Article:
Goodacre, S. orcid.org/0000-0003-0803-8444, Horspool, K., Shephard, N. et al. (8 more authors) (2018) Selecting pregnant or postpartum women with suspected pulmonary embolism for diagnostic imaging: the DiPEP diagnostic study with decision-analysis modelling. Health Technology Assessment, 22 (47). ISSN 1366-5278

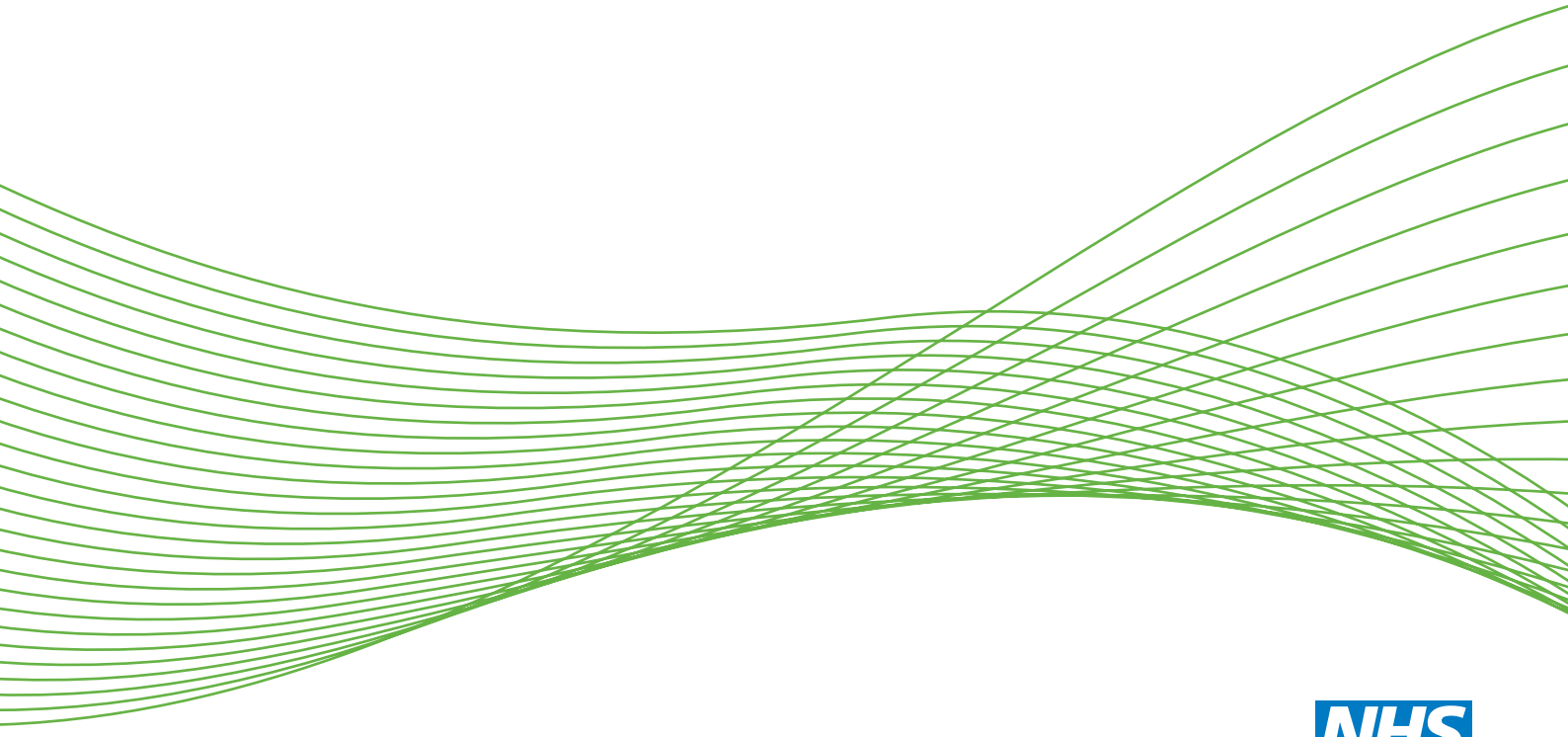
https://doi.org/10.3310/hta22470

HEALTH TECHNOLOGY ASSESSMENT

VOLUME 22 ISSUE 47 AUGUST 2018
ISSN 1366-5278

Selecting pregnant or postpartum women with suspected pulmonary embolism for diagnostic imaging: the DiPEP diagnostic study with decision-analysis modelling

Steve Goodacre, Kimberley Horspool, Neil Shephard, Daniel Pollard, Beverley J Hunt, Gordon Fuller, Catherine Nelson-Piercy, Marian Knight, Steven Thomas, Fiona Lecky and Judith Cohen on behalf of the DiPEP research group



NHS
National Institute for Health Research

DOI 10.3310/hta22470

Selecting pregnant or postpartum women with suspected pulmonary embolism for diagnostic imaging: the DiPEP diagnostic study with decision-analysis modelling

Steve Goodacre,[1]* Kimberley Horspool,[1] Neil Shephard,[1] Daniel Pollard,[1] Beverley J Hunt,[2] Gordon Fuller,[1] Catherine Nelson-Piercy,[2] Marian Knight,[3] Steven Thomas,[4] Fiona Lecky[1] and Judith Cohen[1] on behalf of the DiPEP research group

[1]School of Health and Related Research, University of Sheffield, Sheffield, UK
[2]Guy's and St Thomas' NHS Foundation Trust, London, UK
[3]National Perinatal Epidemiology Unit, University of Oxford, Oxford, UK
[4]Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK

*Corresponding author

Declared competing interests of authors: Steve Goodacre is the chairperson of the National Institute for Health Research (NIHR) Health Technology Assessment (HTA) programme Clinical Evaluation and Trials Board and a member of the HTA Funding Boards Policy Group. Fiona Lecky is a member of the NIHR HTA Emergency and Hospital Care Panel. Catherine Nelson-Piercy has received personal fees from Leo Pharma (Leo Pharma A/S, Copenhagen, Denmark) and personal fees from Sanofi-Aventis (Sanofi SA, Paris, France) outside the submitted work.

Published August 2018
DOI: 10.3310/hta22470

This report should be referenced as follows:

Goodacre S, Horspool K, Shephard N, Pollard D, Hunt BJ, Fuller G, *et al*. Selecting pregnant or postpartum women with suspected pulmonary embolism for diagnostic imaging: the DiPEP diagnostic study with decision-analysis modelling. *Health Technol Assess* 2018;**22**(47).

Health Technology Assessment is indexed and abstracted in *Index Medicus*/MEDLINE, *Excerpta Medica*/EMBASE, *Science Citation Index Expanded* (SciSearch®) and *Current Contents®*/ Clinical Medicine.

Health Technology Assessment

HTA/HTA TAR

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 4.513

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, The Cochrane Library and the Clarivate Analytics Science Citation Index.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: journals.library@nihr.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

The HTA programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

For more information about the HTA programme please visit the website: http://www.nets.nihr.ac.uk/programmes/hta

This report

The research reported in this issue of the journal was funded by the HTA programme as project number 13/21/01. The contractual start date was in October 2014. The draft report began editorial review in June 2017 and was accepted for publication in November 2017. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health and Social Care. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health and Social Care.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).

Abstract

Selecting pregnant or postpartum women with suspected pulmonary embolism for diagnostic imaging: the DiPEP diagnostic study with decision-analysis modelling

Steve Goodacre,[1]* Kimberley Horspool,[1] Neil Shephard,[1] Daniel Pollard,[1] Beverley J Hunt,[2] Gordon Fuller,[1] Catherine Nelson-Piercy,[2] Marian Knight,[3] Steven Thomas,[4] Fiona Lecky[1] and Judith Cohen[1] on behalf of the DiPEP research group

[1]School of Health and Related Research, University of Sheffield, Sheffield, UK
[2]Guy's and St Thomas' NHS Foundation Trust, London, UK
[3]National Perinatal Epidemiology Unit, University of Oxford, Oxford, UK
[4]Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK

*Corresponding author s.goodacre@sheffield.ac.uk

Background: Pulmonary embolism (PE) is a leading cause of death in pregnancy and post partum, but the symptoms of PE are common in normal pregnancy. Simple diagnostic tests are needed to select women for diagnostic imaging.

Objective: To estimate the accuracy, effectiveness and cost-effectiveness of clinical features, decision rules and biomarkers for selecting pregnant or postpartum women with a suspected PE for imaging.

Design: An expert consensus study to develop new clinical decision rules, a case–control study of women with a diagnosed PE or a suspected PE, a biomarker study of women with a suspected PE or diagnosed deep-vein thrombosis (DVT) and decision-analysis modelling.

Setting: Emergency departments and consultant-led maternity units.

Participants: Pregnant/postpartum women with a diagnosed PE from any hospital reporting to the UK Obstetric Surveillance System research platform and pregnant/postpartum women with a suspected PE or diagnosed DVT at 11 prospectively recruiting sites.

Interventions: Clinical features, decision rules and biomarkers.

Main outcome measures: Sensitivity, specificity, area under receiver operating characteristic (AUROC) curve, quality-adjusted life-years (QALYs) and health-care costs.

Results: The primary analysis involved 181 women with PE and 259 women without PE in the case–control study and 18 women with DVT, 18 with PE and 247 women without either in the biomarker study. Most clinical features showed no association with PE. The AUROC curves for the clinical decision rules were as follows: primary consensus, 0.626; sensitive consensus, 0.620; specific consensus, 0.589; PE rule-out criteria, 0.621; simplified Geneva score, 0.579; Wells's PE criteria (permissive), 0.577; and Wells's PE criteria (strict), 0.732. The sensitivities and specificities of the D-dimer measurement were 88.4% and 8.8%, respectively, using a standard threshold, and 69.8% and 32.8%, respectively, using a pregnancy-specific threshold. Previous venous thromboembolism, long-haul travel, multiple pregnancy, oxygen saturation, recent surgery,

temperature and PE-related chest radiograph abnormality were predictors of PE on multivariable analysis. We were unable to derive a rule through multivariable analysis or recursive partitioning with adequate accuracy. The AUROC curves for the biomarkers were as follows: activated partial thromboplastin time – 0.669, B-type natriuretic peptide – 0.549, C-reactive protein – 0.542, Clauss fibrinogen – 0.589, enzyme-linked immunosorbent assay D-dimer – 0.668, Innovance D-dimer (Siemens Healthcare Diagnostics Products GmbH, distributed by Sysmex UK Ltd, Milton Keynes, UK) – 0.651, mid-regional pro-atrial natriuretic peptide (MRproANP) – 0.524, prothrombin fragment 1 + 2 – 0.562, plasmin-antiplasmin – 0.639, Prothombin time – 0.613, thrombin generation lag time – 0.702, thrombin generation endogenous potential – 0.559, thrombin generation peak – 0.596, thrombin generation time to peak – 0.655, tissue factor – 0.531 and troponin – 0.597. The repeat analysis excluding women who had received anticoagulation was limited by the small number of women with PE ($n = 4$). The health economic analysis showed that a strategy of scanning all women with a suspected PE accrued more QALYs and incurred fewer costs than any selective strategy based on a clinical decision rule and was therefore the dominant strategy.

Limitations: The findings apply specifically to the diagnostic assessment of women with a suspected PE in secondary care.

Conclusions: Clinical features, decision rules and biomarkers do not accurately, effectively or cost-effectively select pregnant or postpartum women with a suspected PE for diagnostic imaging.

Future work: New diagnostic technologies need to be developed to detect PE in pregnancy.

Trial registration: Current Controlled Trials ISRCTN21245595.

Funding details: This project was funded by the National Institute for Health Research (NIHR) Health Technology Assessment programme and will be published in full in *Health Technology Assessment*; Vol. 22, No. 47. See the NIHR Journals Library website for further project information.

Contents

List of tables

List of figures

List of supplementary material

Report Supplementary Material 1 Diagnosed pulmonary embolism data collection form

Report Supplementary Material 2 Suspected pulmonary embolism screening form

Report Supplementary Material 3 Suspected pulmonary embolism data collection form

Report Supplementary Material 4 Non-recruited suspected pulmonary embolism screening form

Report Supplementary Material 5 Non-recruited suspected pulmonary embolism data collection form

Report Supplementary Material 6 Diagnosed deep-vein thrombosis screening form

Report Supplementary Material 7 Diagnosed deep-vein thrombosis data collection form

Supplementary material can be found on the NIHR Journals Library report project page (www.journalslibrary.nihr.ac.uk/programmes/hta/132101/#/documentation).

Supplementary material has been provided by the authors to support the report and any files provided at submission will have been seen by peer reviewers, but not extensively reviewed. Any supplementary material provided at a later stage in the process may not have been peer reviewed.

List of abbreviations

A–a	Alveolar–arterial
AIC	Akaike information criterion
APTT	activated partial thromboplastin time
AUC	area under the curve
AUROC	area under the receiver operating characteristic
BIC	Bayesian information criterion
BMI	body mass index
BNP	B-type natriuretic peptide
CDR	clinical decision rule
CI	confidence interval
CRF	case report form
CRP	C-reactive protein
CT	computerised tomography
CTEPH	chronic thromboembolic pulmonary hypertension
CTPA	computerised tomography pulmonary angiography
CTRU	Clinical Trials Research Unit
DiPEP	Diagnosis of Pulmonary Embolism in Pregnancy
DVT	deep-vein thrombosis
ECG	electrocardiogram
ELISA	enzyme-linked immunosorbent assay
EQ-5D-5L	EuroQol-5 Dimensions, five-level version
GP	general practitioner
GSTT	Guy's and St Thomas' NHS Foundation Trust
ICER	incremental cost-effectiveness ratio
LASSO	least absolute shrinkage and selection operator
LMWH	low-molecular-weight heparin
MAICER	maximum acceptable incremental cost-effectiveness ratio
MRProANP	mid-regional pro-atrial natriuretic peptide
NGT	nominal group technique
NICE	National Institute for Health and Care Excellence
ONS	Office for National Statistics
PE	pulmonary embolism
PERC	pulmonary embolism rule-out criteria
PF 1 + 2	prothrombin fragment 1 + 2
PPI	patient and public involvement
PPP	platelet-poor plasma
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PSA	probabilistic sensitivity analysis
PT	prothrombin time
QALY	quality-adjusted life-year
RCOG	Royal College of Obstetricians and Gynaecologists
ROC	receiver operating characteristic
SE	standard error
SNTA	scanning no women, but treating all
SNTN	scanning no women and treating no women
TG	thrombin generation
UKOSS	UK Obstetric Surveillance System
VKA	vitamin K antagonist
VQ	ventilation–perfusion
VQ SPECT	ventilation–perfusion single photon emission computed tomography
VTE	venous thromboembolism

Plain English summary

A blood clot in the lung is a potentially fatal complication of pregnancy that can be difficult to diagnose. Symptoms that suggest a blood clot, such as chest pain or breathlessness, are common in pregnancy. Diagnosis usually requires a scan that involves giving a small dose of radiation to the mother and possibly to the baby.

A clinical decision rule uses information from the woman's medical history and examination to estimate the risk that she has a blood clot. Blood tests that are abnormal in people with blood clots can perform a similar role. We wanted to find out whether or not clinical decision rules or blood tests could be used to decide which women with a suspected blood clot should have a scan.

We collected information from 181 pregnant or recently pregnant women with blood clots in their lungs and 259 women without blood clots who had been investigated in hospital for a suspected blood clot. We also collected blood samples from 36 women with blood clots in their lungs or legs, and 247 with no blood clot. We found that the blood clots were very difficult to diagnose without a scan. None of the clinical decision rules or blood tests was able to reliably determine which women had a blood clot. The economic analysis showed that scanning every woman with a suspected blood clot was a worthwhile use of NHS resources. This is because the risks of scanning are very small, whereas the benefits of detecting and treating blood clots are very large.

Clinical decision rules and blood tests should not be used to select which women with a suspected blood clot in pregnancy have a scan. Future research needs to develop new ways of diagnosing blood clots in pregnancy.

Scientific summary

Background

Pulmonary embolism (PE) is a leading cause of death in pregnancy and post partum. Symptoms suggesting PE are very common in pregnancy and post partum. As a consequence, many pregnant and postpartum women undergo radiological investigation for a suspected PE with a low yield of positive diagnosis. Clinical decision rules use features of the patient history and examination in a structured manner to estimate the probability of disease. A number of biomarkers are known to be increased in the presence of PE. Clinical decision rules or biomarkers could be used to select women with suspected PE for radiological investigation or discharge without imaging.

Objectives

We aimed to estimate the diagnostic accuracy, effectiveness and cost-effectiveness of strategies (including clinical decision rules) for selecting pregnant or postpartum women with a suspected PE for imaging, and determine the feasibility and value of information of further prospective research.

Our specific objectives were to:

- use expert consensus to derive three new clinical decision rules (with different trade-offs between sensitivity and specificity) for pregnant and postpartum women with a suspected PE
- estimate the diagnostic accuracy of clinical variables, our expert-derived clinical decision rules, existing clinical decision rules [Wells's PE criteria, Geneva score and a PE rule-out criteria (PERC)] and the D-dimer measurement in pregnant and postpartum women with suspected PE
- use a statistical analysis of women with a diagnosed or suspected PE to derive a new clinical decision rule for pregnant and postpartum women with suspected PE
- explore the potential diagnostic value of biomarkers for PE in pregnant and postpartum women
- determine the feasibility of using a prospective cohort design to validate a new clinical decision rule or biomarker
- estimate the effectiveness of different strategies, in terms of adverse outcomes from venous thromboembolism (VTE), bleeding and radiation exposure, and cost-effectiveness, measured as the incremental cost per quality-adjusted life-year (QALY)
- estimate the value of information associated with further research.

Methods

The study involved (1) an expert consensus study to develop three new clinical decision rules; (2) a case–control study of women with a diagnosed PE identified through the UK Obstetric Surveillance System (UKOSS) research platform and women with a suspected PE recruited from emergency departments and maternity units at 11 prospectively recruiting sites; (3) a biomarker study involving the prospectively recruited women and additional women with diagnosed deep-vein thrombosis (DVT); and (4) decision-analysis modelling of effectiveness, cost-effectiveness and value of information.

The study population included (1) any pregnant or postpartum women with a diagnosed PE who had presented with suspected PE to a hospital reporting to the UKOSS research platform; (2) pregnant and postpartum women presenting with suspected PE to 11 prospectively recruiting sites; and (3) women with DVT diagnosed at the prospectively recruiting sites. We excluded women who required resuscitation at

presentation from all groups and those who were unable to consent or who had an existing diagnosis of PE from the prospectively recruited group.

The nominated clinician for UKOSS and the research nurse/midwife at prospectively recruiting sites collected data detailing potential clinical predictors, blood tests results, diagnostic imaging, treatment and adverse events. Research nurses/midwives also collected a blood sample from women with suspected PE or diagnosed DVT at the prospectively recruiting sites, and reviewed hospital records at 30 days. Prospectively recruited women were then sent a questionnaire to record adverse events, health-care use and health utility. Two independent assessors, blind to clinical predictors and blood results, classified participants as having PE using diagnostic imaging results and details of treatments and adverse events. The primary analysis was limited to women with PE diagnosed by imaging or post-mortem examination, and women with PE ruled out after imaging. Secondary analyses explored the impact of including women with clinically diagnosed PE or PE ruled out without imaging, and the impact of excluding subsegmental PE.

Blood samples were centrifuged, stored and then transported to Guy's and St Thomas' NHS Foundation Trust for analysis using the following assays: D-dimer [enzyme-linked immunosorbent assay (ELISA)], D-dimer [Innovance (Siemens Healthcare Diagnostics Products GmbH, distributed by Sysmex UK Ltd, Milton Keynes, UK)], plasmin–antiplasmin, prothrombin fragment 1 + 2 (PF 1 + 2), thrombin generation, prothrombin time, activated partial thromboplastin time (APTT), Clauss fibrinogen, soluble tissue factor, troponin I, B-type natriuretic peptide (BNP), C-reactive protein (CRP) and mid-regional pro-atrial natriuretic peptide (MRproANP).

The sample size was ultimately determined by the incidence of a diagnosed and suspected PE, but we estimated that over 18 months we would identify 150 women with a diagnosed PE and 250 women with a suspected PE, resulting in around 155 patients and 245 controls. This would allow the estimation of sensitivity or specificity of 90% with standard errors (SEs) of around 2.5% and 2.0%, respectively. We increased the planned sample size after starting recruitment to ensure that adequate numbers would be included in the primary analysis.

Logistic regression was used to identify associations between clinical predictors and a PE diagnosis. The diagnostic performance of existing clinical decision rules (Wells's PE criteria, simplified revised Geneva score and PERC rule) and those developed by expert consensus was assessed by constructing receiver operating characteristic (ROC) curves, calculating the area under the curve (AUC) and calculating the sensitivity and specificity at key decision-making thresholds. The diagnostic performance of biomarkers was assessed by comparing distributions in women with and without PE, constructing ROC curves, calculating the AUC and calculating sensitivity and specificity at a predefined threshold based on the 99th percentile for a normal population.

Decision-analysis modelling was used to estimate the costs incurred and the expected outcomes from thromboembolism, bleeding and radiation exposure if a hypothetical cohort of pregnant or postpartum women based on the study population was investigated for suspected PE using different strategies, including no imaging, selective imaging and imaging for all. Outcomes were modelled to estimate the QALYs accrued by each strategy and the incremental cost per QALY gained by each strategy compared with the next most effective alternative.

Results

The expert consensus study derived three clinical decision rules for use in pregnant and postpartum women with a suspected PE: a primary rule that provided an optimal balance of sensitivity and specificity, a sensitive rule that maximised sensitivity at the expense of specificity and a specific rule that maximised specificity at the expense of sensitivity.

We identified 198 women with a diagnosed PE who met our inclusion criteria, of whom 163 had a PE confirmed by imaging or post-mortem examination and were included in the primary analysis. We identified 324 women with suspected PE, of whom 18 had PE confirmed by imaging and 259 had PE ruled out after imaging. The primary analysis therefore involved 181 women with PE and 259 women without PE.

Univariable logistic regression showed that the number of previous pregnancies beyond 24 weeks' gestation ($p = 0.017$), surgery (including caesarean section) in the previous 4 weeks ($p = 0.001$), no history of varicose veins ($p = 0.045$), no long-haul travel during pregnancy ($p = 0.006$), receiving thromboprophylaxis ($p < 0.001$), higher temperature ($p = 0.003$), lower oxygen saturation ($p < 0.001$), overall diagnostic impression, suggesting PE using a strict interpretation ($p < 0.001$), PE-related chest radiograph abnormality ($p = 0.01$) and non-PE-related chest radiograph abnormality ($p = 0.001$) were associated with PE. All other clinical features showed no significant association with PE.

The AUC and sensitivity and specificity at the usual recommended threshold for the clinical decision rules were 0.626, 60.9% and 58.5% for the primary consensus rule; 0.620, 95.9% and 3.5% for the sensitive consensus rule; 0.589, 36.1% and 78.3% for the specific consensus rule; 0.621, 67.5% and 51.9% for the PERC score; 0.579, 44.4% and 63.6% for the simplified Geneva score; 0.577, 49.0% and 61.7% for Wells's PE criteria using a permissive interpretation of diagnostic impression; and 0.732, 37.6% and 89.5% for Wells's PE criteria using a strict interpretation of diagnostic impression.

D-dimer measurements were recorded as part of routine care for 44 out of 198 (22%) women with a diagnosed PE and 156 out of 324 (48%) women with a suspected PE. The primary analysis, using results from 43 women with PE and 125 without PE, showed that sensitivity and specificity were 88.4% [95% confidence interval (CI) 74.1% to 95.6%] and 8.8% (95% CI 4.7% to 15.6%) using the hospital laboratory threshold, and 69.8% (95% CI 53.7% to 82.3%) and 32.8% (95% CI 24.8% to 41.9%) using predefined gestation-specific thresholds.

Multivariable analysis showed that the most accurate model used previous VTE, long-haul travel during pregnancy, multiple pregnancy, oxygen saturation (as a continuous variable), surgery in the previous 4 weeks, temperature (as a continuous variable) and PE-related chest radiograph abnormality to predict PE with an AUC of 0.724 (95% CI 0.669 to 0.779). The ROC curve shows that specificity would have to be as low as 20% to achieve a level of sensitivity (> 95%) that was acceptable to allow imaging to be avoided. We therefore did not proceed to internal validation or attempt to make the model more clinically credible or usable.

The optimal model developed by recursive partitioning used body mass index (BMI), trimester, oxygen saturation and heart rate. The AUC was 0.657 (95% CI 0.611 to 0.703) and the threshold that provided a level of sensitivity of > 95% had a corresponding specificity of 5%.

Usable blood samples were taken from 18 women with diagnosed DVT and 310 women with suspected PE, of whom 18 had PE confirmed by imaging and 247 had PE ruled out after imaging and were included in the primary analysis. Mean biomarker levels significantly differed between women with and without PE only for Clauss fibrinogen ($p = 0.007$), ELISA D-dimer ($p = 0.001$), Innovance D-dimer ($p = 0.004$), thrombin generation lag time ($p < 0.001$), thrombin generation time to peak ($p = 0.001$) and plasmin antiplasmin ($p = 0.004$). The AUC for each biomarker was as follows: 0.669 (95% CI 0.570 to 0.768) for APTT, 0.549 (95% CI 0.453 to 0.645) for BNP, 0.542 (95% CI 0.445 to 0.639) for CRP, 0.589 (95% CI 0.476 to 0.701) for Clauss fibrinogen, 0.668 (95% CI 0.561 to 0.776) for the ELISA D-dimer, 0.651 (95% CI 0.545 to 0.758) for the Innovance D-dimer, 0.524 (95% CI 0.418 to 0.630) for MRproANP, 0.562 (95% CI 0.462 to 0.661) for PF 1 + 2, 0.639 (95% CI 0.536 to 0.742) for plasmin–antiplasmin, 0.613 (95% CI 0.508 to 0.718) for prothombin time, 0.702 (95% CI 0.598 to 0.806) for thrombin generation lag time, 0.559 (95% CI 0.437 to 0.681) for thrombin generation endogenous potential, 0.596 (95% CI 0.478 to 0.715) for thrombin generation peak, 0.655 (95% CI 0.541 to 0.769) for thrombin generation time to peak, 0.531 (95% CI 0.424 to 0.638) for tissue factor and 0.597 (95% CI 0.499 to 0.695) for troponin. The ROC curve analysis showed that only

thrombin generation lag time had any potential to rule out PE with sufficient sensitivity while achieving meaningful specificity, with a sensitivity of 97% and a specificity of 25% at the threshold that optimised sensitivity. The repeat analysis excluding women who had received anticoagulation was limited by the small number of women who had PE ($n = 4$).

The study recruited women with suspected PE (prevalence of 7.1%) at a rate of 1.7 women per site per month. This suggests that a prospective cohort study would require 50 sites to recruit for 2 years to achieve a sample size of 2040, including 145 women with PE, which would be sufficient to estimate the sensitivity with acceptable precision.

The health economic analysis showed that a strategy of scanning all women with suspected PE accrued more QALYs and incurred fewer costs than any selective strategy based on a clinical decision rule, and was therefore the dominant strategy. This finding was robust in the sensitivity analysis and the scenario analysis exploring assumptions in the model. A threshold analysis showed that a clinical decision rule to select women for imaging would need to have a sensitivity exceeding 97.5% to be cost-effective compared with the non-selective use of scanning. The value-of-information analysis showed that the value of conducting further research into parameters used in the economic model was likely to be below the cost of conducting further research into any subset of feasible parameters.

Conclusions

We were unable to identify any clinical decision rule or biomarker that could be used to rule out PE in pregnant and postpartum women with acceptable sensitivity while achieving worthwhile specificity. Decision-analysis modelling showed that a strategy of non-selective scanning for all women dominated selective strategies based on decision rules. We found that many clinical features thought to be diagnostically useful for PE showed either no association or a counter-intuitive association with the absence of PE. This may be explained by the selection of women for investigation in secondary care. Those with risk factors for PE or clinical features suggesting PE may be more likely to be referred or to self-present for investigation. The prevalence of PE in those with suspected PE (7.1% overall and 6.5% in the primary analysis population) was higher than suggested by previous data, indicating that, potentially, the NHS is already selecting an appropriate population for hospital investigation.

The accuracy of the biomarkers is likely to have been undermined by the receipt of anticoagulation prior to sampling, but the removal of samples from women who had received anticoagulation left too few women with PE for a meaningful analysis. This highlights a significant practical problem in testing and using biomarkers when guidelines recommend thromboprophylaxis for many women and early use of anticoagulation if PE is suspected.

Our findings do not support the use of clinical decision rules and biomarkers (including D-dimer) in selecting women with suspected PE for imaging. We cannot conclude that all women should receive imaging, as a proportion of the study cohort with suspected PE did not receive imaging and we found no evidence of missed PE. However, a low threshold for scanning is likely to be appropriate given the costs and risks of misdiagnosis highlighted in the decision-analysis modelling.

We have shown that a prospective cohort study to derive or validate a clinical decision rule or biomarker would be feasible, albeit would require a large number of sites (more than one-quarter of all maternity units in the UK) and substantial resources. However, the accuracy of decision rules and biomarkers reported in our study is insufficient to justify a large prospective cohort study to derive a new decision rule or test existing decision rules or biomarkers. Future research efforts would be better directed at developing new biomarkers or alternative diagnostic techniques.

The current Royal College of Obstetricians and Gynaecologists guidance suggests that women should be given information about the risks and benefits of investigation and involved in decision-making. Our

decision-analysis model has identified data sources and methods for weighing the relative risks and benefits of imaging, but has also highlighted the complexity of decision-making. Future research could be used to develop better ways of presenting information regarding the relative risks and benefits of investigation for suspected PE in pregnancy and post partum.

Trial registration

This trial is registered as ISRCTN21245595.

Funding

Funding for this study was provided by the Health Technology Assessment programme of the National Institute for Health Research.

Chapter 1 Introduction

Background and rationale

Pregnant and postpartum women are at an increased risk of developing venous thromboembolism (VTE), which may involve deep-vein thrombosis (DVT) (a blood clot in the veins of a limb, which may be clinically silent or cause limb pain and/or swelling) or a pulmonary embolism (PE) (a blood clot in the artery of the lungs, causing respiratory symptoms such as shortness of breath and chest pain).

Pulmonary embolism is a leading cause of death in pregnancy and post partum, affecting women who would otherwise expect to have a long life expectancy in full health. The outcome for the fetus is dependent on the outcome for the mother, so maternal mortality, which is currently estimated at 0.85 [95% confidence interval (CI) 0.52 to 1.32] per 100,000 maternities,[1] and morbidity associated with PE has inevitable consequences for fetal mortality and morbidity. Women with an appropriately diagnosed and treated PE are at a low risk of experiencing adverse outcomes, so accurate diagnosis can result in substantial benefits. However, the investigations used to diagnose PE [diagnostic imaging with ventilation–perfusion (VQ) scanning or computerised tomography (CT) pulmonary angiography (CTPA)] carry risks of radiation exposure, risk of reaction to contrast media and false-positive diagnoses, are inconvenient for patients and incur costs for health services. Clinicians investigating suspected PE in pregnant and postpartum women therefore need to choose between risking the potentially catastrophic consequences of a missed diagnosis if imaging is withheld and risking iatrogenic harm to women without PE if imaging is overused.

Pregnant and postpartum women with symptoms suggesting PE could be selected for diagnostic imaging on the basis of clinical features or blood tests (biomarkers). A previous history or family history of VTE, immobilisation, surgery and a number of medical and obstetric complications are known to be associated with an increased risk of VTE.[2] Abnormal observations, such as a rapid heart rate, rapid respiratory rate or reduced peripheral oxygen saturation, may be caused by PE, although these may be caused by other pathologies or a normal physiological response to pregnancy.

Individual clinical features are unlikely to have sufficient accuracy to select women for diagnostic imaging, but could be combined to form a clinical decision rule (CDR). This uses a number of clinical features in a structured manner to generate an estimate of the clinical risk of PE or a rule to determine whether or not PE should be investigated. In the general (non-pregnant) population with suspected PE, Wells's score[3] and revised Geneva score[4] have been developed to estimate the risk of developing PE, whereas the PE rule-out criteria (PERC) rule[5] has been developed to select patients for investigation (details of the scores and the rule are provided in *Chapter 3*). These scores and the rule have been extensively validated in the general population with a suspected PE, but the differences between the pregnant and non-pregnant populations mean that findings cannot be automatically extrapolated to the pregnant or postpartum population.

A number of biomarkers have been suggested for use in the PE diagnosis but, to date, only the D-dimer measurement has been used in routine clinical practice. Plasma D-dimers are specific cross-linked fibrin derivatives produced when fibrin is degraded by plasmin, with elevated levels indicating thrombolysis. They are elevated in VTE but also in other conditions such as pregnancy, pre-eclampsia, infections, malignancy and surgery. The D-dimer threshold for positivity is usually set to optimise sensitivity (> 95%) at the expense of specificity. In the general population with a suspected PE, the D-dimer measurement has been recommended alongside a clinical risk score (such as Wells's score) as a way of ruling out PE in low-risk patients without the need for diagnostic imaging. The lack of specificity in the pregnant and postpartum population means that separate validation in this population is required, perhaps using a pregnancy-specific threshold for positivity. There is some evidence that using a higher threshold for positivity can improve the D-dimer specificity in pregnancy without compromising sensitivity.[6]

In summary, although clinical features (structured as a CDR) and the D-dimer measurement are widely used to select patients with a suspected PE for diagnostic imaging in the general population with a suspected PE, evidence of their performance in the relevant population is required before they can be advocated for use in pregnant or postpartum women with a suspected PE.

Literature review

Diagnostic studies of pregnant or postpartum women undergoing imaging for a suspected PE could provide evidence to support the use of clinical features, decision rules or biomarkers to select women for imaging if they compare these index tests to an imaging reference standard. They could also provide estimates of the prevalence of PE in the investigated population to inform the design of future studies.

In January 2014, we systematically searched electronic databases for diagnostic studies of pregnant or postpartum women undergoing imaging for a suspected PE[7] and identified 11 relevant articles, along with a conference abstract and a paper in press. We have since updated the literature searches and have identified an additional four papers, along with the published version of the paper in press.[8] These are outlined in *Table 1*.

In addition to these studies of pregnant and postpartum women with a suspected PE, Kline *et al.*[25] undertook a systematic review of studies of people with suspected PE, which included pregnant and postpartum women. The authors identified 17 studies including 25,399 patients, of whom 506 (2%) were pregnant, with a 4.1% (95% CI 2.6% to 6.0%) prevalence of PE.

The analysis reported by Kline *et al.*[25] and 10 of the studies identified by our review reported the overall prevalence of PE, which was generally found to be low when compared with the non-pregnant population, but did not examine the diagnostic accuracy of clinical features, CDRs or the D-dimer measurement. The remaining seven studies were mostly small and had a low prevalence of PE, and thus had limited power to estimate diagnostic accuracy (especially sensitivity) or detect an association with a reference standard diagnosis of PE.

Cahill *et al.*[12] found that chest pain and low oxygen saturation were associated with a diagnosis of PE, but other features [dyspnoea, tachycardia, Alveolar–arterial (A–a) gradient] showed no evidence of association. Deutsch *et al.*[15] also found that chest pain showed some association with a diagnosis of PE, while other features (dyspnoea, heart rate, respiratory rate, blood pressure, oxygen saturation, A–a gradient) did not. Bourjeily *et al.*[18] found no association between dyspnoea, chest pain, pleuritic chest pain, haemoptysis, cough, DVT signs, wheeze, pleural rub, heart rate, respiratory rate or systolic blood pressure and a diagnosis of PE.

Two studies have suggested that the modified Wells's score may be useful in pregnant or postpartum women. O'Connor *et al.*[17] reported that a modified Wells's score of ≥ 6 units (meaning that PE is likely) has a sensitivity of 100% and a specificity of 90% for PE, whereas Cutts *et al.*[8] reported a sensitivity of 100% (95% CI 40% to 100%) and a specificity of 60% (52% to 67%). The wide CIs for sensitivity mean that further research is required. Other CDRs, such as the Geneva score and the PERC rule, have not yet been tested in pregnant or postpartum women with a suspected PE.

Studies of the D-dimer measurement in pregnant and postpartum women[8,13,16,17] suggest that high levels of positivity at conventional thresholds limit the diagnostic value of this test. However, indirect evidence from studies of the D-dimer measurement for suspected DVT in pregnancy suggests potential diagnostic value. Chan *et al.*[26] reported 100% sensitivity (95% CI 77% to 100%) and 60% specificity (95% CI 52% to 68%) for the qualitative SimpliRED (Agen Biomedical, Brisbane, QLD, Australia) D-dimer in suspected DVT and, although another study of five commercially available assays[6] reported specificities ranging from 6% to 23%, further analysis suggested that using a higher threshold for positivity could improve sensitivity without compromising specificity.

