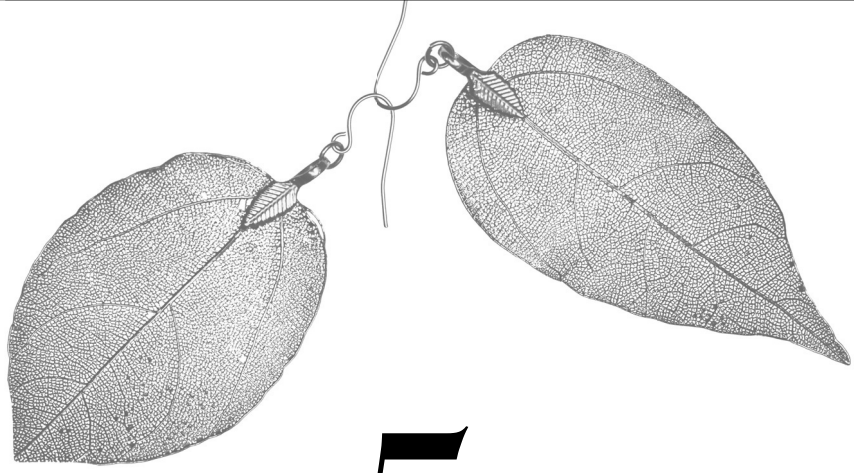
Adjustments in the diagnostic work-up, treatment and prognosis of pulmonary embolism

van Es, J.

Publication date
2013



CHAPTER 5

The accuracy of D-dimer testing in suspected pulmonary embolism varies with the Wells score

J. van Es
L.F.M. Beenen
V.E.A. Gerdes
S. Middeldorp
R.A. Douma
P.M.M. Bossuyt

Journal of Thrombosis and Haemostasis 2012

Abstract

Background: In patients suspected for pulmonary embolism (PE), D-dimer testing is often used in combination with the Wells score to exclude pulmonary embolism, with a fixed D-dimer positivity threshold of 500 μg/L. We evaluated whether its diagnostic accuracy varies with the Wells score.

Methods: Data were collected in a multicentre cohort study of consecutive patients with clinically suspected PE. CT-scan results and follow-up were used as the clinical reference standard. We estimated D-dimer sensitivity, specificity and area under the ROC curve (AUC) for subgroups of patients, defined by having no, one or two items (including the subjective item: 'alternative diagnosis less likely than PE') positive on the Wells score.

Results: Data from 723 patients could be analysed, of which 177 (24%) patients had zero items on the Wells score, 300 (41%) had one item positive, and 136 (19%) had two items positive, including the subjective item. The overall diagnostic accuracy of the D-dimer test, as expressed by the AUC, was 0.92 (95% CI: 0.82 to 0.99) in the subgroup of patients with a zero Wells score. This AUC was significantly higher than the AUC in the subgroup with one item positive (0.78; 95% CI 0.72 to 0.84) and the subgroup with two items positive (0.78; 95% CI 0.70 to 0.86). The estimated sensitivity for the D-dimer test at the 500 μg/L positivity threshold was 1.00, 0.98 and 1.00 in the three subgroups respectively. Specificity was significantly different, at 0.49, 0.30 and 0.16 respectively.

Conclusion: The diagnostic accuracy of D-dimer testing varies significantly across subgroups defined by the Wells score. The D-dimer test positivity threshold could be adapted to keep specificity more comparable across all values on the Wells score.

Introduction

Although the incidence of pulmonary embolism (PE) increases with age from 1 per 10,000 in childhood to nearly 8 per 1000 in older-aged patients (1), the prevalence of objectively proven PE in patients with suspected PE is relatively low: between 20-30% (2;3). Computerized tomographic (CT-) scanning in suspected PE has a high sensitivity, ranging between 64-100%, with a specificity between 89-100% (4), however CT-scans have also adverse effects, such as an increased lifetime risk of malignancy from radiation exposure and the risk of contrast nephropathy (5;6). These concerns force physicians to use a diagnostic strategy in suspected PE that results in the lowest possible number of CT-scans.

