How should we diagnose suspected deep-vein thrombosis?

Goodacre S, Stevenson M, Wailoo A, Sampson F, Sutton A J, Thomas S

Record Status
This is a critical abstract of an economic evaluation that meets the criteria for inclusion on NHS EED. Each abstract contains a brief summary of the methods, the results and conclusions followed by a detailed critical assessment on the reliability of the study and the conclusions drawn.

Health technology
The present study compared 18 available diagnostic strategies for suspected deep-vein thrombosis (DVT) with a "no testing, no treatment" alternative. The 18 algorithms are outlined below.

Algorithm 1: venography for all patients.
Algorithm 2: above-knee ultrasound (US), repeat if negative.
Algorithm 3: full-leg US, repeat if distal found.
Algorithm 4: above-knee US, no repeat.
Algorithm 5: Wells score and above-knee US. If low, discharge if US negative, venogram if positive. If moderate, repeat US if negative, treat if positive. If high, venogram if US negative, treat if US positive.
Algorithm 6: SimpliRED D-dimer (DD) and above-knee US. If US positive then treat. If both are negative then discharge. If DD positive and US negative, repeat US.
Algorithm 7: Wells. High or intermediate: above-knee US, treat if positive, venogram if negative. Low: above-knee US, treat if positive, discharge if negative.
Algorithm 8: Wells. High or intermediate: full-leg US, treat if positive, venogram if negative. Low: full-leg US, treat if positive, discharge if negative.
Algorithm 9: latex DD. If positive above-knee US and repeat, if negative do Wells score. If high US and repeat. If intermediate or low, discharge.
Algorithm 10: latex DD. If positive above-knee US and repeat, if negative do Wells score. If high US, if intermediate or low, discharge.
Algorithm 11: Wells. High: above-knee US, treat if positive, SimpliRED DD if negative. If DD positive venogram, if negative repeat US. Intermediate: US, treat if positive, DD if negative. If DD positive repeat US, if negative discharge. Low: DD, US if positive, discharge if negative.
Algorithm 12: Wells and SimpliRED DD. If Wells high or intermediate, or DD positive, do full-leg US. If Wells low and DD negative then discharge.
Algorithm 13: enzyme-linked immunosorbent assay (ELISA) DD. If negative discharge, if positive do above-knee US. Treat if US positive, do Wells if negative. High Wells: venogram. Intermediate or low Wells: discharge.
Algorithm 14: Wells. If high or intermediate: above-knee US. If positive, treat. If negative, SimpliRED DD. Repeat US if DD positive, discharge if DD negative. Low: US, discharge if negative, treat if positive.
Algorithm 15: Wells. High or intermediate: above-knee US. If positive, treat. If negative, SimpliRED DD. Repeat US if DD positive, discharge if DD negative. Low: DD, discharge if negative, US if positive.
Algorithm 16: Wells. High: above-knee US. If positive, treat. If negative, SimpliRED DD. Repeat US if DD positive, discharge if DD negative. Intermediate or low: DD, discharge if negative, US if positive.

Algorithm 17: Wells. High or intermediate: above-knee US. If positive, treat. If negative, repeat US. Low: US, treat if positive, discharge if negative.

Algorithm 18: Wells. High: above-knee US. If positive, treat. If negative, repeat US. Intermediate and low: US, treat if positive, discharge if negative.

**Type of intervention**  
Diagnosis.

**Economic study type**  
Cost-utility analysis.

**Study population**  
The hypothetical population comprised a cohort of 1,000 outpatients with suspected DVT. Prevalence, mean age and gender distribution were taken from a DVT registry.

**Setting**  
The setting was secondary and tertiary care. The economic study was carried out in Sheffield, UK.

**Dates to which data relate**  
Studies providing effectiveness evidence were from 1975 to 2005, while those providing cost data were from 2003 to 2004. The price year was 2003 to 2004.

**Source of effectiveness data**  
The evidence was derived from a review or synthesis of completed studies and estimates based on authors’ assumptions and expert opinion.