TABLE 1 Diagnostic studies of pregnant or postpartum women with a suspected PE

Study (year of publication)	Country	Population	Index tests	Reference standard	Main findings
Balan *et al.* (1997)[9]	UK	82 pregnant women, one hospital, 5 years	None	VQ scan	VQ scan: 31 (38%) normal, 19 (23%) low probability, 14 (17%) intermediate, 18 (22%) high probability
Chan *et al.* (2002)[10]	Canada	113 pregnant women, two hospitals, 4 and 10 years	None	VQ scan	VQ scan: 83 (73.5%) normal, 28 (24.8%) non-diagnostic, two (1.8%) high probability
Scarsbrook *et al.* (2007)[11]	UK	94 pregnant women, one hospital, 5 years	None	VQ scan	VQ scan: 89 (92%) normal, seven (7%) non-diagnostic, one (1%) high probability
Cahill *et al.* (2009)[12]	USA	304 pregnant or postpartum women, one hospital, 5 years	Clinical features[a]	108 CTPA and 196 VQ scan	CTPA: 18 (5.9%) diagnosed PE Clinical features: low oxygen saturation and chest pain predicted PE, other features did not
Damodaram *et al.* (2009)[13]	UK	37 pregnant women, one hospital, 4 years	D-dimer	VQ scan	VQ scan: 13 (35%) low probability, 24 (65%) intermediate or high probability D-dimer: 73% sensitivity, 15% specificity
Shahir *et al.* (2010)[14]	USA	199 pregnant women, one hospital, 8 years	None	106 CTPA and 99 VQ scan	CTPA: 4/106 (3.7%) PE VQ scans: zero high probability, two (2%) intermediate probability, 19 (19%) low probability, 14 (14%) very low probability, 63 (64%) normal, one (1%) inconclusive
Deutsch *et al.* (2010)[15]	USA	102 pregnant or postpartum women, one hospital, 7 years	Clinical features[b]	CTPA	CTPA: 13/102 (13%) PE Clinical features: only chest pain predicted PE
Hassanin *et al.* (2011)[16]	Egypt	60 postpartum women, one hospital, years not reported	D-dimer	CTPA	CTPA: four (6.6%) PE D-dimer: positive in all patients
O'Connor *et al.* (2011)[17]	Ireland	125 pregnant or postpartum women, one hospital, 5 years	Modified Wells's score D-dimer Arterial blood gas measurement with PE ECG	CTPA	CTPA: 5/103 (5%) PE Modified Wells's score: 100% sensitivity, 90% specificity D-dimer: 0% sensitivity, 74% specificity

continued

TABLE 1 Diagnostic studies of pregnant or postpartum women with a suspected PE (*continued*)

Study (year of publication)	Country	Population	Index tests	Reference standard	Main findings
Bourjeily *et al.* (2012)[18]	USA	343 pregnant women, one hospital, 5 years	Clinical features[c]	CTPA	CTPA: eight (2.3%) PE Clinical features: no association found between clinical features and PE
Abele and Sunner (2013)[19]	Canada	74 pregnant women, three hospitals, 1.5 years	None	Perfusion scan and CTPA if abnormal	Perfusion scan: 61 (82.4%) normal perfusion, 13 (17.6%) abnormal – one (1.4%) PE on CTPA
Nijkeuter [(2013) abstract][20]	The Netherlands	149 pregnant women, three hospitals, 9 years	None	CTPA	CTPA: six (4.2%) PE, eight (5.6%) inconclusive, 129 (90.2%) normal
Cutts *et al.* (2014)[8]	UK and Australia	183 pregnant women, two hospitals, 4 years	Modified Wells's score D-dimer	VQ scan	VQ scan: four (2%) high probability, six (3%) non-diagnostic, 173 (95%) normal D-dimer: 48/51 positive Modified Wells's score predicted PE
Browne *et al.* (2014)[21]	Ireland	124 pregnant and postpartum women, one hospital, 3 years	None	CTPA	CTPA: 1/70 (1.4%) PE in pregnant women, 5/54 (9.3%) PE in postpartum women
Bajc *et al.* (2015)[22]	Sweden	127 pregnant women, one hospital, 5 years	None	VQ SPECT	VQ SPECT: 11/127 (9%) PE
Jordan *et al.* (2015)[23]	USA	50 pregnant or postpartum women, one hospital, 4 years	None	CTPA	CTPA: 1/50 (2%) PE
Ramsay *et al.* (2015)[24]	UK	127 pregnant women, one hospital, 3 years	None	VQ scan	VQ scan: 2/127 (1.6%) PE

ECG, electrocardiogram; VQ SPECT, ventilation–perfusion single photon emission computed tomography.
a Chest pain, dyspnoea, heart rate, oxygen saturation and Alveolar–arterial gradient.
b Chest pain, dyspnoea, heart rate, respiratory rate, blood pressure, oxygen saturation and Alveolar–arterial gradient.
c Chest pain, dyspnoea, pleuritic chest pain, haemoptysis, cough, DVT signs, wheeze, pleural rub, heart rate and respiratory rate, systolic blood pressure.

In summary, diagnostic studies of pregnant and postpartum women with a suspected PE currently provide insufficient evidence to support their use as a way of selecting women for diagnostic imaging.

Risk factors for pulmonary embolism in pregnancy and post partum

Stronger evidence exists relating to predicting the risk of a pregnant or postpartum woman developing PE (as opposed to diagnosing PE in a pregnant or postpartum woman with suspected PE). Epidemiological studies have compared women who developed PE in pregnancy or post partum with a control group without PE to identify the risk factors for developing PE in pregnancy. Knight *et al.*[27] compared women with antenatal PE identified through the UK Obstetric Surveillance System (UKOSS) research platform with pregnant control group participants and showed that multiparity and body mass index (BMI) were independent predictors of developing PE. Kane *et al.*[28] used patients identified by the Scottish Morbidity Record 2 to show that women aged > 35 years, with previous VTE, pre-eclampsia, antenatal haemorrhage or postnatal haemorrhage were more likely to develop PE than those without these characteristics. Henriksson *et al.*[29] showed that VTE is associated with pregnancy following in vitro fertilisation. Sultan *et al.*[30] linked primary (Clinical Practice Research Datalink) and secondary (Hospital Episode Statistics) care records to show that BMI, complications of pregnancy (pre-eclampsia, antenatal or postnatal haemorrhage, diabetes mellitus, hyperemesis), comorbidities (varicose veins, cardiac disease, hypertension) and recent hospital admission were associated with an increased risk of developing PE. A similar analysis in postpartum women[30] showed that smoking, varicose veins, comorbidities, pre-eclampsia/eclampsia, diabetes mellitus, parity, postpartum haemorrhage, caesarean section, stillbirth, postpartum infection, maternal age, BMI and infant birthweight were predictors of VTE included in a clinical prediction model.

These risk factors for developing PE in pregnancy and post partum could be used to select women with suspected PE for imaging. However, there are two reasons why risk factors may not be diagnostically useful. First, guidelines[2] recommend using thromboprophylaxis in pregnancy and post partum to attenuate the thromboembolic risk associated with recognised risk factors. Second, public and professional awareness of risk factors may prompt a lower threshold for presentation and referral to diagnostic services when risk factors are present. The use of risk factors to select women for imaging therefore needs evaluation in a population with suspected PE.

Current practice

Guidelines from the Royal College of Obstetricians and Gynaecologists (RCOG)[2] recommend that pregnant or postpartum women with a suspected PE should receive diagnostic imaging with VQ scan or CTPA. The guidelines recommend against the use of D-dimer testing and highlight the lack of evidence to support the use of clinical probability assessment in pregnancy. Guidelines from the American Thoracic Society[31] also recommend the non-selective use of diagnostic imaging, whereas guidelines from the European Society of Cardiology[32] suggest a possible role for D-dimer in selecting patients.

These recommendations for pregnant and postpartum women contrast with guidelines from the National Institute for Health and Care Excellence (NICE)[33] and the American College of Chest Physicians[34] for the general (non-pregnant) population with a suspected PE, for whom diagnostic imaging is selectively used based upon structured clinical assessment and D-dimer measurement.

The differences in thresholds for investigation are reflected in the differences in the prevalence of PE in the investigated populations. In a review of studies of patients investigated for a suspected PE, Kline *et al.*[25] reported a prevalence of 4.1% for pregnant patients compared with 12.4% for non-pregnant patients. Most of the studies in our review reported a prevalence of PE below 10%.

Need for further research

Existing research suggests that clinical assessment, CDRs and/or D-dimer measurement could be used to select women for imaging, but more precise estimates of diagnostic value are needed before a selective strategy can be recommended. Furthermore, the appropriate use of clinical assessment or biomarkers to select women for imaging can be determined only by explicitly weighing the risks, costs and benefits of different strategies.

Research is required to improve our estimates of the diagnostic accuracy of clinical assessment and biomarkers. A prospective cohort study is in theory the best method for measuring accuracy and developing and validating a CDR, but it would be severely limited by the low incidence of PE in pregnancy and the low prevalence of PE in women presenting with a suspected PE. The incidence of VTE (DVT and PE combined) is cited to be 1 in 1000 pregnancies.[2] A recent meta-analysis[35] supports this estimate and reports a pooled incidence of PE of 0.4 per 1000 pregnancies, with individual study estimates[35] ranging from 0.1 to 0.67 per 1000 pregnancies. Estimates from UK studies are at the lower end of this range, with data from UKOSS[27] suggesting an incidence of 0.13 antenatal PE per 1000 pregnancies and data from the Scottish Morbidity Record 2[28] suggesting 0.2 antenatal and postnatal PE per 1000 pregnancies. Meanwhile, most of the studies in our literature review reported a rate of one or two patients with PE per hospital per year.

Diagnostic sensitivity is the key determinant of the acceptability of any strategy to select women for diagnostic imaging. Patients and clinicians need to know that sensitivity has been estimated with sufficient precision to ensure that women with a negative diagnostic assessment have a low risk of developing PE. The precision of estimates of sensitivity depends on the number of patients recruited with PE. Using a cohort design to accrue sufficient numbers of participants with PE to estimate sensitivity with sufficient precision would be prohibitively expensive and difficult to deliver.

A case–control design offers an alternative when the low prevalence of disease makes a cohort design unfeasible or unacceptably inefficient. The identification of women with the diagnosis of interest (PE in pregnancy or post partum) allows us to recruit sufficient numbers with PE to make reasonably precise estimates of sensitivity. The case–control design carries an increased risk of bias compared with that of the cohort design,[36] but this can be reduced by ensuring that the control group is a representative sample of women with suspected PE who have negative imaging and that the patients are a representative sample of women presenting with a suspected PE who are diagnosed and treated for PE.

Existing decision rules may not be appropriate to the pregnant and postpartum population, but can be tested in a case–control or cohort study. A decision rule for the pregnant and postpartum population could be derived from a case–control or cohort study, but would need validation in a new study. Expert consensus provides a relatively quick and cheap method for deriving a CDR that could then be validated in a case–control or cohort study.

Secondary research in the form of decision-analysis modelling is required to explicitly weigh the costs, risks and benefits of different strategies for selecting women for diagnostic imaging. This allows us to estimate how diagnostic tests lead to differences in clinically meaningful outcomes. Decision-analysis modelling is particularly important in this situation, when the best method of estimating diagnostic parameters (a cohort study) may not be feasible. Decision-analysis modelling allows us to explore the potential impact of uncertainty on our findings. A value-of-information analysis can then be undertaken to determine whether or not further research would be worthwhile to obtain more accurate or precise estimates of diagnostic accuracy.

Aims and objectives

We aimed to estimate the diagnostic accuracy, effectiveness and cost-effectiveness of strategies (including CDRs) for selecting pregnant or postpartum women with a suspected PE for imaging, and to determine the feasibility and value of information of further prospective research.

Our specific objectives were to:

1. use expert consensus to derive three new CDRs (with different trade-offs between sensitivity and specificity) for pregnant and postpartum women with a suspected PE
2. estimate the diagnostic accuracy of clinical variables, our expert-derived CDRs, existing CDRs (Wells's score, Geneva score and PERC score) and D-dimer in pregnant and postpartum women with a suspected PE
3. use a statistical analysis of women with a diagnosed or suspected PE to derive a new CDR for pregnant and postpartum women with a suspected PE
4. explore the potential diagnostic value of biomarkers for PE in pregnant and postpartum women
5. determine the feasibility of using a prospective cohort design to validate a new CDR or biomarker
6. estimate the effectiveness, in terms of adverse outcomes from VTE, bleeding and radiation exposure, and the cost-effectiveness, measured as the incremental cost per quality-adjusted life-year (QALY), of different strategies
7. estimate the value of information associated with further research.

Overview of the study design

We undertook an expert consensus study to address objective 1. A Delphi study was used to identify and select potential clinical predictors and a consensus group was used to create the CDRs.

We undertook a case–control study to address objectives 2 and 3. The design should strictly be described as a prospective cohort study of pregnant or postpartum women with a suspected PE augmented with a retrospective cohort of pregnant or postpartum women with a diagnosed PE. However, for the purposes of brevity and to avoid concerns that using the term 'cohort study' may under-represent the risk of bias associated with the design, we will use the term 'case–control study' throughout this report.

Women with a suspected PE were identified through a prospective study of pregnant or postpartum women presenting to hospital with a suspected PE. Women with a diagnosed PE were retrospectively identified through UKOSS, a UK-wide obstetric surveillance system that has been set up to conduct research on uncommon disorders of pregnancy. Details of the UKOSS methods are available at www.npeu.ox.ac.uk/ukoss/methodology. The cases involved women with a diagnosed PE and a small number of women with a suspected PE who ultimately had a diagnosed PE, whereas the control participants were women with a suspected PE who had PE ruled out.

We undertook a biomarker study to address objective 4. The women with a suspected PE from the case–control study and any pregnant or postpartum woman diagnosed with DVT at the participating hospitals were asked to provide blood samples for analysis. The inclusion of women with diagnosed DVT was planned as an efficient way of increasing the number in the cohort with VTE. There are good pathophysiological reasons for expecting candidate biomarkers to have similar sensitivity for DVT and PE, and studies of D-dimer measurement in the non-pregnant population have shown similar sensitivity and specificity for DVT and PE.[37]

We addressed objective 5 by determining recruitment rates in the prospective study of women with a suspected PE and determining the prevalence of PE in this population. This would allow us to estimate the potential size and duration of a cohort study powered to estimate the sensitivity of a CDR or biomarker with adequate precision.

We developed a decision-analysis model to address objectives 6 and 7. It was designed to estimate the costs incurred and the expected outcomes from thromboembolism, bleeding and radiation exposure if a hypothetical cohort of pregnant or postpartum women were investigated for a suspected PE using different strategies with a range of sensitivities and specificities, varying from no testing/treatment to imaging for all. The diagnostic accuracy of the strategies was estimated from the case–control study and other parameters were estimated from the systematic literature reviews. Clinical outcomes were modelled to estimate the costs and QALYs accrued by each strategy. We then estimated the value of information associated with further prospective research.

Health technologies being assessed

The focus of our evaluation was any health technology that can be used to select pregnant or post partum women with a suspected PE for diagnostic imaging, including clinical features, risk factors and biomarkers, either alone or combined to form a CDR. An initial literature review undertaken during proposal development identified a number of potential clinical features, risk factors, biomarkers and CDRs that could be used to select women for diagnostic imaging. These formed the basis for data collection and analysis. We specifically ensured that we included the constituent elements of CDRs validated for use in the non-pregnant population with a suspected PE (Wells's score, Geneva score, PERC score and D-dimer).

We also used expert opinion in the project team to identify potential clinical predictors that were not identified by our literature review, including other symptoms (chest pain, dyspnoea, syncope, palpitations), other risk factors (gestational age, smoking status, family history, thrombophilia, varicose veins, intravenous recreational drug use), examination findings (respiratory rate, blood pressure, temperature) and routine investigations [electrocardiogram (ECG), chest radiograph].

We used the following methods to structure clinical variables into a CDR:

- existing CDRs (PERC rule, Wells's score, Geneva score) modified to be appropriate to the pregnant and postpartum population
- expert opinion to create up to three CDRs with varying trade-off between sensitivity and specificity that could be tested in the case–control study population
- statistical derivation of a CDR using the case–control study population.

In terms of biomarkers, D-dimer is the only biomarker currently used in routine practice to select patients for VTE imaging, but it is unlikely to have adequate specificity in pregnancy at conventional thresholds for positivity. We therefore planned to examine the accuracy of D-dimer [enzyme-linked immunosorbent assay (ELISA) and Innovance (Siemens Healthcare Diagnostics Products GmbH, distributed by Sysmex UK Ltd, Milton Keynes, UK)] with a higher (pregnancy-specific) threshold for positivity. We also planned to evaluate the following biomarkers: cardiac troponin I, B-type natriuretic peptide (BNP), prothrombin fragment 1 + 2 (PF 1 + 2), plasmin–antiplasmin complexes, prothrombin time (PT), activated partial thromboplastin time (APTT), Clauss fibrinogen levels, thrombin generation, soluble tissue factor, C-reactive protein (CRP), and mid-regional pro-atrial natriuretic peptide (MRProANP).

Chapter 2 Expert consensus clinical decision rule study

Introduction

Clinical decision rules combine a number of symptoms, signs and simple investigation findings into assessment tools to guide therapeutic or diagnostic decisions at a patient's bedside. Effective CDRs have the potential to standardise care, improve outcomes and increase efficiency. Consensus methodological guidelines[38] recommend that the variables in CDRs are initially determined statistically using data from a representative sample of relevant patients. Identified variables, which appear to provide satisfactory diagnostic accuracy for the target condition, are then tested in external samples during validation studies to confirm their performance. Finally, the impact of CDRs on patient outcomes is evaluated in impact studies.[39] Although many statistically derived CDRs have shown excellent results, there are examples of when clinician gestalt or CDRs based on expert clinical opinion have demonstrated equivalent or superior accuracy.

Aims and objectives

We undertook an expert consensus study to derive three CDRs to select pregnant or postpartum women with suspected PE to receive diagnostic imaging. We intended that one rule would be developed with what we anticipated would be an optimal balance of sensitivity and specificity (the primary rule), another would optimise sensitivity at the expense of specificity (the sensitive rule) and another would optimise specificity at the expense of sensitivity (the specific rule). The consequences of the rule being false negative (failure to diagnose PE) are clearly more serious than the consequences of the rule being false positive (unnecessary imaging), so we anticipated that the optimal balance of sensitivity and specificity in the primary rule would be high sensitivity with modest specificity. The sensitive rule would therefore aim to further reduce the risk of false-negative assessments, whereas the specific rule would be specific only relative to the primary rule.

Methods

A two-stage consensus process, guided by best-practice guidelines, was conducted to reduce biases arising from the subjectivity of expert views, and to maximise content and face validity.[40–42]

In the first stage, a modified Delphi survey was conducted to identify candidate predictors of PE. Purposive sampling was used to recruit a heterogeneous group of subjects encompassing the full spectrum of expertise relevant to UK PE management in pregnancy. A sample size of 20 panel members was chosen in accordance with guidelines from the Research And Development (RAND) Corporation.[43] Self-completed questionnaires were subsequently administered using a pre-piloted web-hosted questionnaire. The survey was conducted between January and October 2016.

The classical Delphi approach was modified slightly with the replacement of an open first round with a systematic literature review to identify possible PE predictors.[7] Participants were asked to rate the predictive value of each variable on a 1 (not predictive) to 5 (very strongly predictive) Likert scale and to justify their

opinion in free text. A mixed-methods approach was taken to summarise each round's findings. In subsequent Delphi iterations, participants were provided with quantitative (percentage results and frequency histograms) and qualitative (free-text answers grouped by theme) results of the previous round and a summary of their previous opinions.

Up to four Delphi rounds were planned, dependent on when group agreement or stability of opinion was achieved on at least 80% of variables. Judgements on the consensus and stability of opinion on each variable were guided by quantitative measures of agreement (> 70% agreement for weak/strong prediction) and changes in responses between rounds. Patient and public representatives commented on the patient acceptability of each predictor variable generated through Delphi consultation.

Similarly, a series of four face-to-face consensus meetings of clinical Diagnosis of Pulmonary Embolism in Pregnancy (DiPEP) co-investigators were planned. In the initial meeting, the nominal group technique (NGT) was used to formulate the content of the three expert CDRs. This meeting was facilitated by an independent researcher experienced in the NGT, which followed recommended principles for best practice in developing consensus,[41,42] consisting of the following steps:

- presentation of a summary of Delphi survey results
- facilitated group discussion of individual candidate variables
- initial rating round for inclusion of each variable
- facilitated discussion of rating results
- further rating and discussion rounds
- confirmation of variables with group consensus for inclusion in CDRs.

Online surveys were designed and implemented using the SmartSurvey web application (SmartSurvey Ltd, Tewkesbury, UK). Data analysis was performed in Microsoft Excel® 2013 (Microsoft Corporation, Redmond, WA, USA). Diagnostic accuracy metrics were calculated using R, version 3.3.2 (The R Foundation for Statistical Computing, Vienna, Austria). Ethics approval was obtained from the University of Sheffield. No financial incentives were offered and participants remained anonymous throughout the Delphi survey.

Results

The systematic literature review identified 45 potential variables for evaluation in the Delphi Survey, comprising demographic, obstetric and medical characteristics; symptoms; clinical signs; and bedside investigations.

All 20 experts invited to participate in the Delphi survey completed the first-round questionnaire. At this stage, consensus was adjudged to have been reached on 12 out of the initial 45 variables (27%). Of these, 17 experts also participated in second and third rounds, during which consensus was obtained on a further 26 variables. At this stage, it was apparent that opinions diverged widely on the remaining seven variables and had not been significantly influenced by feedback between rounds. Free-text responses indicated that further convergence of opinion with additional rounds of surveying was unlikely to occur. It was considered that further sampling would lead to declining response rates rather than more relevant findings and the Delphi study was therefore terminated after round 3. *Table 2* summarises the results of the Delphi survey.

Twenty-four variables that were felt to be moderately or strongly predictive of PE were carried forward for consideration in the consensus meetings. Two further variables from the Delphi survey were additionally considered after the initial presentation of results at the request of the DiPEP clinical investigators: pleuritic

TABLE 2 Consensus results from the Delphi survey

Consensus for strong prediction of PE	Consensus for moderate prediction of PE	Consensus for weak prediction of PE	No consensus reached
• Previous history of DVT/PE • Clinical signs suggestive of lower limb DVT (e.g. unilateral swelling) • Chest radiograph showing alternative non-PE diagnosis • Recent significant injury • Oxygen saturations of < 95% on room air • Active i.v. drug use • Family history of VTE • Tachycardia • D-dimer level > 2× upper limit of normal range	• The current pregnancy has had medical or obstetric complications • Post partum vs. third trimester vs. second trimester vs. first trimester • Unstructured clinical gestalt is that PE is the most likely diagnosis • Raised BMI of > 30 kg/m^2 • Prolonged bed rest • Multiple pregnancy • Low systolic blood pressure for gestational age • Increased respiratory rate for gestational age • Multiple pregnancy • Taking VTE prophylaxis • Haemoptysis • Hyperemesis • Lower leg pain/discomfort • SOB at rest • Syncope	• The patient is a current smoker • The patient has varicose veins • The patient is pregnant rather than post partum • The patient has had previous pregnancies • Previous pregnancies have had medical or obstetric complications • Non-pleuritic chest pain • Palpitations • Diastolic blood pressure lower than the gestational age norm • D-dimer level greater than the threshold for positivity • Recent long-haul travel • Aged > 35 years • Pleuritic chest pain • SOB on exertion • Reduced *P*aCO_2	• IVF conception • ECG changes • Productive cough • Temperature of > 38 °C • Raised troponin • Reduced *P*aO_2 • Major comorbidity

i.v., intravenous; IVF, in vitro fertilisation; *P*aCO_2, partial pressure of arterial carbon dioxide; SOB, shortness of breath.

chest pain (rated as weakly predictive) and active medical comorbidity (no consensus). During the initial NGT meeting, consensus was subsequently achieved for the inclusion of 13 predictors in the final CDRs. Variable weightings and the CDR cut-off point for each CDR (balanced, sensitive, specific) were agreed in two subsequent facilitated roundtable meetings. The scope of the CDRs, in terms of which patients these could be applied to, was also confirmed. The final CDRs developed are presented in *Table 3*.

Discussion

We have used expert consensus to develop three CDRs for the purpose of selecting pregnant or postpartum women with a suspected PE for imaging. The primary rule is intended to achieve an appropriate balance between sensitivity and specificity, whereas the sensitive and specific rules prioritise sensitivity and specificity, respectively. These rules have been developed for testing in the DiPEP study and are not ready for clinical use.

To our knowledge, there have been no other CDRs developed specifically for this purpose. Wells's PE criteria, the Geneva score and the PERC rule were developed to assess the clinical probability of PE in the general population with a suspected PE, but were not developed for pregnant or postpartum women. Our consensus-derived rules share a number of criteria with these rules (haemoptysis, previous VTE, clinical symptoms or signs of DVT, recent injury or surgery) and have adapted others by using pregnancy-specific thresholds (heart rate, oxygen saturation). Our expert consensus group drew on pre-existing rules for the general population, but adapted and added criteria to make the rules relevant to the pregnant population.

Consensus development provides a relatively quick and efficient way of developing a CDR, but has some inevitable limitations. Experts should base their judgements on empirical data, but, as *Chapter 1* has highlighted, there are very limited data relating to the clinical prediction of PE in pregnancy. In the absence of empirical evidence, experts may base their judgements on personal experience, which is known to be subject

TABLE 3 Consensus-derived CDRs to guide advanced imaging decisions for PE in pregnancy and the postpartum period

Included variables	Variable weighting[a]		
	Primary CDR	Sensitive CDR	Specific CDR
Haemoptysis	3	1	4
Pleuritic chest pain	0	1	0
Previous VTE	3	1	4
Family history of VTE in first-degree relative	0	1	0
Hospital admission, surgery or significant injury within 90 days (excluding NVD or caesarean section)	2	1	1
Obstetric complication[b]	1	1	0
Active medical comorbidities[c]	2	1	1
Post partum or third trimester	1	1	0
Raised BMI of ≥ 30 kg/m^2	1	1	0
Clinical symptoms or signs of DVT[d]	3	1	4
Oxygen saturation of $< 94\%$ on room air	3	1	3
Tachycardia of > 100 b.p.m. (in the first or second trimester, or post partum)/ tachycardia of > 110 b.p.m. (in third trimester)	2	1	2
Increased respiratory rate of > 24 breaths per minute	2	1	2
CDR cut-off point	3	1	4

b.p.m., beats per minute; NVD, normal vaginal delivery.

a The scoring systems for the primary and specific rules allow the three rules to be presented alongside each other. In practice, these scores can be simplified by removing zero-scoring variables.

b Obstetric complications – apply once if any of the following are present: pre-eclampsia in current pregnancy assisted reproductive technology/in vitro fertilisation (antenatal only), multiple pregnancy, caesarean section in labour, elective caesarean section, mid-cavity or rotational operative delivery, prolonged labour (> 24 hours), postpartum haemorrhage (> 1 litre or transfusion), preterm birth at $< 37^{+0}$ weeks in current pregnancy, stillbirth in current pregnancy, hyperemesis, ovarian hyperstimulation syndrome (first trimester only).

c Active medical comorbidities – apply once if any of the following are present: cancer, heart failure, systemic lupus erythematosus, inflammatory polyarthropathy or inflammatory bowel disease, nephrotic syndrome, type 1 or type 2 diabetes mellitus with nephropathy, sickle cell disease.

d Patients presenting with symptoms and/or signs of DVT and suspicion of PE would initially undergo duplex ultrasound of the leg(s). If positive, patients would be treated for DVT and presumed PE. Negative leg imaging does not rule out DVT and these patients would still be considered to be at a higher risk of developing PE.

Note

The CDRs apply to pregnant or postpartum women presenting with symptoms that prompt consideration of PE (e.g. chest pain, shortness of breath). The rule does not apply if critically ill and/or in need of resuscitation; a clear non-PE diagnosis is identified by clinical assessment, including ECG, chest radiograph and blood tests when appropriate (e.g. chest infection); or an uncommon but powerful VTE risk factor exists (e.g. thrombophilia, intravenous illicit drug misuse).

to cognitive biases, such as the availability heuristic and the Dunning–Kruger effect. These may lead to overestimation of the importance of atypical but memorable observations, as well as overestimation of diagnostic certainty.

The consensus methods used are intended to reduce the risk of domination by a single expert opinion, but can have the opposite risk of discouraging legitimate questioning of commonly held beliefs. We deliberately restricted the number of experts involved in the final phase of developing the rules to ensure that this process was manageable. This carries the risk of supressing dissenting views in the interests of achieving a practical output.

Conclusion

We have developed three CDRs through expert consensus that need to be tested to determine their ability to discriminate between pregnant or postpartum women with and without PE.

Chapter 3 Case–control study

Aims and objectives

The case–control study was intended to compare participants with PE with control participants without PE to allow for the estimation of the diagnostic accuracy of clinical features, CDRs and biomarkers, and statistical derivation of a new CDR. To minimise bias, we tried to ensure that the case participants and control participants were selected in a similar way (i.e. that case participants presented to hospital with suspected PE were investigated accordingly). The postpartum period was defined as the 6 weeks (42 days) at the end of a pregnancy beyond the first trimester. As noted in *Chapter 1*, the design could more accurately be described as a prospective cohort study of women with a suspected PE augmented with a retrospective cohort of women with a diagnosed PE, but, for the reasons previously outlined, we will use the term case–control study.

Methods

Study population

Diagnosed PE: the UKOSS research platform was used to identify a sample of pregnant or postpartum women who were diagnosed with PE in the UK after presentation with a suspected PE. We identified and collected data from any woman diagnosed with PE at a hospital participating in the UKOSS platform between 1 March 2015 and 30 September 2016.

Suspected PE: we recruited a sample of pregnant or postpartum women investigated for a suspected PE across 11 participating sites. We anticipated that 98% of women in the sample would have no confirmed diagnosis of PE and would constitute the control group. Those with a diagnosis of PE confirmed would be analysed with the patients with PE.

Inclusion/exclusion criteria

Diagnosed PE: we included women with PE if they met any of the following definitions:

1. pulmonary embolism confirmed using imaging (angiography, CT, magnetic resonance imaging or VQ scan showing a high probability of PE)
2. pulmonary embolism confirmed at surgery or post mortem
3. clinical diagnosis of PE resulting in a course of anticoagulation therapy for > 1 week.

Women meeting criterion 1 or 2 were included as patients with PE in the primary analysis. Women with a clinical diagnosis of PE (criterion 3) were excluded from the primary analysis, but were included as patients with PE in the secondary analysis. This was because of the risk of incorporation bias if the clinical reference standard diagnosis of PE was based on the clinical variables or biomarkers being used as index tests.

We excluded women who did not present with a suspected PE prior to diagnosis (i.e. with PE identified as an incidental finding). We collected data from women who required life support at presentation (chest compressions or assisted ventilation) to allow for the estimation of the incidence of PE in pregnancy, but did not include these women in the analyses in this study.

Suspected PE: we included any pregnant or postpartum woman presenting to the participating hospitals who was considered to require diagnostic imaging for a suspected PE. Women were recruited once the clinician had decided that imaging would be required. However, not all women received lung imaging; in a proportion of patients, the decision that imaging was required was reversed by a more senior clinician,

and some women received imaging only for DVT (e.g. venous ultrasound). Women who had PE ruled out clinically (i.e. without diagnostic imaging of the lungs for PE) were excluded from the primary analysis, but included in the secondary analysis. This was because of the risk of bias if the PE was missed as a result of a lack of adequate imaging and the decision not to undertake imaging was based on the clinical variables or biomarkers that constituted the index tests.

We excluded women who needed life support on presentation to hospital (chest compressions or assisted ventilation), women who had been diagnosed with PE earlier in the current pregnancy, women who were unable or unwilling to provide informed consent, women aged < 16 years and women previously recruited to the study. The form used for screening is shown in *Report Supplementary Material 1* (the suspected PE screening form).

Setting/context

Diagnosed PE

UKOSS collects data from all UK hospitals with a consultant-led maternity unit. Patients for this study presented through a variety of routes, depending on local practice, but were ultimately the responsibility of the obstetric services, and thus women who had PE at any stage in gestation were identified, provided that their pregnancy was ongoing. Postpartum women were identified if they were still under obstetric care, but inevitably this meant that patient identification became less reliable towards the end of the postpartum period.

Suspected PE

Pregnant and postpartum women with a suspected PE are investigated in secondary care, but may follow a variety of different pathways, depending on local practice. At each hospital, patient recruitment was targeted at the location at which the decision to undertake diagnostic imaging was made – this included the emergency department, the maternity unit or both.

Sampling

Diagnosed PE

Nominated clinicians in each consultant-led maternity unit in the UK were sent a card each month and asked to report all patients with antenatal or postnatal PE, thus covering the entire cohort of UK births. In addition, the ascertainment of any maternal deaths from PE occurring during the study period was checked through MBRRACE-UK (Mothers and Babies: Reducing Risk through Audits and Confidential Enquiries across the UK), the collaboration responsible for the UK Confidential Enquiries into Maternal Death. When a patient was identified, the UKOSS clinician was contacted and asked to complete a data collection form if appropriate.

It was not practical to obtain consent for data collection from individual women with a diagnosed PE. The Confidentiality Advisory Group of the Health Research Authority and equivalent bodies in the devolved nations consider that organisations seeking to use NHS information for research purposes without consent should seek anonymised or pseudonymised data only, and not any personally identifiable information. Accordingly, names, addresses, postcodes, dates of birth and NHS or hospital numbers were not collected in the UKOSS research platform.

Suspected PE

Clinicians in the participating hospitals prospectively identified pregnant or postpartum women with a suspected PE considered to require diagnostic imaging. They contacted the research nurse/midwife or recruiting clinician, who provided women with information about the study, and checked the eligibility criteria. Informed consent to participate was sought prior to discharge, which at some hospitals included consenting women who returned for outpatient appointments for diagnostic imaging.

Data collection and follow-up

Diagnosed PE

UKOSS clinicians who reported a patient were asked to complete a data collection form detailing the clinical variables, diagnostic test results, management and outcomes (see *Report Supplementary Material 2*, the diagnosed PE data collection form). Up to five reminders were sent if completed forms were not returned. On receipt of the data collection forms, patients were checked to confirm that they met the patient definition (see *Inclusion/exclusion criteria*). Duplicate reports were identified by comparing the woman's year of birth, hospital, the date of a suspected PE and the expected date of delivery, or the date of birth for postpartum women.

Suspected PE

The research nurse/midwife completed a data collection form incorporating clinical variables, diagnostic test results and management, using information from patient records (see *Report Supplementary Material 3*, the suspected PE data collection form). Participants provided a blood sample and underwent diagnostic imaging in accordance with local protocols. Ideally, the research nurse/midwife would collect clinical data prior to diagnostic imaging being performed and would thus be blinded to the results of the diagnostic imaging. However, some patients were recruited after diagnostic imaging had been performed. For these patients, we asked the research nurse/midwife to record whether or not they were aware of the results of the diagnostic imaging when they collected clinical data.

At 30 days after recruitment, the research nurse/midwife reviewed hospital records and recorded details of any adverse events and the results of any additional diagnostic investigations for PE. When the research nurse/midwife was aware of follow-up care outside the hospital NHS trust, attempts were made to complete follow-up data using hospital records from the relevant location. All participants who provided contact details, except for those who had died or withdrawn from the study, were sent a questionnaire by mail, e-mail or telephone to record any additional adverse events, details of the health care received, health utility and standardised quality of life using the EuroQol-5 Dimensions, five-level version [(EQ-5D-5L), www.euroqol.org/, accessed 13 June 2017]. Participants received up to three reminders to complete the 30-day questionnaire. One of the two reminders used an alternative method of contact in accordance with their stated preference (e.g. telephone or e-mail if there was no response to posted mail). When insufficient information or no information was obtained on the data collection form, the research nurse/midwife follow-up or the patient follow-up questionnaire, the woman's general practitioner (GP) was contacted to rule out serious adverse events of additional diagnostic investigations for PE using primary care records.

Women recruited with a suspected PE who were subsequently diagnosed with PE were cross-checked with the UKOSS patients to avoid duplication. If duplication was found, data collected by the research nurse/midwife were used.

Non-recruited women with suspected PE: women presenting to the participating hospitals with suspected PE who were eligible but not approached to request participation, were retrospectively identified from hospital systems, radiology records and communication between clinicians (see *Report Supplementary Material 4*, non-recruited suspected PE screening form). The research nurse/midwife then extracted anonymised data from the hospital records. The anonymised data included, when possible, the clinical variables and the imaging, treatment, and follow-up data used to diagnose PE (see *Report Supplementary Material 5*, non-recruited suspected PE data collection form). No blood sample was taken and no follow-up questionnaire was administered, and any data not available in the case notes were recorded as missing. These data were used to explore whether or not the recruited sample was representative.

