Clinical decision rules for excluding pulmonary embolism: a meta-analysis
Lucassen W, Geersing GJ, Erkens PM, Reitsma JB, Moons KG, Buller H, van Weert HC

CRD summary
This review concluded that clinical decision rules or Gestalt could safely exclude pulmonary embolism, when combined with sensitive D-dimer testing. Standardised decision rules were recommended over Gestalt. The review was generally well conducted, but the conclusions and implications for practice seem to be too strong for the evidence presented.

Authors' objectives
To compare the failure rates of standard clinical assessment (Gestalt) and clinical decision rules, when used in combination with D-dimer testing, to diagnose pulmonary embolism in adults with suspected pulmonary embolism.

Searching
MEDLINE and EMBASE were searched for articles published in English, French, German, Italian, Spanish, or Dutch, from 1966 to June 2011; search terms were reported. Reference lists of selected articles were scanned.

Study selection
Prospective studies of adults (aged 16 years or older) with the signs or symptoms of acute pulmonary embolism were eligible for inclusion. Studies had to use Gestalt (unstructured clinical review of patient history and physical examination with or without laboratory tests, electrocardiography, or chest radiography) or a clinical decision rule, based on a multivariate logistic regression, to estimate the clinical probability of pulmonary embolism. Those evaluating a clinical decision rule had to have at least 50 patients with confirmed pulmonary embolism or a diagnosis of deep vein thrombosis. Studies had to provide sufficient data to construct two-by-two tables and the reference standard had to be an appropriate composite (ventilation–perfusion lung scanning, computed tomography, pulmonary angiography, or autopsy). D-dimer testing had to be performed in all patients with an assessment of low probability; in patients with a negative D-dimer, clinical follow-up of at least 45 days was required.

Five clinical decision rules were evaluated: Wells, with a cut-off of less than two, or four or less, and simplified, Geneva (original and revised), Pisa, Charlotte, and the Pulmonary Embolism Rule-out Criteria (PERC). The original Geneva and Pisa rules included electrocardiography and chest radiography. For Gestalt, the cut-off for the lowest probability category ranged from under 10% to under 40%. Studies were published between 1990 and 2011, the prevalence of pulmonary embolism ranged from 4% to 44%, and the mean age ranged from 45 to 72 years.

Two reviewers assessed each paper for inclusion; disagreements were resolved by discussion, with a third reviewer.

Assessment of study quality
Inclusion was restricted to studies that: recruited consecutive patients; blinded interpreters of the D-dimer or imaging tests; and used an appropriate reference standard. Each study was evaluated using the 14-criterion QUADAS tool, by two independent reviewers; disagreements were resolved by discussion, with a third reviewer.

Data extraction
Two reviewers independently extracted the data into two-by-two tables from which sensitivity, specificity, failure (false negative) rates, and efficiency (true negative rates) were calculated, along with 95% confidence intervals. Where data were not dichotomised, the lowest probability category, for each rule, versus the other probability categories was used. Disagreements were resolved by discussion, with a third reviewer.

Methods of synthesis
Pooled estimates of sensitivity and specificity with 95% confidence intervals were calculated, using the bivariate random-effects model, where at least four studies were available. Each decision rule was analysed separately and an overall estimate was produced. Pooled estimates with 95% confidence intervals for failure rates and efficiency were calculated using logit-transformed proportions in a random-effects model. The tests used to investigate heterogeneity
were not reported. The impact of study design, reference standard used, and prevalence of pulmonary embolism was investigated.

**Results of the review**

Fifty-two studies met the inclusion criteria (55,268 patients; range 77 to 8,138). Of these, 63% were subject to differential verification bias and 37% did not report uninterpretable test results. Approximately 75% recruited an appropriate patient spectrum; 95% avoided partial verification bias; and 50% reported blinding of interpreters of the index test.