**Modelling**  
A decision analysis model was developed to compare algorithms in a hypothetical cohort of patients with suspected DVT. Estimates of sensitivity and specificity for each algorithm were applied to the population to determine the proportions of patients with and without DVT or treatment over the minimum treatment period of 3 months. Subsequent lifetime health outcomes and costs accrued by testing and treatment were estimated.

**Outcomes assessed in the review**  
The outcomes in the baseline model included:

- the prevalence of proximal and distal DVT;
- the sensitivity and specificity of different tests for proximal and distal DVT;
- the overall sensitivity and specificity;
- the probability of fatal haemorrhage due to anticoagulant treatment; and
- the probability of fatal pulmonary embolus (PE), nonfatal PE and post-thrombotic syndrome (PTS) in treated and
untreated patients.

**Study designs and other criteria for inclusion in the review**
The studies used were systematic reviews, meta-analysis and other published literature of varying designs, such as randomised trials and cohort studies.

**Sources searched to identify primary studies**
Not reported.

**Criteria used to ensure the validity of primary studies**
Not reported.

**Methods used to judge relevance and validity, and for extracting data**
It would appear that the validity of the primary studies was not assessed.

**Number of primary studies included**
The authors reported that at least 44 primary studies were included in the review for different purposes (prevalence rates, sensitivity and specificity, probabilities of events, and so on).

**Methods of combining primary studies**
The authors used meta-analysis and systematic reviews undertaken in four of their own published studies (Goodacre et al. 2005 and 2006, see ‘Other Publications of Related Interest’ below for bibliographic details).

**Investigation of differences between primary studies**
The authors do not appear to have investigated differences between the primary studies.

**Results of the review**
The mean values of the parameters used in the model were as follows:

- For the prevalence of proximal DVT, 0.147 for the ratio of distal to proximal DVT, and 0.214 for the probability that distal DVT propagates to proximal;
- For treated proximal DVT, the probability was 0.004 for fatal PE, 0.008 for nonfatal PE and 0.053 for PTS;
- For untreated proximal DVT, the probability was 0.019 for fatal PE, 0.093 for nonfatal PE and 0.33 for PTS;
- The probability of fatal intracranial haemorrhage due to anticoagulant treatment was 0.003;
- The probability of nonfatal intracranial haemorrhage was 0.001; and
- The probability of nonfatal non-intracranial haemorrhage was 0.021.

The sensitivity and specificity of the tests were as follows.

For US, the sensitivity was 0.95 for proximal DVT and 0.65 for distal DVT, and the specificity was 0.94.

For ELISA DD, the sensitivity was 0.98 for proximal DVT and 0.86 for distal DVT, and the specificity was 0.34 for Wells high, 0.45 for Wells moderate and 0.52 for Wells low.
For latex DD, the sensitivity was 0.94 for proximal DVT and 0.79 for distal DVT, and the specificity was 0.42 for Wells high, 0.55 for Wells moderate and 0.64 for Wells low.

For SimpliRED DD, the sensitivity was 0.84 for proximal DVT and 0.64 for distal DVT, and the specificity was 0.52 for Wells high, 0.68 for Wells moderate and 0.79 for Wells low.

For the Wells score:

the proportions of proximal DVT were 0.68, 0.25 and 0.07 for high-, moderate- and low-risk categories, respectively;
the proportions of distal DVT were 0.34, 0.48 and 0.18 for high-, moderate- and low-risk categories, respectively; and
the proportions without DVT were 0.11, 0.41 and 0.48 for high-, moderate- and low-risk categories, respectively.

Details of all the parameters were reported (mean value, probability distribution used and reference source).

**Methods used to derive estimates of effectiveness**
This analysis was based on published data, expert opinion and authors’ assumptions.

**Estimates of effectiveness and key assumptions**
Assumptions in the model were that DD specificity was dependent upon Wells score, while sensitivity was independent. US sensitivity and specificity were independent of both the Wells score and DD. If the algorithm defined US as being above-knee only, sensitivity for distal DVT was assumed to be zero. Also, repeated US results were assumed to be entirely dependent upon initial ultrasound for proximal DVT, and that the results of repeat scanning only differed from initial scanning if the patient initially had a distal DVT that then propagated proximally. It was assumed that contrast venography had perfect sensitivity and specificity but would not be feasible in 10%, would cause DVT in 1%, and carried a 1:55,000 risk of fatal anaphylaxis. It was also assumed that distal DVT carried a 21% probability of propagating proximally, where it would then carry the same risks as proximal DVT. An expert panel estimated the probability of developing PTS to be 33% in untreated patients.

