Comparison of Diagnostic Accuracies in Outpatients and Hospitalized Patients of D-Dimer Testing for the Evaluation of Suspected Pulmonary Embolism

John E. Schrecengost, Robin D. LeGallo, James C. Boyd, Karel G.M. Moons, Steven L. Gonias, C. Edward Rose, Jr., and David E. Bruns

Background: The ability of various D-dimer assays to exclude the diagnosis of thromboembolic diseases is controversial. We examined the diagnostic accuracy of two D-dimer methods in hospitalized patients and outpatients.

Methods: We studied consecutive patients for whom D-dimer testing was ordered for investigation of suspected pulmonary embolism. We measured D-dimer by an ELISA (VIDAS D-dimer) and an enhanced micro-latex immunoassay method (Diagnostica Stago STA Liatest D-di). Patient diagnoses were based on imaging studies or, when these were not performed, on follow-up by review of medical records 3 months later.

Results: We examined 233 hospitalized patients and 234 outpatients with a mean age of 58 years (range, 1–92 years) and a female-to-male ratio of 1.4 to 1. Thromboembolism was present in 8% of outpatients and 12% of hospitalized patients. In outpatients, the negative predictive values were 98% [95% confidence interval (CI), 93–100%] and 99% (94–100%) for the microlatex and ELISA methods, respectively, at the recommended cut-offs. Areas under the ROC curves were similar for the two methods [0.77 (95% CI, 0.67–0.87) and 0.81 (0.73–0.89), respectively]. By contrast, in hospitalized patients, the confidence intervals for the areas under the ROC curves included 0.5 [0.60 (95% CI, 0.50–0.71) and 0.56 (0.44–0.67)].

Conclusions: For hospitalized patients, in contrast to outpatients, the diagnostic accuracy of D-dimer testing for pulmonary embolism is poor in a tertiary care setting, presumably reflecting thrombosis and comorbidities, other than pulmonary embolism, that increase the D-dimer concentrations in these patients. The patient population studied appears more important than assay method in studies of the diagnostic accuracy of D-dimer testing.

D-Dimer testing has been used to aid in the diagnosis of disseminated intravascular coagulation (1–3). Recently, however, these tests have been studied for the exclusion of thromboembolic disease in patients suspected of having pulmonary embolism (PE) or deep vein thrombosis (DVT) (4–6). In these conditions, D-dimer, an end-stage product of fibrin degradation by plasmin, is increased as a result of fibrinolysis, but it does not result from fibrinogenolysis. When plasma D-dimer is low, venous thromboembolism (VTE) is likely to be absent (4–6).

Because the 3-month mortality rate of untreated PE can be as high as 17.5% (7), the cost of missing the diagnosis is high. The cutpoint for D-dimer must be low if the test is to be used to exclude PE. False-positive results are thus frequent and are particularly common among individuals who have experienced recent trauma and those with pregnancy or malignancy. These states may be associated with accelerated coagulation and fibrinolysis without true...
large vessel thrombosis (1,8–10). Because the prevalence of these conditions is high among hospitalized patients, D-dimer testing may be of limited utility in hospitals. Moreover, false-positive results are seen in patients with a variety of other serious medical conditions that lead to hospitalization and can affect D-dimer metabolism, e.g., by decreasing its clearance from the circulation. Patients with false-positive D-dimer tests are commonly referred for additional expensive imaging tests; thus, high false-positive rates can increase healthcare costs (11). Nonetheless, in our health center, we observed that the D-dimer assay was used to investigate suspected VTE in hospitalized patients as often as it was used for ambulatory outpatients, in whom the test is less likely to be affected by comorbidities.

To compare the potential utility of D-dimer testing in outpatients and inpatients, we studied its diagnostic accuracy in these groups separately. Because no consensus exists on which D-dimer assay to use, we used two methods that represent the two classes of D-dimer assays that have been reported to achieve high negative predictive values [see, for examples, Refs. (4,5,12)].

**Materials and Methods**

**Patients**

We included all patients seen between May and December 2001 in whom there was a clinical suspicion of PE (which consisted mainly of shortness of breath, chest pain, or syncopal episode) and for whom D-dimer testing was requested to investigate possible VTE. An ELISA D-dimer method was offered, as part of routine testing, in this medical center specifically for the investigation of suspected VTE. Medical records were reviewed to determine the suspicion of PE and the indication for the test. All treating physicians had available to them a clinical probability algorithm to guide their requests for further diagnostic testing, including D-dimer assays, in each patient (13). This algorithm dictates use of the D-dimer assay when clinical suspicion is sufficient to warrant further investigation but is not high enough to warrant immediate use of emergent diagnostic imaging studies. We did not attempt to document use of the clinical probability algorithm. For all ELISA tests during the study period, a portion of the sample was used for simultaneous testing by a comparison microlatex method (see below).