Sample size

The sample size for the UKOSS data were inevitably determined by the incidence of PE during the data collection period. Based on a previous similar study,[27] we anticipated that we would identify 150 patients with diagnosed PE over the 18 months of the study. We aimed to recruit 250 women with suspected PE over the same time period, resulting in around 155 women with PE, and 245 women without PE in the

control group, assuming that the prevalence of PE is 2% in those with a suspected PE. This would allow for the estimation of sensitivity or specificity of 90% with a SE of around 2.5% and 2.0%, respectively. Assuming that the ratio of women with PE to women without PE in the control group would be around 0.4, this sample size would be sufficient to identify an odds ratio of a clinical predictor of around 2, with 90% power and 5% two-sided significance.[44]

With a limited sample size, the complexity of any statistical model was limited in terms of the number of predictor variables that could be included. This was addressed by selecting only those variables that were felt to be clinically important in the model and by aiming for maximum parsimony in the final model. The statistical analysis outlined in the study proposal stipulated modelling outcomes by splitting the cohort into a training data set and a validation data set. This effectively reduces power, so, in developing the statistical analysis plan, an alternative approach of leave-one-out cross-validation was planned instead.

Data management

Diagnosed PE data and suspected PE data were collected on study-specific case report forms (CRFs) and entered onto a secure electronic data capture system at UKOSS and the Clinical Trials Research Unit (CTRU) at the University of Sheffield. The prospective data were managed via the CTRU Prospect system (epiGenesys, University of Sheffield, Sheffield, UK), which used inbuilt validation to promote high-quality data and security management features to ensure data confidentiality. On completion of the study, pseudonymised UKOSS data were encrypted and uploaded to the Prospect system. Physical data were stored in accordance with good clinical practice and local standard operating procedures.

Ethics and research and development approvals

Ethics approval for the study was obtained from the London Brent Research Ethics Committee (reference 14/LO/1695). A feasibility assessment for each participating site was undertaken by the research team and in accordance with local research governance procedures. Permission to undertake the research study was granted and reviewed in response to project amendments.

Patient and public involvement

Four members of the public, representing interests in obstetric care and emergency medicine, were actively involved in the oversight of the study via membership of the Study Steering Committee. Patient and public involvement (PPI) representatives influenced the development of study documents and data collection, shaped the dissemination strategy and commented on outputs. PPI representatives also engaged with external organisations such as the Sheffield Emergency Care Forum and Thrombosis UK.

Analysis populations

The primary analysis population consisted of women recruited with a suspected PE and women identified through UKOSS with a diagnosed PE. Women who presented to hospital needing life support were excluded from the suspected PE sample and were used in the UKOSS sample only to estimate the overall PE incidence. They were therefore excluded from all analyses reported in this study. We also excluded any women from the UKOSS sample if it was not recorded whether or not they presented to hospital needing life support, unless presenting physiological data showed that life support would not have been needed.

We planned a priori that the primary analysis should be limited to participants with diagnostic imaging, surgery or post-mortem confirmation of PE or in whom PE had been ruled out by diagnostic imaging, and thus included only patients without diagnostic imaging, surgery or post-mortem confirmation in the secondary analyses. We also planned a priori to undertake the secondary analysis excluding isolated subsegmental PE, as the identification of subsegmental PE on imaging may be unreliable and the need for treatment may be uncertain.

We identified any duplicates between the UKOSS data set and the women with a suspected PE. If the woman was recruited with a suspected PE, then the corresponding UKOSS data were removed from the overall data set. If the woman was identified but not recruited with a suspected PE, then the UKOSS data

were retained. The anonymised data relating to women's presentation with a suspected PE were used in a descriptive analysis of the non-recruited patients.

Reference standard classification

We planned that the classification of participants as having PE (PE present) or not having PE (PE absent) should be based on the results of imaging, thromboembolic events and the evidence of treatment for PE, regardless of whether or not participants were recruited with a suspected PE or identified as having a diagnosed PE through UKOSS. Two independent assessors (SG and CNP), who were blind to the clinical predictors and the blood results, used a structured process to classify the diagnostic imaging results, the details of adverse events and the details of treatments given, and thus to classify all participants and non-recruited participants as PE present (women with PE) or PE absent (control group participants). Details of the structured process are provided in *Appendix 1*. Disagreements were resolved through adjudication by a third assessor (FL). This process also classified how participants would be handled in the primary and secondary analysis.

We structured the process of classification for primary and secondary analysis around the following principles:

Primary analysis –

- Pulmonary embolism was present if lung imaging was reported as showing PE or if venous imaging showed DVT in the presence of symptoms indicating suspected PE (i.e. if the patient met the eligibility criteria), if surgery or a post-mortem examination revealed PE or if the 30-day follow-up identified a subsequent diagnosis of PE.
- Pulmonary embolism was absent if lung imaging was reported as negative for PE, unless the 30-day follow-up identified subsequent PE.
- Pulmonary embolism was absent if lung imaging was non-diagnostic, no treatment was given for PE (defined as therapeutic anticoagulation for at least 1 week) and no subsequent PE was identified on follow-up.
- Women with clinically diagnosed PE were excluded (i.e. if lung imaging was non-diagnostic, but treatment was given for PE).
- Women with clinically ruled-out PE were excluded (i.e. if lung imaging was not done).

The secondary analyses examined the following reclassifications –

- the inclusion of women with clinically diagnosed PE as PE present (i.e. women with no lung imaging or non-diagnostic lung imaging who received treatment for PE)
- the inclusion of women with clinically ruled out PE as PE absent (i.e. women with no lung imaging who did not receive treatment for PE)
- the exclusion of women with isolated subsegmental PE.

Clinical variable classification

Clinical variables that could be diagnostically useful were classified on the basis of a priori categorisation as to whether the variable was present or absent. For most patients, this was on the basis of an expected association between a variable and the presence or absence of PE. If the variable was present, then it was expected that PE would be more likely to be present. Continuous variables were determined from the expert opinion of DiPEP co-investigators, existing criteria used in the relevant decision tools,[2] or widely acknowledged physiological definitions (e.g. tachycardia) to give clinically meaningful classifications. *Table 4* outlines the classification.

Other previous medical problems and other problems in the current pregnancy were analysed in two ways using the lists above. First, they were analysed using any other previous medical problem or problems with the current pregnancy as the predictor of PE (lists 1 and 2). Then, they were limited to other previous medical problems and problems in the current pregnancy that were known to be associated with an

TABLE 4 Pulmonary embolism predictor variable classification

Clinical variables	Present	Absent
Aged > 35 years	> 35 years	≤ 35 years
BMI of ≥ 30 kg/m^2	≥ 30 kg/m^2	< 30 kg/m^2
Ex-smoker (prior)	Gave up smoking before pregnancy	No
Ex-smoker (during)	Gave up smoking during pregnancy	No
Current smoker	Current smoker	No
Previous pregnancy of < 24 weeks' gestation	One or more	None
Previous pregnancy of > 24 weeks' gestation	One or more	None
Previous pregnancy problems	Any from list 1[a]	None from list 1[a]
History of VTE in first-degree relatives	Yes	No
History of varicose veins	Yes	No
History of i.v. drug use	Yes	No
Known history of thrombophilia	Yes	No
Surgery in the previous 4 weeks	Yes	No
Significant injury in the previous 4 weeks	Yes	No
Previous VTE	Yes	No
Other previous medical problem	Any from list 2[b]	None from list 2[b]
Other previous medical problem with risk of VTE	Any from list 3[c]	None from list 3[c]
Pregnant vs. post partum	Post partum	Pregnant
Trimester	Second	First
Trimester	Third	First
Multiple pregnancy	Yes	No
History of long-haul travel during this pregnancy	Any within 1 month	None within 1 month
Period of immobility/bed rest during this pregnancy	Any within 1 month	None within 1 month
Prior thrombotic event in this pregnancy	Yes	No
Other problem in current pregnancy	Any from list 1[a]	None from list 1[a]
Other problem in current pregnancy with risk of VTE	Any from list 4[d]	None from list 4[d]
Pleuritic chest pain	Yes	No
Other (non-pleuritic) chest pain	Yes	No
SOB on exertion	Yes	No
SOB at rest	Yes	No
Haemoptysis	Yes	No
Other productive cough	Yes	No
Syncope	Yes	No
Palpitations	Yes	No
Other symptoms	Yes	No
Tachycardia	Heart rate of > 100 b.p.m. (in first or second trimester, or post partum) or > 110 b.p.m. (in third trimester)	Other or not recorded

TABLE 4 Pulmonary embolism predictor variable classification (*continued*)

Clinical variables	Present	Absent
Tachypnoea	Respiratory rate of > 24 breaths per minute	Other or not recorded
Hypoxia	SaO_2 on room air of < 94%	Other or not recorded
Low systolic BP	Systolic BP of < 90 mmHg	Other or not recorded
Low diastolic BP	Diastolic BP of < 50 mmHg	Other or not recorded
Fever	Temperature of > 37.5 °C	Other or not recorded
Clinical signs of DVT	Yes	No or not recorded
PE-related ECG abnormality	Yes	Other
PE-related chest radiograph abnormality	Yes	Other
Other chest radiograph abnormality	Yes	Other
Diagnostic impression	PE at least as likely as any other diagnosis	Other
D-dimer	Above the gestational age-specific threshold (see *Clinical variable classification*)	Below gestational age-specific threshold

BP, blood pressure; b.p.m., beats per minute; i.v., intravenous; SaO_2, peripheral oxygen saturation; SOB, shortness of breath.

a List 1: previous or current pregnancy problems (other than VTE) – three or more miscarriages, amniocentesis, amniotic fluid embolism, baby with a major congenital abnormality, dehydration requiring admission, eclampsia, gestational diabetes mellitus, haemorrhage, hyperemesis requiring admission, infant requiring intensive care, large-for-gestational-age infant, neonatal death, ovarian hyperstimulation syndrome, placenta praevia, postpartum haemorrhage requiring transfusion, pre-eclampsia (hypertension and proteinuria), premature rupture of membranes, preterm birth or mid-trimester loss, puerperal psychosis, significant placental abruption, severe infection (e.g. pyelonephritis, small-for-gestational-age infant, stillbirth, surgical procedure in pregnancy.

b List 2: previous medical problems – autoimmune diseases, cancer, cardiac disease (congenital or acquired), diabetes mellitus, endocrine disorders (e.g. hypothyroidism or hyperthyroidism), epilepsy, essential hypertension, gross varicose veins, haematological disorders (e.g. sickle cell disease), hypertension, inflammatory disorders (e.g. inflammatory bowel disease), malignancy within 6 months, myeloproliferative disorders (e.g. essential thrombocythaemia), polycythaemia vera, other medical disorders (e.g. nephrotic syndrome, cardiac disease, paraplegia, psychiatric disorders, renal disease).

c List 3: previous VTE-related medical problems – cancer, heart failure; systemic lupus erythematosus, inflammatory polyarthropathy or inflammatory bowel disease; nephrotic syndrome; type 1 or type 2 diabetes mellitus with nephropathy; sickle cell disease.

d List 4: current VTE-related pregnancy problems – assisted reproductive technology/in vitro fertilisation (antenatal only), multiple pregnancy, caesarean section in labour, elective caesarean section, mid-cavity or rotational operative delivery, prolonged labour (> 24 hours), postpartum haemorrhage (> 1 litre or transfusion), preterm birth of $< 37^{+0}$ weeks in current pregnancy, stillbirth in current pregnancy, hyperemesis, ovarian hyperstimulation syndrome (first trimester only).

increased risk of VTE, based on the outcome of developing the expert consensus-derived CDRs. Problems specifically tested as a separate predictor (e.g. previous thrombotic event) were not included.

The following ECG abnormalities were classified as being PE related: SI QIII TIII pattern, complete or incomplete right bundle branch block, right axis deviation, simultaneous T-wave inversions in the inferior (II, III, aVF) and right precordial leads (V1–3), right atrial enlargement (peaked P wave in lead II of > 2 mm in height), clockwise rotation (shift of the R/S transition point towards V6, persistent S wave in V6) or atrial tachyarrhythmia (atrial fibrillation, flutter or tachycardia).

Chest radiographs were classified by reviewing the radiology report produced as part of clinical care. If the report mentioned the PE or pulmonary infarction in describing radiographic changes or identified potentially PE-related findings (such as atelectasis) without providing an alternative explanation for the finding, then we classified it as referring to a PE-related abnormality. All other abnormal radiographs were classified as having other abnormality.

The diagnostic impression was reviewed by one of the investigators (SG) and classified by whether or not PE was at least as likely as any other diagnosis, based on with the criterion used in Wells's PE score.[3]

This was done at two levels: (1) a strict judgement in which only diagnostic impressions that clearly stated that PE was at least as likely as any other diagnosis were included and (2) a permissive judgement in which any diagnostic impression that suggested that PE was as least as likely as any other diagnosis was included.

D-dimer measurements were recorded as part of routine care in a proportion of women with suspected PE and women with diagnosed PE in the UKOSS cohort. The biochemical analysis was undertaken in many different hospitals, using different assays with different diagnostic thresholds. Furthermore, we expected specificity to decline with gestational age. We therefore planned to test a threshold for positivity for D-dimer that varied with gestational age and was defined in relation to the threshold used for that assay at the relevant hospital rather than as an absolute value. We used the standard threshold during the first trimester, 1.5 times the standard threshold for the second trimester and two times the standard threshold for the third trimester and post partum. This was based on data showing how D-dimer levels increase during pregnancy[45] and evidence that a higher threshold may improve specificity for diagnosing VTE in pregnancy without sacrificing sensitivity.[6]

Missing data

The process for classifying the reference standard (PE present or absent) described above outlines how data relating to imaging, follow-up and treatment are handled. In general, if data were missing, it was assumed that imaging was not performed, treatment was not given or follow-up was negative. However, all data were presented to the independent assessors so that they could take the presence or absence of data into account when making their judgement.

For the clinical variables, there was scope for missing data if the attending clinician failed to measure or record the data, or if the UKOSS clinician or research nurse/midwife was unable to access the necessary hospital records. Analyses involving multiple clinical variables (i.e. multivariable regression and analysis of decision rules) have to either impute missing variables or exclude every patient with a missing variable. In these analyses, we included patients with small numbers of missing data, with missing variables imputed as being normal or negative, and excluded patients with large numbers of missing data (i.e. if any of the following criteria were met):

- more than one of heart rate, respiratory rate and oxygen saturation were missing
- more than half of the variables relating to previous medical history were missing
- more than half of the variables relating to the current pregnancy were missing.

Our rationale for this approach was that if large numbers of data were missing, then any assumption about the pattern of missing data would be speculative and it would be best to exclude the patient, whereas, if only a few variables were missing, it would be more likely that these were not recorded because they were expected to be normal or negative. Imputing missing data as normal or negative would tend to overestimate specificity and underestimate sensitivity. We felt that this represented a conservative assumption that would accord with clinician willingness to accept a degree of overinvestigation and unwillingness to accept the risk of missed PE. We felt that imputing abnormal or positive values when variables were missing, especially in validating CDRs, would be met with scepticism by the clinical users of our findings and would undermine the clinical credibility of these findings.

Planned analyses

Demographic and baseline characteristics were presented descriptively for the cohorts with diagnosed PE and suspected PE, and the eligible but non-recruited patients with suspected PE. Demographics, baseline characteristics and prevalence of PE diagnosis were then compared between the recruited women and the non-recruited women with suspected PE to explore whether or not those recruited were a representative cohort.

The cohort with diagnosed PE and the recruited cohort with suspected PE were then combined to form the main data set for analysis. The primary analysis was limited to women with PE confirmed or ruled out

by imaging, surgery or post-mortem examination. Secondary analyses examined the effect of (1) including women with clinically diagnosed PE, (2) including women with clinically ruled-out PE and (3) excluding women with subsegmental PE as outlined above.

Clinical variables were compared between women with PE and those without PE. Univariable logistic regression was then used to determine the association between each variable and the presence or absence of PE.

Summaries of the responses to the assessment questionnaires were tabulated for each time point.

The accuracy of each index test was assessed by reporting and comparing the sensitivity and specificity. The combined sensitivity and specificity was assessed by plotting receiver operating characteristic (ROC) curves and quantified by the area under the curve (AUC). By virtue of the case–control design oversampling the proportion of patients with PE, no attempt was made to estimate the positive and negative predictive values.

We retrospectively applied each of the decision rules derived by expert consensus (see *Chapter 2*) and the following existing decision rules to the data to estimate diagnostic performance:

- the PERC rule[5]
- the Wells's score[3]
- the revised Geneva score.[4]

These rules were not developed for use in the pregnant population, so we removed criteria that were not relevant to the pregnant population and adapted criteria when appropriate to be relevant to the pregnant population. We therefore removed exogenous oestrogen from the PERC rule and used the thresholds developed for our analysis of clinical variables to dichotomise age, oxygen saturation and heart rate (see *Table 4*). The need to design a CRF that would be usable for both prospective and retrospective data collection and that would address the multiple study objectives meant that criteria used in each rule did not always map precisely onto variables in the CRF. Furthermore, data for some of the variables were missing. *Appendix 2* outlines how the criteria in each rule were applied to the study data.

The diagnostic performance of CDRs is normally presented as the proportion in each risk stratum with the outcome of interest (in this case PE). This is similar to reporting positive and negative predictive values for a diagnostic test. This approach would be misleading in the DiPEP analysis, because the prevalence of PE in the analysis cohort has been deliberately inflated using the UKOSS data to increase the precision of estimates of sensitivity. Positive and negative predictive values are dependent on sensitivity, so the proportion of PE in each stratum would be much higher than if the rule was used in a typical clinical cohort with low prevalence. We therefore assessed the accuracy of each index test by calculating sensitivity and specificity at the usual or recommended decision-making threshold, plotting ROC curves and quantifying the AUC. The consensus-derived rules each specified a threshold for the rule being positive or negative. The PERC score was considered positive if any criterion was positive. The Wells's criteria score was considered to be positive if the score was ≥ 4 points (PE likely). The simplified revised Geneva score was considered to be positive if the score was ≥ 4 points (moderate or high risk).

D-dimer (as measured in the hospital laboratory and recorded in the clinical notes) was analysed as a separate index diagnostic test, rather than as one of the clinical variables, and was not included in the CDRs, multivariable analysis or recursive partitioning. The diagnostic accuracy was assessed by calculating the sensitivity and specificity at the hospital laboratory threshold and the pregnancy-specific thresholds outlined previously. We did not use ROC analysis for D-dimer because of the complexity of having to use different thresholds for different assays across the multiple hospitals contributing data to the study.

Methods of statistical modelling

Three statistical modelling approaches were used in the development of new decision rules:

1. Logistic regression.

Univariable logistic regression analyses were undertaken to identify associations between clinical features and diagnosis of PE.

2. The least absolute shrinkage and selection operator (LASSO).

The LASSO regression modelling approach was used to assess the predictor variables and estimate the effect of each variable on the outcome.[46–48] This method is also based on multivariable logistic regression modelling, but addresses some of the problems of overfitting in the presence of multiple correlated covariates. The LASSO adaptation of logistic regression applies a penalisation against higher-dimension models, thereby helping to protect against coefficients being spuriously inflated.[49] We included all clinical variables in the analysis, regardless of their univariable association, but did not include receipt of thromboprophylaxis as a clinical variable. Thromboprophylaxis is targeted at women who are at risk of VTE and is intended to prevent PE. It is therefore likely to have a complex and inconsistent association with PE.

3. Recursive partitioning.

Recursive partitioning[50] is a different approach which is used to create decision tree models. Rather than estimating a formula linking outcome to a combination of covariates, recursive partitioning attempts to identify subgroups with higher incidences of PE based on their characteristics, the aim being to derive a rule with optimal sensitivity (ideally of > 95%). This had the advantage that continuous variables did not need to be dichotomised beforehand, as the partitioning algorithm not only determines which subset of variables provides the optimal classification, but also the cut-off points within predictor variables. As is standard practice when performing recursive partitioning, cross-validation is employed at each partition to ensure that fully fitted trees were pruned based on a function of the complexity parameter that minimises the cross-validation error in order to avoid overfitting.

We intended that clinicians would review the derived models to ensure clinical credibility. We planned to purge variables that were considered to be inappropriate (e.g. if it was something that could not in practice be assessed) and then refit the models and recalculate the coefficients of the remaining covariates. Clinical opinion would then be sought to weight/round the coefficients into simple decision rules and decide upon a threshold for decision-making.

In accordance with the principles of reproducible research, all analyses were performed using a literate programming approach, which allowed the recreation of tables and figures at will by anyone experienced in using the software. The scripts were version controlled using the Git version 2.13 [(2017) Github Inc., San Francisco, CA, USA] control system and self-documenting. The statistical programming language R was used to undertake the statistical analysis.

Results

Women with diagnosed pulmonary embolism

A total of 224 women were identified through the UKOSS between 1 March 2015 and 30 September 2016. We excluded 13 women because they were recorded as presenting with life-threatening features and a further eight women because presentation with life-threatening features was not recorded and the absence of life-threatening features could not be inferred from physiological data. We identified five women who had also been recruited with suspected PE (their data were removed from the UKOSS data set) and three whose characteristics matched women in the eligible but non-recruited data set (their data

were retained in the UKOSS data set). Thus, we included data from 198 women with diagnosed PE in the study analysis.

Women with suspected pulmonary embolism

We originally planned to recruit across eight sites over 18 months at a rate of two per site per month to achieve a sample size of 250 women. However, we realised after developing the detailed statistical analysis plan and examining initial data that the exclusion of those with clinically diagnosed or clinically ruled-out PE from the primary analysis would potentially leave the primary analysis underpowered. We therefore increased the number of participating sites to 11 to ensure that the planned sample size of 250 women would be achieved for the primary analysis.

A total of 324 women were recruited across 11 participating sites between 15 February 2015 and 31 August 2016. The result of diagnostic imaging for PE was known by the research nurse/midwife at the time of consent for 46 out of 324 women (14%). A further 35 were eligible for recruitment but declined to participate and 95 were not asked to participate despite being eligible, usually because of the lack of availability of an appropriate person to undertake recruitment. Non-identifiable data were collected from this latter group of women, who formed the cohort of eligible but non-recruited women. *Appendix 3* provides details of the recruitment process and the flow of participants through the study. The mean monthly recruitment rate per site was 1.7 women per month per site across the study.

Reference standard classification

Full details of the classification process are provided in *Appendix 1*. The 198 patients in the UKOSS data set consisted of 163 women with PE confirmed by imaging or post-mortem examination (160 by imaging, including seven women with subsegmental PE, two by post-mortem examination alone and one by imaging and post-mortem examination) and 35 women with clinically diagnosed PE (29 with equivocal imaging and six with no imaging recorded; all treated). Thus, 163 women were included as having PE in the primary analysis, 198 women were included in the secondary analysis including those with clinically diagnosed PE, and 156 women were included in the secondary analysis excluding those with subsegmental PE.

The 324 women recruited with suspected PE consisted of 18 women with PE confirmed by imaging (including one with subsegmental PE), five women with clinically diagnosed PE (three with equivocal imaging and two with no lung imaging; all treated), 259 women with PE ruled out after imaging (254 with negative lung imaging and five untreated after equivocal lung imaging) and 42 with PE clinically ruled out without lung imaging (none treated). Thus, 18 women with PE and 259 without PE were included in the primary analysis, 23 women with PE and 259 women without PE were included in the secondary analysis including clinically diagnosed PE, 18 women with PE and 301 women without PE were included in the secondary analysis including clinically ruled-out PE, and 17 women with PE and 259 women without PE were included in the secondary analysis excluding subsegmental PE. The prevalence of PE was therefore 7.1% (23/324) across all women with suspected PE and 6.5% (18/277) when women with clinically diagnosed or ruled-out PE were excluded.

The 95 eligible but non-recruited women with suspected PE consisted of six women with PE confirmed by imaging (including three with subsegmental PE), five women with clinically diagnosed PE (four with equivocal lung imaging and one with no lung imaging; all treated), 73 women with PE ruled out with lung imaging (71 with negative imaging and two untreated after equivocal imaging) and 11 women with PE clinically ruled out without lung imaging (none treated). The prevalence of PE was therefore 11.6% (11/95) across all non-recruited women and 7.6% (6/79) when women with clinically diagnosed or ruled-out PE were excluded.

The total numbers for the analysis were therefore:

- primary analysis – 181 women with PE, 259 women without
- secondary analysis including clinically diagnosed PE – 221 women with PE, 259 women without
- secondary analysis including clinically ruled-out PE – 181 women with PE, 301 women without
- secondary analysis excluding subsegmental PE – 173 women with PE, 259 women without.

Predictor variable completeness

Full details of predictor variable completeness are provided in *Appendix 4*. Missing data rates were generally higher in the women with diagnosed PE than in women with suspected PE. Only previous pregnancy problems, temperature and D-dimer were missing in > 5% in the suspected PE cohort, whereas employment, previous pregnancy problems, all physiological variables, ECG, likely diagnosis and D-dimer were missing in > 5% of the diagnosed PE cohort.

All seven physiological variables were recorded in 141 out of 198 women (71.2%) with diagnosed PE, whereas 44 out of 198 women (22.2%) had between one and three variables missing and 13 out of 198 women (6.6%) had four or more missing. The corresponding figures were 286 out of 324 (88.3%), 37 out of 324 (11.4%) and 1 out of 324 (0.3%) for recruited women with suspected PE, and 49 out of 95 (51.6%), 44 out of 95 (46.3%) and 2 out of 95 (2.1%) for non-recruited women with suspected PE.

Previous medical history data were complete for 190 out of 198 (96.0%) women with diagnosed PE, 323 out of 324 (99.7%) recruited women with PE and 88 out of 95 (92.6%) non-recruited women with suspected PE. Only 1 woman out of 198 women with diagnosed PE (0.5%) and 2 out of 95 non-recruited women with suspected PE (2.1%) had more than half of their data missing for this category.

Data relating to the current pregnancy were complete for 189 out of 198 women with diagnosed PE (95.5%), 324 out of 324 recruited women with suspected PE (100%) and 90 out of 95 non-recruited women with suspected PE (94.7%). Only 4 out of 198 women with diagnosed PE (2.0%) and 3 out of 95 non-recruited women with PE (3.2%) had more than half of their data missing in this category.

Overall, 15 out of 198 women with diagnosed PE (7.6%), 2 out of 324 recruited women with suspected PE (0.6%) and 2 out of 95 non-recruited women with suspected PE (2.1%) were excluded from the multivariable analysis and the analysis of CDRs, as they met our criteria for exclusion on the basis of having too many missing data.

Characteristics of the cohorts

Table 5 shows the characteristics of the women with diagnosed PE, the women with suspected PE and the non-recruited women. The mean age was 29.3 years for the women with suspected PE and 30.1 years for women with diagnosed PE. The mean age for mothers of live births in England and Wales was 30.3 years in 2015.[51] Most women were white British, but there were significant minorities of Asian Pakistani and Asian Bangladeshi women in the suspected PE cohort, reflecting the minority ethnic populations of the participating site catchment areas. The incidence of both suspected and diagnosed PE increased from the first trimester to the third trimester.

Table 6 compares the characteristics of the recruited women and the non-recruited women with suspected PE. There were no marked differences between the recruited women and the non-recruited women. In December 2016, the mean age at booking of pregnant women in the UK was 29.6 years and 46% were classified as being overweight or obese (BMI of > 25 kg/m^2), so both groups were similar to the general UK pregnant population.[52]

Follow-up of suspected pulmonary embolism cohort

There were no withdrawals from the prospective data collection and hospital records were viewed at the 30-day follow-up point for all participants. A questionnaire was sent to 321 out of 324 participants (99%)

TABLE 5 Characteristics of the cohorts

Characteristic	Cohort		
	Women with diagnosed PE (*N* = 198)	Recruited women with suspected PE (*N* = 324)	Non-recruited women with suspected PE (*N* = 95)
Mean age (years)	30.1	29.3	30.0
Ethnic group, *n* (%)			
White British	151 (76.3)	204 (63.0)	50 (52.6)
White Irish	4 (2.0)	–	–
White other	9 (4.5)	18 (5.6)	5 (5.3)
Mixed white and black Caribbean	–	3 (0.9)	–
Mixed white and black African	–	2 (0.6)	–
Mixed white and Asian	–	2 (0.6)	–
Mixed other	–	1 (0.3)	–
Asian Indian	4 (2.0)	6 (1.9)	4 (4.2)
Asian Pakistani	7 (3.5)	36 (11.1)	15 (15.8)
Asian Bangladeshi	1 (0.5)	19 (5.9)	3 (3.2)
Asian other	3 (1.5)	5 (1.5)	5 (5.3)
Black Caribbean	4 (2.0)	5 (1.5)	2 (2.1)
Black African	9 (4.5)	17 (5.2)	6 (6.3)
Black other	–	4 (1.2)	1 (1.1)
Chinese	–	–	–
Other	3 (1.5)	2 (0.6)	2 (2.1)
Missing	3 (1.5)	–	2 (2.1)
Marital status, *n* (%)			
Cohabiting	67 (33.8)	87 (26.9)	29 (30.5)
Married	93 (47.0)	171 (52.8)	46 (48.4)
Single	35 (17.7)	65 (20.0)	19 (20.0)
Missing	3 (1.5)	1 (0.3)	1 (1.1)
Employment, *n* (%)			
Unemployed	74 (37.4)	127 (39.2)	46 (48.4)
Employed	111 (56.0)	195 (60.2)	33 (34.7)
Missing	13 (6.6)	2 (0.6)	16 (16.8)
Previous pregnancies lasting for > 24 weeks' gestation, *n* (%)			
None	59 (29.8)	115 (35.5)	30 (31.6)
≥ 1	136 (68.7)	209 (64.5)	63 (66.3)
Missing	3 (1.5)	–	2 (2.1)

continued

TABLE 5 Characteristics of the cohorts (*continued*)

Characteristic	Cohort		
	Women with diagnosed PE (*N* = 198)	Recruited women with suspected PE (*N* = 324)	Non-recruited women with suspected PE (*N* = 95)
Previous pregnancies lasting for < 24 weeks' gestation, *n* (%)			
None	116 (58.6)	199 (61.4)	62 (65.3)
1 or more	74 (37.3)	125 (38.6)	27 (28.4)
Missing	8 (4.0)	–	6 (6.3)
Current pregnancy, *n* (%)			
First trimester	15 (7.6)	21 (6.5)	7 (7.4)
Second trimester	43 (21.7)	110 (34.0)	18 (18.9)
Third trimester	70 (35.4)	138 (42.6)	34 (35.8)
Post partum	60 (30.3)	55 (17.0)	35 (36.8)
Missing	10 (5.1)	–	1 (1.1)

–, no data.

TABLE 6 Characteristics of recruited women and non-recruited women with a suspected PE

Characteristic	Cohort	
	Recruited women with suspected PE (*N* = 324)	Non-recruited women with suspected PE (*N* = 95)
Mean age (years)	29.3	30.0
Mean BMI (kg/m^2)	28.0	28.7
Multiple pregnancy, *n* (%)	12 (3.7)	3 (3.2)
Receiving thromboprophylaxis, *n* (%)	89 (27.5)	29 (30.5)
Mean heart rate (b.p.m.)	95.8	96.2
Mean respiratory rate (breaths per minute)	18.7	18.7
Mean oxygen saturation (%)	97.7	97.8
Mean systolic BP (mmHg)	122	126
Mean diastolic BP (mmHg)	73	75
Mean temperature (°C)	36.6	36.5
PE in the primary analysis, *n*/*N* (%)	18/277 (6.5)	6/79 (7.6)

BP, blood pressure; b.p.m., beats per minute.

and 265 out of 324 participants (82%) received at least one reminder to return the questionnaire. Questionnaires were returned by 135 out of 324 participants (41%) with a mean response time of 52.4 days. There was insufficient follow-up information to rule out a VTE event for 16 out of 324 women (4.9%), which resulted in contact with the GP by local researchers, and 14 out of 16 questionnaires were returned by the GP.

Table 7 summarises the EQ-5D-5L data at the 30-day follow-up point from the 136 participants who completed and returned the questionnaire. Most women were in good health at 30 days, although a substantial proportion had slight problems with mobility, usual activities, anxiety/depression and pain/discomfort, and around 15% had moderate problems with usual activities or pain/discomfort.

TABLE 7 The EQ-5D-5L results at the 30-day follow-up point

	Problems and severity, *n* (%)				
Dimension	None	Slight	Moderate	Severe	Extreme
Mobility	92 (67.6)	33 (24.3)	9 (6.6)	1 (0.7)	1 (0.7)
Self-care	126 (92.6)	8 (5.9)	2 (1.5)	0 (0)	0 (0)
Usual activity	79 (58.0)	36 (26.5)	19 (14.0)	2 (1.5)	0 (0)
Pain/discomfort	59 (43.4)	51 (37.5)	21 (15.4)	4 (2.9)	1 (0.7)
Anxiety/depression	99 (72.8)	28 (20.6)	5 (3.7)	4 (2.9)	0 (0)

Characteristics of women with and without pulmonary embolism

Table 8 compares the characteristics of women with and without PE in the primary analysis data set. Women with PE were more likely to be older, to be post partum, to have given up smoking during the current pregnancy, to have had a previous pregnancy lasting for > 24 weeks' gestation, to have had previous pregnancy problems, to have had surgery in the previous 4 weeks, to have a history of VTE, to have a problem with their current pregnancy or to have received thromboprophylaxis. They were less likely to have a family history of VTE, varicose veins, multiple pregnancy or recent long-haul travel.

TABLE 8 Characteristics of women with and without PE in the primary analysis data set

	Cohort, *n* (%)	
Predictor	Women without PE	Women with PE
Aged > 35 years	40 (15.4)	37 (20.4)
BMI of ≥ 30 kg/m^2	85 (32.8)	60 (34.5)
Smoking status		
Never	171 (66.0)	116 (64.1)
Gave up prior to pregnancy	39 (15.1)	28 (15.5)
Gave up during pregnancy	19 (7.3)	23 (12.7)
Current	30 (11.6)	14 (7.7)
Previous pregnancies		
One or more previous pregnancy lasting for < 24 weeks' gestation	97 (37.5)	68 (37.6)
One or more previous pregnancy lasting for > 24 weeks' gestation	165 (63.7)	126 (69.6)
Previous pregnancy problems	70 (27.0)	55 (30.4)
Previous medical problems		
Family history of VTE	46 (17.8)	24 (13.3)
History of varicose veins	19 (7.3)	5 (2.8)
History of i.v. drug use	1 (0.4)	1 (0.6)
Known thrombophilia	7 (2.7)	4 (2.2)
Surgery in the previous 4 weeks	21 (8.1)	35 (19.3)
Significant injury in the previous 4 weeks	3 (1.2)	2 (1.1)
Previous VTE	15 (5.8)	19 (10.5)
Other previous medical problem	110 (42.5)	75 (41.4)
Other previous medical problem (VTE related)	6 (2.3)	4 (2.2)

continued

TABLE 8 Characteristics of women with and without PE in the primary analysis data set (*continued*)

Predictor	Cohort, *n* (%)	
	Women without PE	Women with PE
Current pregnancy		
First trimester	20 (7.7)	15 (8.3)
Second trimester	79 (30.5)	37 (20.4)
Third trimester	116 (44.8)	60 (33.1)
Post partum	44 (17.0)	63 (34.8)
Multiple pregnancy	12 (4.6)	4 (2.2)
Long-haul travel during pregnancy	21 (8.1)	2 (1.1)
≥ 3 days of immobility/bed rest during pregnancy	21 (8.1)	14 (7.7)
Received thromboprophylaxis	70 (27.0)	88 (48.6)
Previous thrombotic event during this pregnancy	3 (1.2)	5 (2.8)
Other problems with this pregnancy	73 (28.2)	74 (40.9)
Other problems with this pregnancy (VTE related)	19 (7.3)	15 (8.3)
Presenting features		
Presenting feature: pleuritic chest pain	137 (52.9)	94 (51.9)
Presenting feature: non-pleuritic chest pain	47 (18.1)	38 (21.0)
Presenting feature: SOB at rest	157 (60.6)	97 (53.6)
Presenting feature: SOB on exertion	125 (48.3)	93 (51.4)
Presenting feature: haemoptysis	10 (3.9)	13 (7.2)
Presenting feature: productive cough	23 (8.9)	16 (8.8)
Presenting feature: syncope	7 (2.7)	9 (5.0)
Presenting feature: palpitations	30 (11.6)	24 (13.3)
Presenting feature: other	90 (34.7)	62 (34.3)
Temperature of > 37.5 °C	7 (2.7)	14 (7.7)
Diastolic BP of < 50 mmHg	2 (0.8)	4 (2.2)
Systolic BP of < 90 mmHg	1 (0.4)	3 (1.7)
Oxygen saturation of < 94% on room air	10 (3.9)	27 (14.9)
Respiratory rate of > 24 breaths per minute	25 (9.7)	18 (9.9)
Heart rate of > 100 b.p.m. (in first or second trimester, or post partum) or of > 110 b.p.m. (in third trimester)	72 (27.8)	55 (30.4)
Clinical signs of DVT	23 (8.9)	23 (12.7)
PE-related ECG abnormality	8 (3.1)	4 (2.2)
PE-related chest radiograph abnormality	1 (0.4)	9 (5.0)
Other chest radiograph abnormality	18 (6.9)	30 (16.6)
Normal chest radiograph	221 (85.3)[a]	108 (59.7)[a]
PE most likely diagnosis (permissive)	202 (78.0)	144 (79.6)
PE most likely diagnosis (strict)	34 (13.1)	99 (54.7)

BP, blood pressure; b.p.m., beats per minute; i.v., intravenous; SOB, shortness of breath.
a Thirty-six women with PE (19.9%) and 19 women without PE (7.3%) had no chest radiograph report available.

