	90/584 (15, 13-18)	65/499 (13, 10-16)	88/553 (16, 13-19)	95/576 (17, 14-20)
CDR likely, n (%, 95% CI)	223 (28, 25-31)	308 (38, 35-41)	254 (32, 28-35)	231 (29, 26-32)
Prevalence of PE in patients with CDR likely, n (%, 95% CI)	95/223 (43, 36-49)	120/308 (39, 34-44)	97/254 (38, 32-44)	90/231 (39, 32-45)
CDR unlikely and D-dimer normal, n (%, 95% CI)	184 (23, 20-26)	178 (22, 19-25)	185 (23, 20-26)	190 (24, 21-27)
VTE incidence in patients with CDR unlikely and a normal D-dimer, n (%, 95%CI)	1/184 (0.5, 0.0-3.0)	1/178 (0.6, 0.0-3.1)	1/185 (0.5, 0.0-3.0)	1/190 (0.5, 0.0-2.9)

CDR: Clinical decision rule; PE: pulmonary embolism; VTE: venous thromboembolism.

Performance of the four CDRs together with D-dimer

Combined with a normal D-dimer, the four CDRs excluded PE in similar proportions of patients, ranging from 22% to 24% (Table 4). There was no difference between the 3-month VTE failure rates of the different CDR-D-dimer combinations. This failure rate ranged from 0.5% to 0.6% (Table 4). The 95% CI is 0 to 3% for all CDRs (Table 4). When combined with the D-dimer test result, the sensitivities of the various CDRs did not differ, while there were small differences in specificity (Table 5b).

Discordance between the CDRs

Of the 434 patients with all 4 CDRs indicating PE unlikely, 52 (12%) were diagnosed with PE; all patients except one had an abnormal D-dimer test result, the latter which indicated the need for CT-scanning.

In 243 of 807 patients (29%) discordance between CDRs was observed (Figure 1); the D-dimer test result was normal in 29 abnormal in 196, and incorrectly not performed in 18. In the latter 18 patients, CT-scanning was performed which confirmed the diagnosis of PE in one patient.

The number of discordant cases between two scores ranged from 25/807 (3.1%) between the revised Geneva score and the simplified revised Geneva score, to 199/807 (25%) between the Wells rule and the revised Geneva score (Table 6a). The agreement was greatest between the original scores (Wells, revised Geneva score) and their simplified versions (simplified Wells and simplified Geneva score; discordance 11% and 3.1% of the total cohort, respectively, while discordance was above 20% between all other scores.

Table 5a. Accuracy indices the clinical decision rules alone in 807 patients with a suspected event.

	Wells rule	Simplified Wells rule	RGS	Simplified RGS
Sensitivity, n, % (95% CI)	99/192, 52 (45-59)	125/192, 65 (58-72)	101/192, 53 (46-60)	94/192, 49 (42-56)
Specificity, n, % (95% CI)	491/615, 80 (77-83)	432/615, 70 (67-74)	462/615, 75 (72-79)	478/615, 78 (74-81)
NPV, n, % (95% CI)	491/584, 84 (81-87)	432/499, 87 (84-90)	462/553, 84 (81-87)	478/576, 83 (80-86)

RGS: revised Geneva score; CI: confidence interval; NPV: negative predictive value; PE: pulmonary embolism.
Sensitivity: the number of patients correctly identified as having PE by the CDR alone (independent of D-dimer results), divided by the total number of patients with proven PE identified by CT scan at the time of initial evaluation or VTE during 3-month follow-up.
Specificity: the number of patients correctly identified as not having PE by the CDR alone (independent of D-dimer results), divided by the total number of patients in whom PE was excluded by CT scan at the time of initial evaluation or VTE during 3-month follow-up.
NPV: the number of patients correctly identified as not having PE based on the CDR alone (independent of D-dimer results); divided by the total number of patients with CDR unlikely.
VTE: venous thromboembolism (i.e. PE and deep vein thrombosis).