The study population comprised emergency room patients, patients from outpatient clinics, and hospitalized patients. Patients were excluded from the present analysis when either the ELISA or the microlatex agglutination result was not available, when the diagnosis of VTE was made despite negative imaging, and when the clinical presentation indicated the presence of DVT only (e.g., patients with a swollen extremity only). The requesting clinician received the result of only the ELISA assay during the patient evaluation, and not the microlatex assay result. The examined population included patients who were potentially on anticoagulation therapy for various conditions, including atrial fibrillation, cardiac valve replacement, or recent/remote history of PE or DVT (see below).

Clinical data were collected retrospectively by review of electronic medical records. We tabulated each patient’s age, sex, presenting symptoms and signs, significant medical history, and information on the final diagnosis, including the presence or absence of thromboembolic disease (see “Reference Standard” below). The results of the two D-dimer tests were retained in the laboratory and were added to the data collection form. The elapsed time between D-dimer testing and reference testing was <48 h in all cases. This study was approved by the Institutional Review Board (Human Investigation Committee) at the University of Virginia Health System.

The sponsor was not involved in the study design, data collection or analysis, or the decision to submit the manuscript for publication.

**D-Dimer Testing**

The same blood sample of each referred patient was used for the ELISA and for the microlatex method. Blood samples for D-dimer determinations were collected in plastic tubes containing 0.109 mol/L trisodium citrate at a ratio of nine parts blood to one part sodium citrate and were centrifuged for 10 min at 2000g. For any one sample, the same registered laboratory technologist performed both tests within 15 min of centrifugation of the sample. All D-dimer testing was performed by one of three technologists without knowledge of the other test result or the patient’s clinical outcomes or pretest clinical probability. D-Dimer was then assayed by the VIDAS (bioMerieux, Inc.) and LiaTest D-di (Diagnostica Stago) methods.

The VIDAS D-dimer ELISA assay, an enzyme-linked immunofluorescence assay, has been described elsewhere (14–16). D-Dimer results were reported in mg/L of fibrinogen equivalent units (FEU). Results above 1.0 mg/L were reported as >1.0 mg/L. The assay was recalibrated for each new lot of reagents and at least every 2 weeks. Control samples were tested every 8 h or with each instrument run if the time between runs exceeded 8 h. The imprecision (CV) of the assay was 5.8% at a mean concentration of 0.42 mg/L and 6.5% at a mean concentration of 0.69 mg/L.

The STA Liatest D-di assay is an automated microlatex immunnoassay performed on the Stago Star analyzer. The details of the method have been described previously (17,18). The results were reported as mg/L of FEU; values above 4.0 mg/L were initially reported by the instrument as >4.0 mg/L, but a repeat test with a 1:5 dilution extended the reportable range to 20 mg/L. Higher results were reported as >20 mg/L. Precalibrated reagent lots (barcoded) were used, and a new calibration curve was provided with each new lot of reagents. Controls were analyzed every 8 h or with each instrument run if the time between runs exceeded 8 h. The imprecision
(CV) of the assay was 5.0% at a mean concentration of 2.8 mg/L and 14% at a mean concentration of 0.31 mg/L.

Because the microlatex assay provided a wider range of reportable values than the ELISA, the microlatex assay had the potential to produce a misleading larger area under the ROC curve. To adjust for this potential artifact, we used Passing–Bablok regression analysis on the data set for ELISA values up to 0.999 mg/L to determine the approximate value of the microlatex method that corresponded to the upper reporting limit (1.000 mg/L) for the ELISA method. The regression analysis used 243 pairs of results. From the calculated slope and intercept, the microlatex value that corresponded to an ELISA result of 1.000 mg/L was 1.06 mg/L; thus, all higher results from the microlatex method were assigned a value of 1.06 mg/L for ROC analysis. Because the clinical cutoff values for these assays (~0.5 mg/L) were well below this value, the calculation appeared acceptable for all data analyses performed here.