In terms of presenting characteristics, women with PE were more likely to have reported haemoptysis or syncope and less likely to have reported shortness of breath on exertion. There were only small differences in the proportion of women reporting pleuritic or non-pleuritic chest pain, shortness of breath at rest, cough, palpitations or other symptoms. Women with PE were more likely to have an elevated temperature or low peripheral oxygen saturation. There were only small differences in the proportion of women with a high heart rate or respiratory rate, and very few women had a low blood pressure. Women with PE were more likely to have an abnormal chest radiograph than those without PE, regardless of whether or not the abnormality was considered to be PE related. Women with PE were more likely to have an overall diagnostic impression suggesting PE, but only when a strict interpretation rather than a permissive interpretation was used.

Table 9 compares the mean age, BMI and physiological measurements between women with and without PE in the primary analysis data set. The distributions of these measures are shown in the figures in *Appendix 5*. Women with PE were slightly older, had a slightly higher mean BMI and mean heart rate and had a slightly lower mean oxygen saturation. There was little difference between women with and without PE, in terms of both the mean and the distribution. Many women with PE had normal physiological measurements and many without PE had abnormal physiological measurements.

Table 10 shows the results of the univariable logistic regression using the primary analysis data set. The only clinical features significantly associated with a diagnosis of PE ($p < 0.05$) were number of previous pregnancies lasting beyond 24 weeks' gestation, surgery in the previous 4 weeks (including caesarean section), no history of varicose veins, no long-haul travel during pregnancy, higher temperature, lower oxygen saturation, overall diagnostic impression suggesting PE (strict interpretation) and chest radiograph abnormality (both PE related and non-PE related). Women who had received thromboprophylaxis were significantly more likely to have a diagnosis of PE.

Some of these associations suggest that known risk factors for PE are associated with a diagnosis of no PE. This may be explained by women having a lower threshold for seeking medical attention or clinicians having a lower threshold for investigation in the presence of known risk factors. Whatever the explanation, these counter-intuitive associations are unlikely to be useful in diagnostic decision-making, as they will not be clinically credible and any diagnostic value is likely to depend on implicit selection processes during presentation that may vary between settings.

TABLE 9 Comparison of physiological measurements between women with and without PE

	Cohort, mean (SD)	
Measurement	Women without PE	Women with PE
Age (years)	29.4 (5.9)	30.2 (6.2)
BMI (kg/m^2)	28.0 (6.5)	28.7 (7.6)
Heart rate (b.p.m.)	95.5 (17.7)	98.3 (19.7)
Respiratory rate (breaths per minute)	19.0 (4.4)	19.0 (5.0)
Oxygen saturation (%)	97.8 (1.8)	96.5 (4.4)
Systolic BP (mmHg)	123 (15.7)	121 (17.0)
Diastolic BP (mmHg)	72.9 (11.7)	74.2 (12.3)
Temperature (°C)	36.6 (0.8)	36.8 (0.6)

BP, blood pressure; b.p.m., beats per minute; SD, standard deviation.

TABLE 10 Odds ratios, 95% CIs and *p*-values for univariable regression

Term	Odds ratio	95% CI	*p*-value
Age (continuous)	1.02	0.99 to 1.05	0.179
Age > 35 years	1.41	0.86 to 2.31	0.176
BMI (kg/m^2) (continuous)	1.01	0.99 to 1.04	0.372
BMI of ≥ 30 kg/m^2	1.01	0.68 to 1.52	0.942
Ex-smoker (prior)	1.06	0.61 to 1.81	0.837
Ex-smoker (during)	1.78	0.93 to 3.46	0.082
Current smoker	0.69	0.34 to 1.33	0.279
Number of previous pregnancies lasting for < 24 weeks' gestation (continuous)	1.05	0.91 to 1.23	0.509
≥ 1 pregnancy lasting for < 24 weeks' gestation	1.00	0.68 to 1.49	0.98
Number of previous pregnancies lasting for > 24 weeks' gestation (continuous)	1.20	1.04 to 1.30	0.017
≥ 1 pregnancy lasting for > 24 weeks' gestation	1.30	0.87 to 1.97	0.198
Previous pregnancy problems	1.18	0.77 to 1.79	0.442
Family VTE	0.71	0.41 to 1.20	0.205
History of varicose veins	0.36	0.12 to 0.91	0.045
History of i.v. drug use	1.43	0.06 to 36.4	0.8
Known thrombophilia	0.81	0.21 to 2.74	0.745
Surgery in previous 4 weeks (including caesarean section)	2.72	1.53 to 4.92	0.001
Injury in the previous 4 weeks	0.95	0.12 to 5.81	0.959
Previous VTE	1.91	0.95 to 3.92	0.073
Other previous medical problem	0.96	0.65 to 1.41	0.829
Other previous medical problem (VTE related)	1.02	0.26 to 3.62	0.978
Second trimester	0.62	0.29 to 1.37	0.234
Third trimester	0.69	0.33 to 1.46	0.324
Post partum	1.91	0.89 to 4.19	0.101
Multiple pregnancy	0.47	0.13 to 1.36	0.191
Long-haul travel during pregnancy	0.13	0.02 to 0.44	0.006
≥ 3 or more days of immobility/bed rest during pregnancy	0.95	0.46 to 1.91	0.887
Received thromboprophylaxis	2.56	1.72 to 3.82	< 0.001
Previous thrombotic event this pregnancy	2.44	0.59 to 12.0	0.226
Other problem with this pregnancy	1.46	0.97 to 2.20	0.067
Other problem with this pregnancy (VTE related)	1.14	0.56 to 2.31	0.713
Presenting: pleuritic chest pain	0.96	0.66 to 1.41	0.842
Presenting: non-pleuritic chest pain	1.20	0.74 to 1.93	0.457
Presenting: SOB (exertion)	1.13	0.77 to 1.66	0.52
Presenting: SOB (rest)	0.75	0.51 to 1.10	0.142
Presenting: haemoptysis	1.93	0.83 to 4.61	0.129
Presenting: cough	1.00	0.50 to 1.93	0.988
Presenting: syncope	1.88	0.69 to 5.36	0.218

TABLE 10 Odds ratios, 95% CIs and *p*-values for univariable regression (*continued*)

Term	Odds ratio	95% CI	*p*-value
Presenting: palpitations	1.17	0.65 to 2.07	0.598
Presenting: other	0.98	0.65 to 1.46	0.914
Temperature of > 37.5 °C	3.02	1.23 to 8.11	0.02
Temperature (continuous)	1.75	1.22 to 2.57	0.003
Diastolic BP of < 50 mmHg	2.90	0.56 to 21.1	0.221
Diastolic BP (continuous)	1.01	0.99 to 1.03	0.256
Systolic BP of < 90 mmHg	4.35	0.55 to 88.3	0.205
Systolic BP (continuous)	0.99	0.98 to 1.01	0.322
Oxygen saturation of < 94%	4.37	2.12 to 9.71	< 0.001
Oxygen saturation (continuous)	0.85	0.78 to 0.92	< 0.001
Respiratory rate of > 24 breaths per minute	1.03	0.54 to 1.95	0.919
Respiratory rate (continuous)	1.00	0.96 to 1.04	0.948
Heart rate of > 100 b.p.m. (110 b.p.m. in third trimester)	1.13	0.75 to 1.72	0.556
Heart rate (continuous)	1.01	0.99 to 1.02	0.126
Clinical signs of DVT	1.49	0.81 to 2.77	0.199
PE-related ECG abnormality	0.71	0.19 to 2.29	0.579
PE-related chest radiograph abnormality	15.2	2.82 to 282.0	0.01
Other chest radiograph abnormality	2.82	1.53 to 5.33	0.001
PE is most likely diagnosis *or* equally likely (permissive)	1.48	0.85 to 2.62	0.174
PE is most likely diagnosis *or* equally likely (strict)	9.15	5.70 to 15.0	< 0.001

BP, blood pressure; b.p.m., beats per minute; i.v., intravenous; SOB, shortness of breath.
CIs were calculated using methods outlined by Harrell[53] and Dobson and Barnett.[54]

Diagnostic accuracy of clinical decision rules

Table 11 reports the diagnostic accuracy of each CDR and *Figure 1* shows the ROC curve. The sensitivity and specificity are reported for the recommended or usual threshold for decision-making (i.e. any score vs. zero for the PERC score, PE likely vs. unlikely for Wells's criteria and low risk vs. moderate or high risk for the Geneva score). The ROC figure and the area under the ROC (AUROC) curve relate to the performance of the rule across the range of values, rather than just using the recommended or usual threshold.

The diagnostic accuracy of the rules was generally poor. The consensus rules were derived specifically for pregnant and postpartum women and were intended to identify a low-risk group of women who could be discharged without imaging, but performed little better than chance. The sensitive rule had good sensitivity (95%) but very poor specificity (4%), showing that sensitivity was achieved only by setting a very low threshold for positivity.

The existing CDRs were developed for the general population with suspected PE. We adapted the relevant criteria to make them appropriate for a pregnant population, but did not otherwise alter the rules. Wells's criteria may have some modest diagnostic value if the criterion 'PE is the most likely diagnosis or equally likely' is applied in a strict way (i.e. it is only positive if PE is clearly considered to be the most likely or equally most likely diagnosis). The other rules performed little better than chance.

TABLE 11 Diagnostic accuracy of the CDRs

Decision rule	Diagnostic accuracy		
	AUROC using full range of score values (95% CI)	Sensitivity at usual or recommended threshold (95% CI), *n*/*N*	Specificity at usual or recommended threshold (95% CI), *n*/*N*
Primary consensus	0.626 (0.572 to 0.681)	0.609 (0.532 to 0.683), 103/169	0.585 (0.523 to 0.646), 151/258
Sensitive consensus	0.620 (0.566 to 0.675)	0.959 (0.917 to 0.983), 162/169	0.035 (0.016 to 0.065), 9/258
Specific consensus	0.589 (0.537 to 0.642)	0.361 (0.289 to 0.438), 61/169	0.783 (0.728 to 0.832), 202/258
PERC score	0.621 (0.570 to 0.672)	0.675 (0.598 to 0.745), 114/169	0.519 (0.457 to 0.582), 134/258
Simplified revised Geneva score	0.579 (0.526 to 0.632)	0.444 (0.368 to 0.522), 75/169	0.636 (0.574 to 0.694), 164/258
Wells's score (permissive)[a]	0.577 (0.522 to 0.632)	0.490 (0.410 to 0.571), 77/157	0.617 (0.553 to 0.678), 153/248
Wells's score (strict)[a]	0.732 (0.682 to 0.782)	0.376 (0.300 to 0.457), 59/157	0.895 (0.850 to 0.930), 222/248

AUROC, area under the receiver operating characteristic.
a Wells's criteria scores were tested using a liberal (permissive) interpretation of clinical diagnosis text to determine whether PE was the most likely or equally most likely diagnosis and a more strict interpretation.

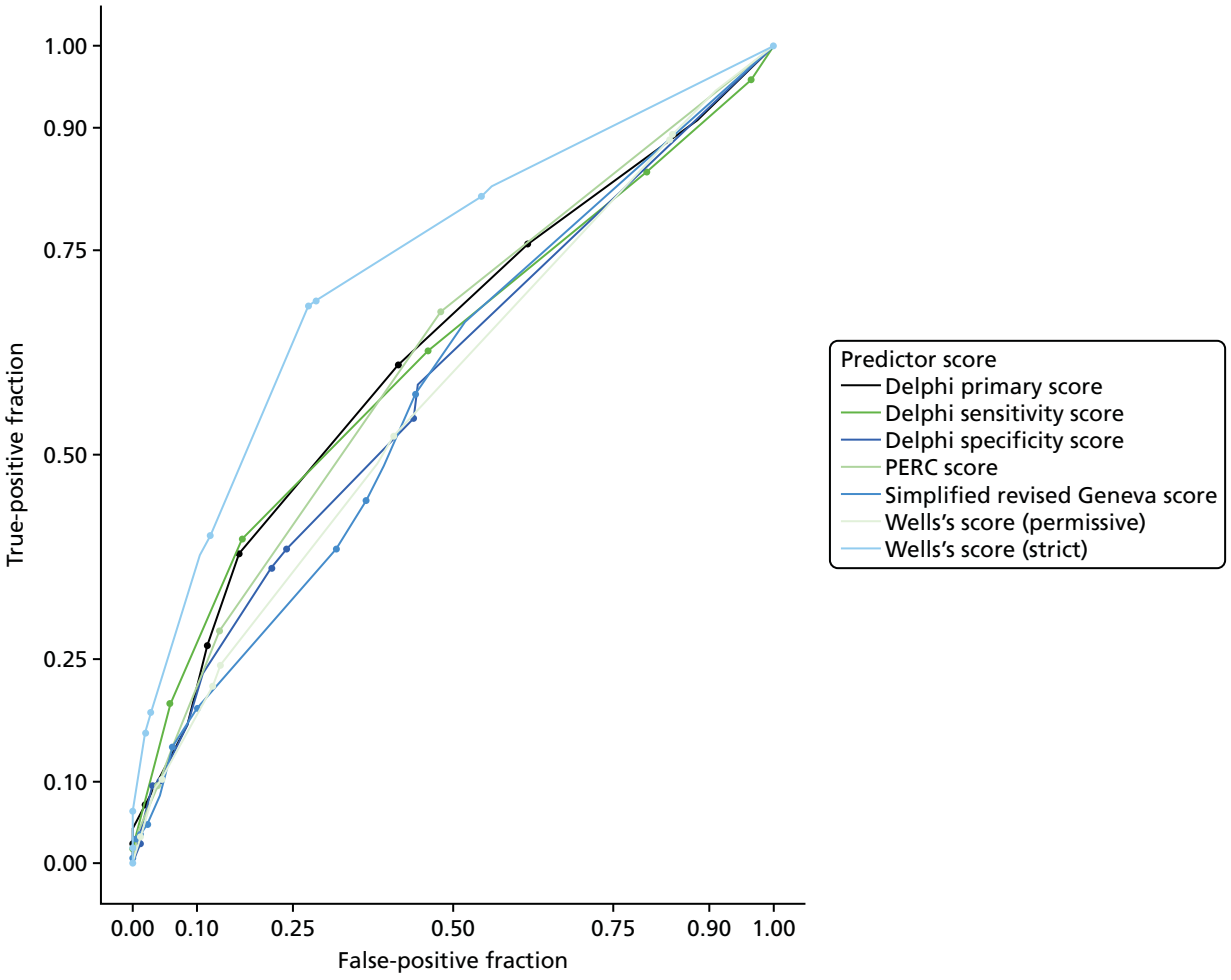


FIGURE 1 Receiver operating characteristic curves for CDRs.

Appendix 6 provides figures showing how each rule performed across the range of scores. These show that a non-negligible proportion of women with PE were categorised in the lowest risk stratum of all of the scores. Only the sensitive consensus rule and Wells's score (permissive application) had < 5% of women with PE in the lowest stratum and both of these methods categorised too few women without PE in the lowest risk category to be clinically useful (i.e. using the rule would not make a practically important difference to patient management).

Wells's score and the simplified revised Geneva score are usually used to select low-risk patients for biomarker testing, whereas the PERC score is intended to allow discharge without testing if it is negative. Our analysis showed that a substantial proportion of women with PE have a negative PERC score. In applying the PERC score, we did not apply the criterion of exogenous oestrogen. If being pregnant or post partum were considered to carry the same risk as exogenous oestrogen, then the PERC score would be positive in all pregnant and postpartum women.

Appendix 7 shows the coefficients from logistic regression for each of the elements of each rule. This gives an indication of the contribution that each element makes to the diagnostic performance of the rule. The consensus-derived rules included elements that the experts thought would be useful, but added little diagnostic value (pleuritic chest pain, injury, medical comorbidities, raised BMI, tachycardia and tachypnoea) and two elements (being in the third trimester, family history of VTE) that had weak associations with the absence of PE. The existing rules had similar problems related to the inclusion of elements with little diagnostic value, but Wells's score benefited from the inclusion of the criterion 'PE is the most likely diagnosis *or* equally likely' when this criterion was recorded as positive only if the diagnostic impression clearly indicated a positive diagnosis of PE rather than mentioning it as a possibility.

D-dimer analysis (using measurements from routine care)

D-dimer measurements were recorded as part of routine care for 44 out of 198 women with diagnosed PE (22%) and 156 out of 324 women with suspected PE (48%). After the exclusion of 22 women with clinically diagnosed or ruled-out PE, the primary analysis data set for those with routine care D-dimer measurements consisted of 53 women with PE and 125 women without PE.

We have not reported absolute D-dimer values because the measurements were made using a variety of assays with different thresholds for positivity. Instead, we simply report the sensitivity and specificity of D-dimer measurements using the threshold specified by the hospital laboratory (the conventional threshold) and the pregnancy-specific thresholds we defined a priori (conventional threshold in the first trimester, 1.5× the conventional threshold in the second trimester and 2× the conventional threshold in the third trimester). Ten women with PE did not have a hospital laboratory threshold reported and thus could not be included in the analysis, which therefore involved 43 women with PE and 125 women without PE.

Using the hospital laboratory threshold, the sensitivity (*n*/*N*) of D-dimer was 88.4% (38/43, 95% CI 74.1% to 95.6%) and the specificity was 8.8% (11/125, 95% CI 4.7% to 15.6%). Using the gestation-specific threshold, the sensitivity (*n*/*N*) of D-dimer was 69.8% (30/43, 95% CI 53.7% to 82.3%) and the specificity was 32.8% (41/125, 95% CI 24.8% to 41.9%).

Multivariable analysis

Details of the multivariable analysis are reported in *Appendix 8*. Leave-one-out cross-validation was utilised internally at each step of the fitting of the LASSO to shrink the point estimate and inform the next iteration. Optimal values for the parameter lambda were identified as the minimum value or the value corresponding to 1 × the SE of the point estimate of the mean squared error. Models were derived for each of these values and the diagnostic parameters were calculated. *Figure 2* shows the ROC curve and *Table 12* reports the coefficients for each predictor that contributed significantly to the model and the diagnostic parameters for the model.

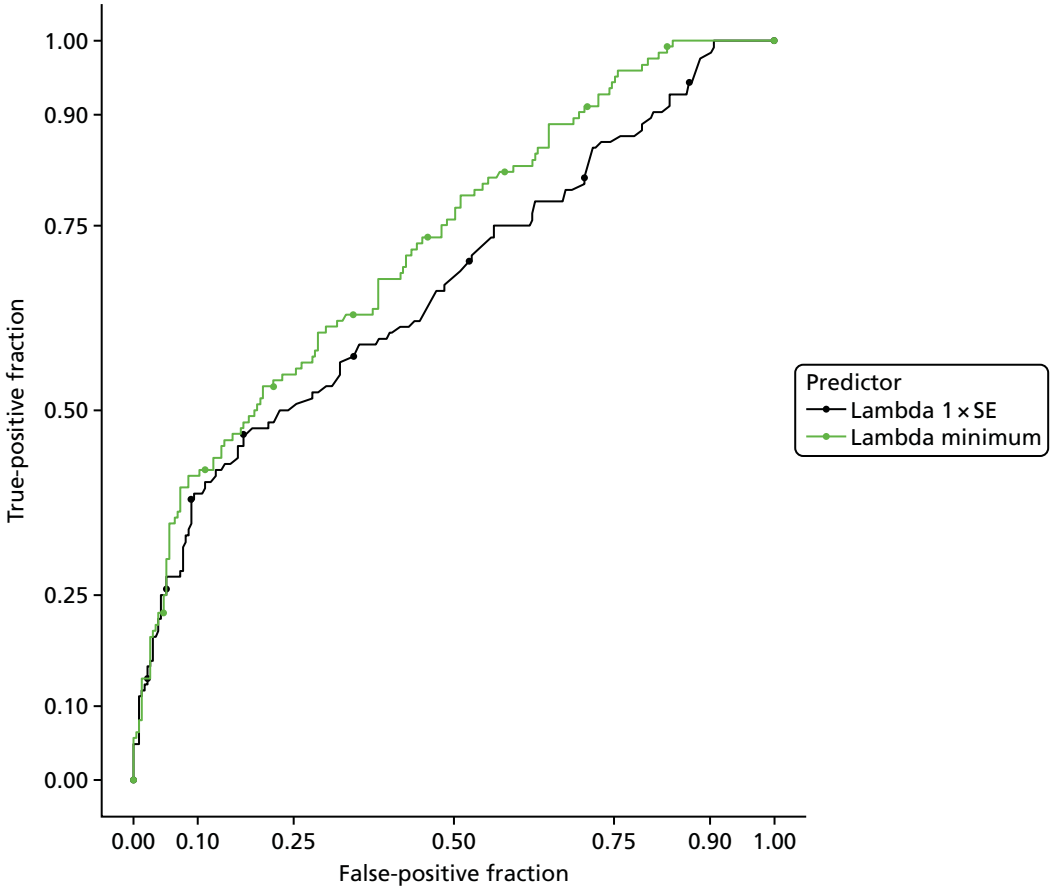


FIGURE 2 Receiver operating characteristic curves for the multivariable models.

TABLE 12 Coefficients for each term and diagnostic parameters for the multivariable models

Term or diagnostic parameter	1 × SE model	Minimum value model
(Intercept)	1.915	–3.987
Previous VTE	0.000	0.256
Long-haul travel during pregnancy	–0.428	–1.225
Multiple pregnancy	0.000	–0.402
Oxygen saturation (continuous)	–0.041	–0.065
Surgery in previous 4 weeks	0.028	0.299
Temperature (continuous)	0.037	0.273
PE-related chest radiograph abnormality	0.413	0.660
AUC (95% CI)	0.668 (0.607 to 0.729)	0.724 (0.669 to 0.779)
Sensitivity, % (95% CI)	1.00 (0.971 to 1.000)	0.831 (0.753 to 0.892)
Specificity, % (95% CI)	0.077 (0.046 to 0.119)	0.391 (0.328 to 0.456)

The analysis suggests that there is little potential for an accurate and usable CDR. The most accurate model used previous VTE, long-haul travel during pregnancy, multiple pregnancy, oxygen saturation (as a continuous variable), surgery in the previous 4 weeks, temperature (as a continuous variable) and PE-related chest radiograph abnormality to predict PE with an AUC of 0.724 (95% CI 0.669 to 0.779). The ROC curve shows that the specificity would have to be as low as 20% to achieve a level of sensitivity that is acceptable to allow imaging to be avoided (> 95%). This estimate would need to be validated in a new cohort and statistical shrinkage would probably result in worse accuracy. Furthermore, the model includes variables with a counter-intuitive negative association with PE (history of long-haul travel and multiple pregnancy) that may be dependent on referral processes and would therefore vary between settings and over time, if referral processes changed.

In view of the poor accuracy of the model in the derivation cohort, we did not proceed to internal validation or attempting to make the model more clinically credible or usable.

Recursive partitioning

Recursive partitioning resulted in a range of models of increasing complexity, accuracy and risk of overfitting. Details are reported in *Appendix 9*. The optimal model is shown in *Figure 3*; it uses BMI, oxygen saturation, heart rate and trimester to categorise women. The percentages show how the study population is split by the partitioning process. The proportions are the proportion of PE in each subgroup. The high proportion of PE in each subgroup reflects the case–control design. The prevalence of PE in the suspected PE group was 6.5%, so the proportion of women with PE in each group is around six times higher than would be expected in a typical population with suspected PE.

Figure 4 shows the ROC curve for the model. The AUC was 0.657 (95% CI 0.611 to 0.703) and the threshold that provided a sensitivity of > 95% had a corresponding specificity of 5%. More complex models with more variables had higher accuracy (see *Appendix 9*), but with an increasing risk of overfitting. Therefore, although some of the very complex models had apparently acceptable accuracy, statistical shrinkage in the validation analysis would be highly likely to result in unacceptable accuracy in a validation cohort. Highly complex

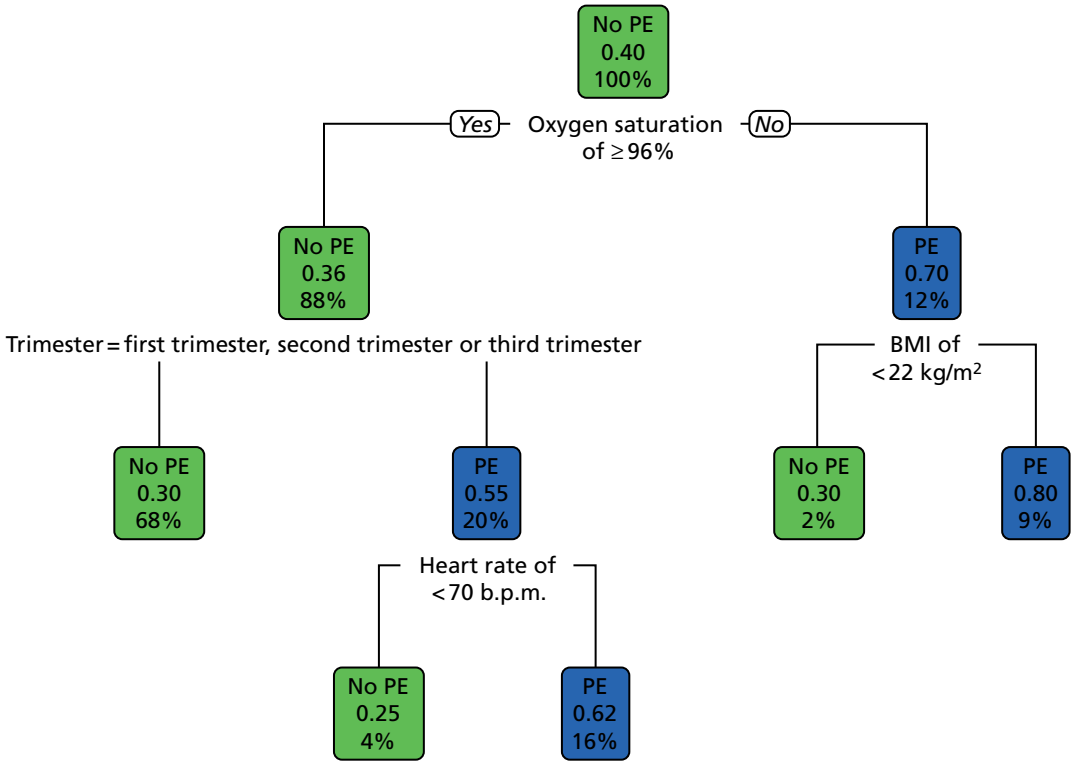


FIGURE 3 Dendrogram of the optimal recursive partitioning model. b.p.m., beats per minute.

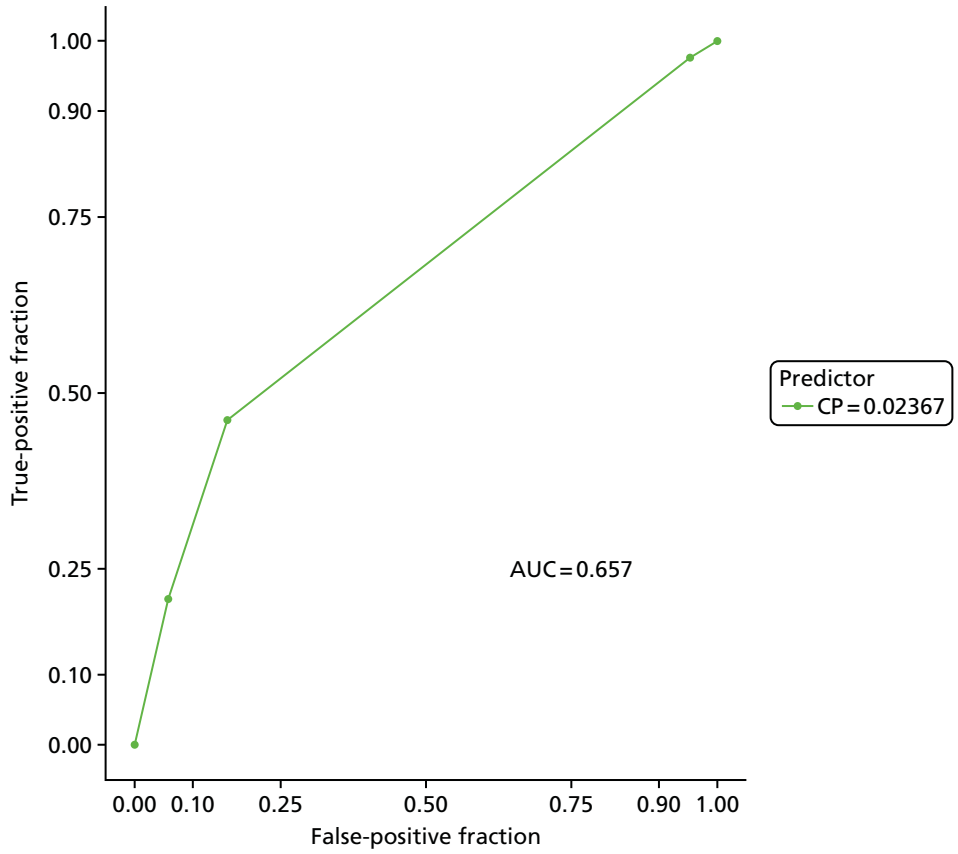


FIGURE 4 The ROC curve for the optimal recursive partitioning model. CP, complexity parameter.

models would be difficult to use in clinical practice without appropriate software and, being based on variables with weak associations with PE diagnosis, would lack clinical credibility.

Secondary analysis

The results of the secondary analysis are summarised in *Appendix 10*. There were no meaningful differences between the results of the primary analysis and the secondary analysis.

Feasibility of a prospective cohort study

We aimed to use the recruitment data for women with suspected PE to determine the feasibility of using a prospective cohort design to validate a new CDR or biomarker. The study recruited women with suspected PE at a rate of 1.7 women per site per month and a prevalence of 23 out of 324 women with PE (7.1%). This suggests that a prospective cohort study involving 50 sites (more than one-quarter of all potential sites in the UK) recruiting for 2 years could achieve a sample size of 2040, including 145 women with PE.

Discussion

Main findings

Our analysis showed that clinical variables have little diagnostic value in the assessment of pregnant and postpartum women with suspected PE. There was very little difference in physiological measures between women with PE and those without PE. Many women with PE had entirely normal physiology while many without PE had abnormal physiology. Higher temperature and lower oxygen saturation were associated with PE, but the association was too weak to be clinically useful. Recent surgery (almost entirely caesarean section) was associated with an increased risk of developing PE, whereas older age, giving up smoking during pregnancy, higher heart rate (measured as a continuous variable), absence of shortness of breath

at rest, haemoptysis, previous thrombosis, being post partum, having a single pregnancy and having obstetric problems during pregnancy showed weak associations with PE.

Important negative findings were that presenting features, with the exception of haemoptysis and possibly syncope, had little diagnostic value. Risk factors for VTE, with the exception of obstetric problems and a past history of VTE, also had little diagnostic value. Some risk factors (recent long-haul travel, varicose veins and possibly multiple pregnancy and family history of VTE) were more common in women presenting with suspected PE but in whom the diagnosis was not confirmed, than in women in whom the diagnosis of PE was confirmed. This is likely to reflect selection processes, whereby women with symptoms suggestive of PE and known risk factors are more likely to present, be referred and be investigated than those without risk factors.

There were some counter-intuitive findings, most notably that abnormal chest radiograph appearances (even those not considered to be related to PE) were associated with an increased risk of PE. Diagnostic guidelines for suspected PE in pregnancy[2] advise a chest radiograph on the basis of identifying alternative causes and determining whether to use CTPA or VQ scanning. Our findings suggest that an abnormal chest radiograph increases rather than decreases the probability of PE.

Given the poor discriminant value of the clinical predictors, it is not surprising that CDRs also had poor discriminant value, with only Wells's criteria (with a strict interpretation of PE likelihood) having an AUC that was > 0.7. Critically for clinical practice, none of the rules was able to identify a meaningfully sized low-risk group that could be selected for discharge without imaging.

The same problems arose when we attempted to derive a diagnostic model for PE using multivariable analysis and recursive partitioning. The models and decision trees we derived had poor accuracy, were unable to achieve clinically useful specificity at acceptable sensitivity and included variables with counter-intuitive associations with PE that lacked clinical credibility. The poor performance of these models in the derivation cohort suggests that attempts at validation are not worthwhile. Our findings suggest that CDRs have no useful value in selecting pregnant and postpartum women with suspected PE for imaging.

Comparison with previous studies

Our systematic review identified few studies evaluating clinical variables or CDRs for diagnosing PE in pregnancy and post partum, and those we identified included few women with PE. Previous studies reporting a lack of association between presenting features or physiological measurements and a diagnosis of PE[12,15,18] may be explained by a lack of statistical power. The DiPEP study was powered to detect clinically important associations. Our findings should therefore convincingly refute the suggestion that clinical features are diagnostically useful for PE in pregnancy and post partum (in the setting of secondary care at least).

Previous studies suggested that Wells's PE criteria may be useful in pregnant or postpartum women.[8,17] Our findings suggest that Wells's PE criteria may have some modest diagnostic value when adapted for the pregnant population and applied with a strict interpretation of the likelihood of PE, but this is unlikely to be clinically useful. The discordance between our study and previous studies may be explained by random error, owing to the small numbers in previous studies or differences in study design.

Strengths and limitations

The DiPEP study is the largest study ever undertaken to evaluate the diagnostic accuracy of clinical predictors and biomarkers for pregnant and postpartum women with suspected PE. In terms of the number of women with PE, it is many times larger than any previous study. This provided us with much greater power to detect associations with PE and allowed us to estimate diagnostic sensitivity with much greater precision. The women with suspected PE were a relatively unselected cohort presenting to a representative group of hospitals, involving both maternity units and emergency departments. The women with diagnosed PE were identified across all UK hospitals and, therefore, were likely to be a representative sample. Data completion rates were generally good and follow-up ensured that the risk of misclassification of women

with PE as having no PE was minimised. The primary analysis was limited to women with imaging confirmation to reduce the risk of misclassification, whereas the secondary analysis including women with clinically diagnosed or ruled-out PE explored the applicability of findings to a wider cohort.

The large number of women in the study with PE was achieved by using a case–control design, which has some inevitable limitations. Strictly speaking, the design was a cohort design in which the cohort was augmented by the inclusion of additional patients with PE. However, it shares potential limitations with the case–control design. Patients may be selected and represent the more severe end of the spectrum with the disease, whereas control patients may not be typical of those investigated without the disease and may not have conditions that lead to a false-positive diagnosis. We reduced these risks by including all women with diagnosed PE as participants and then excluding those requiring immediate resuscitations (or who we could not verify had not required immediate resuscitation), so that they would be representative of women diagnosed with PE and presenting with suspected PE. Women in the control group were not healthy control participants, but women presenting with suspected PE who underwent investigation. Crucially, it should be noted that the bias associated with the case–control design tends to inflate the estimates of sensitivity and specificity. This means that our findings of poor discriminant value for clinical predictors and CDRs are unlikely to be undermined by the risk of bias.

A more significant risk of bias may relate to differences in data collection methods between women with diagnosed PE and women with suspected PE. Data from the former group were retrospectively recorded by UKOSS clinicians using the hospitals records, whereas data from the latter group could be collected by the research nurse/midwife directly questioning patients. We encouraged the research nurse/midwife to rely on hospital records rather than patient interview to collect predictor variables, but we cannot be sure that they did not rely on patient interview. Furthermore, to reduce potential screening and data collection bias, the immediate collection of prospective data would be preferable. Missing data rates were higher for the UKOSS data, which suggests that the identification of clinical predictors from the women with suspected PE may have been more rigorous. This potential bias offers an alternative explanation to the one considered above for the counter-intuitive findings that some risk factors for developing VTE (long-haul travel, varicose veins) were associated with an absence of PE. It does not explain the negative associations with easily identifiable predictors, such as multiple pregnancy, or the limited predictive value of physiological variables.

