Meta-analysis: the value of clinical assessment in the diagnosis of deep venous thrombosis

Goodacre S, Sutton A J, Sampson F C

CRD summary
This review assessed whether physicians’ empirical judgements, clinical findings and risk scores affect the likelihood of detecting deep vein thrombosis (DVT) on definitive testing. The authors concluded that individual clinical features are of limited value, whereas an overall assessment using the Wells score is more useful in the diagnosis of DVT. These conclusions are appropriate and appear reliable.

Authors' objectives
To assess whether physicians' empirical judgements, clinical findings and risk scores affect the likelihood of detecting thrombosis with venography, ultrasonography, or plethysmography in adults with suspected deep venous thrombosis (DVT).

Searching
MEDLINE, EMBASE, CINAHL, Web of Science, the Cochrane Database of Systematic Reviews, the Cochrane Controlled Trials Register, DARE and ACP Journal Club were searched from 1966 to January 2005; the search terms were not reported. In addition, the bibliographies of all retrieved articles were checked. Studies were limited to those published in English, French, Spanish or Italian.

Study selection

Cohort studies were eligible for inclusion. Case-control studies and studies with less than 10 participants were excluded.

Studies that reported physicians' empirical judgements, clinical findings, or a clinical score and then undertook diagnostic testing for DVT were eligible for inclusion.

Studies that used venography, ultrasonography or plethysmography as the reference standard test were eligible for inclusion.

Studies that included participants with suspected DVT were included.

No inclusion criteria were specified for the outcomes. The specific outcomes assessed were the numbers of true-positive, true-negative, false-positive and false-negatives results for both proximal and distal DVT for each clinical feature, and the number of patients with and without DVT for each clinical score. Studies that reported the risk for developing DVT after recording clinical characteristics that did not measure the probability that DVT was present at the time of assessment were excluded.

How were decisions on the relevance of primary studies made?
Two reviewers independently assessed studies for inclusion, with any disagreements being resolved through discussion.

Assessment of study quality
The validity of the primary studies was assessed according to whether: the reference standard was applied independently of the findings of the clinical assessment; the observers who undertook the clinical assessment were blinded to the results of the reference standard test; and observers blinded to the results of the clinical assessment.
interpreted the reference standard. The authors did not state how many reviewers performed the validity assessment.

**Data extraction**
The authors did not state how the data were extracted for the review, or how many reviewers performed the data extraction.

The numbers of true-positive, true-negative, false-positive and false-negatives results for both proximal and distal DVT for each clinical feature, as well as the number of patients with and without DVT for each clinical score, were extracted. The positive and negative likelihood ratios (LRs) for each clinical characteristic within each study were calculated. Where additional information was required, attempts were made to contact the corresponding author of the primary studies.

**Methods of synthesis**
How were the studies combined?
The LRs for the presence and absence of each clinical feature were combined in a meta-analysis using a random-effects model. The Wells score and empirical estimates groups were categorised into three groups (high, intermediate and low risk for DVT), and ordinal regression models were used to estimate the probability of being categorised as having high, intermediate or low risk. Separate models were used for persons with any DVT, proximal DVT, distal DVT and those without DVT. Estimates of sensitivity and specificity were then derived for two decision thresholds for all cases of DVT: high versus intermediate and low, and high and intermediate versus low. The LRs for high and low categories were then pooled using a random-effects meta-analysis.

How were differences between studies investigated?
Differences between the studies for the presence or absence of each clinical feature were assessed using the chi-squared test. Meta-regression models using a receiver operating characteristic curve were used to explore the influence of study level covariates on the diagnostic performance of the Wells score.

**Results of the review**
Fifty-one cohort studies with a total of 10,311 participants were included.

Apart from the studies that augmented an ultrasonography reference standard with further testing based on clinical probability, in most studies the reference standard was applied independently of the results of the clinical assessment. The reporting of blinding of the clinical assessment and the reference standard was generally poor.

Clinical features of DVT.