The overall sensitivity of the D-dimer test in detecting PE, using a cut-off value of 500µg/L is estimated at nearly 100%; its reported specificity, however, varies between 30-40% (7;8). Standardised clinical decision rules (CDRs) are therefore additionally used in suspected PE. One frequently used, is the Wells score, which consists of items obtained from history (risk factors for PE), physical examination (such as heart rate and signs of deep venous thrombosis (DVT)) and a subjective item, where the physician can judge whether an alternative diagnosis is more likely than PE (9). Recently, the simplified version of the Wells score (all items are assigned one point) was found to have a performance similar to that of the original rule in combination with a high sensitive D-dimer test (3;7). An 'unlikely' Wells score combined with a negative D-dimer test safely excludes PE in 20-40% of patients with suspected PE (10). Of the remaining patients in whom a CT-scan should be performed in order to either exclude or confirm PE, only 20-30% has a positive CT-scan for PE (11). CT-scanning is thus indicated in case of a 'likely' Wells score or a positive D-dimer test (3).

It is well known that the diagnostic accuracy of a test can vary with the strength of clinical suspicion (12). In that case, previously reported accuracy estimates may not be generalizable to all subgroups. We investigated if this is also the case in D-dimer testing in suspected PE. We compared the accuracy of the D-dimer in subgroups of patients with 0, 1 or 2 items of the Wells score positive.

Methods

For this purpose we performed an analysis of data from a prospective multicentre cohort study in patients with clinically suspected PE, reported in detail elsewhere (13). In brief, the study population consisted of consecutive in- and outpatients in whom acute PE was suspected. The dichotomized Wells score (cut-off value 4), simplified Wells score (cut-off value 1), and dichotomized Revised Geneva Score (RGS) (cut-off value 5) and simplified RGS (cut-off value 2) were calculated in each patient and combined with a high-sensitivity quantitative D-dimer test. A CT-scan was performed in patients with a 'likely' outcome of at least one of the CDRs, or if the D-dimer was elevated (> 500 μg/L). In case of an unlikely CDR with all scores and a normal D-dimer result, no CT-scan was performed. Patients were followed for 3 months by telephone; they were instructed to return to the hospital if symptoms of PE or DVT occurred and imaging diagnostic tests were done if PE or DVT was suspected (13).

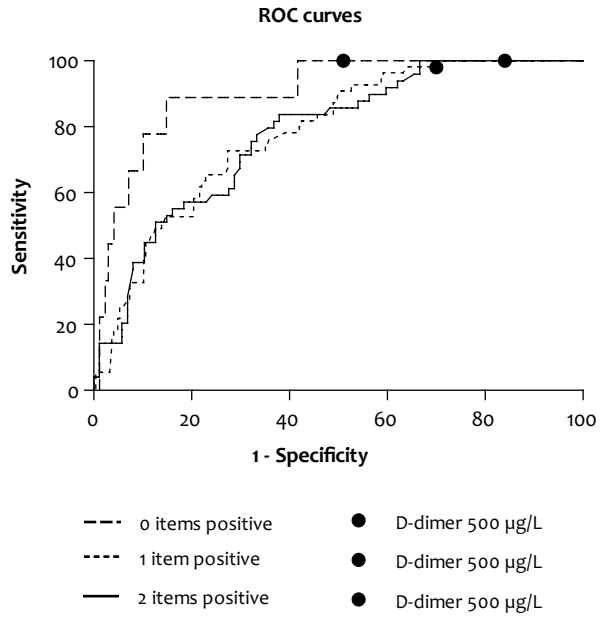
The mean age of the 807 included patients with suspected PE was 52 years; 61% were females. A D-dimer result was available in 723 (90%), of whom 156 patients had PE (22%). Median duration of symptoms was 2 days (inter quartile range 1-7); the mean body mass index was 26 (standard deviation 5).

We defined 3 groups of patients, based on the number of positive items on the Wells score. The first group had none of the Wells items positive. The second group had only 1 item positive; since our analysis focused on patients with a low pre-test probability, it did not matter which item was positive. Finally we assessed patients with 2 positive Wells items, 1 of which was the subjective item, as it is known that when physicians are aware of a positive D-dimer, they are more likely to consider the subjective item 'alternative diagnosis less likely than PE' as positive, regardless of the clinical situation (14). Consequently, in clinical practice, there sometimes is only one item positive officially, combined with a 'false positive' subjective item (14).

To express the diagnostic accuracy of the D-dimer in each subgroup, we calculated the area under the Receiver Operating Characteristics curve (AUC). To evaluate the significance of difference in the AUC, the z-test statistic was used. We also calculated the sensitivity, specificity, positive and negative predictive values, and positive and negative likelihood ratio of the 500 μg/L D-dimer cut-off in each group. All analyses were performed using SPSS software, version 19.0 (SPSS Inc, Chicago, Ill). A significance level of 0.05 was used in all analyses.