Crucially, it should be noted that the participants in our study had all been through some sort of selection process, either by virtue of deciding to self-present to emergency or maternity services, or by referral by a health professional. This process is likely to have been based on many of the clinical variables that we examined in our analysis. It is therefore possible that the diagnostic value of clinical variables is ‘used up’ during the referral process. For example, if women with risk factors for VTE or abnormal physiology are more likely to self-present or be referred, then these factors will have less diagnostic value than if those presenting are a random or unselected sample of women suffering symptoms compatible with PE during pregnancy or post partum.

This does not undermine the relevance of our findings to secondary care, as clinicians working in emergency departments and maternity units will see similarly selected patients, but does mean that we should be cautious about extrapolating our findings to primary care and other settings outside hospital. Clinical variables may be being used to select women for hospital attendance, which explains their lack of value in a hospital cohort.

It is also important to recognise that our findings do not challenge existing knowledge on what constitutes risk factors for VTE in pregnancy. The DiPEP study was designed to determine the diagnostic value of clinical features (including the risk factors for developing VTE) in women with suspected PE in pregnancy and post partum. It was not designed to determine whether or not these features are risk factors for developing VTE. Previous studies comparing women who develop VTE in pregnancy with those who do not have already answered this question.[27–29,55]

Conclusions

Clinical features have limited diagnostic value for PE in pregnant and postpartum women and CDRs are unlikely to have a useful role in selecting women for imaging. We were unable to derive a new decision rule with sufficient diagnostic accuracy and clinical credibility and utility to justify further validation.

Chapter 4 Biomarker study

Introduction

As outlined in *Chapter 1*, a number of biomarkers have a potential role in selecting pregnant or postpartum women for diagnostic imaging, but only D-dimer measurement is currently used in routine clinical practice, with troponin and BNP being used to grade the extent of PE. D-dimer measurements recorded in clinical practice were collected and analysed as part of the case–control study, but we anticipated that the analysis of these data would be limited by missing data, as current guidance[2] advises against using the D-dimer measurement, and by the use of different assays with different thresholds for positivity at different hospitals. We therefore planned to analyse the D-dimer measurement along with a number of other biomarkers, using blood samples obtained from women with suspected PE and women presenting to the prospectively recruiting sites with diagnosed DVT.

Aims and objectives

The aim of the biomarker study was to explore the potential diagnostic value of classical and alternative biomarkers for PE in pregnant and postpartum women.

Methods

The prospective identification of a cohort of pregnant or postpartum women with suspected PE for the case–control study offered the opportunity to collect blood samples for biomarker evaluation for at least a proportion of the study population. This allowed us to evaluate the potential alternative biomarkers and undertake a more detailed analysis of the D-dimer measurement to determine whether or not a pregnancy-specific threshold could optimise specificity without compromising sensitivity.

Patient consent was required to take additional blood samples so that the biomarker study could not include women with diagnosed PE identified through UKOSS. We anticipated that only a small number of women with suspected PE would actually have PE. This would provide very little power to estimate sensitivity with any degree of precision. We therefore augmented the sample with pregnant or postpartum women who had DVT diagnosed during the recruitment period at the participating hospitals, thus including all women with diagnosed VTE. There are good pathophysiological reasons for expecting that biomarkers will have the same sensitivity in PE and DVT, and empirical studies of D-dimer measurement have shown similar sensitivity in DVT and PE.[37]

Target population

Suspected pulmonary embolism

Pregnant or postpartum women with suspected PE who consented to participate in the study were asked to provide an additional blood sample. Details of the study population with suspected PE, the inclusion and exclusion criteria, sampling, consent and the data collection procedures are described in *Chapter 3*.

Diagnosed deep-vein thrombosis

We also recruited any women identified with a DVT diagnosis confirmed by imaging (ultrasound, magnetic resonance, CT or contrast venography) who were willing to provide an additional blood sample. The recruitment of these women was limited to hospitals participating in the recruitment of women with suspected PE. We excluded women with symptoms of suspected PE (who should be recruited as having suspected PE), women who had been diagnosed with PE or DVT earlier in the current pregnancy, women who were unable or unwilling to provide informed consent and women aged < 16 years. The screening form and the data collection form are available in *Report Supplementary Material 6* and *7*.

Sample size

The incidence of DVT in pregnancy and post partum is around four times that of PE,[28] so we anticipated that we would recruit around 20 women with DVT. Thus, the sample for the biomarker substudy was expected to include 245 women with suspected PE but negative diagnostic testing, five women with diagnosed PE, and 20 women with diagnosed DVT (i.e. 25 women with confirmed VTE).

Blood sample collection, handling, storage and analysis

Serum and citrate blood samples were collected by a member of the clinical team or research nurse/midwife using venepuncture technique, ideally while obtaining routine blood samples for standard clinical assessment in diagnostic workup. Sample preparation was conducted by the research nurse/midwife or a member of the hospital laboratory staff. The samples were centrifuged at 2000 g for 15 minutes at room temperature within 4 hours of being obtained. Citrate samples were further processed to obtain platelet-free plasma.

Plasma and serum samples were stored in aliquots labelled with the patient identification and the storage box co-ordinates were recorded on paper and electronic study documentation, in accordance with local protocols. The samples were stored in –70 °C freezers at each participating hospital (with the exception of one location in which a –40 °C freezer was used) for the duration of the study, until all samples were transported for analysis to Guy's and St Thomas' NHS Foundation Trust (GSTT), London, UK.

Biomarker analysis

The biomarkers selected for analysis are outlined in *Table 13*.

Analytic techniques

Citrated plasma was utilised for PT, APTT, Clauss fibrinogen, D-dimer (Innovance), D-dimer ELISA, thrombin generation (TG), plasmin–antiplasmin, prothrombin fragment 1 + 2 (PF 1 + 2), and tissue factor. Serum was utilised for troponin-1, BNP, MRProANP, and CRP assays.

Thrombin generation was measured by the Thrombinoscope (ThermoElectron Corporation, Cambridge, UK). The samples were tested in batches to minimise variability. The frozen-plasma aliquots were placed in a water bath (at 37 °C) to thaw for 5 minutes. Platelet-poor plasma (PPP)-reagent LOW (Diagnostica Stago UK Ltd., Theale, UK) was used because of expected hypercoagulability; PPP-reagent LOW consists of 1 pM of tissue factor with 4 µM of phospholipids. A measurement of 20 µl of PPP reagent was added to each TG well together with 80 µl of platelet-free plasma and 20 µl of fluorogenic substrate and calcium. The fluorogenic substrate consisted of amino-methyl-coumarin. The calibrator wells consisted of 80 µl of platelet-free plasma, 20 µl of calibrator and 20 µl of fluorogenic substrate and calcium. All of the reagents were from Diagnostica Stago, Reading, UK. The analysis was conducted in an ELISA plate (Diagnostica Stago, Reading, UK) that enables the thrombin formation to be followed in a Fluoroskan (Fluoroskan Ascent™, Thermo Scientific, Loughborough, UK). The coefficient of variation was 2.2% to 3.2% intra-assay and 5.1% to 16.7% interassay for lag time, endogenous thrombin potential, peak and time to peak.

TABLE 13 Biomarkers selected for analysis

Biomarker	Description	Reference range
D-dimers (ELISA)	A fibrin degradation product – a small protein fragment present in the blood after a blood clot is degraded by fibrinolysis. Measured by ELISA and a highly sensitive assay	0–400 ng/ml
D-dimers (Innovance)	As above, but near patient testing and fast turn around time allows for day- to-day use. This is a point-of-care test that is used by many routine laboratories in the UK in 2016	0–1.13 mg/l
Plasmin–antiplasmin level	An ELISA assay that measures the level of plasmin–antiplasmin complexes and thus is a very sensitive assay of plasmin activation	150–800 µg/l
PF 1 + 2	A small molecule cleaved from prothrombin when thrombin is generated. It is thus a sensitive marker of thrombin generation i.e. coagulation turnover. It is an ELISA assay	200–1200 pmol/l
Thrombin generation	Thrombin generation can be measured dynamically using the endogenous thrombin potential (ETP), a term introduced by Hemker in 1986 that refers to the total amount of thrombin generated during the test. Commonly measured variables when analysing thrombin generation include the lag time, the time to peak thrombin generation, the endogenous thrombin potential (ETP) – the area under the curve	Lag time: 0.9–3.4 minutes ETP: 696–1533 nM × minutes Peak: 103–475 nM Time to peak: 1.4–7.7 minutes
PT	A routine measure of the extrinsic pathway of coagulation, used to determine the clotting tendency of blood	11.7–15.9 seconds
APTT	A routine measure of the intrinsic and common coagulation pathways, used to detect abnormalities in blood clotting	27–52 seconds
Clauss fibrinogen	A functional measure of fibrinogen	2.03–4.11 g/l
Soluble tissue factor	A marker of tissue factor activation – when tissue factor is upregulated, part of the molecule may be cleaved and enters the systemic circulation	40–300 pg/ml
Troponin I	Part of the troponin complex in cardiac muscle tissue, used to detect myocardial damage resulting from myocardial ischaemia or noncardiac causes such as PE	0.91–2.63 ng/ml
B-type natriuretic peptide	A polypeptide secreted by the ventricles of the heart in response to excessive stretching of heart muscle cells, used to measure heart strain resulting from primary heart disease or noncardiac causes such as PE	107–523 pg/ml
C-reactive protein	CRP is an acute-phase protein, the levels of which rise in response to inflammation. Elevation of CRP has been shown to be associated with a diagnosis of PE	0–3104 ng/ml
MRproANP	MRproANP is an emerging measure of right ventricular strain which occurs as a consequence of pulmonary embolism	0–954 pmol/l

PF 1 + 2, prothrombin fragment 1 + 2.

The PT, APTT, and Clauss fibrinogen were measured on the ACL300R (Werfen UK, Warrington, UK) using PT High Sensitivity Plus reagent for the PT, HemosIL APPT Synthetic Phospholipids liquid for the APTT (Werfen UK, Warrington, UK), and fibrinogen C for the Clauss fibrinogen. All reagents were purchased from Werfen UK (Warrington, UK). The tests were measured in accordance with the manufacturer's instructions for the ACL300R analyser. The coefficient of variation was 3.2% to 3.5% intra-assay and 3.6% to 4.2% interassay for PT, APTT and Clauss fibrinogen.

The latex-based D-dimer was measured on the CA660 analyser from Sysmex UK (Milton Keynes, UK). Innovance D-dimer reagent (Sysmex UK, Milton Keynes, UK) was used to measure the D-dimer in accordance with the manufacturer's instructions. The coefficient of variation was 6.0% intra-assay and 12% interassay for the Innovance D-dimer.

The Zymutest D-dimer ELISA assay (Quadratech Diagnostics Ltd, Epsom, UK) was used to measure the D-dimers by ELISA. D-dimer was measured in accordance with the manufacturer's instructions. The coefficient of variation was 4.6% intra-assay and 10.8% interassay for the Zymutest D-dimer.

The plasmin–antiplasmin ELISA (Immunodiagnostics Systems Ltd, Boldon Colliery, UK) was used to measure the plasmin–antiplasmin. The assay was measured in accordance with the manufacturer's instructions. The coefficient of variation was 4.2% intra-assay and 7.3% interassay for the plasmin–antiplasmin.

The fragment 1 + 2 Micro (Sysmex, Milton Keynes, UK) was used to measure the PF 1 + 2 in the citrated plasmas. The PF 1 + 2 was measured in accordance with the manufacturer's instructions. The coefficient of variation was 6.0% intra-assay and 9.0% interassay for the PF 1 + 2.

The Immubind tissue factor (Invitech Ltd, Huntingdon, UK), was used to measure the tissue factor by ELISA. The tissue factor was measured in accordance with the manufacturer's instructions. The coefficient of variation was 6.0% intra-assay and 5.0% interassay for the tissue factor.

The troponin-1 type 3 ELISA (Bio Techne, Abingdon-on-Thames, UK) was used to measure the troponin-1 levels. The troponin-1 levels were measured in accordance with the manufacturer's instructions. The coefficient of variation was 4.0% intra-assay and 4.6% interassay for the troponin-1 levels.

The BNP ELISA (Bio Techne, Abingdon-on-Thames, UK) was used to measure the BNP levels. The BNP levels were measured in accordance with the manufacturer's instructions. The coefficient of variation was 10% intra-assay and 15% interassay for the BNP assay.

The human midregional pro-atrial natriuretic peptide ELISA (2B Scientific, Upper Heyford, UK) was used to measure the MRProANP levels. The MRProANP levels were measured in accordance with the manufacturer's instructions. The coefficient of variation was 8% intra-assay and 10% interassay for the MRProANP levels.

The human CRP Quantikine assay (Bio Techne, Abingdon-on-Thames, UK) was used to measure the CRP levels. The CRP levels were measured in accordance with the manufacturer's instructions. The coefficient of variation was 5.5% intra-assay and 6.5% interassay for the CRP assay.

Statistical analysis

The biomarker analysis included all women with suspected PE or diagnosed DVT who provided consent and an analysable blood sample. Women with suspected PE were classified as having VTE if they were classified as having PE in accordance with the method described in the case–control study. The primary and secondary analyses were planned along the lines outlined in the case–control study. Women recruited with diagnosed DVT were all classified as having VTE and included in the primary analysis.

Blood samples were analysed at GSTT and the results of the analysis were sent to the Sheffield CTRU. GSTT established normal ranges for the assays using 20 normal plasma/serum samples (depending on the assay), with the 99th percentile used as the top of the normal range.

We calculated the AUC for each biomarker and the sensitivity and specificity at the upper limit of the normal range. We then examined the ROC curve to determine whether or not there was an optimal threshold for clinical practice, whereby sensitivity exceeds 95% but specificity still allows a meaningful proportion of women without PE to have the diagnosis ruled out.

Results

The characteristics of the 324 recruited women with suspected PE are described in *Chapter 3*. Blood samples were taken from 312 out of 324 women. The reasons for failure to take a blood sample were inability to draw blood ($n = 7$), patient refused ($n = 2$), unavailability of blood-handling services ($n = 1$), patient discharged before venepuncture ($n = 1$) and unknown ($n = 1$). Two samples were not labelled correctly (one from a woman with PE clinically ruled out without imaging and one with PE ruled out by negative imaging) and were therefore not analysed, leaving 310 samples for analysis.

We recruited 18 women with diagnosed DVT, nine of whom were recruited at Guy's and St Thomas' Hospital Maternity Unit (a specialist centre); the remaining nine women were recruited from Leeds Teaching Hospitals Maternity Unit ($n = 3$), Queen Alexandra Hospital Portsmouth Maternity Unit ($n = 3$), Royal Berkshire Hospital Maternity Unit ($n = 2$) and Bradford Royal Infirmary Maternity Unit ($n = 1$). A further six were eligible for recruitment but declined to participate.

The women with diagnosed DVT had a mean age of 28.3 years, a mean BMI of 26.3 kg/m^2 and were from the following ethnic groups: white British ($n = 9$), black African ($n = 2$), black Caribbean ($n = 1$), Asian Pakistani ($n = 1$), mixed white and black African ($n = 1$), mixed white and black Caribbean ($n = 1$), other white ($n = 1$) and other ethnic group ($n = 2$). They included seven married, three single and eight cohabiting women; 11 women were employed and seven were unemployed, and by gestational age, one women was in the first trimester, one women was in the second trimester, nine women were in the third trimester and seven women were post partum.

Adding the 18 samples from women with DVT to the 310 samples from women with suspected PE gave 328 samples for analysis. The 310 women recruited with suspected PE consisted of 18 women with PE confirmed by imaging (including one women with subsegmental PE), five women with clinically diagnosed PE (three with equivocal imaging and two with no imaging; all treated), 247 women with PE ruled out after imaging (242 women with negative imaging and five untreated after equivocal imaging) and 40 women with PE clinically ruled out without imaging (none treated). Thus, 36 women with VTE and 247 women without VTE were included in the primary analysis; 41 women with VTE and 247 women without VTE were included in the secondary analysis including clinically diagnosed PE; 36 women with VTE and 287 women without VTE were included in the secondary analysis including clinically ruled-out PE; and 35 women with PE and 247 women without PE were included in the secondary analysis excluding subsegmental PE.

Table 14 compares the mean biomarker levels between women with and without VTE in the primary analysis. D-dimer (both assays), TG (lag time and time to peak), Clauss fibrinogen and plasmin–antiplasmin had significantly higher mean levels in women with VTE than in women without VTE. The mean levels of the other biomarkers did not significantly differ between the groups.

TABLE 14 Mean (standard deviation) biomarker levels for the patient groups in the primary analysis

Biomarker	Mean biomaker level (SD) in women with		*p*-value
	No VTE (*N* = 247)	VTE (*N* = 36)	
APTT (minutes)	39.7 (22.1)	41.4 (13.2)	0.660
Clauss fibrinogen	5.37 (1.69)	6.30 (2.73)	0.007
CRP level (pg/ml)	5348 (1705)	5603 (1646)	0.401
PT (minutes)	16.2 (5.4)	18.7 (13.2)	0.089
D-dimer (ELISA)	1247 (1474)	2401 (2642)	0.001
D-dimer (Innovance)	1.147 (1.269)	2.282 (3.388)	0.004
TG (lag time)	8.70 (4.84)	13.85 (8.30)	< 0.001
TG (endogenous potential)	1217 (558)	1081 (561)	0.241
TG (time to peak)	14.8 (9.1)	21.5 (13.1)	0.001
TG (peak)	162 (116)	130 (124)	0.160
Plasmin–antiplasmin level	688 (251)	915 (647)	0.004
BNP level	372 (900)	385 (731)	0.932
MRproANP	603 (1016)	753 (1159)	0.415
Tissue factor (pg/ml)	291 (320)	488 (1067)	0.065
PF 1 + 2 (pmol/l)	623 (408)	550 (333)	0.298
Troponin level (ng/ml)	1.328 (2.458)	0.762 (0.968)	0.105

SD, standard deviation.

Appendix 11 provides further details for each biomarker, with a box-and-whisker plot showing the distribution for women with DVT, women with PE and women with no PE, and those who were excluded from the analysis. The distributions of all biomarkers overlapped substantially between women with and without VTE.

Figures 5–7 show the ROC curves for the D-dimer, APTT, PF 1 + 2, PT and TG biomarkers, and for the other biomarkers. It was not possible to identify a threshold for any biomarker that would optimise sensitivity (> 98%) while maintaining meaningful specificity.

Table 15 reports the AUROC for the continuous biomarker and diagnostic parameters for the biomarkers at the predefined threshold for positivity and the threshold that optimised sensitivity (> 95%) at the expense of specificity. No biomarker had sufficient sensitivity to rule out VTE while achieving meaningful specificity, with the possible exception of TG (lag time), with an AUC of 0.702, a sensitivity of 97% and a specificity of 25% at the threshold that optimised sensitivity at the expense of specificity.

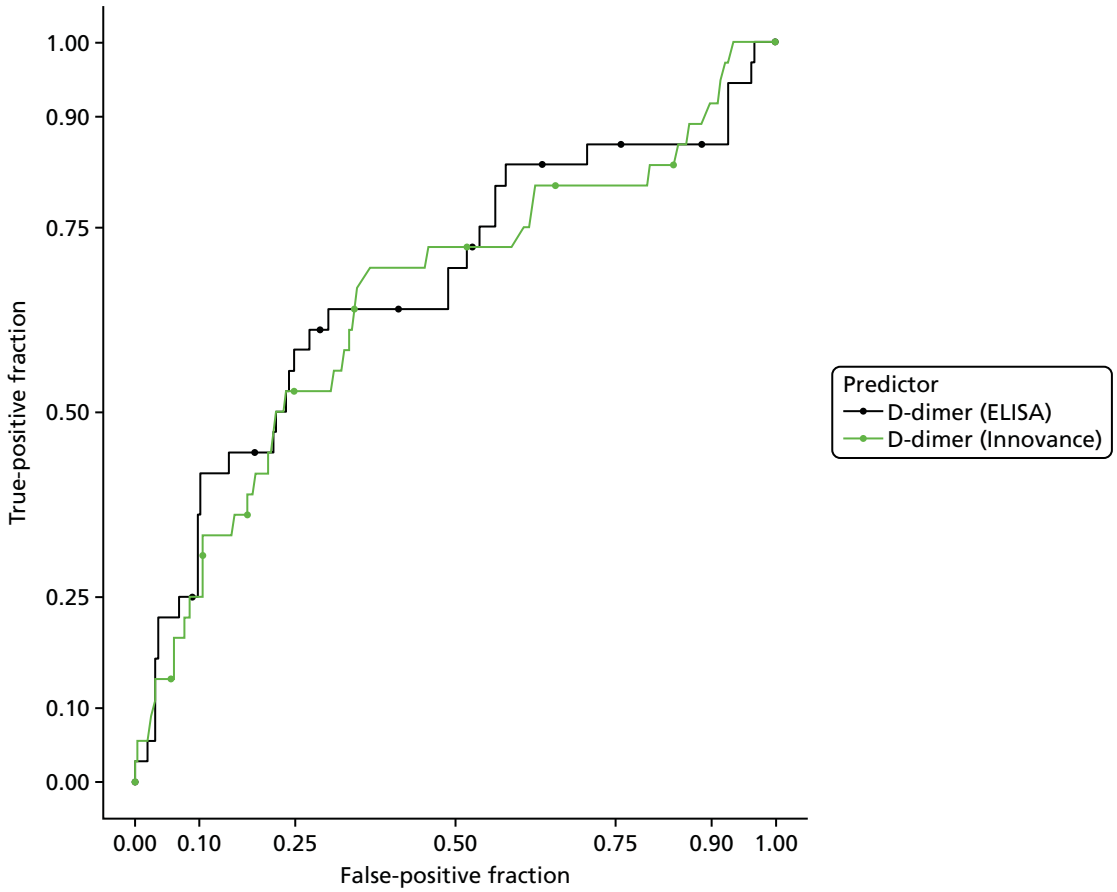


FIGURE 5 Receiver operating characteristic curves for D-dimer biomarkers. Adapted from Hunt *et al.*[56] This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License (CC BY-NC 4.0), which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

Analysis excluding women who had received anticoagulation treatment

Anticoagulation treatment with heparin is known to interfere with biomarker assays. Unfractionated heparin will prolong the APTT and thrombin time and low-molecular-weight heparin (LMWH) may cause a slight prolongation of the APTT. However, suppressing the activation of both Factor Xa and thrombin will decrease all parameters of the TG assay and PF 1 + 2. Furthermore, decreasing the generation of thrombin, which is a major stimulator of fibrinolysis, will reduce the D-dimer and plasmin–antiplasmin values.

We repeated the analysis, having excluded 240 out of 328 women who had received anticoagulation treatment prior to blood sampling. The primary analysis involved only 66 women, of whom only four had VTE, so the findings were limited by small numbers. Details of the findings are provided in *Appendix 11*. The differences in mean biomarker levels observed in the main analysis between women with and without PE disappeared or even reversed when those receiving anticoagulation treatment were removed, but this probably reflects the small numbers. ROC analysis suggested that BNP (AUC 0.774, 95% CI 0.670 to 0.878), PF 1 + 2 (0.795 pmol/l, 95% CI 0.644 to 0.947 pmol/l), TG lag time (0.735, 95% CI 0.531 to 0.940) and troponin (0.742 ng/ml, 95% CI 0.453 to 1.000 ng/ml) may have some potential to rule out VTE with acceptable sensitivity, but the CIs were wide and estimates would need to be validated in a larger cohort with VTE.

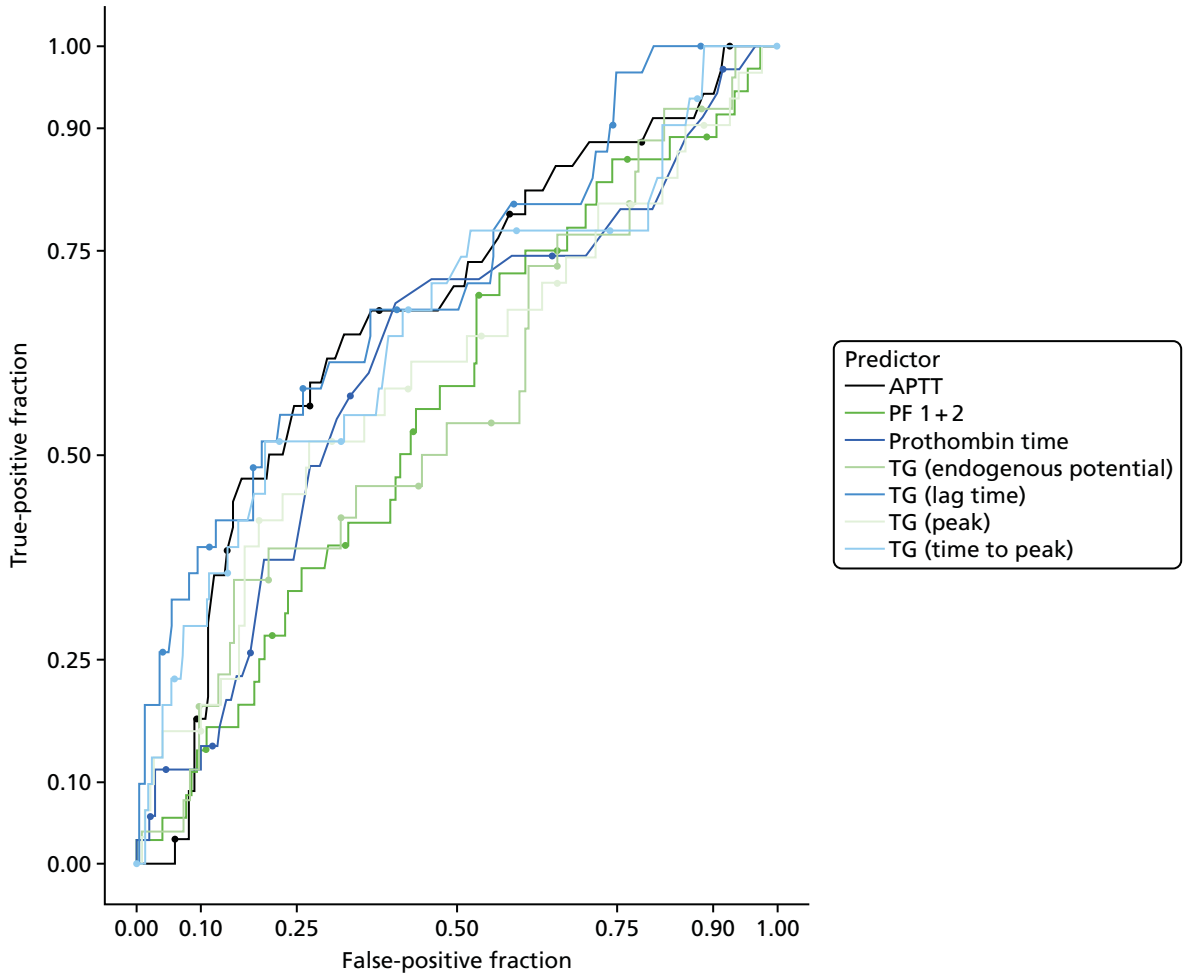


FIGURE 6 Receiver operating characteristic curves for APTT, PF 1 + 2, PT and TG biomarkers. Adapted from Hunt *et al.*[56] This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License (CC BY-NC 4.0), which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

Discussion

We were unable to identify any biomarker that would provide clinically useful discrimination between women with and without VTE. In the analysis involving all patients, Clauss fibrinogen, both D-dimer assays, TG (lag time and time to peak) and plasmin–antiplasmin had significantly higher levels in those with VTE than in those without. Other biomarker levels did not significantly differ between women with and without VTE. The only biomarker with an AUC that was > 0.7 was TG (lag time), with an AUC of 0.702. The ROC curves showed that there was no threshold for sensitivity and specificity that would be useful for clinical decision-making, with the possible exception of TG (lag time), with a sensitivity of 97% and a specificity of 25% at the threshold that optimised sensitivity at the expense of specificity.

Most of the women in the analysis (240/330) had received anticoagulation treatment prior to blood sampling. As a consequence, most of the coagulation biomarkers were affected because all forms of heparin used to treat these women (unfractionated and LMWH) suppress coagulation activation, and thus TG and PF 1 + 2 are suppressed and APTT and thrombin time are prolonged. Furthermore, thrombin is a major stimulator of fibrinolysis and, therefore, when thrombin is reduced, there is less increment in plasmin–antiplasmin and D-dimer.

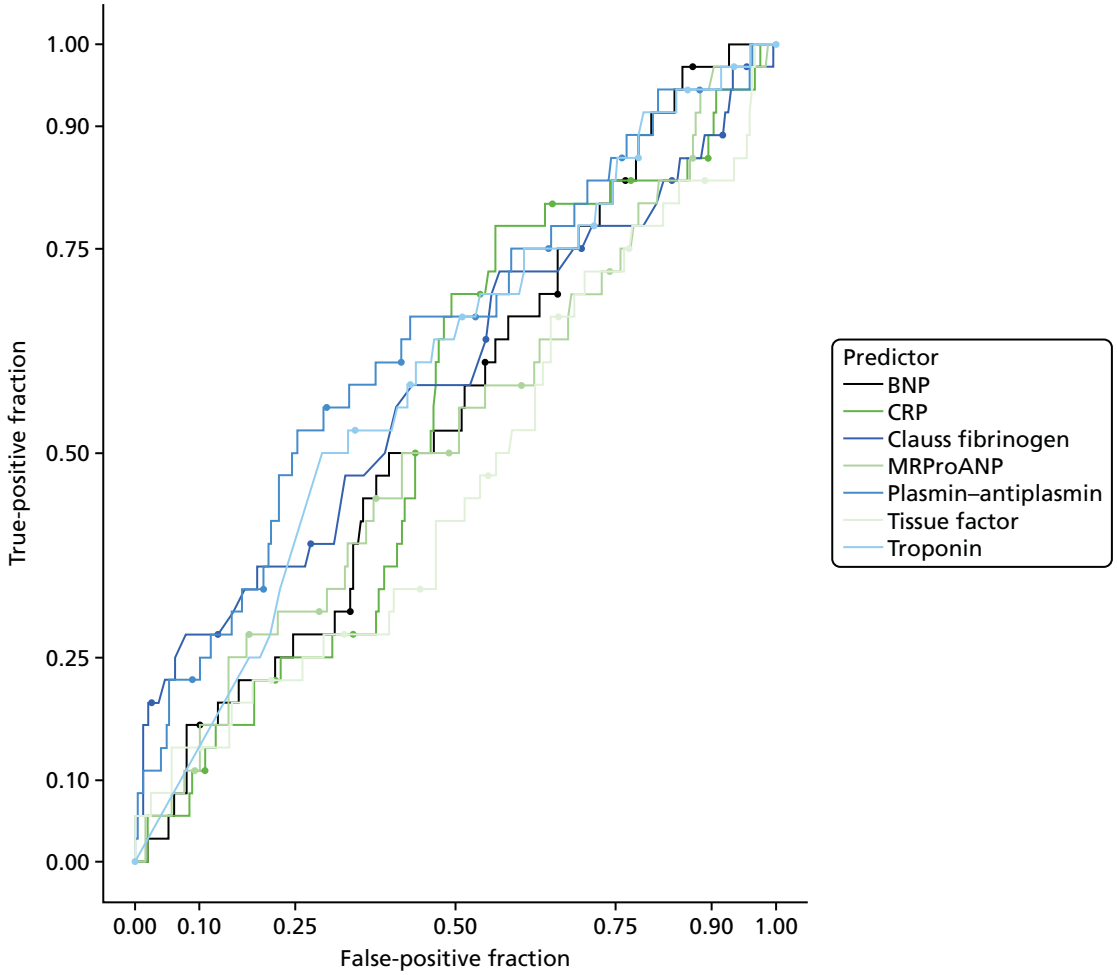


FIGURE 7 Receiver operating characteristic curves for the other biomarkers. Adapted from Hunt *et al.*[56] This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License (CC BY-NC 4.0), which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

To address this issue, we repeated the analysis, excluding those who had received anticoagulation treatment. Unfortunately, this reduced the sample size markedly and included only four women with VTE. No biomarker showed any association with VTE, although this may reflect a lack of statistical power. BNP, PF 1 + 2, TG (lag time) and troponin may have some potential to rule out VTE with acceptable sensitivity, but the CIs are wide, being based on only four women with VTE. Further validation is therefore required in a larger cohort of women with VTE.

We selected the biomarkers for analysis on the basis of previous evidence suggesting that they may be diagnostically useful. Outside pregnancy, within secondary care, the D-dimer measurement has been validated as a useful biomarker to aid in the diagnosis of PE. Indeed, it is chiefly used for its negative predictive value in combination with a low Wells's score to exclude PE.[57] The previous data on the use of D-dimer in pregnancy are of low quality, but some authors have suggested that D-dimer is increased in women with PE during pregnancy.[58] However, there are more substantial data showing that, in normal pregnant women, D-dimer values increase continuously during pregnancy across all gestation periods and that the 'normal range' outside pregnancy cannot be applied to pregnant women.[21,59] Hedengran *et al.*[59] also showed that the D-dimer values in individual healthy pregnant women fluctuated by > 50%, and thus concluded that they may not be of value in the diagnosis of VTE during pregnancy.

TABLE 15 The AUROC, sensitivity and specificity for each biomarker

Biomarker	AUC (95% CI)	At the predefined threshold, % (95% CI)		At the threshold with optimal sensitivity, % (95% CI)	
		Sensitivity	Specificity	Sensitivity	Specificity
APTT (minutes)	0.669 (0.570 to 0.768)	0.088 (0.019 to 0.237)	0.914 (0.870 to 0.947)	0.971 (0.847 to 0.999)	0.086 (0.053 to 0.130)
BNP level	0.549 (0.453 to 0.645)	0.167 (0.064 to 0.328)	0.879 (0.831 to 0.917)	0.972 (0.855 to 0.999)	0.146 (0.104 to 0.196)
CRP level (pg/ml)	0.542 (0.445 to 0.639)	0.861 (0.705 to 0.953)	0.121 (0.083 to 0.169)	0.972 (0.855 to 0.999)	0.032 (0.014 to 0.063)
Clauss fibrinogen	0.589 (0.476 to 0.701)	0.778 (0.608 to 0.899)	0.228 (0.177 to 0.286)	0.972 (0.855 to 0.999)	0.066 (0.038 to 0.106)
D-dimer (ELISA)	0.668 (0.561 to 0.776)	0.861 (0.705 to 0.953)	0.196 (0.148 to 0.251)	0.972 (0.855 to 0.999)	0.037 (0.017 to 0.069)
D-dimer (Innovance)	0.651 (0.545 to 0.758)	0.528 (0.355 to 0.696)	0.727 (0.666 to 0.781)	0.972 (0.855 to 0.999)	0.078 (0.047 to 0.118)
MRproANP	0.524 (0.418 to 0.630)	0.278 (0.142 to 0.452)	0.785 (0.729 to 0.835)	0.972 (0.855 to 0.999)	0.097 (0.063 to 0.141)
PF 1 + 2 (pmol/l)	0.562 (0.462 to 0.661)	0.056 (0.007 to 0.187)	0.935 (0.896 to 0.962)	0.972 (0.855 to 0.999)	0.045 (0.023 to 0.079)
Plasmin–antiplasmin level	0.639 (0.536 to 0.742)	0.472 (0.304 to 0.645)	0.763 (0.705 to 0.815)	0.972 (0.855 to 0.999)	0.041 (0.020 to 0.074)
PT (minutes)	0.613 (0.508 to 0.718)	0.486 (0.314 to 0.660)	0.730 (0.669 to 0.785)	0.971 (0.851 to 0.999)	0.084 (0.052 to 0.127)
TG (lag time)	0.702 (0.598 to 0.806)	1.000 (0.888 to 1.000[a])	0.000 (0.000 to 0.017[a])	0.968 (0.833 to 0.999)	0.251 (0.195 to 0.314)
TG (endogenous potential)	0.559 (0.437 to 0.681)	0.231 (0.090 to 0.436)	0.706 (0.638 to 0.767)	0.962 (0.804 to 0.999)	0.069 (0.038 to 0.112)
TG (peak)	0.596 (0.478 to 0.715)	0.000 (0.000 to 0.097)	0.996 (0.977 to 1.000)	0.968 (0.833 to 0.999)	0.059 (0.032 to 0.099)
TG (time to peak)	0.655 (0.541 to 0.769)	1.000 (0.888 to 1.000)	0.110 (0.071 to 0.159)	1.000 (0.888 to 1.000)	0.114 (0.075 to 0.164)
Tissue factor (pg/ml)	0.531 (0.424 to 0.638)	0.222 (0.101 to 0.392)	0.771 (0.714 to 0.822)	0.972 (0.855 to 0.999)	0.037 (0.017 to 0.069)
Troponin level (ng/ml)	0.597 (0.499 to 0.695)	0.056 (0.007 to 0.187)	0.887 (0.840 to 0.923)	0.972 (0.855 to 0.999)	0.085 (0.053 to 0.127)

a All participants had abnormally high lag time, giving a sensitivity of 1 and a specificity of 0. CIs are one-sided 97.5% binomial intervals.