REFERENCE STANDARD
A final diagnosis of VTE, i.e., a pulmonary embolus or DVT, in these patients, all of whom had symptoms suggestive of PE, required a positive result for VTE by one of the following executed imaging procedures, as documented in the electronic medical record: ventilation-perfusion scan of high probability (19); positive findings on either pulmonary angiogram or computed tomography angiogram (20); or the presence of a DVT on compression ultrasound studies of the extremities (21). In the absence of a positive result (e.g., negative or low-probability result), patients were classified as "negative", i.e., VTE absent. For those patients who did not undergo one of the above procedures or who underwent only lower extremity ultrasound, a 3-month follow-up examination of the medical record was used to determine whether they experienced VTE as indicated by follow-up visits or autopsy results or as suggested by patient death. The pathologist who reviewed the medical records had access to the D-dimer results.

STATISTICAL ANALYSIS
We plotted ROC curves and estimated the areas under the curves (ROC area) and their 95% confidence intervals (95% CIs). All analyses were performed with SPSS, Ver. 11.0. The ROC area was used to provide an overall estimate of the tests' diagnostic performance in differentiating between patients with and without VTE in the setting of suspected PE. Differences in diagnostic accuracy between the two test methods were estimated by use of the difference in ROC area, taking into account the correlation between the two test methods as they were based on the same individuals (22–25). Because the ROC area provides only an overall estimate of test diagnostic accuracy and does not directly indicate the performance of the test as used or perceived by physicians (26), we also tabulated the absolute number of correctly and incorrectly classified patients, calculating negative predictive value, positive predictive value, sensitivity, and specificity (and their corresponding 95% CIs) for each test at recommended cutoffs. We used cutoff values that were recommended in the corresponding manufacturer's product literature and the references therein; these thresholds were 0.5 mg/L FEU for the ELISA and 0.5 mg/L FEU for the microlatex assay.

Discordant results (microlatex-positive, ELISA-negative, or vice versa) were also analyzed over the entire patient population. This analysis was carried out using simple statistics to evaluate the possible introduction of verification bias into the study as a result of physicians' access to the ELISA, but not microlatex, results.

RESULTS
Of 504 patients enrolled in the study, we excluded 15 patients who had no microlatex result, 20 who had a clinical presentation of DVT with no evidence of PE, and 2 who were clinically diagnosed as having thromboembolic disease despite negative results by diagnostic testing (Fig. 1). The remainder of the patients had both D-dimer tests (Fig. 1). The mean age of the 467 patients included in the study (195 males and 272 females) was 58 years (range, 1–92 years). Review of the 3-month follow-ups in patients who had not undergone imaging studies revealed no deaths or evidence of PE at return visits.

Of the 467 patients, 48 (10%) had VTE (Table 1). Considering all 467 patients together, the ROC areas were 0.66 (95% CI, 0.58–0.74) for the microlatex method and 0.70 (0.63–0.77) for the ELISA method. These areas were not statistically significantly different. As shown in Table 2, the sensitivity, specificity, and predictive values were also statistically indistinguishable for the ELISA and microlatex methods. The total numbers of true-negative results at the recommended cutoff values were 149 (32%) for the ELISA method and 143 (31%) for the microlatex method. A total of six falsely negative results were identified by the microlatex method and four falsely negative results by the ELISA method.

Because concurrent anticoagulant therapy can decrease D-dimer concentrations, we reviewed the patients' medical histories to identify those patients whose results may have been affected by anticoagulant use. Because medical records may fail to capture prescription histories, we also identified all patients who had conditions that made them candidates for anticoagulation before the time that the D-dimer samples were collected. The conditions identified included atrial fibrillation, status post cardiac valve replacement, and a history of prior PE and/or DVT.

Records of anticoagulant use or indications for anticoagulation were found in 45 patients (10% of the patients studied). Of these 45 patients, D-dimer was <0.5 mg/L in 13 patients, including 2 of the 4 patients noted above with PE and falsely negative ELISA results. The microlatex results in both patients also were negative. One of these two patients was in the hospital on chronic warfarin...
therapy for atrial fibrillation; the other had been sent from another emergency department and had received heparin before arrival at our emergency department. If these two patients were to be excluded as inappropriate for testing, the negative predictive values of ELISA and microlatex tests would be 99% and 97%, respectively. If all 45 patients were excluded, the negative predictive values would be 99% and 97%.