Sensitivity was significantly different between the Wells rule and the simplified Wells rule ($p<0.001$); the simplified Wells rule and the RGS ($p=0.001$); the simplified Wells rule and the simplified RGS ($p<0.001$); and the RGS and simplified RGS ($p=0.039$). Other differences in sensitivity were not statistically significant.
Specificity was significantly different between the Wells rule and the simplified Wells rule ($p<0.001$); the Wells rule and the RGS ($p=0.020$); the simplified Wells rule and the RGS ($p=0.011$); the simplified Wells rule and the simplified RGS ($p<0.001$); and the RGS and the simplified RGS ($p<0.001$). Other differences in specificity were not statistically significant.

Despite the discordant scores, PE was not missed in any of the patients from the discordant group who had a normal D-dimer (Table 6b). Therefore, the scores performed equally well in excluding PE when combined with a D-dimer.

Inpatients

Both inpatients and outpatients were included in the study. The proportions of inpatients who were categorized as "unlikely" were as follows: 37% using the simplified Wells rule, 48% using the revised Geneva rule, 50% using the simplified revised Geneva rule and 57% using the Wells rule. These proportions were smaller compared to the proportions of outpatients categorized as "unlikely": 68%, 74%, 77% and 76% for the four CDRs, respectively (multiple tests, all with $p< 0.01$).

The failure rate of excluding PE based on an "unlikely" CDR and normal D-dimer test was similar for both inpatients and outpatients with all four CDRs. However, the proportion of inpatients in which PE could be excluded non-invasively was very low: only three inpatients using the simplified Wells rule (3/163; 1.8%); four patients using the Wells rule (2.5%); and 5 patients using the revised Geneva score and the simplified revised Geneva

Table 4/5b. Accuracy indices of the clinical decision rules in combination with a normal D-dimer test in patients with a suspected event.

	Original Wells rule N=796*	Simplified Wells rule N=803*	RGS N=796*	Simplified RGS N=795*
Sensitivity, n, % (95% CI)	190/191 99.5 (97-100)	191/192, 99.5 (97-100)	188/189, 99.5 (97-100)	187/188, 99.5 (97-100)
Specificity, n, % (95% CI)	183/605, 30 (27-34)	177/611, 29 (25-33)	184/607, 30 (27-34)	189/607, 31 (28-34)
NPV, n, % (95% CI)	183/184, 99.5 (97-100)	177/178, 99.4 (97-100)	184/185, 99.5 (97-100)	189/190, 99.5 (97-100)

RGS: revised Geneva rule; CI: confidence interval; NPV: negative predictive value; PE: pulmonary embolism.
*Patients with a CDR indicating "PE unlikely" but in whom the D-dimer result was missing (protocol violation) were not included in this analysis, this number differed between the four CDRs.
Sensitivity: the number of patients correctly identified as having PE by the combination of CDR and D-dimer testing, divided by the total number of patients with proven PE identified by CT scan at the time of initial evaluation or VTE during 3-month follow-up.
Specificity: the number of patients correctly identified as not having PE by the combination of CDR and D-dimer testing, divided by the total number of patients in whom PE was excluded by CT scan at the time of initial evaluation or VTE during 3-month follow-up.
NPV: the number of patients correctly identified as not having PE by the combination of CDR and D-dimer testing, divided by the total number of patients with CDR/D-dimer combination indicating PE excluded.
VTE: venous thromboembolism (i.e. PE and deep vein thrombosis).

Sensitivities did not differ between the four CDRs in combination with D-dimer test. Specificity was significantly different between the Wells rule and the simplified Wells rule (p=0.031) and the simplified Wells rule and simplified RGS (p=0.017). Other differences in specificity were not statistically significant.

score (3.1%). No failures occurred in the inpatients in whom PE was excluded without the need for CTPA.

Stratification by academic versus non-academic hospitals

In total, 5 academic hospitals included 598 (74%) patients, while the 2 non-academic hospitals included 209 patients (26%). The demographic characteristics as described in Table 2 did not differ for patients from academic versus non-academic hospitals, except for malignancy (16% vs. 8.1%, p<0.001) and recent surgery or immobilization (26% vs. 11%, p<0.001). Adjusting the results for academic and non-academic hospitals, a correct categorization of probability (categorization as "unlikely" or "likely" with respect to the outcome of PE) was found more often at non-academic sites. A correct categorization of 66% up to 71% was found at academic hospitals versus 75% up to 79% at non-academic hospitals; p<0.001 for all four CDRs.