During clot formation there is activation of coagulation and, therefore, we did look at markers that measured this, which measured simple coagulation tests, PF 1 + 2 and TG. In actuality, these have been poorly studied in diagnosing VTE outside pregnancy because the D-dimer assay has long been well established, can be done cheaply and, most importantly, is validated in the non-pregnant setting. Whether or not any of these markers might have an ancillary role in the diagnosis of VTE both inside and outside pregnancy remains uncertain because of the effects of anticoagulation.

This analysis has a number of limitations. We were unable to obtain blood samples from the UKOSS cohort with diagnosed PE and so we had to supplement the anticipated low prevalence of PE among women with suspected PE by including women with diagnosed DVT. Women with DVT are likely to have a lower thrombotic load than women with PE and are less likely to have cardiac strain, so biomarkers may be less sensitive for DVT than for PE. Most of the blood samples were taken after anticoagulation treatment was given, which we considered (especially from the effect on TG) to have interfered with the biomarker assays and reduced their diagnostic value. This is an inevitable consequence of the current guidance[33] stating that patients with suspected PE should be given anticoagulation treatment while awaiting diagnostic testing if any delay is anticipated. We repeated the analysis, having excluded women who had received anticoagulation treatment, but this resulted in a small sample size with little statistical power to draw reliable conclusions.

Conclusion

Our analysis suggests that there are currently no biomarkers that can be recommended for clinical use as a way of selecting women with suspected PE in pregnancy or post partum for imaging. The findings for D-dimer in particular suggest that it should not be recommended for use in the diagnostic work-up of PE in pregnancy.

Future research would ideally test biomarkers on a large cohort including a substantial number of women with VTE and would involve blood-sampling before anticoagulation treatment is given. This will be very difficult to achieve, owing to the need to give anticoagulation treatment as soon as the suspicion of VTE is raised. Unless consent is obtained very quickly from such women, it would not be ethical to pursue this. Our study is reflective of this, as we were able to recruit only four women with VTE who had not received anticoagulation treatment before blood sampling, despite recruiting from 11 sites over 18 months.

Chapter 5 Decision-analysis modelling

Introduction

An economic model evaluated the cost-effectiveness of using a CDR in pregnant women with suspected PE. In the model, 10 strategies were tested: (1) scanning all pregnant women with suspected PE (current recommended care), (2–4) applying the three expert-derived clinical consensus decision rules (primary, sensitive and specific), (5) applying a permissive interpretation of Wells's decision rule (Wells's permissive), (6) applying a strict interpretation of Wells's decision rule (Wells's strict), (7) applying the PERC decision rule, (8) applying the simplified Geneva decision rule, (9) scanning no women, but treating all (SNTA) and (10) scanning no women and treating no women (SNTN). The sensitivity and specificity of these strategies were estimated using the primary data. Sensitivities and specificities of strategies based on CDRs were estimated from the women with PE confirmed or ruled out in the case–control study primary analysis. In all of the strategies involving a decision rule, women with a positive decision rule result received a scan and those with a negative result did not. If a woman's scan result was positive, they then received anticoagulation treatment. Strategies 9 and 10 are included as there is little evidence supporting whether or not the benefits of current care outweigh the risks of exposing pregnant/postpartum women to radiation.

Literature review

Study identification

The searches were conducted in MEDLINE, MEDLINE In process and NHS Economic Evaluation Database on the 25 August 2016. The disease-specific search terms were the same as those used in the DiPEP literature review.[7] The Scottish Intercollegiate Guidelines Network economic search filters were added to these search terms. Cost-effectiveness studies on CDRs were identified using the criteria given in *Appendix 12*.

Results

No studies were found that estimated the cost-effectiveness of the use of selective imaging for the diagnosis of PE in pregnant or postpartum women. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram is given in *Appendix 12*. The key finding is that no economic study has previously been conducted in a population of women who were pregnant or post partum.

Methods

Decision problem and perspective

The economic modelling assessed whether or not using CDRs would be cost-effective compared with scanning all (current care) and scanning none for pregnant or postpartum women who had a suspected PE in the UK. In line with NICE guidance, the analyses took a NHS and personal social services perspective and a lifetime horizon, and future costs and QALYs were discounted at a rate of 3.5% per annum.[60]

Model description

A patient-level decision tree was developed in Microsoft Excel 2016 to estimate the cost-effectiveness of the 10 strategies for pregnant or postpartum women with a suspected PE at admission. A decision tree is believed to be an appropriate model for this decision problem, as events do not repeatedly occur over time, representing the likely clinical reality for pregnant/postpartum women with suspected PE. The structure also reflects the fact that the model is estimating the impact of a single decision at a single point in time. A patient-level model was chosen over a cohort model, as this allowed the model to include the

fact that women who were more likely to die if untreated were also more likely to have a positive CDR and hence be referred for a scan.

Patient population

The population included in the economic model was pregnant or post partum (up to 6 weeks after birth) women who presented with a suspected PE at a UK hospital. The fetuses were considered to be outside the scope of the model, apart from any imaging-induced childhood cancers.

The primary statistical classification was used to determine whether or not women had PE. In this analysis group, the population was limited to women with PE diagnosed by imaging or post-mortem examination and women with PE ruled out after imaging. After statistical imputation in the suspected PE cohort, 13 women with PE and 248 without PE had complete data for all decision rules. In line with the statistical analysis of rule sensitivity and specificity, these data were supplemented by 144 women with diagnosed PE and from the UKOSS observation data set.

Estimation of outcomes across the populations

The model estimates the costs and QALYs for each strategy in the PE and no PE populations separately. The costs and QALYs for each strategy in the population with a suspected PE were calculated using the following formula:

$$\text{Mean outcome pop} = \text{mean outcome PE} \times \text{probability of having PE} + \text{mean outcome no PE} \times (1 - \text{probability of having PE}), \quad (1)$$

in which the mean outcomes are derived from the model and the probability of having PE is obtained from the suspected PE data set only.

Determining the number of patients

The model uses a bootstrapping procedure to estimate the lifetime costs and QALYs for each strategy. Within each bootstrap, 157 women with PE and 248 women without PE are sampled from their respective populations with replacement. It was determined that 100 bootstraps were sufficient to produce robust results. Details of how this was determined are provided in *Appendix 13*.

Model structure

The model was structured as an individual-level decision tree; the key aspects of the model are presented in *Figure 8*. The key events included in the decision tree were PE-related death, major bleeding events, deaths associated with major bleeding events, chronic thromboembolic pulmonary hypertension (CTEPH), recurrent VTE and deaths from recurrent VTE.

In the base case, it was assumed that recurrent VTE and CTEPH were modifiable by anticoagulation in pregnant/postpartum women with a suspected PE. Furthermore, it was assumed that the risk of major bleeding events was assumed to be solely attributable to anticoagulation. If a woman experienced a major bleeding event following an initial anticoagulation treatment, then it was assumed that they would not receive anticoagulation for any subsequent VTEs or gain a therapeutic benefit from their current anticoagulation treatment. Minor bleeding events were excluded from the economic model, as the definition of major bleeding events included all clinically overt bleeds that could have resulted in a hospitalisation. Finally, it was assumed that each woman was at risk of only one recurrent VTE, which was assumed on the grounds that the PE for these women was related to their pregnancy rather than other underlying conditions, which would typically be present in the population that presents with a VTE.

For all women who were scanned, a QALY decrement and treatment cost were applied for imaging-induced lung and breast cancers in the mother, and a QALY decrement was applied for childhood cancers induced by imaging in the fetuses who would survive anticoagulation to term. These were identified as the key harms

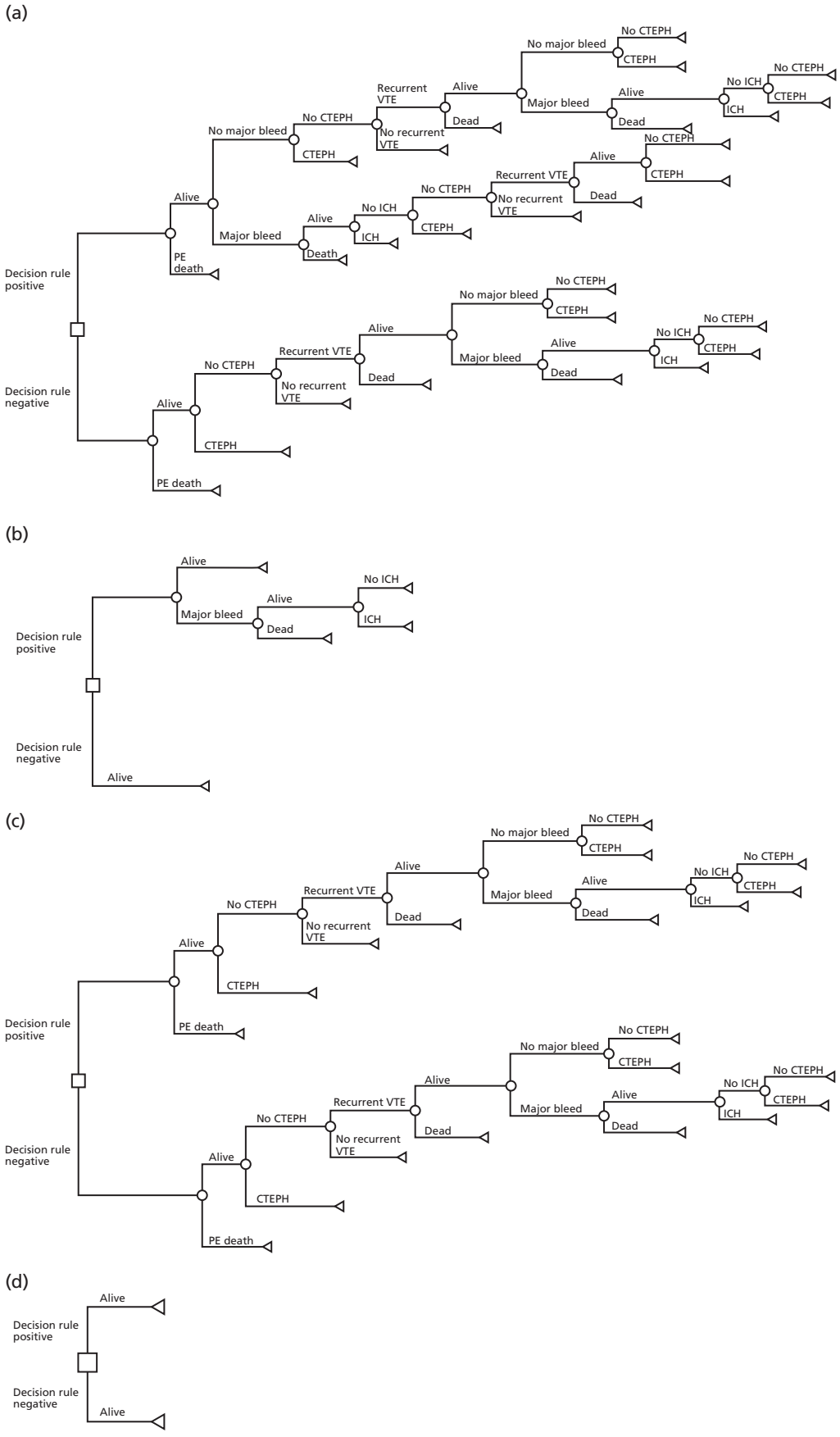


FIGURE 8 The key aspects of the economic model: (a) true-positive scan; (b) false-positive scan; (c) false-negative scan; and (d) true-negative scan. CTEPH, chronic thromboembolic pulmonary hypertension; ICH, intracranial haemorrhage.

from scanning. We did not model any cost or QALY loss associated with teratogenesis, as evidence suggests that this is not a significant risk.[61]

Model data

Clinical parameters

The clinical parameters in the model included the CDRs, the sensitivity and specificity of scanning, the risk of death after PE, the risk of major bleeding, the risk of recurrent VTE and the risk of CTEPH. A summary of all clinical parameters used in the base case are given in *Appendices 15–17*.

The clinical decision rules

Seven CDRs were modelled in line with the statistical analysis. Details are provided on these rules in *Chapter 2* and *Appendix 2*. In brief, all decision rules assigned a score to clinically identifiable factors for women with a suspected PE at presentation. This score was compared to a cut-off value that was specific to each decision rule. If the woman scored above the cut-off value, they were sent to be scanned and if they scored lower than the cut-off value, they did not receive scanning.

Sensitivity and specificity of scanning

In the base case, it was assumed that all scanning techniques allowed a perfect diagnosis of PE. This assumption was made based on the clinical advice that it would not be possible to know which scan results were false positive or false negative from the DiPEP study, but these were unlikely to be random events and were instead likely to be linked to the woman's characteristics and the severity of the PE. Therefore, it was deemed to be clinically implausible that the sensitivity and specificity of the scan results would be independent of the woman's decision rule score. A scenario analysis was conducted, in which it was assumed that the sensitivity and specificity of the scans were independent of the CDRs. In this scenario analysis, the sensitivity and specificity of a CTPA scan was taken from Ohno *et al.*[62] [magnetic resonance angiography with sensitivity encoding (SENSE) for suspected PE: comparison with multidetector computed tomography and VQ scintigraphy] and the sensitivity and specificity of a VQ scan was taken from Gutte *et al.*,[63] [comparison of VQ single photon emission computed tomography (VQ SPECT) and planar VQ lung scintigraphy in diagnosing acute PE], as these values were used in the cost-effectiveness analysis supporting the NICE guidelines on the diagnosis of PE.[33]

The risk of 30-day mortality following a pulmonary embolism

Aujesky *et al.*[64] conducted a cohort study of adults (aged ≥ 18 years old) diagnosed with PE between January 2000 and November 2002 in the USA. No restriction was made to limit the study population to women who were either pregnant or in the postpartum period. The study assessed the risk factors associated with 30-day all-cause mortality for people diagnosed with PE. However, the study could not be used to estimate the risk of 30-day all-cause mortality in the model for three reasons. First, the constant parameter was not presented in the paper, meaning that the mortality risk could not be estimated. Second, it was unclear whether or not a multivariate logistic regression had been fitted or rather a series of univariate analyses. Finally, the variance–covariance matrix was not given, meaning that the parameters could not be correlated if they were from a multivariate analysis.

To overcome these limitations in the literature, an expert elicitation exercise was conducted. Data were extracted for each woman with PE in the primary analysis on their heart rate, respiratory rate, oxygen saturation, temperature, blood pressure, whether they were post partum or pregnant and, if they were pregnant, how many weeks into the pregnancy they were. Risk factors that were not modifiable by treatment, such as age and cancer status, should be excluded from the elicitation exercise on the advice of clinical experts. Data were also presented to the experts on the risk of death in the UKOSS cohort (2.82%, 95% CI 0.92% to 6.47%), to ensure that each expert did not believe that the average elicited risk of death was implausible to them. These data were presented to the four experts and they were asked to estimate the risk of death for each woman if they had received anticoagulation treatment. Based on these data,

four experts in the project management group (SG, GF, FL, CNP) were asked to estimate the probability of death within 30 days for each woman if they had received anticoagulation treatment.

To combine the expert's answers, the average value across all four expert answers for each woman's risk of death was taken. If more than one expert believed that they could not estimate a probability of death for a woman given her characteristics, then these data were treated as missing.

These data were analysed into the model by preforming a beta regression.[65] The risk of death from PE within 30 days was the predicted variable. All variables presented to the experts were included as explanatory variables. Scenario analyses were conducted in the secondary, tertiary and quaternary statistical analysis populations. Further scenario analyses were conducted in which the answers for each expert individually were used in the primary population data set. The results of the beta regression and the scenario analyses are presented in *Appendix 15*.

The risk of major bleeding and the split of bleeding types

The probability of a major bleeding event occurring was obtained from Carrier *et al.*[66] Carrier *et al.*[66] was a systematic review of case fatality rates of recurrent VTE and major bleeding events among patients treated for an initial VTE. Studies published between 1950 and September 2008 were identified.[66] For people who received anticoagulation treatment for 3 months, the probability of a major bleeding event was 1.8% (95% CI 1.1% to 2.6%) and the probability of a fatal major bleeding event was 0.2% (95% CI 0.1% to 0.4%). For people who received anticoagulation treatment for 6 months, the probability of a major bleeding event was 2.1% (95% CI 1.5% to 2.7%) and the probability of a fatal major bleeding event was 0.6% (95% CI 0.01% to 2.5%).

The proportion of the different types of major bleeding events was taken from Ensor *et al.*,[67] which reported this parameter for a population that had a recurrent VTE in a large European data set ($n = 15{,}041$).[67] The split of major bleeding events in this database was 499 gastrointestinal bleeds, 245 intracranial haemorrhages and 622 other bleeds. These data were used to estimate the split of bleeding events in the base case.

The risk of recurrent venous thromboembolism

The probability of experiencing a recurrent VTE while on anticoagulation treatment was obtained from Carrier *et al.*[66] For people with an initial PE, the risk of a recurrent VTE while on anticoagulation treatment was 3.6% (95% CI 2.3 to 5.0) for 3 months of anticoagulation treatment and 4.9% (95% CI 1.5 to 10.15) for 6 months of anticoagulation treatment.

The risk of death following a recurrent venous thromboembolism

The probability of a recurrent VTE being fatal while a person is on anticoagulation treatment was obtained from Carrier *et al.*[66] For people with an initial PE, the case fatality rate for a recurrent VTE was 30.1% (95% CI 12.3 to 51.8) for 3 months of anticoagulation treatment and 20.6% (95% CI 8.9 to 35.5) for 6 months of anticoagulation treatment. A sensitivity analysis was conducted in which the case fatality rate for a recurrent VTE was set to 0, as the case fatality rate will reflect some underlying comorbidities that are present in an older population, but are unlikely to be present in pregnant women.

The risk of chronic thromboembolic pulmonary hypertension

The risk of CTEPH following PE has been estimated to vary widely across studies, with studies estimating the cumulative risk of CTEPH following PE to be 0.4–9.1%.[68] The risk of CTEPH was estimated in the base case by assuming that the probability of CTEPH was the same as the risk of death from PE obtained from the expert elicitation exercise. This approach was preferred, as it meant that the risk of CTEPH was proportional to the likely size of the embolus. A scenario analysis was conducted in which the risk of CTEPH after PE was estimated to be 0.5% (4 out of 866 patients in the cohort were diagnosed with CTEPH) based on the data presented in Klok *et al.*[69] This study was chosen for the scenario analysis as, out of the studies referenced in Lang *et al.*,[68] it had adequate numbers of patients with PE ($n = 866$) who were relatively young (mean age 56 years),

a high proportion of women (52.7%), a relatively long follow-up period of 34 months, and patients that were recruited between 2001 and 2007 (which is relatively recent compared with the other studies).

Adjustments if a women with a pulmonary embolism did not receive anticoagulation treatment

Barrit and Jordan[70] conducted a small randomised controlled trial in one hospital that showed that the odds ratios for untreated people with PE compared with people treated with anticoagulation treatment were 6.923 (95% CI 0.734 to 65.259) for all-cause mortality (6/19 vs. 1/16), 16.667 (95% CI 1.818 to 152.770) for recurrent VTE or death (1/16 vs. 10/19) and 12.517 (95% CI 0.636 to 246.384) for odds ratio for fatal PE (0/16 vs. 5/19). For people who were untreated, their risk of death from PE was modified using the odds ratio estimated for the risk of death from a fatal PE, their risk of a recurrent VTE was modified using the odds ratio estimated for the risk of recurrent VTE or death and their risk of a CTEPH event was modified using the odds ratio estimated for the risk of recurrent VTE or death. The assumption that the risk of CTEPH was modifiable by anticoagulation was made on the basis of the expert clinical input of the four clinicians who took part in the expert elicitation exercise. For women who had their anticoagulation treatment discontinued as a result of bleeding events, it was assumed that they received no benefit from their initial anticoagulation treatment, as this is the worst-case scenario.

The percentage of fetuses that survive to term

The percentage of fetuses that survived to term was estimated using the data on the number of fetuses that were either terminated, miscarried, died or were a still birth at the 30-day follow-up point for women who were pregnant and did not die from PE. The percentage of fetuses that survived to term was estimated separately for women who had a confirmed PE and women who had a suspected PE, but not an actual PE. For the surviving women with PE, 96.5% of fetuses (167/173) survived until the 30-day follow-up point. For the surviving women with suspected PE, 100% of fetuses (258/258) survived until the 30-day follow-up point.

Life expectancies

All-cause life expectancy

Women who had no adverse events or survived their adverse events and did not end up having a disabling bleed or a CTEPH were assumed to have a normal life expectancy for their age at diagnosis. This age-specific life expectancy was calculated for each woman by age at diagnosis using a Markov model. The model took a time horizon of up to 100 years old and had a yearly time cycle. In each cycle, the woman's age-specific risk of death was obtained from the Office for National Statistics (ONS) all-cause mortality statistics for the UK.[71] A half-cycle correction was applied, assuming that all deaths would occur, on average, half-way through a year.

Pulmonary embolism and recurrent venous thromboembolism

Other than the case fatality rates, no long-term increased risk of death for women who experienced PE or recurrent VTE was modelled. However, a short-term risk of death was applied based on the expert elicitation exercise and the case fatality rates from Carrier *et al.*[66] (see *The risk of 30-day mortality following a pulmonary embolism* and *The risk of death following a recurrent venous thromboembolism*). If the women survived the decision tree with no intracranial bleeds or CTEPH, then they were assumed to receive the same life expectancy as the general population.

Life expectancy after chronic thromboembolic pulmonary hypertension

Delcroix *et al.*[72] conducted a retrospective analysis of long-term outcomes in people with newly diagnosed CTEPH from 27 centres in Europe and Canada between 2007 and 2009. They presented Kaplan–Meier curves for two groups: those who were surgically treated and those who were medically treated. Quasi-patient-level data were obtained from the Kaplan–Meier curves using the Guyot *et al.*[73] method. Exponential, Weibull, Gompertz, log-normal and log-logistic parametric survival curves were fitted to these data.[74] The goodness of fit of the curves was assessed using visual assessment and the Bayesian information criterion [(BIC) see

Appendix 16]. On this basis, a log-normal curve was selected to model post-CTEPH mortality in the surgically treated group and an exponential curve was selected to model post-CTEPH mortality in the medically treated group. When the curves were extrapolated for use in the model, it was assumed that age did not influence the extrapolated curve and that the probability of death in each year would be constrained so that it could not be less than the age-matched risk of death for a woman in the ONS life tables.[71] In the study by Delcroix *et al.*,[72] 404 patients were treated surgically and 275 patients were treated medically. In the base case, it was assumed that 59.4% (404/679) were treated surgically. A scenario analysis was also conducted in which 100% of women were treated surgically, as it was noted that the surgically treated group were much younger than the medically treated group (mean age 60 years vs. 67 years). The life expectancy was calculated using parametric survival curves and a yearly time cycle and deaths were assumed to occur half-way through the year.

Life expectancy after an intracranial haemorrhage

In the economic model, it was assumed that all patients who had an intracranial haemorrhage suffered a stroke. The base case used data from the study by Fogehom *et al.*,[75] which suggested that the annual risk of dying after a stroke compared with the general Finish population was 4.5-fold in the first year, 2.2-fold in years 2–6 and 0.9-fold in years 7–16.[75] The population analysed had a mean age of 67.3 years, and 48.4% of the population were male. No SEs were presented. In the base case, it was assumed that women who suffered an intracranial haemorrhage had a 4.5-fold increase in their general population mortality in the first year and a 2.2-fold increase in their general population mortality in the years 2–6. On the grounds of clinical plausibility, it was assumed that women returned to their baseline risk of mortality from year 7 onwards rather than experiencing a reduction in their mortality risk. Other than the inclusion of the mortality ratio in the all-cause risk of death, life expectancies were calculated using the same method as was used to calculate the life expectancy in the general population.

Quality of life

The quality-of-life parameters used in the economic model, along with the distributions used in the probabilistic sensitivity analysis (PSA), are given in *Appendix 17*.

General population

The utility of the general population was calculated using the formula from Ara and Brazier.[76] The formula adjusts each individual's utility score for their gender and age, as given in the formula below:

$$\begin{aligned}\text{General population utility} = {} & 0.9508566 + 0.0212126(1 = \text{male},\ 0 = \text{female}) - 0.0002587 \\ & \times \text{age} - 0.0000332 \times \text{age}^2.\end{aligned} \tag{2}$$

Quality of life following a pulmonary embolism

Locadia *et al.*[77] conducted the time trade-off technique to value health states (on the scale of 1 being equivalent to perfect health and 0 being equivalent to death) related to VTE in 159 patients who had either experienced a VTE, a bleeding event related to receiving a vitamin K antagonist (VKA), or had a post-thrombotic syndrome. They reported the median valuations and interquartile ranges of no VKA treatment, own current health, VKA treatment, post-thrombotic syndrome, DVT, muscular bleeding, gastrointestinal bleeding, PE and non-fatal haemorrhagic stroke. The mean valuations were not reported. Locadia *et al.*[77] found that PE had a median valuation of 0.63, with an interquartile range of 0.36–0.86. On the basis of clinical input, this median value was applied multiplicatively to each patient's baseline utility for a period of 4 weeks.

Quality of life following a recurrent venous thromboembolism

The ratio of PEs and DVT for people experiencing a recurrent VTE was taken from Carrier *et al.*,[66] who found that out of the people with initial PE, 3.0% (95% CI 2.5% to 3.7%) had recurrent PE and 3.6% (95% CI 2.3% to 5.0%) had a recurrent VTE after 3 months of anticoagulation.[66] The associated utility

value for a DVT event was obtained from Locadia *et al.*,[77] who found that the median valuation for a DVT was 0.84, with an interquartile range of 0.64–0.98. On the basis of clinical input, this median value was applied multiplicatively to each patient's baseline utility for a period of 4 weeks.

Quality of life following a bleeding event

Locadia *et al.*[77] reported utility losses related to bleeding associated with the treatment of VTE. The utility for people with a gastrointestinal bleed was 0.65 with an interquartile range of 0.49–0.86 and the utility for people with a non-fatal haemorrhagic stroke was 0.33 with an interquartile range of 0.14 to 0.53. No utility values were reported that included all other types of bleeding; therefore, it was assumed that women who experienced another bleeding event had the same utility value as women who had a gastrointestinal bleeding event. In the model, women who experienced an intracranial haemorrhage had a permanently reduced quality of life and women who experienced another type of bleeding had a reduced quality of life for 4 weeks.

Quality-of-life losses resulting from cancers induced by imaging

There were two elements to the quality-of-life losses resulting from an imaging-induced cancer. First, there was a quality-of-life loss associated with cancer, and second, there was the estimated risk of a cancer being induced by a scan.

A decrement was applied to those individuals who were estimated to have cancer. Ara and Brazier[78] estimated the utility of an individual with cancer as being 0.697 (95% CI 0.657 to 0.736) and the utility for equivalent people without cancer as being 0.795 (95% CI 0.754 to 0.836).[78] This gave a utility multiplier of 0.8767. This utility multiplier was applied to each individual's baseline utility for the remainder of their lifetime after they developed a type of cancer.

The lifetime attributable risks of developing radiation-induced cancers were obtained from the literature.[61] Based on the available data, imaging-induced breast and lung cancers in the mother and childhood cancers in the fetus were included in the analysis. No information was available on the yearly risk of developing a type of cancer as a result of a single scan. Owing to this lack of data, it was assumed that the incidence of radiation-induced cancers would follow that of the general population.[79] The life expectancy of individuals with and without cancer were therefore normalised based on the age-specific incidence of cancer in the whole population, which will include some radiation-induced cancers and cancers that develop as a result of other causes. For the mothers, if they were older than the lowest age observed in the incidence statistics, the incidence of their cancer prior to the woman's age was ignored and the distribution was renormalised based on the remaining data. Relative survival statistics in the UK population with each type of cancer and ONS all-cause mortality statistics were used to calculate the life expectancy of people with cancer.[71,79–86] A summary of the discounted cancer decrements by age for the mother is given in *Appendix 18*. The mean discounted QALY decrement for the induction of childhood cancer per scan was –0.000037.

Chronic thromboembolic pulmonary hypertension

There was a lack of evidence on the quality of life for CTEPH patients. Owing to this lack of evidence, it was assumed that CTEPH patients would have the same utility as people with heart failure. The utility data were sourced from Ara and Brazier,[76] with people who had a heart problem other than a myocardial infarction having a utility of 0.672 (95% CI 0.649 to 0.694) and a matched person without heart problems having a utility of 0.802 (95% CI 0.771 to 0.831).[78] This utility multiplier of 0.838 was applied to the general population utility for the first year for those who could be treated surgically, and was applied for a lifetime for those who could not be treated surgically.

Costs

This section provides a detailed description of the cost parameters used in the economic model. A summary of all cost parameters used in the base case, and the distributions around them, are provided in *Appendix 19*. The price year used for the costs was 2015–16. All costs from previous price years were inflated to 2015–16 values using the Hospital and Community Health Services pay and prices index.[87]

Performing the clinical decision rule

In the base case, it was assumed that collecting all of the necessary information to perform the CDRs would be performed by a registrar and that this would take them on average 5 minutes longer than the measurements that they would collect in standard practice for a woman with a suspected PE. The assumption about the length of time was tested in a scenario analysis, in which it was assumed that it would take the registrar 10 additional minutes to collect the information. The cost of registrar time was obtained from the *Unit Costs of Health and Social Care 2016*.[87]

Anticoagulation treatment

The split of the different types of anticoagulation drugs was taken from the UKOSS and suspected PE data set. For the women with PE, the three most common LWMHs used were enoxaparin (Clexane®; Sanofi, Paris, France; $n = 88$), dalteparin (Fragmin®; Pfizer Inc., New York, NY, USA; $n = 54$) and tinzaparin (Innohep®; Leo Pharma, Ballerup, Denmark; $n = 30$). A weighted average of the cost of using each of the three drugs was used to calculate the drug costs in the model. The dose of these LWMHs was calculated in accordance with information in the *British National Formulary*.[88] In the base case, it was assumed that the current weight of pregnant women was used to calculate the dose, despite doses having been based on early pregnancy weight. In a scenario analysis, 12.5 kg was removed from a pregnant woman's weight for the calculation of the dose if they were > 20 weeks pregnant. This scenario was conducted as this is the upper range of the typical weight gain for a pregnant woman in the UK and this weight is typically gained in the second half of the pregnancy.[89] In line with recommendations made by the RCOG, in the model, women received anticoagulation treatment for whichever was the greater of 3 months or the remaining length of their pregnancy plus 6 weeks.[2] In the base case, it was assumed that women received the same duration and dose of anticoagulant drugs for the initial PE as for a recurrent VTE. This assumption was tested in a scenario analysis in which it was assumed that there was no additional cost of a second VTE, if the woman was treated for her initial PE.

Pulmonary embolism

The cost of PE event was taken from the *NHS Reference Costs 2015 to 2016*.[90] PE was assumed to be any pulmonary embolus with a complication score of 0–8 or 9+ (currency code DZ09K and DZ09J) that was a non-elective stay. Excess bed-days were included in the calculation of this cost. The mean cost of treating PE was £4778 with a SE of £224.7.

Recurrent venous thromboembolism

The ratio of DVT-related events (17%) and PE-related events (83%) for recurrent VTE was calculated as described in the clinical parameters section. The cost of DVT was obtained from the *NHS Reference Costs 2015 to 2016*.[90] The cost of a DVT was assumed to be the weighted average of the cost of a DVT score of 0–2, 3–5, 6–8, 9–11 and 12+ (currency codes YQ51A to YQ51E) using the number of finished consultant episodes that were non-elective inpatient stays. The mean cost of treating a DVT was £2612 with a SE of £68.6.

Bleeding

The cost of an intracranial haemorrhage was obtained from Luengo-Fernandez *et al.*,[91] as there would be an initial hospitalisation cost and an ongoing cost for rehabilitation.[91] This study had up to 5 years of follow-up of 729 stroke patients in Oxfordshire, UK. The cost in the first year was taken from the first year of follow-up and an ongoing cost of treating stroke patients was calculated by taking a weighted average based on the number of remaining patients and the mean costs in each year following the first. The lifetime costs of intracranial bleeding were calculated by attaching the health state costs to the estimated life expectancy of people with an intracranial haemorrhage and assuming that individuals who died did so half-way through the year. No information on the SE of these costs was presented, so it was assumed that the SE was 20% of the mean cost. The cost of intracranial bleeding in year 1 was £11,707 with a SE of 2341 and the cost of intracranial bleeding in subsequent years was £1524 with a SE of 305.

Chronic thromboembolic pulmonary hypertension

The cost of a pulmonary endarterectomy was taken from NICE guidelines[33] on the use of rivaroxaban for the treatment of DVT.[33] This was applied as a one-off cost for everyone who was eligible to receive surgery for their CTEPH. This was found to be £6558.11 after inflation to 2015–16 prices. The ongoing cost of CTEPH was also taken from the same guidelines and was found to be £15,968 per annum after inflation to 2015–16 prices. The lifetime costs of CTEPH were calculated by attaching the health state costs to the estimated life expectancy of people with CTEPH and by assuming that individuals who died did so half-way through the year.

Cancer

The lifetime attributable cost of breast cancer was obtained from a 15-month follow-up study of women diagnosed with breast cancer at one UK site by Hall *et al.*[92] They found that the cost of breast cancer was £13,241 in 2015–16 prices and that at least 75% of the costs were incurred within 6 months; as such, no upwards adjustment to these costs was applied. The lifetime attributable cost of lung cancer was obtained from a report by Incisive Health.[93] This report provided costs of non-small-cell lung cancer by stage at diagnosis. These stage-specific costs were weighted by the stage distribution of non-small-cell lung cancer by Cancer Research UK.[79] This gave a lifetime cost of a lung cancer to be £16,095 in 2015–16 prices. The cost of childhood cancers to health-care systems was found to be poorly understood. A study by van Listenburg *et al.*[94] compared the cost-effectiveness of two treatment regimens for childhood acute lymphoblastic leukaemia with chemotherapy in one Dutch centre between 2002 and 2006. The cost of the most recent regimen [ALL-10, the Dutch Childhood Oncology Group protocol for the treatment of children with acute lymphoblastic leukaemia (ALL)] was converted into pounds sterling using purchasing power parity rates and then inflated to 2015–16 prices. This gave the cost of treating a childhood cancer to be £126,273.

Outcome measures

The main outcome measure for this analysis is the incremental cost-effectiveness ratio (ICER) for each strategy. A full incremental analysis will be conducted, which will compare all of the strategies with each other. The ICER will be compared with a maximum acceptable ICER (MAICER) of £20,000 per QALY gained. This is in line with decision-making processes by NICE.[60]

A value-of-information analysis will also be conducted to determine if further research into the cost-effectiveness of the strategies is an efficient use of resources from the health-care system perspective. In the value-of-information analysis, it was expected that the number of women affected per year was 2231. This was based on there being 723,913 live births in England and Wales in 2011 and data from the Scottish Morbidity Record suggesting a combined incidence of 2.0 per 10,000 maternities for antenatal and postnatal PE.[28] This value was then uplifted by assuming that for women with suspected PE, 18 out of 277 women would have PE and 259 out of 277 women would not. On the basis of clinical input, it is expected that any information generated by a future study would be useful for 10 years.

The value of research into particular questions will be assessed by calculating the partial value of information for a particular set of parameters. Previously, this has not been feasible because of time constraints, but recently developed techniques now allow for the calculation of the partial value of information from the model outputs of the PSA,[95] estimating multiparameter partial expected value of perfect information from a PSA sample (a non-parametric regression approach). The value of particular study designs to address any research question will not be explored, as performing this calculation would require a known study design to address a particular research question, neither of which was developed in this project.[96]

Summary of assumptions

- Long-term evidence from the population with PE is valid for pregnant/postpartum women who have PE.
- The women were only at risk of VTE for 1 year because their suspected PE was caused by their pregnancy. As such, the long-term risk factors that have been shown for other older populations with VTE were not relevant.
- Fetuses are outside the decision problem, apart from the adverse effects of any imaging-induced cancers.
- When initiating anticoagulation treatment, it was assumed that, in the UK, the guidelines by the Royal College of Obstetricians and Gynaecologists would be followed.
- For the base case, it was assumed that an accurate diagnosis of PE/no PE would be made when a woman was referred to imaging.
- In the base case, it was assumed that CTEPH was a risk factor that depended on an individual's characteristics and was modifiable by anticoagulation treatment.
- The utility of a woman with CTEPH would likely be the same as someone with heart failure.
- The cost of anticoagulation treatment for a second VTE event would be the same as the cost of anticoagulation treatment for the initial VTE.
- All doses are based on a woman's current weight.