**Measure of benefits used in the economic analysis**
The measure of benefit was the quality-adjusted life-years (QALYs). The authors estimated utility values on the basis of published literature and expert opinion. Individuals who died from an initial event were assigned 0 QALYs. It was assumed that initial event-free survival was followed by a normal quality-adjusted life-expectancy of 11.58 QALYs for an individual aged 60 years, based on interim life tables. The mean value for a PTS was 0.977, for nonfatal intracranial haemorrhage 0.29, and for a nonfatal PE 0.94. Discounting was performed at a rate of 3.5%. A beta-distribution was used for the probabilistic analysis.

**Direct costs**
Health service direct costs of medical care were included. These were appropriately discounted at a rate of 3.5% per year. These included drugs, tests, nursing and general practitioner (GP) visits, therapeutic costs, treatment of adverse events, and lifetime costs for PTS. The quantities and the costs were analysed separately and were derived using modelling. The DD assay costs were estimated using NHS Trust data. NHS reference costs were used to estimate costs for US, venography, fatal and nonfatal PE, and nonfatal and non-intracranial bleeding. The drug costs were taken from the 2004 British National Formulary. The treatment costs of proximal DVT, GP and nursing costs, and costs of fatal bleeding and nonfatal intracranial bleeding were taken from published medical literature. All direct costs were reported in 2003 to 2004 UK pounds sterling.

**Statistical analysis of costs**
The incorporation of the costs depended on the nature of the cost. Costs related to laboratory and SimpliRED DD, clinical risk stratification, anticoagulant clinic review, warfarin treatment, GP and nursing visits, and lifetime PTS were
treated deterministically. Costs treated stochastically included treatment of fatal PE and nonfatal PE, vascular surgery and follow-up, treatment of non-intracranial bleeding, full-leg and above-knee US, and venogram. Normal distributions were defined for these parameters and incorporated for the probabilistic analysis. A log normal distribution was used for the number of days of heparin.

**Indirect Costs**
No indirect costs were reported.

**Currency**
UK pounds sterling (€).

**Sensitivity analysis**
Probability distributions were assigned to parameters used in the model. A Monte Carlo simulation was used to sample randomly from these distributions, the model being recalculated for each simulation. Several one-way sensitivity analyses were performed in addition to the probabilistic sensitivity analysis. Details were not reported but were available from the authors.

**Estimated benefits used in the economic analysis**
QALYs accrued were reported for each algorithm.

The strategy "no test, no treatment" showed 11,523 QALYs accrued for 1,000 patients.

Algorithm 1 (venography for all patients) showed the highest value with 11,560 QALYs accrued.

Algorithms 5, 7 and 8 showed 11,559 QALYs accrued.

Algorithms 2, 9, 10, 11 and 13 showed 11,558 QALYs accrued.

Algorithms 3, 6, 12, 14, 15 and 17 showed 11,557 QALYs accrued.

Algorithms 4, 16 and 18 showed 11,556 QALYs accrued.

**Cost results**
The total costs ranged from 144,040 for the "no test, no treatment" strategy to the highest value of 399,733 for algorithm 8, (Wells. High or intermediate: full-leg US, treat if positive, venogram if negative. Low: full-leg US, treat if positive, discharge if negative).

Algorithm 1 (venography for all patients) showed a total cost of 358,864.

**Synthesis of costs and benefits**
Strategy 16 (Wells. High: above-knee US. If positive, treat. If negative, SimpliRED DD. Repeat US if DD positive, discharge if DD negative. Intermediate or low: DD, discharge if negative, US if positive) showed the highest mean net benefit of 255,673 for a 10,000 willingness-to-pay.

Algorithm 9, (latex DD. If positive, above-knee US and repeat; if negative do Wells score. If high, US and repeat. If intermediate or low, discharge.) showed the highest mean net benefits of 597,675 and 948,356 for 20,000 and 30,000 willingness-to-pay, respectively.