Table 6a. Discordances between the categorization in "unlikely" and "likely" clinical probability groups according to four clinical decision rules in 807 patients with suspected pulmonary embolism. In total, 243 patients had discordant results.

	Wells "likely" (n=223)		**SW "likely"** (n=308)		**RGS "likely"** (n=254)		**SRGS "likely"** (n=231)	
	n	n with PE	n	n with PE	n	n with PE	n	n with PE
Wells "unlikely" (n=584)	X		85	25	115	26	100	23
SW "unlikely" (n=499)	0	0	X		65	14	51	11
RGS "unlikely" (n=553)	84	24	119	37	X		1	1
SRGS "unlikely" (n=576)	92	28	128	41	24	8	X	

Wells: Original Wells rule; SW: simplified Wells Rule; RGS: Revised Geneva score; SRGS: simplified revised Geneva score; PE: pulmonary embolism.
The number of patients with discordant CDR results when two CDRs are compared can be calculated by adding the number of patients with an "unlikely" score according to one CDR, but a "likely" score according to the other CDR, to the number of patients with a "likely" score according to the first CDR but an "unlikely" score according to the second CDR. For instance: to find the number of patients with discordant results comparing the RGS with the simplified RGS: There are 24 patients with a "likely" RGS result who have an "unlikely" simplified RGS result. Also, there is one patient with an "unlikely" RGS results but with a "likely" simplified RGS result. This means there is a total 24 + 1 = 25 patients with discordances when the RGS and simplified RGS are compared, out of a total of 807 patients (3.1%).

Table 6b. Discordances between the categorization in "unlikely" and "likely" clinical probability groups according to four clinical decision rules in 205 patients with suspected pulmonary embolism and a normal D-dimer test result. In total, 29 patients had discordant clinical decision rule results.

	Wells "likely" (n=21)		**SW "likely"** (n=27)		**RGS "likely"** (n=20)		**SRGS "likely"** (n=15)	
	n	n with PE	n	n with PE	n	n with PE	n	n with PE
Wells "unlikely" (n=184)	X		6	0	12	0	8	0
SW "unlikely" (n=178)	0	0	X		9	0	5	0
RGS "unlikely" (n=185)	13	0	16	0	X		0	0
SRGS "unlikely" (n=190)	14	0	17	0	5	0	X	

Wells: Original Wells rule; SW: simplified Wells Rule; RGS: Revised Geneva score; SRGS: simplified revised Geneva score; PE: pulmonary embolism.

DISCUSSION

This accuracy study directly compared four CDRs for the probability assessment of PE and showed that the CDRs are similar in 1) their ability to categorize patients in an "unlikely" and "likely" clinical probability group; 2) the proportion of patients in whom CTPA was not indicated on the basis of an "unlikely" CDR result and a normal D-dimer test and 3) the 3-month failure rate for VTE in the patients in whom PE was excluded by CDR and D-dimer testing. Importantly, although discordances in the categorization of patients in an "unlikely" or "likely" group by the scores were present in 30% of the

patients, this did not result in a difference in failure rates when the CDR was combined with the D-dimer.

Our results are important and relevant for clinical practice. Despite the debate on the subjective variable in the Wells rule, in this direct comparison, the Wells rule and simplified Wells rule showed to be equivalent in performance compared to the fully objective revised Geneva score. Additional to the comparison of these two rules, we were able to validate the performance of the recently introduced simplifications of the Wells rule and revised Geneva score. Both simplified scores had similar diagnostic performance to their original and extensively validated versions. Despite discordances between the CDR outcomes in 30% of patients, there was no difference in safety using a management strategy based on any of the CDRs combined with D-dimer testing. This equal performance could be explained by the use of a highly sensitive D-dimer test in patients with a CDR indication of "PE unlikely".