Analysis

Base case

A deterministic and probabilistic sensitivity analysis was performed for the base-case analysis. In the PSA, 1000 runs were conducted and the model produced stable estimates of the total costs and QALYs after approximately 800 runs. Full details of the stability of the model results with respect to the number of PSA runs are given in *Appendix 20*. A full incremental analysis of all 10 decision options was performed in all analyses.

The robustness of the base-case results was assessed by conducting scenario analyses and threshold analyses.

Threshold analyses

Two threshold analyses were conducted. Both threshold analyses were based on the base-case deterministic analysis.

In the first threshold analysis, 'scan all' was compared with a series of hypothetical decision rules with given sensitivity and specificity values. A necessary assumption in this threshold analysis was that these hypothetical decision rules were unrelated to each woman's characteristics. This analysis provides information on how good a decision rule would have to be before it would be cost-effective compared with the current standard of care for pregnant women with PE.

In the second threshold analysis, the QALY-maximising strategy out of 'scan all', SNTA and SNTN was determined with respect to the probability that a woman had PE. To do this analysis, the base-case deterministic analysis was adapted so that bootstrapped mean QALYs for women with PE and women without PE were produced for each strategy. These QALYs were then combined using the threshold probability that a woman has PE.

List of scenario analyses

(1) Assuming that the sensitivity and the specificity of the scanning tests are taken from the same sources as the NICE guidelines.
(2) Assuming that the risk of CTEPH is estimated from data presented in Klok *et al.*[69]
(3) Assuming that the risk of CTEPH is not modifiable by anticoagulation.

(4) Assuming that there is no risk of death from a recurrent VTE.
(5) Assuming that if the woman experiences a recurrent VTE and is treated for her initial PE then the cost of anticoagulation treatment is £0.
(6) Assuming that the estimated risk of death from the expert elicitation exercise was conducted for:

(6.1) the combination of expert's answers in the secondary statistical population
(6.2) the combination of expert's answers in the tertiary statistical population
(6.3) the combination of expert's answers in the quaternary statistical population
(6.4) expert 1's answers in the primary statistical population
(6.5) expert 2's answers in the primary statistical population
(6.6) expert 3's answers in the primary statistical population
(6.7) expert 4's answers in the primary statistical population.

(7) Using cohort data on the risk of death from PE in pregnant women from the UKOSS patient.
(8) Assuming that for the calculation of the drug costs, 12.5 kg is removed from the woman's weight if she is in the second half of her pregnancy.
(9) Assuming that the following costs for CTEPH are used:

(9.1) £24,000 for CTEPH surgery
(9.2) the costs of managing CTEPH are taken from Schweikert *et al.*[97]
(9.3) both 9.1 and 9.2.

(10) Women are not at risk of bleeding, recurrent VTE or CTEPH. In this scenario, the only risk is immediate death associated with PE.

(10.1) When the risk of PE-related death is obtained from the expert elicitation exercise.
(10.2) When the risk of PE-related death is obtained from the UKOSS patient.

(11) All scanning-induced cancers present within 15 years of a scan rather than the woman's lifetime.

All scenario analyses were based upon the model being run deterministically. In scenario 1, instead of assuming perfect scanning, it was assumed that the scans would have a probability of producing both false-positive and false-negative results, which were calculated from the sensitivity and specificity of the scan that the patient received. The model structure diagram in *Figure 8* provides the structure of these aspects of the decision tree. It was assumed that those with a false-positive result were at risk of bleeding and it was assumed that those with a false-negative result would be discharged without anticoagulation treatment regardless of the result of the decision rule. All other scenario analyses assumed that false-positive and false-negative scans were not possible, but used different assumptions or data in the base-case model.

Results

Base-case analysis

The results of the base-case health economic analysis are given in *Table 16*. In the deterministic analysis, 'scan all' was the dominant strategy, as it produced the most QALYs (20.3855) at the lowest cost (£1359). In the PSA, the 'scan all' strategy also dominated all of the other 10 strategies, as again it produced the most QALYs (20.3832) at a cost of £1360. As well as being dominant on average, in the 1000 PSA replications, the 'scan all' strategy was the most cost-effective option in 98.9% of PSA replications when a MAICER of £20,000 per QALY gained was used. This indicates that there is only a very small probability (1.1%) that scanning all women would give an incorrect decision, given the parameter distributions used in the base-case analysis. A cost-effectiveness acceptability curve showing the probability that each intervention is the most cost-effective at different MAICERs is given in *Appendix 21* (see *Figure 62*).

TABLE 16 The results of the base-case health economic analysis

Strategy	Costs (£)								QALYs	Incremental		ICER	Probability of being the most cost-effective strategy at £20,000 per QALY gained
	Decision rule	Scans	Drugs	VTE	Bleeds	Induced cancers	CTEPH	Total		Costs (£)	QALYs		
Deterministic analysis													
No scan, treat none	0	0	15	56	1	0	2757	2830	19.8589	–	–	Dominated by scan all	N/A
Wells's score (strict)	4	27	73	155	4	1	1970	2233	20.0575	–	–	Dominated by scan all	N/A
Delphi specificity score	4	49	71	157	4	2	1979	2265	20.0630	–	–	Dominated by scan all	N/A
Geneva score	4	82	85	173	5	3	1825	2175	20.0952	–	–	Dominated by scan all	N/A
Wells's score (permissive)	4	88	89	186	5	3	1727	2101	20.1213	–	–	Dominated by scan all	N/A
Delphi primary score	4	96	104	222	6	3	1462	1896	20.1914	–	–	Dominated by scan all	N/A
PERC score	4	109	109	236	6	4	1351	1818	20.2164	–	–	Dominated by scan all	N/A
No scan, treat all	0	0	1260	322	122	0	647	2352	20.3013	–	–	Dominated by scan all	N/A
Delphi sensitivity score	4	215	145	313	9	7	733	1424	20.3663	–	–	Dominated by scan all	N/A
Scan all	0	223	151	322	9	7	647	1359	20.3855	–1470	0.5266	Dominant	N/A

continued

TABLE 16 The results of the base-case health economic analysis (*continued*)

Strategy	Costs (£)									Incremental			Probability of being the most cost-effective strategy at £20,000 per QALY gained
	Decision rule	Scans	Drugs	VTE	Bleeds	Induced cancers	CTEPH	Total	QALYs	Costs (£)	QALYs	ICER	
PSA													
No scanning, treat none	0	0	16	59	1	0	2173	2249	19.8157	–	–	Dominated by scan all	0.5%
Delphi specificity score	4	49	70	157	4	1	1629	1914	20.0291	–	–	Dominated by scan all	0.0%
Wells's score (strict)	4	27	73	157	4	1	1599	1865	20.0308	–	–	Dominated by scan all	0.4%
Geneva score	4	83	84	175	5	2	1504	1855	20.0695	–	–	Dominated by scan all	0.0%
Wells's score (permissive)	4	88	89	188	5	2	1430	1805	20.0984	–	–	Dominated by scan all	0.0%
Delphi primary score	4	95	103	222	6	2	1250	1682	20.1710	–	–	Dominated by scan all	0.0%
PERC score	4	109	108	237	7	3	1161	1628	20.2006	–	–	Dominated by scan all	0.1%
No scanning, treat all	0	0	1260	321	123	0	650	2353	20.2937	–	–	Dominated by scan all	0.0%
Delphi sensitivity score	4	217	143	312	9	5	714	1403	20.3618	–	–	Dominated by scan all	0.1%
Scan all	0	225	150	321	9	5	650	1360	20.3832	–888	0.5675	Dominant	98.9%

N/A, not applicable; –, no data.

Value-of-information analysis

Table 17 shows the results of value-of-information analysis. The value-of-information analysis found that reducing all uncertainty in the economic model for this decision problem would have a maximum value of £4.57 per person per year. Assuming that 2231 women per year are able to benefit from any research over a period of 10 years, this gives a maximum value of conducting research into this decision problem of £101,952. However, using one piece of research to reduce all of the parameters used in the economic model is highly optimistic, so we conducted an analysis of the value of conducting further research into a group of parameters for which a single study could be designed. The group with the highest value was the effectiveness of anticoagulation treatment, as this had a value of £72.77 for 2231 women with a suspected PE over a 10-year time horizon. As this is the maximum value of future research into this parameter, conducting any research into these values that costs more than this is not indicated in this decision analysis. Given the value-of-information analysis, further research to reduce uncertainty in the parameter estimates used in the economic model does not appear to be indicated for this decision problem. It should be noted that the accuracy of the decision rules could not be included in this analysis, as they were applied to each woman's modelled characteristics. Therefore, they were not random variables in the PSA, which is necessary to be able to conduct a value-of-information analysis. Instead, the value of conducting the analysis was assessed by conducting the threshold analysis that compared decision rules with hypothetical sensitivity and specificity values to scanning all women. It should also be noted that much of the data in the base case came from the general population with PE and, therefore, these data may not be representative of pregnant/postpartum women with PE. This uncertainty was addressed in the scenario analysis in which the probability of recurrent VTE, CTEPH and bleeding were set to 0%.

Scenario analyses

A summary of the scenario analysis results are provided in *Table 18*; detailed results for each scenario are provided in *Appendix 20*. When using a MAICER of £20,000 per QALY gained, scanning all women was the most cost-effective strategy in all scenario analyses. In fact, scanning all women dominated all of the other strategies in every scenario, except when CTEPH was not modifiable by anticoagulation treatment, when the

TABLE 17 The results of the value-of-information analysis for the diagnosis of pregnant or postpartum women with PE

Parameters	Value (£) per person per year[a]	Approximate SE (£)	Value (£) for 2231 pregnant and postpartum women over 10 years[a]
All parameters			
Overall	3.64	N/A	81,114
Parameter groups			
The effectiveness of anticoagulation treatment	0.00	1.31	0.00
The risk of recurrent VTE	0.00	1.33	0.00
Death resulting from recurrent VTEs	0.00	0.00	0.00
Incidence of major bleeding events	0.00	0.00	0.00
Life expectancy of women with CTEPH	0.22	7.68	4828
Risk of death from PE/risk of experiencing a CTEPH	0.00	0.62	0.00
All utility parameters	0.00	0.00	0.00
All cost parameters	0.09	10.86	1920

N/A, not applicable.
a At a MAICER of £20,000 per QALY gained.

TABLE 18 A summary of the scenario analysis results

Scenario analysis	Most cost-effective strategy using a MAICER of £20,000 per QALY gained	Location of the table with the full results
(1) Scanning is assumed to result in an imperfect diagnosis	Scan all	*Appendix 21, Table 83*
(2) Risk of CTEPH from Klok *et al.*[69] is used	Scan all	*Appendix 21, Table 84*
(3) CTEPH is not modifiable by anticoagulation treatment	Scan all	*Appendix 21, Table 85*
(4) 100% of patients with CTEPH were treated surgically	Scan all	*Appendix 21, Table 86*
(5) A Weibull curve is used to estimate the life expectancy of women with a surgically treated CTEPH	Scan all	*Appendix 21, Table 87*
(5) A Gompertz curve is used to estimate the life expectancy of women with a surgically treated CTEPH	Scan all	*Appendix 21, Table 88*
(5) A log-logistic curve is used to estimate the life expectancy of women with a surgically treated CTEPH	Scan all	*Appendix 21, Table 89*
(5) A gamma curve is used to estimate the life expectancy of women with a surgically treated CTEPH	Scan all	*Appendix 21, Table 90*
(5) A generalised gamma curve is used to estimate the life expectancy of women with a surgically treated CTEPH	Scan all	*Appendix 21, Table 91*
(6) A Weibull curve is used to estimate the life expectancy of women with a medically treated CTEPH	Scan all	*Appendix 21, Table 92*
(6) A Gompertz curve is used to estimate the life expectancy of women with a medically treated CTEPH	Scan all	*Appendix 21, Table 93*
(6) A log-logistic curve is used to estimate the life expectancy of women with a medically treated CTEPH	Scan all	*Appendix 21, Table 94*
(6) A gamma curve is used to estimate the life expectancy of women with a medically treated CTEPH	Scan all	*Appendix 21, Table 95*
(6) A generalised gamma curve is used to estimate the life expectancy of women with a medically treated CTEPH	Scan all	*Appendix 21, Table 96*
(7) There is no risk of death following a recurrent VTE	Scan all	*Appendix 21, Table 97*
(8) There is no anticoagulation treatment cost for recurrent VTEs	Scan all	*Appendix 21, Table 98*
(9.1) The expert elicitation exercise on the risk of mortality from PE was conducted for women with PE as defined in the secondary statistical population	Scan all	*Appendix 21, Table 99*
(9.2) The expert elicitation exercise on the risk of mortality from PE was conducted for women with PE as defined in the tertiary statistical population	Scan all	*Appendix 21, Table 100*
(9.3) The expert elicitation exercise on the risk of mortality from PE was conducted for women with PE as defined in the quaternary statistical population	Scan all	*Appendix 21, Table 101*
(9.4) The expert elicitation exercise on the risk of mortality from PE was conducted for expert 1's answers with PE as defined in the primary statistical population	Scan all	*Appendix 21, Table 102*
(9.5) The expert elicitation exercise on the risk of mortality from PE was conducted for expert 2's answers with PE as defined in the primary statistical population	Scan all	*Appendix 21, Table 103*
(9.6) The expert elicitation exercise on the risk of mortality from PE was conducted for expert 3's answers with PE as defined in the primary statistical population	Scan all	*Appendix 21, Table 104*

TABLE 18 A summary of the scenario analysis results (*continued*)

Scenario analysis	Most cost-effective strategy using a MAICER of £20,000 per QALY gained	Location of the table with the full results
(9.7) The expert elicitation exercise on the risk of mortality from PE was conducted for expert 4's answers with PE as defined in the primary statistical population	Scan all	*Appendix 21*, *Table 105*
(10) The risk of PE-related mortality is taken from UKOSS	Scan all	*Appendix 21*, *Table 106*
(11) 12.5 kg reduction in the weight of pregnant women who were > 20 weeks pregnant for the purpose of calculating their anticoagulant drug dose	Scan all	*Appendix 21*, *Table 107*
(12.1) The cost of CTEPH surgery is £24,000	Scan all	*Appendix 21*, *Table 108*
(12.2) The cost of CTEPH management is taken from Schweikert *et al.*[97]	Scan all	*Appendix 21*, *Table 109*
(12.3) 12.1 and 12.2	Scan all	*Appendix 21*, *Table 110*
(13.1) Women are not at risk of bleeding, recurrent VTE or CTEPH, and the risk of PE-related death is from the expert elicitation base case	Scan all	*Appendix 21*, *Table 111*
(13.2) Women are not at risk of bleeding, recurrent VTE or CTEPH, and the risk of PE-related death is from UKOSS patients	Scan all	*Appendix 21*, *Table 112*
(14) All scanning-induced cancers are present within 15 years	Scan all	*Appendix 21*, *Table 113*

risk of CTEPH was estimated from data presented in the literature and when women were not at risk of bleeding, recurrent VTE or CTEPH. The final scenario is of particular note, as this is the most unfavourable scenario to the 'scan all' strategy, as the only benefit of anticoagulation treatment is the prevention of death associated with a woman's initial PE and there are no harms associated with anticoagulation treatment. This indicates that the results of the economic model are relatively robust to the model assumptions that have been tested.

Threshold analyses

The threshold analyses explored which strategy out of 'scan all', 'scan none, treat all' and 'scan none, treat none' would be the most effective strategy, conditional on the probability that a woman with suspected PE actually had PE. In the first threshold analysis, the model base case was used, and in the second threshold analysis, the scenario in which CTEPH was not modifiable by anticoagulation treatment was used. In the analysis that used the base case, scanning all women is optimal if the probability that they have PE ranges from 0.1% to an upper limit of between 96.5% and 97.0%. At a probability of having PE of 97.0% or above, a woman should be given treatment directly, as the small chance of inducing a bleed in women without PE is outweighed by the QALY losses associated with radiation-induced cancers for the mother and the fetus (if applicable). The results in the scenario in which anticoagulation treatment does not have an impact on a woman's risk of CTEPH are broadly similar to those of the base-case analysis. However, the threshold at which all women are treated without scanning increases slightly to 98.0%, compared with 97.0% in the base case.

The second threshold analysis calculated pairwise ICERs comparing scanning all women with a series of hypothetical CDRs in which the rules being positive or not is independent of each woman's characteristics. The sensitivity and 1 – specificity provide the probabilities that the rules are positive for pregnant and postpartum women with a suspected PE, who actually have PE and actually do not have PE, respectively. The results of this threshold analysis are given in *Tables 19* and *20*. The tables show that even if a decision rule could be developed that had 97.5% sensitivity and 90% specificity, the ICER would be £13,392, which is

TABLE 19 The ICERs of a hypothetical decision rule to selectively scan women with a suspected PE compared with scanning all women

	Sensitivity					
Specificity	97.5	95	90	85	80	75
90	£13,392	£4345	£414	Dominated	Dominated	Dominated
80	£11,193	£3425	Dominated	Dominated	Dominated	Dominated
70	£9089	£2523	Dominated	Dominated	Dominated	Dominated
60	£7072	£1639	Dominated	Dominated	Dominated	Dominated

TABLE 20 The incremental QALY gains of a hypothetical decision rule to selectively scan women with a suspected PE compared with scanning all women

	Sensitivity					
Specificity	97.5	95	90	85	80	75
90	–0.0110	–0.0242	–0.0507	–0.0772	–0.1036	–0.1301
80	–0.0113	–0.0245	–0.0510	–0.0774	–0.1039	–0.1303
70	–0.0115	–0.0247	–0.0512	–0.0777	–0.1041	–0.1306
60	–0.0118	–0.0250	–0.0514	–0.0779	–0.1044	–0.1308

well below the £20,000-per-QALY-gained threshold. This indicates that scanning all women would probably be considered to be cost-effective, even if such a decision rule could be developed for pregnant and postpartum women with a suspected PE. It should be noted that all of the hypothetical decision rules lead to fewer expected QALYs than scanning all women.

Discussion

We found that scanning all women dominated all other strategies considered for pregnant and postpartum women with a suspected PE. In all scenario analyses and in the PSA, scanning all women either was the dominant strategy or had the highest ICER below £20,000 per QALY gained of the 10 strategies being considered. The value of conducting further research into parameters used in the economic model was likely to be below the cost of conducting further research into any subset of feasible parameters. The threshold analyses indicated that, if a CDR was to be developed for the diagnosis of PE in pregnant or postpartum women with a suspected PE, the rule would require a sensitivity value in excess of 97.5% for the ICER of 'scan all' versus the decision rule to exceed £20,000 per QALY gained. It should also be noted that the sensitivity of the decision rule would need to exceed 97.5% for the hypothetical decision rule to result in expected QALY gains compared with scanning all women. Another point to note is that our threshold analyses demonstrated that clinicians must believe that the probability of a pregnant or postpartum woman having PE is $< 0.1\%$ for discharging the women to result in a higher number of lifetime QALYs than sending the woman for a scan. To our knowledge, this study is the first to assess the outcomes associated with selectively scanning pregnant or postpartum women with a suspected PE.

The key implication of this study is that any diagnostic strategy other than scanning all pregnant and postpartum women with suspected PE is unlikely to be cost-effective. Furthermore, threshold analyses suggest that a clinician must be highly certain that a pregnant/postpartum woman with suspected PE does not have PE to make scanning the woman result in fewer expected lifetime QALYs than discharging her immediately would.

A key strength of this study is that this is the first mathematical model to assess the long-term outcomes of selectively scanning pregnant and postpartum women with a suspected PE. However, the study does have some limitations. First, much of the long-term evidence on outcomes for pregnant and postpartum women with PE comes from a much older population with PE, who typically have experienced some comorbidities. Although the older population provides the best available evidence on the long-term outcomes for pregnant and postpartum women with suspected PE, the event rates may differ in the population of pregnant and postpartum women with PE, as they are much younger. This limitation was addressed with a scenario analysis in which all risks that were sourced from the general population (bleeding, CTEPH or recurrent VTE) were set to 0%. This scenario was unfavourable to the 'scan all' strategy, but even then it remained the most cost-effective strategy at a MAICER of £20,000 per QALY gained. Second, it was structurally assumed on the grounds of clinical plausibility that, in the population of pregnant and postpartum women, it was their pregnancy that caused their increased risk of PE. Consequently, it was assumed that these women were at risk of complications resulting from their initial PE only during the anticoagulation treatment period. This assumption could potentially be verified or refuted if data on long-term outcomes (> 1 year) were to be collected on a large number of pregnant and postpartum women diagnosed with PE. Currently, the best available evidence on the outcomes of pregnant and postpartum women with PE is from the UKOSS patients used in this study, and these have only a 30-day follow-up period. Finally, as each woman's DiPEP data were directly used in the model, with the decision rules being applied to their individual characteristics, the sensitivity and specificity values closely matched those of the statistical analyses. There could be some design-related bias in the estimates of the sensitivity and specificity resulting from the design of the clinical study. Given the results of the threshold analysis, which indicated that a decision rule would need a proven sensitivity in excess of 97.5% in order to be cost-effective in the UK, it is highly unlikely that this limitation would alter the conclusions of the health economic analysis.

The value-of-information analysis and the threshold analyses did not identify any promising areas for conducting future research on this decision problem. A well-designed study on the long-term follow-up of pregnant and postpartum women would be able to inform the modelled risks of bleeding, recurrent VTE and CTEPH. However, designing sufficiently well-designed long-term follow-up studies to address the limitations in the evidence on the long-term risks associated with having PE in pregnant and postpartum women would probably be infeasible. This would be because of difficulties in maintaining a long enough follow-up and recruiting enough women, as we estimated that there would be only 145 cases per year of PE in pregnant or postpartum women in the UK. Based on our current evidence, the most promising avenue for future research is probably research into reducing the radiation exposure to the mother and the fetus from diagnostic imaging and hence reducing the QALY losses associated with the imaging technique used. This research would have benefits for everyone who receives a diagnostic imaging scan, not just pregnant and postpartum women with a suspected PE.

In conclusion, scanning all pregnant/postpartum women with suspected PE is likely to be cost-effective. A CDR would need to have a high sensitivity value (> 97.5%) and a high specificity value (> 90%) to be cost-effective compared with scanning all women. Future research into reducing the radiation dose associated with scanning or developing new diagnostic technologies would probably be a more promising way of providing cost-effective care to pregnant/postpartum women with a suspected PE than developing a decision rule.

Chapter 6 Implications for policy, practice and future research

The main findings of the DiPEP study were as follows:

- In pregnant and postpartum women presenting to hospital with suspected PE, many of the recognised clinical features of PE and risk factors for PE do not have diagnostic value and in a few patients may be associated with the absence of PE. This may reflect selective referral or self-presentation of women with clinical features or risk factors for PE (i.e. women with symptoms suggesting PE and recognised risk factors or clinical features may be more likely to present for investigation than those with symptoms suggesting PE but no other features).
- Expert-derived CDRs, existing CDRs (Wells's score, Geneva score and PERC score) and D-dimer measurement have poor diagnostic accuracy for PE in pregnant and postpartum women presenting to hospital with suspected PE.
- We were unable to derive a new CDR using multivariate analysis or recursive partitioning that achieved acceptable sensitivity and specificity without being overfitted. This probably reflects the limited diagnostic value of the clinical features in this cohort.
- Biomarkers for VTE showed poor diagnostic accuracy, with only TG lag time achieving high sensitivity with potentially worthwhile specificity. This may reflect the widespread use of anticoagulation treatment prior to blood sampling. When women who had received anticoagulation treatment were removed from the sample, the number of women with PE was too small for meaningful analysis.
- A prospective cohort design to validate a new CDR or biomarker is potentially feasible, but would require recruitment across a large number of sites (≈50) for a substantial period of time (2 years). Our data suggest that current decision rules and biomarkers show insufficient promise to justify such a study.
- A non-selective strategy of CT scan for all women with suspected PE was cheaper and more effective than strategies that selected women for scanning on the basis of a CDR.
- The value of conducting further research into parameters used in the economic model is likely to be below the cost of conducting further research into any subset of feasible parameters.

Implications for policy and practice

Our findings do not support the use of CDRs and biomarkers in selecting women with suspected PE in pregnancy or post partum to receive imaging. This does not necessarily mean that all women with suspected PE in pregnancy or post partum must receive imaging, given the recognised limitations of imaging. The suspected PE cohort included 42 women who had PE ruled out without imaging. We did not include these women in the primary analysis because without imaging it is uncertain whether or not PE may have been missed, but we found no evidence of missed pathology on follow-up. We do not know exactly why these women, who had been identified as requiring imaging, did not ultimately receive imaging. Responses to enquiries to sites and comments on the CRF suggested that, for most patients, a more senior or specialist clinician had decided that imaging was not necessary after a more junior or generalist clinician had initially decided that it was. The DiPEP study was not designed to explore the reasons for or the appropriateness of decisions to perform imaging, so we are unable to determine whether or not the decision to forgo imaging for these patients was reasonable.

Our findings regarding the limited diagnostic value of clinical features suggest that there is little objective evidence to guide decision-making regarding imaging. Many of the women with PE had normal physiology and no risk factors for PE, so normal physiology and the absence of risk factors should provide little reassurance that PE can be ruled out without imaging. In general, risk factors, clinical features and physiology showed little difference between women with PE and women without PE.

We found that chest radiograph abnormality is associated with PE even when the reported abnormality is not thought to be PE related. RCOG guidance[2] recommends performing chest radiography prior to investigation for PE. Our findings suggest that non-specific chest radiograph abnormalities should increase rather than reduce the suspicion of PE.

It is important to recognise that our findings apply specifically to the diagnostic assessment of women with suspected PE in secondary care. The DiPEP study was not designed to evaluate the risk of developing PE in pregnancy. Epidemiological studies comparing women who develop VTE in pregnancy with those who do not have established the important risk factors for VTE in pregnancy.[27–29,55] The failure of the risk factors to be diagnostically useful in the DiPEP study probably reflects the process of selection into a population undergoing diagnostic evaluation in hospital. Risk factors for VTE should continue to be used to guide the provision of thromboprophylaxis in pregnancy and other relevant decisions. The findings may also not apply to primary care or other out-of-hospital settings. Clinicians working in the community or patients themselves may be successfully using clinical features to select women for hospital assessment, thus explaining why these clinical features have little value in a selected hospital population.

The decision-analysis modelling showed that a strategy of non-selective scanning for all women was likely to be cost-effective compared with strategies that used decision rules to select women for scanning or the alternatives of treating all or treating none without the use of scanning. This probably reflects the substantial QALY gains associated with accurately identifying and treating PE in pregnancy, compared with the small QALY decrement associated with scanning. The non-selective use of scanning was not only more effective than selective scanning, it was also cheaper in the base-case analysis. This was probably because identifying and treating PE reduced the substantial long-term costs of treating CTEPH. A scenario analysis in which it was assumed that treatment did not influence the probability of developing CTEPH showed that non-selective scanning would be more expensive than selective scanning, but still had the highest ICER below the £20,000-per-QALY-gained MAICER used by NICE.

The conclusion that scanning was likely to be cost-effective compared with using a decision rule was robust in all sensitivity analyses and scenario analysis. A threshold analysis comparing scanning for all with selection based on hypothetical decision rules with high sensitivity and specificity found that sensitivity would need to exceed 97.5% and specificity would need to exceed 90% before a decision rule would be cost-effective. This is far more accurate than the decision rules we developed and tested. It suggests that there is little possibility of developing a decision rule with sufficient accuracy to be cost-effective.

Finally, our threshold analysis showed that scanning is not cost-effective compared with discharge without treatment only when the probability of PE is $< 0.1\%$. This again reflects the substantial health gains associated with identifying and treating PE in pregnancy compared with the small health risks and modest costs. It suggests that clinicians should have a low threshold for providing imaging and discharge without imaging only when the risk of PE is negligible. If women are to be involved in decision-making, as recommended by the RCOG,[2] then the comparative risks and benefits of scanning need to be accurately presented.

Implications for future research

We undertook a case–control study instead of a cohort study because of concerns about the high cost, long duration and unknown feasibility of the latter design. The suspected PE cohort showed that recruitment across a variety of sites was achievable and that the prevalence of PE was higher than suggested in previous studies. As a consequence, we conclude that a prospective cohort study of pregnant and postpartum women with suspected PE is feasible, but will require a large number of sites ($\approx$50) and around 2 years of recruitment to provide a sample with a sufficient number of women with PE to estimate sensitivity with acceptable precision. However, a study of this size would require substantial funding and could be justified only if there was a promising diagnostic technology with good evidence to suggest that it could be validated

in a large cohort. Decision-analysis modelling suggested that a hypothetical decision rule would need a far higher level of accuracy than any of the decision rules we evaluated in order for it to be cost-effective compared with scanning all women. The value-of-information analysis showed that the value of conducting further research into parameters used in the economic model was likely to be below the cost of conducting further research into any subset of feasible parameters.

The biomarker study was limited by the small number of patients with PE and the fact that most women received anticoagulation treatment prior to blood sampling. The former limitation was inevitable given the low prevalence of PE, whereas the latter was an unavoidable consequence of a challenging research environment and a clinical imperative to provide treatment if there is any delay in diagnosis. Further study of biomarkers may be worthwhile, but only if these issues can be addressed.

Future research probably needs to go 'back to the drawing board'. The physiological changes associated with pregnancy clearly undermine the potential of many clinical features and biomarkers that could otherwise be useful for diagnosing PE. An additional problem in developing new biomarkers is the widespread use of anticoagulation treatment (either as thromboprophylaxis or as treatment during diagnostic assessment) that interferes with biomarker assays. New diagnostic technologies for PE need to be developed specifically for the pregnant and postpartum population. This may include imaging techniques that are more convenient for women and/or do not involve ionising radiation.

Guidance from the RCOG[2] suggests that women should be informed of the benefits and risks of imaging and involved in the decision to undertake imaging. Our decision analysis has identified data sources that could be used to inform decision-making and has developed a model that could be used to weigh the risks and benefits. These data suggest that the benefits of imaging substantially outweigh the risks. However, understanding these risks, applying them to the individual patient and weighing them against each other is an extremely complex process. Future research is therefore required to develop tools, such as decision aids, that could be used to present information to women and to help women to participate in a decision. Research is also needed to develop ways of equipping clinicians to provide information and involve women in decision-making.

Acknowledgements

We would like to extend our thanks to the women who participated in the DiPEP study and the research nurses and midwives who worked diligently to collect data. We would also like to thank the UKOSS reporting clinicians who contributed to the study and without whom the study would not have been possible.

Thanks also go to the project steering committee members for their contribution.

The Diagnosis of Pulmonary Embolism in Pregnancy research group

Report authors

Professor Steve Goodacre, University of Sheffield.

Miss Kimberley Horspool, University of Sheffield.

Mr Neil Shephard, University of Sheffield.

Mr Daniel Pollard, University of Sheffield.

Professor Beverley Hunt, Guy's and St Thomas' NHS Foundation Trust.

Dr Gordon Fuller, University of Sheffield.

Professor Catherine Nelson-Piercy, Guy's and St Thomas' NHS Foundation Trust.

Professor Marian Knight, University of Oxford.

Dr Steven Thomas, Sheffield Teaching Hospitals NHS Foundation Trust.

Professor Fiona Lecky, University of Sheffield.

Dr Judith Cohen, University of Sheffield.

Project management group

Amanda Loban, University of Sheffield.

Professor Beverley J Hunt, Guy's and St Thomas' NHS Foundation Trust.

Professor Catherine Nelson-Piercy, Guy's and St Thomas' NHS Foundation Trust.

Daniel Pollard, University of Sheffield.

Ellen Bradley, University of Sheffield.

Professor Fiona Lecky, University of Sheffield.

Dr Gordon Fuller, University of Sheffield.

Dr Judith Cohen, University of Sheffield.

Kimberley Horspool, University of Sheffield.

Kiran Parmar, Guy's and St Thomas' NHS Foundation Trust.

Professor Marian Knight, University of Oxford.

Professor Matthew Stevenson, University of Sheffield.

Professor Michael Campbell, University of Sheffield.

Neil Shephard, University of Sheffield.

Professor Steve Goodacre, University of Sheffield.

Dr Steven Thomas, Sheffield Teaching Hospitals NHS Foundation Trust.

Dr Wee-Shian Chan, BC Women's Hospital & Health Centre.

Trial Steering Committee

Dr Andrew Thomson, Royal Alexandra Hospital.

Franchesca Cullinane, PPI lay representative.

Kimberley Horspool, University of Sheffield.

Shan Bennett, PPI lay representative.

Dr Steve Crane, York District General Hospital.

Professor Steve Goodacre, University of Sheffield.

Dr Sue Pavord, John Radcliffe Hospital.

Dr Trevor Cox, Liverpool Cancer Trials Unit.

Local investigators

Professor Steve Goodacre, Sheffield Teaching Hospitals NHS Foundation Trust.

Professor Derek Tuffnell, Bradford Teaching Hospitals NHS Foundation Trust.

Dr Richard Parris, Bolton NHS Foundation Trust.

Dr William Townend, Hull and East Yorkshire Hospitals NHS Trust.

Dr Etienne Ciantar, Leeds Teaching Hospitals NHS Trust.

Dr Richard Body, Central Manchester University Hospitals NHS Foundation Trust.

Professor Catherine Nelson-Piercy, Guy's and St Thomas' NHS Foundation Trust.

Professor Tim Harris, Bart's Health NHS Trust.

Dr Chris Vorwerk, Portsmouth Hospitals NHS Trust.

Dr Gillian Swallow/Dr Judith Moore, Nottingham University Hospitals NHS Trust.

Dr Andy Walden, Royal Berkshire NHS Foundation Trust.

Contributions of authors

Steve Goodacre, **Kimberley Horspool**, **Neil Shephard**, **Daniel Pollard**, **Beverley J Hunt**, **Gordon Fuller**, **Catherine Nelson-Piercy** and **Marian Knight** contributed to the first draft, and all authors commented on the final report.

Steve Goodacre, Marian Knight, Beverley Hunt, Catherine Nelson-Piercy, **Steven Thomas**, **Fiona Lecky** and **Judith Cohen** were applicants on the Health Technology Assessment grant and contributed to the study design.

Beverley Hunt, Catherine Nelson-Piercy, Daniel Pollard, Fiona Lecky, Gordon Fuller Judith Cohen, Kimberley Horspool, Marian Knight, Neil Shephard, Steve Goodacre and Steven Thomas were members of the project management group and contributed to the delivery of the trial and revisions of the protocol.

Contributions of others

Matthew Stevenson (School of Health and Related Research, University of Sheffield, Sheffield, UK), Wee-Shian Chan (British Columbia Women's Hospital, Vancouver, BC, Canada), Michael Campbell (School of Health and Related Research, University of Sheffield, Sheffield, UK) were applicants on the Health Technology Assessment grant and contributed to the study design.

Amanda Loban (School of Health and Related Research, University of Sheffield, Sheffield, UK), Diana Papaioannou (School of Health and Related Research, University of Sheffield, Sheffield, UK), Ellen Bradley (School of Health and Related Research, University of Sheffield, UK), Kiran Parmar (Guy's and St Thomas' NHS Foundation Trust, London, UK), Matthew Stevenson, Michael Campbell (School of Health and Related Research, University of Sheffield, UK) were members of the project management group and contributed to the delivery of the trial and revisions of the protocol.

Publications

Fuller GW, Nelson-Piercy C, Hunt BJ, Lecky FE, Thomas S, Horspool K, Goodacre S. Consensus-derived clinical decision rules to guide advanced imaging decisions for pulmonary embolism in pregnancy and the postpartum period. *Eur J Emerg Med* 2018;**25**:221–2.

Hunt BJ, Parmar K, Horspool K, Shephard N, Nelson-Piercy C, Goodacre S, on behalf of the DiPEP research group. The DiPEP (Diagnosis of PE in Pregnancy) biomarker study: an observational cohort study augmented with additional cases to determine the diagnostic utility of biomarkers for suspected venous thromboembolism during pregnancy and puerperium. *Br J Haematol* 2018;**180**:694–704.

Data-sharing statement

All data requests should be submitted to the corresponding author for consideration. Access to anonymised data may be granted following review.

Patient data

This work uses data provided by patients and collected by the NHS as part of their care and support. Using patient data is vital to improve health and care for everyone. There is huge potential to make better use of information from people's patient records, to understand more about disease, develop new treatments, monitor safety, and plan NHS services. Patient data should be kept safe and secure, to protect everyone's privacy, and it's important that there are safeguards to make sure that it is stored and used responsibly. Everyone should be able to find out about how patient data are used. #datasaveslives You can find out more about the background to this citation here: https://understandingpatientdata.org.uk/data-citation.