A total of 177 patients had zero positive items on the Wells score, of whom 9 patients had PE (5%). The AUC of the D-dimer test in this subgroup was 0.91 (95% CI 0.82-0.99) (Figure 1). Another 300 patients had one item positive: this was PE as the most likely diagnosis in 64%, being immobilized within 4 weeks prior to enrolment in 9%, malignancy in 7%, heart rate above 100 beats per minute in 11%, clinical suspicion of DVT in 3%, and 3% had haemoptysis. In this subgroup of 300 patients 55 had PE (18%). The AUC of the D-dimer test in this subgroup was significantly lower than in the group with zero items positive: 0.78 (95% CI 0.72-0.84; p=0.020) (Figure 1).

Figure 1. Receiver operating characteristic curves illustrating the diagnostic performance of the D-dimer in combination with Wells scores of 0, 1 and 2 items scored as positive.



The areas under the curve were 0.91 (95% confidence interval (CI) 0.82-0.99), 0.78 (95% CI 0.72-0.84) and 0.78 (95% CI 0.70-0.86) respectively.

A total of 136 patients had 2 items positive, including the subjective item. Of these patients 49 had PE (36%). The AUC in this subgroup was also significantly lower than in the subgroup with a zero score: 0.78 (95% CI 0.70-0.86; p=0.031). The AUC of the D-dimer test was not significantly different between the patients with one and those with 2 Wells items positive (p=0.42) (Figure 1).

The sensitivity of D-dimer at the conventional 500 μg/L cut-off in patients with a zero Wells score was 100% (95% CI 69%-100%). This estimated sensitivity was comparable with that in patients with 1 (98%, 95% CI 90%-100%) or 2 items positive (100%, 95% CI 93%-100%). However, the specificity was lower with increasing number of items positive: 49% (95% CI 42%-75%), 30% (95% CI 25%-37%) and 16% respectively (95% CI 10%-25%). In addition, the positive predictive value differed significantly between the 0, 1 and 2 item groups: 10%, 24% and 33% ($p < 0.0001$). The negative predictive value did not: 100%, 99% and 100% ($p=0.5$).The positive likelihood ratio in patients with 0, 1 and 2 Wells items positive was 2.0, 1.4 and 1.2, respectively. The negative likelihood ratio, however, was 0.0 for all groups.

Discussion

We showed that there is a significant difference in the diagnostic accuracy of the D-dimer test depending on the Wells score. The overall accuracy is significantly lower in subgroups with higher Wells score. At the current cut-off value, D-dimer specificity is more than halved in patients with two items positive compared to those with a zero Wells score.

Since the accuracy of the D-dimer test varies, one could consider adapting the D-dimer positivity threshold to make the behaviour similar in subgroups defined by the Wells score. Previous studies have demonstrated that a D-dimer cut-off value, adapted to the clinical probability category of the patient, has a better accuracy in excluding PE than using a single D-dimer cut-off value, regardless of the clinical probability (15-17). In those studies, the proposed cut-off value was kept at 500 μg/L for patients with an intermediate clinical suspicion of PE, but doubled (1000 μg/L) in patients with a low clinical suspicion, and halved (250 μg/L) in patients with a high clinical suspicion of PE. In a study by Kabrhel and colleagues, for example, the conventional cut-off of 500 μg/L had an overall sensitivity of 94% (95% CI 91-97) at a specificity of 58% (95% CI 56-60), and a NPV of 99.5% (95% CI 99.1- 99.7). When probability-dependent cut-offs were used, the overall sensitivity became 88% (95% CI 83-92), specificity was 75% (95% CI 74-76), and NPV was 99.1% (95% CI 98.7- 99.4) (15).

Despite the statistical significance of the differences in accuracy, the subgroups in the current analysis were of moderate size, especially the one of patients with zero positive items, making estimates in subgroups somewhat imprecise. In 84 patients no D-dimer was performed (13). The reason for these missing D-dimer tests is unclear, but

it is possible that these patients had a 'likely' result on the Wells rule, in which case a D-dimer test was not necessary to indicate CT-scan as the next diagnostic step (7). Since we also included patients with 2 items positive (including the subjective item counting for 3 points), we have also included patients with a 'likely' clinical probability according to the Wells rule in this group of patients (3).