The positive and negative LRs were, respectively: for calf pain, 1.08 (95% confidence interval, CI: 0.96, 1.20) and 0.90 (95% CI: 0.78, 1.03) (12 studies); for calf swelling, 1.45 (95% CI: 1.25, 1.69) and 0.67 (95% CI: 0.58, 0.78) (16 studies); for history of DVT, 2.25 (95% CI: 1.57, 3.23) and 0.90 (95% CI: 0.85, 0.95) (11 studies); for malignant disease, 2.71 (95% CI: 2.16, 3.39) and 0.89 (95% CI: 0.85, 0.93) (20 studies); for recent immobilisation 1.98 (95% CI: 1.70, 2.30) and 0.90 (95% CI: 0.85, 0.94) (17 studies); for recent surgery, 1.76 (95% CI: 1.40, 2.20) and 0.94 (95% CI: 0.91, 0.97) (17 studies); for obesity, 0.85 (95% CI: 0.59, 1.23) and 1.04 (95% CI: 0.96, 1.13) (5 studies); for difference in calf diameter, 1.80 (95% CI: 1.48, 2.19) and 0.57 (95% CI: 0.44, 0.72) (8 studies); for Homan sign, 1.40 (95% CI: 1.18, 1.66) and 0.87 (95% CI: 0.79, 0.96) (11 studies); for warmth, 1.29 (95% CI: 1.07, 1.54) and 0.97 (95% CI: 0.78, 0.98) (12 studies); for tenderness, 1.27 (5% CI: 1.11, 1.45) and 0.80 (5% CI: 0.72, 0.89) (14 studies); for erythema, 1.30 (95% CI: 1.02, 1.67) and 0.88 (95% CI: 0.80, 0.98) (6 studies); and for oedema, 1.35 (95% CI: 1.05, 1.74) and 0.86 (95% CI: 0.79, 0.94).

Based on an LR of greater than 2 being useful for ruling in DVT and a ratio of less than 0.5 being useful for ruling out DVT, then only a history of DVT and malignancy are useful for ruling in DVT. No clinical features were useful for ruling out DVT. Recent immobilisation, recent surgery, or a difference in calf diameter were of borderline value for ruling in DVT. The absence of calf swelling or a difference in calf diameter were of borderline significance for ruling out DVT.
Wells clinical probability score (22 studies).

A high Wells score increased the probability of DVT (LR=5.2), while a low Wells score decreased the probability of DVT (LR=0.25).

Physicians' empirical judgements (8 studies).

Four studies categorised patients into low, intermediate or high risk for DVT, while four dichotomised assessment as low or high risk or dichotomised DVT as present or absent. For the 4 studies that categorised patients into low, intermediate or high risk, the LRs for the empirical judgements of high and low risk were similar to those of the Wells scores. However, there was evidence of significant heterogeneity across these studies. The meta-analysis of the dichotomised empirical judgements gave a sensitivity of 86.6% (95% CI: 80.7, 91.2), a specificity of 69.3% (95% CI: 64.4, 73.9), a positive LR of 6.2 (95% CI: 1.0, 40.0) and a negative LR of 0.18 (95% CI: 0.13, 0.26). Again, there was evidence of significant heterogeneity across these studies.

Authors' conclusions

Individual clinical features are of limited diagnostic value. An overall assessment of clinical probability using the Wells score is more useful in the diagnosis of DVT.

CRD commentary

The review question was clearly defined in terms of the interventions, reference standards, participants and study designs. A number of sources were searched for relevant studies, and efforts were made to minimise language bias. However, since no efforts were made to locate unpublished studies, some potentially relevant studies might have been missed. Two reviewers were involved in selecting the studies, and the quality of the included studies was adequately assessed. It was unclear though how the validity assessment and data extraction processes were conducted, and whether any efforts were made to minimise reviewer bias and errors. The use of meta-analyses to combine the studies where possible was appropriate, and differences between the studies were thoroughly explored. Overall, the authors' conclusions are appropriate given the evidence reviewed and are likely to be reliable.

Implications of the review for practice and research

Practice: The authors stated that the use of the Wells score appears to be the most valuable element of clinical assessment for the patient with suspected DVT.

Research: The authors stated that further research is needed to determine how the Wells score performs in different clinical settings and when used by different observers.

Funding

United Kingdom Health Technology Assessment R&D Programme, grant number 02/03/01.

Bibliographic details


PubMedID

16027455

Original Paper URL

http://www.annals.org/cgi/content/full/143/2/129

Other publications of related interest
These additional published commentaries may also be of interest. Stevens SM, Ageno W. Review: the Wells rule is more useful than individual clinical features for predicting risk for deep venous thrombosis. ACP J Club 2006;144:46-7. Stevens SM, Ageno W. Review: the Wells rule is more useful than individual clinical features for predicting risk for deep venous thrombosis. Evid Based Med 2006;11:56-7.

Indexing Status
Subject indexing assigned by NLM

MeSH
Humans; Likelihood Functions; Predictive Value of Tests; ROC Curve; Risk Factors; Venous Thrombosis /diagnosis

AccessionNumber
12005008370

Date bibliographic record published
28/02/2006

Date abstract record published
28/02/2006

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.