Cost-effectiveness acceptability curves were reported. These showed the probability that an algorithm would be the most cost-effective at each value of willingness-to-pay from 0 to 100,000 per QALY.
Up to the 30,000 threshold, algorithms 16, 9 and 13 were the most likely to be optimal. For thresholds of 40,000 to 70,000 per QALY, algorithm 5 was most likely to be optimal. For thresholds of 80,000 to 100,000 per QALY, a strategy of venography for all patients was most likely to be optimal.

Authors’ conclusions
Diagnostic strategies for deep vein thrombosis (DVT) that involve radiological testing for all patients are unlikely to be cost-effective at currently recommended thresholds of willingness-to-pay. The authors recommended widespread adoption throughout the National Health Service (NHS) of a diagnostic strategy that uses Wells score and D-dimer (DD) to exclude DVT in low- and intermediate-risk patients.

CRD COMMENTARY - Selection of comparators
The choice of the algorithms was justified, mainly because combining simple diagnostic tests might provide an alternative of reducing the need for expensive, definite tests, and because published studies had not explicitly weighed the costs and benefits of different diagnostic approaches. The comparators were extensive, which helps the study to be relevant in many settings. Nevertheless, you should judge whether these algorithms are relevant in your own setting, or whether other diagnostic and treatment strategies could have been relevant as well.

Validity of estimate of measure of effectiveness
The authors used meta-analysis and systematic reviews carried out in their own studies (Goodacre et al. 2005 and 2006). They derived estimates of effectiveness from published literature, expert opinion and their own assumptions (which they justified). The estimates were investigated using a probabilistic sensitivity analysis. All input parameters of the model were reported extensively and, in some cases, they were available from authors upon request.

Validity of estimate of measure of benefit
The authors used QALYs as a measure of benefits. The quality-adjustment utility weights were estimated on the basis of published literature and expert opinion. The estimation of benefits was modelled through a decision analysis model to simulate DVT diagnosis and treatment. The model outputs were also compared for external validity with studies that reported cost-effectiveness calculations.

Validity of estimate of costs
The authors reported that the study had been conducted from the UK NHS perspective. Most of the relevant costs were included. Some stochastic analysis of the costs and sensitivity analyses were conducted, but the results were not reported in detail. The costs and the quantities were reported separately. Discounting was appropriately carried out since the time horizon of the model was longer than 2 years. The price year was reported, which will aid any future reflation exercises.

Other issues
The authors compared their findings with those from other studies, generally finding them to be concordant. The issue of the generalisability of the results to other settings was addressed. The authors' conclusions reflected the scope of the analysis.

The authors recognised potential limitations to their study. First, the assumption that US results were independent of Wells score and DD, which could favour strategies using US. Second, numerous potential combinations of tests were excluded because only algorithms supported by empirical data showing that they work in practice were included. Third, the model did not allow the potential impact of the diagnostic strategy upon the selection of patients to be determined, and whether this influenced cost-effectiveness. Finally, the findings might not apply to certain patient groups, such as inpatients developing symptoms of DVT, patients with suspected recurrent DVT, pregnant patients, intravenous drug abusers, or those with prolonged symptoms.
Implications of the study
Future research is needed to evaluate whether patient selection is influenced by the diagnostic tests used and to determine whether it has consequences for cost-effectiveness. Also, there is very little empirical data on algorithms involving plethysmography and how it interacts with other tests. The authors recommended the widespread adoption of a diagnostic strategy that uses Wells score and DD in low- and intermediate-risk patients.

Source of funding
Funded by the NHS Health Technology Assessment R&D Programme.

Bibliographic details

DOI
10.1093/qjmed/hcl051

Other publications of related interest


Indexing Status
Subject indexing assigned by NLM

MeSH
Cost-Benefit Analysis; Decision Support Techniques; Diagnostic Techniques, Cardiovascular /economics; Great Britain; Humans; Quality-Adjusted Life Years; State Medicine /economics; Venous Thrombosis /diagnosis

AccessionNumber
22006001416

Date bibliographic record published
31/01/2007

Date abstract record published
31/01/2007