The outpatient group consisted of 234 patients from the emergency room and outpatient clinics. Their mean age...
was 54 years (range, 15–90 years), with a female-to-male ratio of 1.98 to 1. The VTE prevalence in this group was 8%. The areas under the ROC curves (Fig. 2A) were 0.81 (95% CI, 0.73–0.89) for the ELISA and 0.77 (0.67–0.87) for the microlatex test, which were not statistically different. As shown in the 2×2 contingency table (Table 3B), one ELISA result was falsely negative, as were two microlatex assay results at the usually recommended cutpoint of 0.5 mg/L. A total of 109 (47%) patients were correctly “ruled-out” (negative D-dimer and absence of VTE) by the ELISA method and 103 (44%) by the microlatex method. The sensitivity, specificity, and predictive values for the two assays were not significantly different (Table 3A).

The group of 233 hospitalized patients included individuals on medical and surgical floors, as well as intensive care unit patients. Their mean age was 62 years (range, 1–92 years), and the female-to-male ratio was 1.02 to 1. The prevalence of thromboembolic disease in the hospitalized patient group was 12%. The 95% CI for the area under the ROC curve for each of the two test methods (Fig. 2B) included 0.5 [0.60 (95% CI, 0.50–0.71) for the ELISA and 0.56 (95% CI, 0.44–0.67) for the microlatex]. The areas under the ROC curves were recalculated after the exclusion of two patients who were <1 year of age. This exclusion did not alter the areas under the curves, nor did it alter the CIs. As shown in the 2×2 contingency table (Table 3B), at the usually recommended cutpoint of 0.5 mg/L for each assay, three patients were falsely negative for the ELISA and microlatex tests. The sensitivity, specificity, and predictive values for the hospitalized patients (Table 3A) were similar for the two methods. The false-positive rate was ~80% for each assay method.

**ANALYSIS OF Discordant RESULTS**

Among the 467 examined patients, the ELISA and microlatex results were discordant in 42 patients, with 19 ELISA-positive/microlatex-negative data pairs and 23 ELISA-negative/microlatex-positive results. Of the first group (ELISA-positive), 17 of 19 (89%) went on to reference imaging, whereas only 7 of the 23 (30%) with a negative ELISA received reference imaging, suggesting that positive ELISA results (which were seen by the physicians) prompted further testing. Positive microlatex results were not seen by the physicians and thus could not have this effect. As a result, the study was potentially biased against the microlatex test. Two of the 17 ELISA-positive patients in whom imaging tests were performed...
had positive imaging results; none of the 7 patients in the microlatex-positive group who went on to imaging had positive results. This difference was not statistically significant (Fisher exact test).

Discussion

The value of the D-dimer assay in the evaluation of thromboembolic disease lies in its high negative predictive value and high sensitivity, which provide the possibility to avoid expensive or invasive tests for PE, such as ventilation-perfusion lung scanning and angiography, in patients with negative D-dimer test results (4, 5, 16, 26). High sensitivity and high negative predictive values are critical to minimize the possibility of missing the diagnosis of PE because the mortality of untreated PE approaches 26% (7). On the other hand, the high false-positive rate that has been noted for D-dimer testing may serve only to increase the number of patients evaluated for VTE, leading to increased morbidity and expense of clinical testing (27).

Recently, several enhanced microlatex agglutination assays have become available commercially, and results of several studies of these assays appear promising (e.g., 12, 17, 18). Our study is one of the few to compare the diagnostic performance of the new Stago microlatex assay with ELISA D-dimer testing for inpatients and outpatients with suspected PE. We conclude that differences, if any, between the two assays are numerically small in both inpatients and outpatients.

In this study, both D-dimer assays had high negative predictive values and sensitivities in the outpatient population, with relatively low false-positive rates. By contrast, there seems to be little justification for the use of the D-dimer assay in the hospitalized population, as indicated by the poor performance of these assays in this group compared with the outpatient group. The difference in the utility of the D-dimer test to rule out PE in inpatients vs patients presenting in the emergency room probably reflects the typical presentation of these patient populations. D-dimer may be increased equally in PE or in DVT without embolism. Patients with extensive inflammation and wound healing or malignancy also may have increased plasma D-dimer concentrations (28–30). Furthermore, because D-dimer is cleared principally in the liver, patients with liver disease may have an increased D-dimer concentration (30, 31). In our hospital population, patients with suspected PE are frequently those who are at high risk for venous thrombosis, including postoperative patients and other immobilized individuals (data not shown), and they are more likely than outpatients to have comorbidities that increase D-dimer. Because of these comorbid disorders, which may affect the D-dimer concentration, the usefulness of this clinical test in this patient population is limited.