Overall in 23 studies (embolism prevalence of 14%), using Gestalt or decision rules, combined with qualitative or quantitative D-dimer testing, the failure rate was 0.7 (95% CI 0.5 to 1.0) and the efficiency was 35% (95% CI 30 to 41).

Using Gestalt (15 studies), sensitivity was 85% (95% CI 78 to 90) and specificity was 51% (95% CI 39 to 63). With qualitative D-dimer testing (two studies), the failure rate was 0.7 (95% CI 0.4 to 1.2) and the efficiency was 52% (95% CI 40 to 64).

Using Wells with cut-off of less than two (19 studies), sensitivity was 84% (95% CI 78 to 89) and specificity was 58% (95% CI 52 to 65). With qualitative D-dimer testing (five studies), the failure rate was 0.9 (95% CI 0.6 to 1.5) and the efficiency was 40% (95% CI 33 to 48).

Using Wells with a cut-off of four or less (11 studies), sensitivity was 60% (95% CI 49 to 69) and specificity was 80% (95% CI 75 to 84). With quantitative D-dimer testing (four studies), the failure rate was 0.5 (95% CI 0.2 to 0.9) and the efficiency was 39% (95% CI 31 to 47). With qualitative D-dimer testing (three studies), the failure rate was 1.7 (95% CI 1.0 to 2.8) and the efficiency was 42% (95% CI 32 to 52).

Using Geneva (five studies), sensitivity was 84% (95% CI 81 to 87) and specificity was 50% (95% CI 29 to 72). With quantitative D-dimer testing (two studies), the failure rate was 0.0 (95% CI 0.0 to 1.3) and the efficiency was 21% (95% CI 14 to 31).

Using the revised Geneva (four studies), sensitivity was 91% (95% CI 73 to 98) and specificity was 37% (95% CI 22 to 55). With quantitative D-dimer testing (two studies), the failure rate was 0.3 (95% CI 0.0 to 1.7) and the efficiency was 23% (95% CI 15 to 33).

The results for all 52 studies were presented in figures and discussed in a narrative. The results of sensitivity analyses were presented.

**Authors' conclusions**

Clinical decision rules or Gestalt could safely exclude pulmonary embolism, when combined with sensitive D-dimer testing.

**CRD commentary**

The review addressed a clear question, supported by appropriate inclusion criteria. Relevant sources were searched, but the search was restricted, and publication and language bias cannot be ruled out. Each stage of the review was conducted in duplicate, reducing the potential for error and bias. Despite the use of quality criteria for study selection, some studies failed these apparently essential criteria. Several of the included studies seemed to have methodological limitations. The results of the quality assessment were only summarised, making it difficult to determine which analyses were based on the strongest evidence. There was substantial heterogeneity across studies, but robust methods of synthesis were used and the potential causes of this heterogeneity were investigated and discussed.

This was generally a well-conducted review, but the conclusions and implications for practice seem to be too strong for the evidence presented.

**Implications of the review for practice and research**
**Practice**: The authors recommended standardised decision rules because Gestalt had lower specificity; the choice of a particular rule and D-dimer test depended on the prevalence and setting.

**Research**: The authors recommended the evaluation of a diagnostic strategy to exclude pulmonary embolism in primary care, without imaging, before its implementation.

**Funding**
Dutch Heart Foundation, project number 2006B237.

**Bibliographic details**

**PubMedID**
21969343

**DOI**
10.7326/0003-4819-155-7-201110040-00007

**Original Paper URL**
http://www.annals.org/content/155/7/448.abstract

**Indexing Status**
Subject indexing assigned by NLM

**MeSH**
Decision Support Techniques; Fibrin Fibrinogen Degradation Products /analysis; Gestalt Theory; Humans; Pulmonary Embolism /diagnosis; Referral and Consultation; Sensitivity and Specificity; Triage; Venous Thrombosis /diagnosis

**AccessionNumber**
12011005953

**Date bibliographic record published**
05/10/2011

**Date abstract record published**
12/10/2011

**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.