The importance of clinical probability estimation has been emphasized on many occasions.[14-18] Although the D-dimer is a sensitive assay in the diagnosis of PE, false negative results are more likely to occur when the pre-test clinical probability is high.[18] This prospective validation of the simplified CDRs has relevant practical implications, for they enable easier computation of the clinical probability score, which in turn could lead to better implementation of CDR use in daily clinical care. Our findings are in line with previous studies using these CDRs in a two-category scheme.[14] With the Wells rule, 51-84% of patients were categorized as "unlikely" in previous reports, with prevalence of PE ranging from 3.4-12%, compared to 72% categorized as "unlikely" in this study with a prevalence of PE of 15%. Using the simplified Wells rule, the proportion of "PE unlikely" patients was slightly lower in our cohort (62%) compared to a previous validation study (78%), but prevalence of PE in this "unlikely" group was comparable (13% in both studies).[10] Similarly, in an earlier retrospective study, the simplified revised Geneva score classified 65% of the patients as "unlikely" compared to 62% in the current analysis, with PE prevalence of 13% in the previous study and 16% in our study.[8] As this is the first study to report a two-category scheme for the revised Geneva score, we cannot compare our data of the two-category revised Geneva score with previous findings. However, the 69% patients with an "unlikely" CDR result according to the revised Geneva score as well as the 16% prevalence of PE in this group overlap well with the results from the other three decision rules.

Four highly sensitive but different D-dimer assays were used. Because the CDRs were determined in all patients, the types of D-dimers assays were equally represented in the different CDR groups, which enable comparison of the CDRs, irrespective of the D-dimer assay. Type of assay was not based on randomization, but was dependent on the prefer-

ence of the study center. Unfortunately, therefore, comparisons between the various D-dimer assays are limited by the sample sizes.

After several retrospective or small prospective comparisons, this is the first large study to directly compare the most widely used CDRs (original Wells rule and original Revised Geneva score) in the diagnostic management of PE. Furthermore, this study prospectively validated the performance of the recently introduced simplified Wells rule and simplified revised Geneva score. Calculation of the scores in all patients allowed a direct comparison of the CDRs in a single patient population. Also, due to the computer-aided design of the study, calculation errors were minimized. Likewise, this use of a computer program to guide the physician on the next step in the diagnostic algorithm excluded the possibility of allowing the physician's preference for a certain CDR to influence the management of a patient. There are several arguments for the results of our study to be applicable in a wide range of clinical settings. First, the clinical characteristics of the patients in the study are comparable with those in other population-based studies,[2,4] and the 23% prevalence of PE in this cohort is comparable to other reports.[2,4,6] Second, consecutive patients were included from both academic and non-academic medical centers.

Several potential limitations of our study require comment. First, a randomized controlled trial between the four CDRs is an alternative study design, but in view of the reasonably high concordance rates would likely have been very inefficient. In addition, by study design, CT scans were performed in all patients with discordant CDRs and ensured that an imaging diagnosis was available in all those patients. The diagnostic protocol was violated in four patients, in whom CT was not performed despite discordance of the CDRs. Three-month follow-up, however, was uneventful in these patients. Second, we did not manage on one of the separate CDRs in combination with D-dimer testing but the combination of the four CDRs and D-dimer testing. According to the protocol, all patients with discordant CDR results underwent CT scanning. The majority of these patients had elevated D-dimer levels and would have an indication for CT-scanning, even if just one of the CDRs was used for decision making. Only patients with discordant results and a normal D-dimer level (29; 3.6% of the included patients), did not have an indication for a CT scan using one of the separate rules combined with D-dimer testing; they underwent CT scanning because at least one of the other rules indicated PE "likely". Third, use of a computerized decision-support system improves the diagnosis of pulmonary embolism.[19] However, in daily clinical practice, these systems may not be widely available and our results may therefore differ from a setting in which more miscalculations are possible.

Finally, although both inpatients and outpatients were included in this study and no failures occurred in the patients in whom the diagnosis could be excluded, we are

unable to validate that any of the CDRs/D-dimer combinations can safely exclude the diagnosis in inpatients.

Further research may include an outcome study using one of the simplified CDRs in combination with D-dimer testing. Since patients with suspected recurrent PE were not included, the performance of the CDRs in this group will need additional research.

In conclusion, the Wells rule, the revised Geneva score, the simplified Wells rule as well as the simplified revised Geneva score, in combination with a D-dimer test, all performed similarly in the exclusion of acute PE. This prospective validation indicates that the simplified, more straightforward CDRs may be used in clinical practice. Which rule a physician will use should depend on local preference and acquaintance, in order to accomplish correct use of the CDR and prevent miscalculations.