In conclusion, the diagnostic accuracy of D-dimer testing varies significantly across subgroups defined by the Wells score. In particular, specificity is significantly lower in subgroups with one or more positive items of the Wells score. At the current cut-off value, D-dimer specificity is more than halved in patients with two items positive compared to those with none of the Wells items positive. These findings could be taken into account in improving the diagnostic strategy in patients with suspected PE.

Reference List

(1) Silverstein MD, Heit JA, Mohr DN, Petterson TM, O'Fallon WM, Melton LJ, III. Trends in the incidence of deep vein thrombosis and pulmonary embolism: a 25-year population-based study. Arch Intern Med 1998 Mar 23;158(6):585-93.

(2) Le Gal G, Righini M, Roy PM, Meyer G, Aujesky D, Perrier A, et al. Differential value of risk factors and clinical signs for diagnosing pulmonary embolism according to age. J Thromb Haemost 2005 Nov;3(11):2457-64.

(3) van Belle A, Buller HR, Huisman MV, Huisman PM, Kaasjager K, Kamphuisen PW, et al. Effectiveness of managing suspected pulmonary embolism using an algorithm combining clinical probability, D-dimer testing, and computed tomography. JAMA 2006 Jan 11;295(2):172-9.

(4) Rathbun SW, Raskob GE, Whitsett TL. Sensitivity and specificity of helical computed tomography in the diagnosis of pulmonary embolism: a systematic review. Ann Intern Med 2000 Feb 1;132(3):227-32.

(5) Einstein AJ, Henzlova MJ, Rajagopalan S. Estimating risk of cancer associated with radiation exposure from 64-slice computed tomography coronary angiography. JAMA 2007 Jul 18;298(3):317-23.

(6) Hurwitz LM, Reiman RE, Yoshizumi TT, Goodman PC, Toncheva G, Nguyen G, et al. Radiation dose from contemporary cardiothoracic multidetector CT protocols with an anthropomorphic female phantom: implications for cancer induction. Radiology 2007 Dec;245(3):742-50.

(7) 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 Oct;134(4):789-93.

(8) Douma RA, van Sluis GL, Kamphuisen PW, Sohne M, Leebeek FW, Bossuyt PM, et al. Clinical decision rule and D-dimer have lower clinical utility to exclude pulmonary embolism in cancer patients. Explanations and potential ameliorations. Thromb Haemost 2010 Oct;104(4):831-6.

(9) Wells PS, Ginsberg JS, Anderson DR, Kearon C, Gent M, Turpie AG, et al. Use of a clinical model for safe management of patients with suspected pulmonary embolism. Ann Intern Med 1998 Dec 15;129(12):997-1005.

(10) Carrier M, Righini M, Djurabi RK, Huisman MV, Perrier A, Wells PS, et al. VIDAS D-dimer in combination with clinical pre-test probability to rule out pulmonary embolism. A systematic review of management outcome studies. Thromb Haemost 2009 May;101(5):886-92.

(11) 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 Apr 28;352(17):1760-8.

(12) Irwig L, Bossuyt P, Glasziou P, Gatsonis C, Lijmer J. Designing studies to ensure that estimates of test accuracy are transferable. BMJ 2002 Mar 16;324(7338):669-71.

(13) Douma RA, Mos IC, Erkens PM, Nizet TA, Durian MF, Hovens MM, et al. Performance of 4 Clinical Decision Rules in the Diagnostic Management of Acute Pulmonary Embolism: A Prospective Cohort Study. Ann Intern Med 2011 Jun 7;154(11):709-18.

(14) Douma RA, Kessels JB, Buller HR, Gerdes VE. Knowledge of the D-dimer test result influences clinical probability assessment of pulmonary embolism. Thromb Res 2010 Oct;126(4):e271-e275.

(15) Kabrhel C, Mark Court, Camargo CA, Jr., Moore CL, Richman PB, Plewa MC, et al. Potential impact of adjusting the threshold of the quantitative D-dimer based on pretest probability of acute pulmonary embolism. Acad Emerg Med 2009 Apr;16(4):325-32.

(16) Linkins LA, Bates SM, Ginsberg JS, Kearon C. Use of different D-dimer levels to exclude venous thromboembolism depending on clinical pretest probability. J Thromb Haemost 2004 Aug;2(8):1256-60.

(17) Kline JA, Hogg MM, Courtney DM, Miller CD, Jones AE, Smithline HA. D-dimer threshold increase with pretest probability unlikely for pulmonary embolism to decrease unnecessary computerized tomographic pulmonary angiography. J Thromb Haemost 2012 Apr;10(4):572-81.