Physicians ordering the ELISA on inpatients were from several medical and surgical services, and we found little evidence that these physicians followed a standard testing pathway. Testing may have been performed on patients who had very high or very low clinical probability of having VTE. This may have led to the low rule-out rate, with exclusion of only 40 (17%) of the 233 inpatients evaluated, giving a false-positive rate >80%. Such a high false-positive rate and low rate of excluding the diagnosis may lead to increased and unnecessary use of expensive imaging tests as has been described elsewhere (11).

A potential threat to the validity of this study is verification bias. The study was based on routinely documented medical data, which commonly carries the potential of verification bias (27, 32–36). This bias commonly leads to an overestimation of the sensitivity and positive predictive value and underestimation of the specificity and negative predictive value. In addition, the ELISA result was available to the clinicians during the patient evaluation in our study, but the microlatex assay result was not. Our analysis of the discordant test results indicated that patients with positive ELISA results were three times as likely to undergo one or more of the reference tests as were those with a negative ELISA result and a positive microlatex result (94% vs 30%). This additional source of verification bias may have led to the underdiagnosis of VTE in the ELISA-negative group, falsely underestimating the number of both false and true negatives (37). Consequently, the relative diagnostic accuracy of the ELISA test may have been overestimated compared with that of the microlatex test.

It does not appear likely that verification bias obscured a significant difference in accuracy between the two assays. Statistical modeling (not shown) suggests that correction for verification bias might narrow the difference between the tests further. Verification bias may, however, explain why the specificities and negative predictive values in our study are somewhat lower than those reported in previous studies (5, 12, 15, 17, 38). Selective D-dimer testing in only the low clinical probability group could have artificially increased the exclusion rates. Nonetheless, the potential effects of both patient selection and verification bias were similar for both tests, and they do not affect our conclusion that the tests behave similarly.

A further, potential source of bias is that not all patients underwent objective reference testing or underwent only lower extremity ultrasound, and we were forced to rely on clinical follow-up to set the final diagnosis. Although small numbers of patients may have been misclassified through our use of clinical follow-up information, there is no reason to assume that this misclassification differed between the two assays. Because differential verification commonly leads to overestimation of the diagnostic accuracy, this threat to validity would not be expected to change our conclusion that D-dimer testing has low diagnostic accuracy in inpatients. Moreover, outcome-based standards have been considered appropriate for this type of evaluation, and a 3-month follow-up period has been suggested to be a "valid outcome mea-
Anticoagulation therapy is a known confounder of D-dimer assay results, producing lower D-dimer concentrations, and the use of D-dimer testing in the anticoagulated population remains a point of continued controversy. Mean plasma D-dimer concentrations of untreated patients with chronic atrial fibrillation are higher than those of control individuals and decrease to concentrations similar to those of the controls on anticoagulation therapy (41, 42). The D-dimer assay appears to retain utility in select patient groups, but the results need to be more carefully considered (43, 44). Although anticoagulation can lower D-dimer concentrations, in 32 of 45 of our anticoagulation group patients, D-dimer was >0.5 mg/L. A higher result would not have changed their classification. Of the 13 patients with negative D-dimer results, 2 had evidence of VTE. Exclusion of these 2 patients or of all 45 patients in the anticoagulant group only slightly changed assay sensitivity and negative predictive value.

In summary, we conclude that the diagnostic accuracies of the ELISA and microlatex methods were statistically indistinguishable for inpatients, outpatients, and all patients combined in this study. Not surprisingly, we observed greater numbers of patients with false-positive D-dimer tests in the inpatient group, likely because of an increased prevalence of comorbid disease in these patients and of inconspicuous thromboses without PE. In contrast, the outpatient group had fewer patients with false-positive results, suggesting that outpatients are more suitable for D-dimer testing to exclude PE. However, in both inpatients and outpatients, we observed a small number of negative D-dimer results in patients with VTE, with four patients having falsely negative results by both methods. Thus, even with the suggestion of superior D-dimer performance in outpatients when used to screen for VTE, there must continue to be consideration that a negative D-dimer result does not exclude the possibility of VTE.

Supported, in part, by a grant from Stago. We gratefully acknowledge the technical assistance of Gary Manuel.

References
Limitations of sensitivity, specificity, likelihood ratio and Bayes’ theorem in assessing diagnostic probabilities: a clinical example. Epidemiology 1997;8:12–7.


