Variation in the diagnostic performance of D-dimer for suspected deep vein thrombosis
Goodacre S, Sampson F C, Sutton A J, Mason S, Morris F

CRD summary
This review assessed the ability of D-dimer tests to diagnose suspected deep vein thrombosis (DVT) and investigated reasons for differing results between studies. The authors concluded that a D-dimer test may be sufficient to rule-out proximal DVT, while its ability to rule-in DVT appears to depend on pre-test clinical probability (the test performs better in low-risk patients). The review was well conducted and reported and its conclusions are likely to be reliable.

Authors’ objectives
To estimate the diagnostic accuracy of D-dimer testing for suspected deep vein thrombosis (DVT); to compare accuracy for the detection of proximal and distal DVT; and to investigate sources of between-study heterogeneity using meta-regression.

Searching
MEDLINE, EMBASE, CINAHL, Web of Science, the Cochrane Controlled Trials Register, the Cochrane Database of Systematic Reviews, DARE and ACP Journal Club were searched; a full search strategy was reported. In addition, the bibliographies of included studies were screened for further relevant articles. Manufacturers of D-dimer assays were contacted for unpublished studies. Studies published in English, French, Spanish, Italian or German were included.

Study selection
Study designs of evaluations included in the review
Diagnostic cohort studies that included at least 10 patients were eligible for inclusion. Diagnostic case-control studies were excluded.

Specific interventions included in the review
Studies that assessed D-dimer for the detection of DVT were eligible for inclusion. Studies that measured the risk of DVT developing after D-dimer testing were excluded, as were management studies that ruled out DVT on the basis of a negative D-dimer. The assay methods used in the included studies were enzyme-linked immunosorbent assay (ELISA), latex agglutination and whole-blood agglutination. The review included a large number of studies and the D-dimer diagnostic thresholds used in individual included studies were not reported; D-dimer threshold and its definition (whether or not it was defined before the study) were included as variables in the meta-regression.

Reference standard test against which the new test was compared
Studies that used a reference standard of venography, ultrasound, or plethysmography were eligible for inclusion. The reference standards used were venography, ultrasound alone, ultrasound with clinical follow-up, serial ultrasound, either ultrasound or venography, and a combination of ultrasound and plethysmography.

Participants included in the review
Studies of patients with clinically suspected DVT were eligible for inclusion. Studies of patients with suspected pulmonary embolism were excluded. The mean or median age of the participants ranged from 51 to 69 years (with the exception of 1 study that included only over 70s). The median proportion of males was 42% (range: 17 to 62). The median prevalence of DVT was 36% (range: 2 to 78). Where reported, the proportion of DVT which was proximal ranged from 27 to 100%.

Outcomes assessed in the review
No inclusion criteria for the outcome measures were specified. The outcomes reported in the review were the sensitivity and specificity.

How were decisions on the relevance of primary studies made?
Two reviewers independently screened titles and abstracts to identify potentially relevant articles. Full copies of all selected articles were retrieved. The same two reviewers then independently selected articles for inclusion. Kappa scores were calculated to measure agreement at both stages and any disagreements were resolved by consensus.

**Assessment of study quality**

Two reviewers independently assessed study quality using the following criteria: application of the reference standard independently of the D-dimer result; measurement of D-dimer blind to the result of the reference standard; conduct of the reference standard blind to the result of D-dimer testing; and use of a pre-determined diagnostic threshold for D-dimer rather than a threshold derived from the study data. An independent reviewer resolved any disagreements.

**Data extraction**

Two reviewers independently extracted the data using a standardised form. A third reviewer checked and resolved any discrepancies. Data were extracted on study and patient characteristics, the D-dimer assays and diagnostic thresholds used, reference standards used, true positives (proximal and distal), true negatives, false positives and false negatives (proximal and distal).

**Methods of synthesis**

**How were the studies combined?**

The results of the individual included studies were plotted in receiver operating characteristic space (true-positive rate versus false-positive rate). A random-effects model was used to estimate the overall sensitivity and specificity, with 95% confidence intervals (CIs). Where zero values occurred in the study data, a continuity correction of 0.5 was used. Analyses were conducted using MetaDiSc software.

Publication bias was assessed using funnel plots of the log odds of sensitivity and specificity against their corresponding standard errors.

**How were differences between studies investigated?**

Between-study heterogeneity was assessed using a chi-squared test. Initially, all studies were pooled and meta-regression was used to identify potential causes of heterogeneity in the sensitivity and specificity. Any covariate that showed an association with either sensitivity or specificity (p<0.1) was used to define subgroups for separate meta-analyses. In addition, subgroup analyses were conducted for: D-dimer assay method (ELISA, latex agglutination, whole-blood agglutination), proximal and distal DVT, individual D-dimer assays (by product name), patients with malignancy, and studies reporting results by Wells clinical risk stratification (see Other Publications of Related Interest).

**Results of the review**

Ninety-seven studies, with a total of 99 patient groups and 198 analyses of D-dimer assays, were included in the review.

