Systematic review and meta-analysis of strategies for the diagnosis of suspected pulmonary embolism

Roy P M, Colombet I, Durieux P, Chatellier G, Sors H, Meyer G

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 determine the diagnostic accuracy of tests used for suspected pulmonary embolism (PE) and, hence, to estimate the range of pre-test probabilities over which each test can be used to rule-in or rule-out PE.

Searching
MEDLINE, EMBASE and Pascal Biomed were searched from January 1990 to September 2003 for studies reported in the English language; the search terms were reported. In addition, published bibliographies and the authors' own libraries were manually searched. Studies published as abstracts only were excluded.

Study selection
Study designs of evaluations included in the review
Prospective studies with consecutive participant recruitment, in which the index test and the reference standard were interpreted independently, were eligible for inclusion. Retrospective studies were excluded, as were follow-up studies where more than 5% of the population were lost to follow-up.

The review included both studies using composite reference standards and studies conducted in selected populations; these features were used as criteria for classifying study quality.

Specific interventions included in the review
Studies that evaluated test strategies for the confirmation or exclusion of PE were eligible for inclusion. The included studies used a variety of imaging techniques and methods of measuring D-dimer.

Reference standard test against which the new test was compared
The included studies were required to use pulmonary angiography (confirmation strategies), and follow-up or pulmonary angiography (exclusion strategies) as the reference standard. Where follow-up was used as the reference standard, participants with a negative test result were considered to have a false-negative result if they developed deep vein thrombosis or PE within a 3-month follow-up period. Studies of exclusion strategies that used additional imaging to pulmonary angiography in participants with a negative test result were excluded.

Participants included in the review
The included studies were required to have used consecutive participant recruitment. Studies of specific populations (not defined) were excluded.

Outcomes assessed in the review
Studies were excluded if data for the calculation of positive and negative likelihood ratios (LRs) could not be extracted. Positive LRs, with 95% confidence intervals (CIs), were calculated and reported for the confirmation strategies; negative LRs, also with 95% CIs, were calculated and reported for the exclusion strategies.

How were decisions on the relevance of primary studies made?
Two reviewers independently selected potentially relevant studies.
Assessment of study quality
The included studies were classified as grade A, B, or C. Grade A studies were cohort studies on unselected participants in which pulmonary angiography was used as the reference standard for all participants. Grade B studies were cohort studies on consecutive, unselected participants in which a composite reference standard was used to confirm either the absence (confirmation studies) or presence (exclusion studies) of PE. Grade C studies were performed in selected participants (e.g. those referred for pulmonary angiography) and had the same reference standard criteria as grade B studies. The authors did not state how many reviewers performed the validity assessment.

Data extraction
Two reviewers independently extracted data on the study design, participant characteristics, and numbers of true-positive, false-positive, false-negative and true-negative test results. A third author resolved any discrepancies.

Methods of synthesis
How were the studies combined?
The pooled estimates of LRs were calculated using the random-effects method of DerSimonian and Laird. Bayes theorem was used to calculate the probability of PE, conditioned by the LR as a function of pre-test probability; post-test versus pre-test probability curves were presented, along with 95% CIs. Based on published criteria, the authors considered that a confirmation strategy was sufficiently accurate to diagnose PE when the post-test probability was greater than 85% and, conversely, that an exclusion strategy was sufficiently accurate to rule-out PE when the post-test probability was less than 5% (see Other Publications of Related Interest).

How were differences between studies investigated?
Cochran's Q statistic and I-squared were used to assess the extent of between-study heterogeneity. Where the I-squared value was greater than 0%, potential sources of heterogeneity were explored using subgroup analyses based upon the three categories for study quality.

Results of the review
Forty-eight studies involving 11,004 participants with suspected PE were included; PE was confirmed in 3,329 participants. Studies that evaluated electron beam computed tomography were not analysed as this technique is no longer used.

Confirmation strategies.
The pooled positive LRs were as follows: high probability ventilation perfusion lung scan (1 study), 18.3 (95% CI: 10.3, 32.5); spiral computed tomography (6 studies), 24.1 (95% CI: 12.4, 46.7); ultrasonography of leg veins (4 studies), 16.2 (95% CI: 5.6, 46.7); perfusion lung scan compatible with PE (1 study), 7.1 (95% CI: 4.6, 11.0); echocardiography (2 studies), 5.0 (95% CI: 2.3, 10.6); and magnetic resonance angiography (5 studies; significant heterogeneity), 11.7 (95% CI: 3.6, 37.8).