REFERENCES

1. Wells PS, Anderson DR, Rodger M, Stiell I, Dreyer JF, Barnes D, 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.
2. The Christopher study investigators, van Belle A, Buller HR, Huisman MV, Huisman PM, Kaasjager K, et al. Effectiveness of managing suspected pulmonary embolism using an algorithm combining clinical probability, D-dimer testing, and computed tomography. JAMA. 2006;295:172-9.
3. Perrier A, Roy PM, Sanchez O, Le Gal G, Meyer G, Gourdier AL, et al. Multidetector-row computed tomography in suspected pulmonary embolism. N Engl J Med. 2005;352:1760-8.
4. Ten Cate-Hoek AJ, Prins MH. Management studies using a combination of D-dimer test result and clinical probability to rule out venous thromboembolism: a systematic review. J Thromb Haemost. 2005;3:2465-70.
5. 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. Thromb Haemost. 2000;83:416-20.
6. Le Gal G, Rigini M, Roy PM, Sanchez O, Aujesky D, Bounameaux H, Perrier A. Prediction of pulmonary embolism in the emergency department: the revised Geneva score. Ann Intern Med. 2006; 144:165-71.
7. Gibson NS, Sohne M, Kruip MJ, Tick LW, Gerdes VE, Bossuyt PM, et al. Further validation and simplification of the Wells clinical decision rule in pulmonary embolism. Thromb Haemost. 2008; 99:229-34.
8. Klok FA, Mos IC, Nijkeuter M, Righini M, Perrier A, Le Gal G, Huisman MV. Simplification of the revised Geneva score for assessing clinical probability of pulmonary embolism. Arch Intern Med. 2008;168:2131-6.
9. Klok FA, Kruisman E, Spaan J, Nijkeuter M, Righini M, Aujesky D, et al. Comparison of the revised Geneva score with the Wells rule for assessing clinical probability of pulmonary embolism. J Thromb Haemost. 2008;6:40-4.
10. Douma RA, Gibson NS, Gerdes VE, Büller HR, Wells PS, Perrier A, Le Gal G. Validity and clinical utility of the simplified Wells rule for assessing clinical probability for the exclusion of pulmonary embolism. Thromb Haemost. 2009;101:197-200.
11. Calisir C, Yavas US, Ozkan IR, Alatas F, Cevik A, Ergun N, Sahin F. Performance of the Wells and Revised Geneva scores for predicting pulmonary embolism. Eur J Emerg Med. 2009;16:49-52.
12. Wong DD, Ramaseshan G, Mendelson RM. Comparison of the Wells and revised Geneva score for the diagnosis of pulmonary embolism: an Australian experience. Intern Med J. 2010.
13. Gardner, M.J., & Altman, D.G. Statistics with confidence –Confidence intervals and statistical guidelines. London: British Medical Journal, 1989 (version 1.0)
14. Ceriani E, Combescure C, Le Gal G, Nendaz M, Perneger T, Bounameaux H, et al. Clinical prediction rules for pulmonary embolism: a systematic review and meta-analysis. J Thromb Haemost. 2010; 8:957-70.

15. Roy PM, Meyer G, Vielle B, Le Gall C, Verschuren F, Carpentier F, et al. Appropriateness of diagnostic management and outcomes of suspected pulmonary embolism. Ann Intern Med. 2006;144: 157-64.

16. Kelly J, Hunt BJ. A clinical probability assessment and D-dimer measurement should be the initial step in the investigation of suspected venous thromboembolism. Chest. 2003;124:1116-9.

17. Righini M, Aujesky D, Roy PM, Cornuz J, de Moerloose P, Bounameaux H, Perrier A. Clinical usefulness of D-dimer depending on clinical probability and cutoff value in outpatients with suspected pulmonary embolism. Arch Intern Med. 2004;164:2483-7.

18. Gibson NS, Sohne M, Gerdes VE, Nijkeuter M, Buller HR. The importance of clinical probability assessment in interpreting a normal d-dimer in patients with suspected pulmonary embolism. Chest. 2008;134:789-93.

19. Roy PM, Durieux P, Gillaizeau F, Le Gall C, Armand-Perroux A, Martino L, et al. A computerized handheld decision-support system to improve pulmonary embolism diagnosis: a randomized trial. Ann Intern Med. 2009;151:677-86.