**Study quality.**

The reference standard was applied independently of the results of D-dimer testing in 86 studies. D-dimer was measured blind to the results of the reference standard in 43 studies, and the reference standard was conducted blind to the results of D-dimer testing in 50 studies. The diagnostic threshold for D-dimer was defined a priori in 82 studies.

**Diagnostic accuracy.**

For all included studies, the sensitivity of the D-dimer test ranged from 48 to 100% and its specificity ranged from 5 to 100%. There was significant heterogeneity in both parameters (p<0.001).

**Investigation of heterogeneity.**

A number of covariates showed significant association with variation in sensitivity and/or specificity. These included methodological characteristics of the studies, participant characteristics, and aspects of the reference standard and index.
test methodology; they were reported in full in the paper. Subgroup analyses using groups identified by meta-regression indicated that studies using more selective patient groups tended to have higher sensitivity and specificity, while higher quality studies tended to have higher specificity and studies deriving the diagnostic threshold for D-dimer after data analysis tended to have higher sensitivity. However, significant heterogeneity remained after stratification. The results were reported in full.

Subgroup analyses.

For ELISA (91 analyses), the pooled sensitivity was 94% (95% CI: 93, 95) and the pooled specificity was 45% (95% CI: 44, 46). For latex agglutination assays (74 analyses), the pooled sensitivity was 89% (95% CI: 88, 90) and the pooled specificity was 55% (95% CI: 54, 56). For whole-blood agglutination assays (29 analyses), the pooled sensitivity was 87% (95% CI: 85, 88) and the pooled specificity was 68% (95% CI: 67, 69). Heterogeneity was not reported for these subgroups. An analysis of studies that reported proximal and distal DVT separately showed that all assays had higher sensitivity for proximal DVT. An analysis of studies that stratified participants by Wells clinical risk score indicated that the specificity of D-dimer may be dependent on the clinical probability of DVT (increasing with decreasing clinical risk). Significant heterogeneity remained within the subgroups, which could not be eliminated by further stratification to individual D-dimer assays; the results were reported in full.

Publication bias.

The funnel plots appeared symmetrical for specificity, but markedly asymmetrical for sensitivity. Few small studies reported low sensitivity, indicating that small studies with low sensitivities may be less likely to be published (specificity is less likely to influence publication in this way as a D-dimer assay is unlikely to be used to rule-in DVT).

Authors' conclusions

D-dimer has good sensitivity (which is higher for proximal than distal DVT), but poor specificity for the diagnosis of DVT. Specificity appeared to be dependent upon pre-test clinical probability, increasing in patients with lower clinical probability of DVT. All analyses showed significant between-study heterogeneity.

CRD commentary

This was a well conducted and clearly reported review. It addressed a number of clearly stated objectives and appropriate inclusion criteria were defined. Extensive searches of the published literature were made and the search strategy was reported. However, some language restrictions were applied and this might have resulted in the loss of some relevant data. Some attempts to identify unpublished studies were reported and the potential for publication bias was assessed; the authors fully discussed the implications of this assessment. The review methodology was rigorous and clearly reported, and included appropriate measures to minimise error and bias. The methodological quality of the included studies was assessed using criteria appropriate for diagnostic accuracy studies, and individual quality criteria were included as covariates in the meta-regression (an appropriate method of investigating the impact of study quality upon estimates of diagnostic accuracy).

The statistical methods used to generate estimates of diagnostic accuracy and to investigate between-study heterogeneity were appropriate and clearly reported. The value of pooled estimates of diagnostic accuracy is limited by the presence of significant residual heterogeneity, even within subgroups. However, the authors discussed fully the limitations of the available data and its implications for their study. The authors' conclusions follow broadly from the data presented and are likely to be reliable.

Implications of the review for practice and research

Practice: Data suggest that D-dimer may have sufficient sensitivity to rule-out DVT where the detection and treatment of distal DVT are not a concern. It may be appropriate to use different diagnostic thresholds for different pre-test probabilities (e.g. lower threshold in low-risk patients to minimise the risk of inappropriate discharge with DVT).

Research: Future studies should collect and report data on patient selection, exclusion criteria and the presence of co-morbidities. Stratification of patients by clinical probability of DVT is particularly important (the Wells score is...
recommended for this purpose).

**Funding**
UK Health Technology Assessment R&D Programme, grant number 02/03/01.

**Bibliographic details**

**Other publications of related interest**

**Indexing Status**
Subject indexing assigned by NLM

**MeSH**
Antifibrinolytic Agents /diagnostic use; Enzyme-Linked Immunosorbent Assay /methods; Fibrin Fibrinogen Degradation Products /diagnostic use; Regression Analysis; Research Design; ROC Curve; Sensitivity and Specificity; Venous Thrombosis /diagnosis; Whole Blood Coagulation Time /methods

**AccessionNumber**
12005004200

**Date bibliographic record published**
30/04/2007

**Date abstract record published**
30/04/2007

**Record Status**
This is a critical abstract of a systematic review that meets the criteria for inclusion on DARE. Each critical abstract contains a brief summary of the review methods, results and conclusions followed by a detailed critical assessment on the reliability of the review and the conclusions drawn.