In patients with moderate (prevalence of PE 35% or greater) or high (prevalence of PE 70% or greater) pre-test probability, a positive finding on high probability ventilation perfusion lung scan, spiral computed tomography, or ultrasonography of the leg veins was associated with a greater than 85% post-test probability of PE.

Exclusion strategies.
The pooled negative LRs were as follows: normal or near normal lung scan (9 studies), 0.05 (95% CI: 0.03, 0.10); negative spiral computed tomography and leg vein ultrasonography (3 studies), 0.04 (95% CI: 0.03, 0.06); D-dimer concentration of less than 500 microg/L as measured by quantitative enzyme-linked immunosorbent assay (11 studies; significant heterogeneity, quantitative latex agglutination tests excluded), 0.08 (95% CI: 0.04, 0.18); low probability ventilation perfusion lung scan (1 study), 0.36 (95% CI: 0.25, 0.50); perfusion lung scan not compatible with PE (1 study), 0.09 (95% CI: 0.06, 0.15); negative spiral computed tomography (9 studies; significant heterogeneity), 0.11.
(95% CI: 0.06, 0.19); negative leg vein ultrasonography (6 studies; significant heterogeneity), 0.67 (95% CI: 0.50, 0.89); negative magnetic resonance angiography (5 studies), 0.20 (95% CI: 0.12, 0.34); negative D-dimer test measured by semi-quantitative latex agglutination (2 studies), 0.29 (95% CI: 0.03, 2.46); negative D-dimer test measured in whole blood (3 studies), 0.31 (95% CI: 0.18, 0.56).

In patients with low (prevalence of PE 10% or less) or moderate pre-test probability, a normal or near normal appearance on lung scan, a negative result on both spiral computed tomography and ultrasound, or a D-dimer concentration of 500 microg/L or less (measured by quantitative enzyme-linked immunosorbent assay) was associated with a less than 5% post-test probability of PE.

Authors' conclusions
Whilst the accuracy of tests for suspected PE varies greatly, it is possible to estimate the range of pre-test probabilities over which tests or strategies can perform to rule-in or rule-out PE. Negative lung scan, spiral computed tomography in combination with ultrasound, or quantitative D-dimer assay were considered adequate to rule-out PE where the population prevalence was 35% or less. Positive high probability ventilation perfusion lung scan, spiral computed tomography, or ultrasound were considered adequate to rule-in PE where the population prevalence was 35% or greater.

The authors further concluded that the performances of several diagnostic tests require further evaluation.

CRD commentary
The review addressed a clearly stated question and employed appropriate and well-defined inclusion criteria. Some studies that did not meet all of the inclusion criteria relating to study design (grade C studies) were included, and used in subgroup analyses based on study quality. The review methods were adequately reported and appropriate measures were taken to minimise the introduction of bias and error. The restriction of the literature search to published, English language studies might have resulted in the omission of some relevant data. Similarly, the use of diagnostic search terms is known to reduce search sensitivity. No attempt to assess the presence of publication bias was reported.

The pooling of LRs for groups of studies where significant between-study heterogeneity remained is of questionable value and, although subgroup analyses were conducted, the results were not presented in a manner which clarified the impact of study quality on diagnostic accuracy. The authors' conclusions follow from the data presented, but should be treated with caution in view of the limitations outlined.

Implications of the review for practice and research
Practice: The authors stated that, as a general rule, discordance between clinical (pre-test) probability and diagnostic test result requires further investigation.

Research: The authors stated that the performances of several diagnostic tests require further evaluation.

Bibliographic details

PubMedID
16052017

DOI
10.1136/bmj.331.7511.259

Original Paper URL
http://bmj.bmjournals.com/cgi/content/full/331/7511/259
Other publications of related interest
Kearon C. Diagnosis of pulmonary embolism. CMAJ 2003;290:2849-58.

Indexing Status
Subject indexing assigned by NLM

MeSH
Echocardiography /standards; Fibrin Fibrinogen Degradation Products /analysis; Humans; Leg /blood supply;
Likelihood Functions; Magnetic Resonance Angiography /standards; Pulmonary Embolism /diagnosis /radionuclide
imaging /ultrasonography; Sensitivity and Specificity; Tomography, Spiral Computed /standards; Veins /physiology

AccessionNumber
12005008320

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.