References

1. Knight M, Nair M, Tuffnell D, Kenyon S, Shakespeare J, Brocklehurst P, *et al. Saving Lives, Improving Mothers' Care – Surveillance of Maternal Deaths in the UK 2012–14 and Lessons Learned to Inform Maternity Care from the UK and Ireland Confidential Enquiries into Maternal Deaths and Morbidity 2009–14*. Oxford: National Perinatal Epidemiology Unit, University of Oxford; 2016.
2. Royal College of Obstetricians and Gynaecologists. *Thromboembolic Disease in Pregnancy and the Puerperium: Acute Management*. Green-top Guideline No. 37b. 2015. URL: www.rcog.org.uk/globalassets/documents/guidelines/gtg-37b.pdf (accessed 27 October 2017).
3. Wells PS, Anderson DR, Rodger M, Stiell I, Dreyer JF, Barnes D, *et al.* Excluding pulmonary embolism at the bedside without diagnostic imaging: management of patients with suspected pulmonary embolism presenting to the emergency department by using a simple clinical model and D-dimer. *Ann Intern Med* 2001;**135**:98–107. https://doi.org/10.7326/0003-4819-135-2-200107170-00010
4. Le Gal G, Righini M, Roy PM, Sanchez O, Aujesky D, Bounameaux H, Perrier A. Prediction of pulmonary embolism in the emergency department: the revised Geneva score. *Ann Intern Med* 2006;**144**:165–71. https://doi.org/10.7326/0003-4819-144-3-200602070-00004
5. Kline JA, Mitchell AM, Kabrhel C, Richman PB, Courtney DM. Clinical criteria to prevent unnecessary diagnostic testing in emergency department patients with suspected pulmonary embolism. *J Thromb Haemost* 2004;**2**:1247–55. https://doi.org/10.1111/j.1538-7836.2004.00790.x
6. Chan WS, Lee A, Spencer FA, Chunilal S, Crowther M, Wu W, *et al.* D-dimer testing in pregnant patients: towards determining the next 'level' in the diagnosis of deep vein thrombosis. *J Thromb Haemost* 2010;**8**:1004–11. https://doi.org/10.1111/j.1538-7836.2010.03783.x
7. Goodacre S, Nelson-Piercy C, Hunt B, Chan WS. When should we use diagnostic imaging to investigate for pulmonary embolism in pregnant and postpartum women? *Emerg Med J* 2015;**32**:78–82. https://doi.org/10.1136/emermed-2014-203871
8. Cutts BA, Tran HA, Merriman E, Nandurkar D, Soo G, DasGupta D, *et al.* The utility of the Wells clinical prediction model and ventilation-perfusion scanning for pulmonary embolism diagnosis in pregnancy. *Blood Coagul Fibrinolysis* 2014;**25**:375–8. https://doi.org/10.1097/MBC.0000000000000054
9. Balan KK, Critchley M, Vedavathy KK, Smith ML, Vinjamuri S. The value of ventilation-perfusion imaging in pregnancy. *Br J Radiol* 1997;**70**:338–40. https://doi.org/10.1259/bjr.70.832.9166067
10. Chan WS, Ray JG, Murray S, Coady GE, Coates G, Ginsberg JS. Suspected pulmonary embolism in pregnancy: clinical presentation, results of lung scanning, and subsequent maternal and pediatric outcomes. *Arch Intern Med* 2002;**162**:1170–5. https://doi.org/10.1001/archinte.162.10.1170
11. Scarsbrook AF, Bradley KM, Gleeson FV. Perfusion scintigraphy: diagnostic utility in pregnant women with suspected pulmonary embolic disease. *Eur Radiol* 2007;**17**:2554–60. https://doi.org/10.1007/s00330-007-0607-0
12. Cahill AG, Stout MJ, Macones GA, Bhalla S. Diagnosing pulmonary embolism in pregnancy using computed-tomographic angiography or ventilation-perfusion. *Obstet Gynecol* 2009;**114**:124–9. https://doi.org/10.1097/AOG.0b013e3181a99def

13. Damodaram M, Kaladindi M, Luckit J, Yoong W. D-dimers as a screening test for venous thromboembolism in pregnancy: is it of any use? *J Obstet Gynaecol* 2009;**29**:101–3. https://doi.org/10.1080/01443610802649045

14. Shahir K, Goodman LR, Tali A, Thorsen KM, Hellman RS. Pulmonary embolism in pregnancy: CT pulmonary angiography versus perfusion scanning. *AJR Am J Roentgenol* 2010;**195**:W214–20. https://doi.org/10.2214/AJR.09.3506

15. Deutsch AB, Twitty P, Downes K, Parsons MT. Assessment of the alveolar-arterial oxygen gradient as a screening test for pulmonary embolism in pregnancy. *Am J Obstet Gynecol* 2010;**203**:373.e1–4. https://doi.org/10.1016/j.ajog.2010.04.049

16. Hassanin IM, Shahin AY, Badawy MS, Karam K. D-dimer testing versus multislice computed tomography in the diagnosis of postpartum pulmonary embolism in symptomatic high-risk women. *Int J Gynaecol Obstet* 2011;**115**:200–1. https://doi.org/10.1016/j.ijgo.2011.05.024

17. O'Connor C, Moriarty J, Walsh J, Murray J, Coulter-Smith S, Boyd W. The application of a clinical risk stratification score may reduce unnecessary investigations for pulmonary embolism in pregnancy. *J Matern Fetal Neonatal Med* 2011;**24**:1461–4. https://doi.org/10.3109/14767058.2011.614652

18. Bourjeily G, Khalil H, Raker C, Martin S, Auger P, Chalhoub M, *et al.* Outcomes of negative multidetector computed tomography with pulmonary angiography in pregnant women suspected of pulmonary embolism. *Lung* 2012;**190**:105–11. https://doi.org/10.1007/s00408-011-9329-9

19. Abele JT, Sunner P. The clinical utility of a diagnostic imaging algorithm incorporating low-dose perfusion scans in the evaluation of pregnant patients with clinically suspected pulmonary embolism. *Clin Nucl Med* 2013;**38**:29–32. https://doi.org/10.1097/RLU.0b013e31827088f6

20. Nijkeuter M. *Safety of Ruling Out Pulmonary Embolism (PE) in Pregnancy by Computed Tomography Pulmonary Angiography (CTPA)*. The XXIV Congress of the International Society on Thrombosis and Haemostasis, Amsterdam, 29 June–4 July 2013.

21. Browne AM, Cronin CG, NiMhuircheartaigh J, Donagh C, Morrison JJ, Lohan DG, Murphy JM. Evaluation of imaging quality of pulmonary 64-MDCT angiography in pregnancy and puerperium. *AJR Am J Roentgenol* 2014;**202**:60–4. https://doi.org/10.2214/AJR.12.9917

22. Bajc M, Olsson B, Gottsäter A, Hindorf C, Jögi J. V/P SPECT as a diagnostic tool for pregnant women with suspected pulmonary embolism. *Eur J Nucl Med Mol Imaging* 2015;**42**:1325–30. https://doi.org/10.1007/s00259-015-3056-z

23. Jordan EJ, Godelman A, Levsky JM, Zalta B, Haramati LB. CT pulmonary angiography in pregnant and postpartum women: low yield, high dose. *Clin Imaging* 2015;**39**:251–3. https://doi.org/10.1016/j.clinimag.2014.11.006

24. Ramsay R, Byrd L, Tower C, James J, Prescott M, Thachil J. The problem of pulmonary embolism diagnosis in pregnancy. *Br J Haematol* 2015;**170**:727–8. https://doi.org/10.1111/bjh.13322

25. Kline JA, Richardson DM, Than MP, Penaloza A, Roy PM. Systematic review and meta-analysis of pregnant patients investigated for suspected pulmonary embolism in the emergency department. *Acad Emerg Med* 2014;**21**:949–59. https://doi.org/10.1111/acem.12471

26. Chan WS, Chunilal S, Lee A, Crowther M, Rodger M, Ginsberg JS. A red blood cell agglutination D-dimer test to exclude deep venous thrombosis in pregnancy. *Ann Intern Med* 2007;**147**:165–70. https://doi.org/10.7326/0003-4819-147-3-200708070-00005

27. Knight M, UKOSS. Antenatal pulmonary embolism: risk factors, management and outcomes. *BJOG* 2008;**115**:453–61. https://doi.org/10.1111/j.1471-0528.2007.01622.x

28. Kane EV, Calderwood C, Dobbie R, Morris C, Roman E, Greer IA. A population-based study of venous thrombosis in pregnancy in Scotland 1980-2005. *Eur J Obstet Gynecol Reprod Biol* 2013;**169**:223–9. https://doi.org/10.1016/j.ejogrb.2013.03.024

29. Henriksson P, Westerlund E, Wallén H, Brandt L, Hovatta O, Ekbom A. Incidence of pulmonary and venous thromboembolism in pregnancies after in vitro fertilisation: cross sectional study. *BMJ* 2013;**346**:e8632. https://doi.org/10.1136/bmj.e8632

30. Sultan AA, West J, Grainge MJ, Riley RD, Tata LJ, Stephansson O, *et al.* Development and validation of risk prediction model for venous thromboembolism in postpartum women: multinational cohort study. *BMJ* 2016;**355**:i6253. https://doi.org/10.1136/bmj.i6253

31. Leung AN, Bull TM, Jaeschke R, Lockwood CJ, Boiselle PM, Hurwitz LM, *et al.* An official American Thoracic Society/Society of Thoracic Radiology clinical practice guideline: evaluation of suspected pulmonary embolism in pregnancy. *Am J Respir Crit Care Med* 2011;**184**:1200–8. https://doi.org/10.1164/rccm.201108-1575ST

32. Torbicki A, Perrier A, Konstantinides S, Agnelli G, Galiè N, Pruszczyk P, *et al.* Guidelines on the diagnosis and management of acute pulmonary embolism: the Task Force for the Diagnosis and Management of Acute Pulmonary Embolism of the European Society of Cardiology (ESC). *Eur Heart J* 2008;**29**:2276–315. https://doi.org/10.1093/eurheartj/ehn310

33. National Institute for Health and Care Excellence. *Venous Thromboembolic Diseases: Diagnosis, Management and Thrombophilia Testing*. Clinical Guideline [CG144]. London: NICE; 2012. URL: www.nice.org.uk/guidance/CG144 (accessed 12 May 2017).

34. Bates SM, Jaeschke R, Stevens SM, Goodacre S, Wells PS, Stevenson MD, *et al.* Diagnosis of DVT: Antithrombotic Therapy and Prevention of Thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. *Chest* 2012;**141**(Suppl. 2):e351–418.

35. Kourlaba G, Relakis J, Kontodimas S, Holm MV, Maniadakis N. A systematic review and meta-analysis of the epidemiology and burden of venous thromboembolism among pregnant women. *Int J Gynaecol Obstet* 2016;**132**:4–10. https://doi.org/10.1016/j.ijgo.2015.06.054

36. Lijmer JG, Mol BW, Heisterkamp S, Bonsel GJ, Prins MH, van der Meulen JH, Bossuyt PM. Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA* 1999;**282**:1061–6. https://doi.org/10.1001/jama.282.11.1061

37. Stein PD, Hull RD, Patel KC, Olson RE, Ghali WA, Brant R, *et al.* D-dimer for the exclusion of acute venous thrombosis and pulmonary embolism: a systematic review. *Ann Intern Med* 2004;**140**:589–602. https://doi.org/10.7326/0003-4819-140-8-200404200-00005

38. Fink A, Kosecoff J, Chassin M, Brook RH. Consensus methods: characteristics and guidelines for use. *Am J Public Health* 1984;**74**:979–83.

39. Knottnerus JA1, van Weel C, Muris JW. Evaluation of diagnostic procedures. *BMJ* 2002;**324**:477–80.

40. Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. *J Adv Nurs* 2000;**32**:1008–15.

41. Jones J, Hunter D. Consensus methods for medical and health services research. *BMJ* 1995;**311**:376–80. https://doi.org/10.1136/bmj.311.7001.376

42. McMillan SS, King M, Tully MP. How to use the nominal group and Delphi techniques. *Int J Clin Pharm* 2016;**38**:655–62. https://doi.org/10.1007/s11096-016-0257-x

43. Dalkey N, Brown B, Cochran S. *The Delphi Method*. Santa Monica, CA: RAND Corporation; 1969.

44. Machin D, Campbell MJ, Tan SB, Tan SH. *Sample Size Tables for Clinical Studies*. Hoboken, NJ: John Wiley & Sons, Inc.; 2009.

45. Murphy N, Broadhurst DI, Khashan AS, Gilligan O, Kenny LC, O'Donoghue K. Gestation-specific D-dimer reference ranges: a cross-sectional study. *BJOG* 2015;**122**:395–400. https://doi.org/10.1111/1471-0528.12855

46. Tibshirani R. Regression shrinkage and selection via the lasso. *J R Stat Soc Ser B* 1996;**58**:267–88.

47. James G, Witten D, Hastie T, Tibshirani R. *An Introduction to Statistical Learning: With Applications in R*. New York, NY: Springer-Verlag; 2013. https://doi.org/10.1007/978-1-4614-7138-7

48. Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York, NY: Springer-Verlag; 2003.

49. Steyerberg EW, Eijkemans MJ, Harrell FE, Habbema JD. Prognostic modeling with logistic regression analysis: in search of a sensible strategy in small data sets. *Med Decis Making* 2001;**21**:45–56. https://doi.org/10.1177/0272989X0102100106

50. Strobl C, Malley J, Tutz G. An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychol Methods* 2009;**14**:323–48. https://doi.org/10.1037/a0016973

51. Office for National Statistics. *Statistical Bulletin: Births by Parents' Characteristics in England and Wales: 2015*. Newport: Office for National Statistics; 2015. URL: www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/livebirths/bulletins/birthsbyparentscharacteristicsinenglandandwales/2015 (accessed 30 March 2017).

52. NHS Digital. *Maternity Services Maternity Services Monthly Statistics, England – December 2016, Experimental Statistics*. NHS Digital; 2016. URL: www.content.digital.nhs.uk/catalogue/PUB23944/msms-dec16-exp-rep.pdf (30 March 2017).

53. Harrell FE. *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis*. New York, NY: Springer-Verlag; 2001. https://doi.org/10.1007/978-1-4757-3462-1

54. Dobson AJ, Barnett AG. *An Introduction to Generalized Linear Models*. Boca Raton, FL: CRC Press; 2008.

55. Abdul Sultan A, West J, Tata LJ, Fleming KM, Nelson-Piercy C, Grainge MJ. Risk of first venous thromboembolism in pregnant women in hospital: population based cohort study from England. *BMJ* 2013;**347**:f6099. https://doi.org/10.1136/bmj.f6099

56. Hunt BJ, Parmar K, Horspool K, Shephard N, Nelson-Piercy C, Goodacre S, on behalf of the DiPEP research group. The DiPEP (Diagnosis of PE in Pregnancy) biomarker study: an observational cohort study augmented with additional cases to determine the diagnostic utility of biomarkers for suspected venous thromboembolism during pregnancy and puerperium. *Br J Haematol* 2018;**180**:694–704. https://doi.org/10.1111/bjh.15102

57. Crawford F, Andras A, Welch K, Sheares K, Keeling D, Chappell FM. D-dimer test for excluding the diagnosis of pulmonary embolism. *Cochrane Database Syst Rev* 2016;**8**:CD010864. https://doi.org/10.1002/14651858.CD010864.pub2

58. Kovac M, Mikovic Z, Rakicevic L, Srzentic S, Mandic V, Djordjevic V, *et al.* The use of D-dimer with new cutoff can be useful in diagnosis of venous thromboembolism in pregnancy. *Eur J Obstet Gynecol Reprod Biol* 2010;**148**:27–30. https://doi.org/10.1016/j.ejogrb.2009.09.005

59. Hedengran KK, Andersen MR, Stender S, Szecsi PB. Large D-dimer fluctuation in normal pregnancy: a longitudinal cohort study of 4,117 samples from 714 healthy Danish women. *Obstet Gynecol Int* 2016;**2016**:3561675. https://doi.org/10.1155/2016/3561675

60. National Institute for Health and Care Excellence. *Guide to the Methods of Technology Appraisal 2013*. Process and Methods [PMG9]. London: NICE; 2013. URL: www.nice.org.uk/process/pmg9/chapter/foreword (accessed 30 March 2017).

61. Wall B, Meara J, Muirhead C, Bury R, Murray M. *Protection of Pregnant Patients During Diagnostic Medical Exposures to Ionising Radiation*. London: Health Protection Agency, The Royal College of Radiologists and the College of Radiographers; 2009.

62. Ohno Y, Higashino T, Takenaka D, Sugimoto K, Yoshikawa T, Kawai H, *et al.* MR angiography with sensitivity encoding (SENSE) for suspected pulmonary embolism: comparison with MDCT and ventilation-perfusion scintigraphy. *AJR Am J Roentgenol* 2004;**183**:91–8. https://doi.org/10.2214/ajr.183.1.1830091

63. Gutte H, Mortensen J, Jensen CV, von der Recke P, Petersen CL, Kristoffersen US, Kjaer A. Comparison of V/Q SPECT and planar V/Q lung scintigraphy in diagnosing acute pulmonary embolism. *Nucl Med Commun* 2010;**31**:82–6. https://doi.org/10.1097/MNM.0b013e3283336747

64. Aujesky D, Obrosky DS, Stone RA, Auble TE, Perrier A, Cornuz J, *et al.* Derivation and validation of a prognostic model for pulmonary embolism. *Am J Respir Crit Care Med* 2005;**172**:1041–6. https://doi.org/10.1164/rccm.200506-862OC

65. Ferrari SLP, Cribari-Neto F. Beta regression for modelling rates and proportions. *J Appl Stat* 2004;**31**:799–815. https://doi.org/10.1080/0266476042000214501

66. Carrier M, Le Gal G, Wells PS, Rodger MA. Systematic review: case-fatality rates of recurrent venous thromboembolism and major bleeding events among patients treated for venous thromboembolism. *Ann Intern Med* 2010;**152**:578–89. https://doi.org/10.7326/0003-4819-152-9-201005040-00008

67. Ensor J, Riley RD, Jowett S, Monahan M, Snell KIe, Bayliss S, *et al.* Prediction of risk of recurrence of venous thromboembolism following treatment for a first unprovoked venous thromboembolism: systematic review, prognostic model and clinical decision rule, and economic evaluation. *Health Technol Assess* 2016;**20**(12). https://doi.org/10.3310/hta20120

68. Lang IM, Madani M. Update on chronic thromboembolic pulmonary hypertension. *Circulation* 2014;**130**:508–18. https://doi.org/10.1161/CIRCULATIONAHA.114.009309

69. Klok FA, van Kralingen KW, van Dijk AP, Heyning FH, Vliegen HW, Huisman MV. Prospective cardiopulmonary screening program to detect chronic thromboembolic pulmonary hypertension in patients after acute pulmonary embolism. *Haematologica* 2010;**95**:970–5. https://doi.org/10.3324/haematol.2009.018960

70. Barritt DW, Jordan SC. Anticoagulant drugs in the treatment of pulmonary embolism. A controlled trial. *Lancet* 1960;**1**:1309–12. https://doi.org/10.1016/S0140-6736(60)92299-6

71. Office for National Statistics. *National Life Tables, UK: 2013–15*. London: Office for National Statistics; 2016. URL: www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/bulletins/nationallifetablesunitedkingdom/20132015 (accessed 30 March 2017).

72. Delcroix M, Lang I, Pepke-Zaba J, Jansa P, D'Armini AM, Snijder R, *et al.* Long-term outcome of patients with chronic thromboembolic pulmonary hypertension: results from an international prospective registry. *Circulation* 2016;**133**:859–71. https://doi.org/10.1161/CIRCULATIONAHA.115.016522

73. Guyot P, Ades AE, Ouwens MJ, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. *BMC Med Res Methodol* 2012;**12**:9. https://doi.org/10.1186/1471-2288-12-9

74. Collett D. *Modelling Survival Data in Medical Research*. 2nd edn. Boca Raton, FL: Chapman & Hall/CRC Press; 2003.

75. Fogelholm R, Murros K, Rissanen A, Avikainen S. Long term survival after primary intracerebral haemorrhage: a retrospective population based study. *J Neurol Neurosurg Psychiatry* 2005;**76**:1534–8. https://doi.org/10.1136/jnnp.2004.055145

76. Ara R, Brazier JE. Populating an economic model with health state utility values: moving toward better practice. *Value Health* 2010;**13**:509–18. https://doi.org/10.1111/j.1524-4733.2010.00700.x

77. Locadia M, Bossuyt PM, Stalmeier PF, Sprangers MA, van Dongen CJ, Middeldorp S, *et al.* Treatment of venous thromboembolism with vitamin K antagonists: patients' health state valuations and treatment preferences. *Thromb Haemost* 2004;**92**:1336–41. https://doi.org/10.1160/TH04-02-0075

78. Ara R, Brazier JE. Using health state utility values from the general population to approximate baselines in decision analytic models when condition-specific data are not available. *Value Health* 2011;**14**:539–45. https://doi.org/10.1016/j.jval.2010.10.029

79. Cancer Research UK. *Lung Cancer Incidence Statistics: Lung Cancer Incidence By Age*. London: Cancer Research UK; 2016. URL: www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/lung-cancer/incidence#heading-One (accessed 30 March 2017).

80. Office for National Statistics. *Deaths Registered in England and Wales (Series DR)*. London: Office for National Statistics; 2015. URL: www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/datasets/deathsregisteredinenglandandwalesseriesdrreferencetables (accessed 17 March 2016).

81. Northern Ireland Statistics and Research Agency. *Microdata for Deaths in Northern Ireland, 2001–2014*. Belfast: Northern Ireland Statistics and Research Agency; 2014. URL: www.nisra.gov.uk/archive/demography/publications/births_deaths/Pivot_Table_2001_2014.xlsx (accessed 17 March 2016).

82. National Records of Scotland. *Vital Events Reference Tables*. Edinburgh: National Records of Scotland; 2015. URL: www.nrscotland.gov.uk/statistics-and-data/statistics/statistics-by-theme/vital-events/general-publications/vital-events-reference-tables/2014 (accessed 17 March 2016).

83. Cancer Research UK. *Lung Cancer Survival Statistics: One-, Five- and Ten-Year Survival for Lung Cancer*. London: Cancer Research UK; 2014. URL: www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/lung-cancer/survival#heading-Zero (accessed 21 March 2016).

84. Cancer Research UK. *Breast Cancer Survival Statistics: One-, Five- and Ten-Year Survival for Breast Cancer*. London: Cancer Research UK; 2014. URL: www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer/survival#heading-Zero (accessed 30 March 2017).

85. Cancer Research UK. *Children's Cancers Survival Statistics?: Survival Trends Over Time in Children's Cancers*. London: Cancer Research UK; 2015. URL: www.cancerresearchuk.org/health-professional/cancer-statistics/childrens-cancers/survival#heading-Two (accessed 30 March 2017).

86. Cancer Research UK. *Breast Cancer Incidence (Invasive) Statistics: Breast Cancer Incidence (Invasive) Statistics*. London: Cancer Research UK; 2016. URL: www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer/incidence-invasive#heading-One (accessed 30 March 2017).

87. Curtis L, Burns A. *Unit Costs of Health and Social Care 2016*. Canterbury: PSSRU, University of Kent; 2016. URL: www.pssru.ac.uk/project-pages/unit-costs/2016/ (accessed 30 March 2017).

88. Joint Formulary Committe. *British National Formulary*. 72nd edn. London: BMJ Group and Pharmaceutical Press; 2016.

89. NHS Choices. *How Much Weight Will I Put On During My Pregnancy?* London: NHS Choices; 2015. URL: www.nhs.uk/chq/Pages/2311.aspx?CategoryID=54 (accessed 30 March 2017).

90. Department of Health and Social Care. *NHS Reference Costs 2015 to 2016*. London: Department of Health and Social Care; 2016. URL: www.gov.uk/government/publications/nhs-reference-costs-2015-to-2016 (accessed 30 March 2017).

91. Luengo-Fernandez R, Gray AM, Rothwell PM, Oxford Vascular Study. A population-based study of hospital care costs during 5 years after transient ischemic attack and stroke. *Stroke* 2012;**43**:3343–51. https://doi.org/10.1161/STROKEAHA.112.667204

92. Hall PS, Hamilton P, Hulme CT, Meads DM, Jones H, Newsham A, *et al.* Costs of cancer care for use in economic evaluation: a UK analysis of patient-level routine health system data. *Br J Cancer* 2015;**112**:948–56. https://doi.org/10.1038/bjc.2014.644

93. Incisive Health. *Saving Lives, Averting Costs: An Analysis of the Financial Implications of Achieving Earlier Diagnosis of Colorectal, Lung and Ovarian Cancer*. London: Incisive Health; 2014. URL: www.incisivehealth.com/uploads/Saving%20lives%20averting%20costs.pdf (accessed 21 November 2016).

94. van Litsenburg RR, Uyl-de Groot CA, Raat H, Kaspers GJ, Gemke RJ. Cost-effectiveness of treatment of childhood acute lymphoblastic leukemia with chemotherapy only: the influence of new medication and diagnostic technology. *Pediatr Blood Cancer* 2011;**57**:1005–10. https://doi.org/10.1002/pbc.23197

95. Strong M, Oakley JE, Brennan A. Estimating multiparameter partial expected value of perfect information from a probabilistic sensitivity analysis sample: a nonparametric regression approach. *Med Decis Making* 2014;**34**:311–26. https://doi.org/10.1177/0272989X13505910

96. Strong M, Oakley JE, Brennan A, Breeze P. Estimating the expected value of sample information using the probabilistic sensitivity analysis sample: a fast, nonparametric regression-based method. *Med Decis Making* 2015;**35**:570–83. https://doi.org/10.1177/0272989X15575286

97. Schweikert B, Pittrow D, Vizza CD, Pepke-Zaba J, Hoeper MM, Gabriel A, *et al.* Demographics, clinical characteristics, health resource utilization and cost of chronic thromboembolic pulmonary hypertension patients: retrospective results from six European countries. *BMC Health Serv Res* 2014;**14**:246. https://doi.org/10.1186/1472-6963-14-246

98. Hurwitz LM, Reiman RE, Yoshizumi TT, Goodman PC, Toncheva G, Nguyen G, Lowry C. Radiation dose from contemporary cardiothoracic multidetector CT protocols with an anthropomorphic female phantom: implications for cancer induction. *Radiology* 2007;**245**:742–50. https://doi.org/10.1148/radiol.2453062046

99. Kass RE, Raftery AE. Bayes factors. *J Am Stat Assoc* 1995;**90**:773–95.

100. National Institute for Health and Care Excellence. *Rivaroxaban for the Treatment of Deep Vein Thrombosis and Prevention of Recurrent Deep Vein Thrombosis and Pulmonary Embolism*. Technology Appraisal Guidance [TA261]. London: NICE; 2012.

101. Cancer Research UK. *Lung Cancer Incidence*. London: Cancer Research UK; 2016. URL: www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/lung-cancer/incidence#heading-Three

102. Joint Formulary Committee. *British National Formulary*. 74 ed. London: BMJ Group and Pharmaceutical Press; 2017.

Appendix 1 Reference standard classification

The project description stated that two independent assessors, blind to clinical predictors and blood results, will classify participants as having PE using diagnostic imaging results, details of adverse events and details of treatments given. Disagreements will be resolved through adjudication by a third assessor.

We structured the process of classification to ensure that it was transparent and reproducible. We provided each independent assessor with details of the diagnostic imaging report, anticoagulant treatment given and 30-day follow-up events, and asked them to classify the information using the following codes.

Classify imaging as:

I1a – imaging is reported as showing PE (including any qualified statement suggesting probable PE, but excluding isolated subsegmental PE) or DVT
I1b – imaging shows only isolated subsegmental PE
I2 – imaging is reported as being equivocal, uninterpretable or indeterminate
I3 – imaging is reported as being negative for PE
I4 – lung imaging is not done.

Classify treatment as:

T1 – therapeutic anticoagulation treatment for > 1 week
T2 – anything less than T1.

Classify follow-up as:

F1 – subsequent PE diagnosis during 30-day follow-up or PE is confirmed at surgery or post-mortem examination
F2 – no subsequent PE diagnosis during 30-day follow-up.

We then applied the principles outlined in *Chapter 3*, *Reference standard classification*, to these codes to determine whether PE was diagnosed or ruled out, or whether or not the woman was excluded, in the primary and secondary analyses. *Table 21* summarises the process.

The process resulted in the classifications shown in *Table 22*.

This resulted in the classifications for the primary analysis and the secondary analysis, as shown in *Table 23*.

TABLE 21 Use of imaging, treatment and follow-up to classify the reference standard for primary and secondary analysis

			Analysis			
Imaging	Treatment	Follow-up	Primary	Secondary 1	Secondary 2	Secondary 3
I1a	T1 or T2	F1 or F2	PE	PE	PE	PE
I1b	T1 or T2	F1 or F2	PE	PE	PE	Exclude
I2	T1	F1	PE	PE	PE	PE
I2	T1	F2	Exclude	PE	Exclude	Exclude
I2	T2	F1	PE	PE	PE	PE
I2	T2	F2	No PE	No PE	No PE	No PE
I3	T1	F1	PE	PE	PE	PE
I3	T1	F2	No PE	No PE	No PE	No PE
I3	T2	F1	PE	PE	PE	PE
I3	T2	F2	No PE	No PE	No PE	No PE
I4	T1	F1	PE	PE	PE	PE
I4	T1	F2	Exclude	PE	Exclude	Exclude
I4	T2	F1	PE	PE	PE	PE
I4	T2	F2	Exclude	Exclude	No PE	Exclude

Secondary analyses: (1) include clinically diagnosed PE; (2) include clinically ruled-out PE; (3) exclude subsegmental PE.

TABLE 22 Classification of participants on the basis of imaging, treatment and follow-up

			Cohort		
Imaging	Treatment	Follow-up	Diagnosed PE	Suspected PE recruited	Suspected PE non-recruited
I1a	T1	F1	0	0	0
I1a	T1	F2	146	17	3[a]
I1a	T2	F1	1	0	0
I1a	T2	F2	7	0	0
I1b	T1	F1	0	0	0
I1b	T1	F2	7	1	3[b]
I1b	T2	F1	0	0	0
I1b	T2	F2	0	0	0
I2	T1	F1	0	0	0
I2	T1	F2	29	3	4
I2	T2	F1	0	0	0
I2	T2	F2	0	5	2
I3	T1	F1	0	0	0
I3	T1	F2	0	19	4
I3	T2	F1	0	0	0
I3	T2	F2	0	235	67
I4	T1	F1	1	0	0

TABLE 22 Classification of participants on the basis of imaging, treatment and follow-up (*continued*)

Imaging	Treatment	Follow-up	Cohort		
			Diagnosed PE	Suspected PE recruited	Suspected PE non-recruited
I4	T1	F2	6	2	1
I4	T2	F1	1	0	0
I4	T2	F2	0	42	11

a Includes two patients duplicated in the diagnosed PE cohort.
b Includes one patient duplicated in the diagnosed PE cohort.

TABLE 23 Reference standard classifications in the primary and secondary analyses

Analysis	Cohort					
	Diagnosed PE		Suspected PE recruited		Suspected PE non-recruited	
	PE	No PE	PE	No PE	PE	No PE
Primary	163	0	18	259	6[a]	73
Secondary 1	198	0	23	259	11[a]	73
Secondary 2	163	0	18	301	6[a]	84
Secondary 3	156	0	17	259	3[b]	73

a Includes three patients duplicated in the diagnosed PE cohort.
b Includes one patient duplicated in the diagnosed PE cohort.

Appendix 2 Mapping of clinical decision rules to Diagnosis of Pulmonary Embolism in Pregnancy study data

In practice, CDRs are applied prospectively by clinicians obtaining relevant information directly from patients. In the DiPEP study, the CDRs were tested by being applied to data collected on the CRF. The CRF was designed to meet a number of purposes and to be consistent and usable for both research nurses collecting data prospectively from consented patients and UKOSS clinicians collecting anonymised data retrospectively from case notes. This meant that the CRF data did not map precisely onto all criteria in the CDRs. This appendix describes how we mapped the CDRs to CRF data.

Furthermore, as described in the main report, we excluded exogenous oestrogen use from the PERC score and used the thresholds developed for our analysis to dichotomise age, oxygen saturation and heart rate, so that the rules would be more applicable to a pregnant population.

Table 24 shows the consensus-derived CDRs. Most of the items map directly onto CRF variables. The time period for the rules was intended to be 90 days, but the CRF identified hospital admission, surgery or significant injury within only 28 days. Obstetric complications and active medical comorbidities were identified from a number of sources on the CRF. The CRF items 'Other problems in this pregnancy' and 'Other previous or pre-existing medical problems' provided lists of specific problems on a drop-down list, including those on the decision rule definition, along with a free-text box that was searched by a clinical expert for relevant terms. Hospital admissions were not specifically identified on the CRF, although those related to obstetric complications or active medical comorbidities were included under these criteria. Caesarean section was recorded both as surgery and an obstetric complication, so we limited the criterion including surgery within 90 days to operations other than caesarean section to avoid double-counting. Clinical signs of DVT were identified by a specific CRF question, but clinical symptoms could be identified only if 'Presenting feature – other' was ticked and the free text recorded a symptom such as lower limb pain or swelling.

Table 25 shows the PERC score. The criterion 'exogenous oestrogen' was removed from the rule, as it was felt to be inappropriate to apply in pregnant women. All other criteria mapped directly onto CRF items. Age, heart rate and oxygen saturation were dichotomised using the thresholds developed for our analysis rather than those in the original rule.

Table 26 shows Wells's criteria for PE. Heart rate was dichotomised using the threshold developed for our analysis rather than that in the original rule. As with the consensus-derived decision rules, the clinical symptoms of DVT were identified only if recorded as free text after 'Presenting features – other' had been ticked. There was insufficient detail to determine whether or not pre-existing cancer had been treated within 6 months, so we included any women with previous cancer under this criterion. The criterion 'PE is the most likely diagnosis or equally likely' was determined by clinical expert review of the free-text response to the CRF question 'What was considered the most likely diagnosis after initial clinical assessment?' In the main analysis, this was identified as positive if there was any mention of PE as a likely diagnosis, unless it was in the context of ruling out PE or stating that PE was unlikely. A secondary analysis was undertaken in which the criterion was positive only if the free text clearly indicated that PE was the most likely or equally likely diagnosis.

TABLE 24 Consensus-derived CDRs

Criterion	CRF item and dichotomisation used (when relevant)
Haemoptysis	Presenting feature haemoptysis ticked
Pleuritic chest pain	Presenting feature pleuritic chest pain ticked
Previous VTE	Does the woman have a past history of thrombosis (either in previous pregnancies or when not pregnant)?
Family history of VTE in a first-degree relative	Is there a history of thrombosis in first-degree relatives?
Hospital admission, surgery or significant injury within 90 days (excluding NVD or caesarean section)	Either (1) did the woman have surgery in the 4 weeks prior to PE in this pregnancy or (2) did the woman have a significant injury in the 4 weeks prior to PE in this pregnancy? Excluding women who had a caesarean section
Obstetric complication[a]	Any of (1) was this a multiple pregnancy? (2) was delivery by caesarean section? (3) were there any other problems in this pregnancy? Selected if any of the following were recorded: pre-eclampsia, ART/IVF, prolonged labour (> 24 hours), PPH (> 1 litre or transfusion), preterm birth at < 37^{+0} weeks in current pregnancy, stillbirth in current pregnancy, hyperemesis or OHSS
Active medical comorbidities[b]	Did the woman have any other previous or pre-existing medical problems? Selected if any of the following were recorded: cancer, heart failure, systemic lupus erythematosus, inflammatory polyarthropathy, inflammatory bowel disease, nephrotic syndrome, diabetes mellitus with nephropathy, sickle cell disease
Post partum or third trimester	Calculated from expected or actual date of delivery
Raised BMI of ≥ 30 kg/m^2	Calculated from booking weight and height
Clinical signs or symptoms of DVT	Either (1) presenting feature other – lower limb pain or (2) clinical signs of DVT?
Oxygen saturation of < 94% on room air	Oxygen saturation on room air, dichotomised using a threshold of < 94%
Tachycardia of > 100 b.p.m. (in first or second trimester, or post partum)/tachycardia of > 110 b.p.m. (in third trimester)	Heart rate, dichotomised using > 100 b.p.m. in first or second trimester or post partum and < 110 b.p.m. in the third trimester
Increased respiratory rate of > 24 breaths per minute	Respiratory rate, dichotomised using a threshold of > 24 breaths per minute

ART, assisted reproductive technology; b.p.m., beats per minute; IVF, in vitro fertilisation; NVD, normal vaginal delivery; OHSS, ovarian hyperstimulation syndrome; PPH, postpartum haemorrhage.

a Obstetric complications: apply once if any of the following are present – pre-eclampsia in current pregnancy assisted reproductive technology/in vitro fertilisation (antenatal only), multiple pregnancy, caesarean section in labour, elective caesarean section, mid-cavity or rotational operative delivery, prolonged labour (> 24 hours), postpartum haemorrhage (> 1 litre or transfusion), preterm birth at < 37^{+0} weeks in current pregnancy, stillbirth in current pregnancy, hyperemesis, ovarian hyperstimulation syndrome (first trimester only).

b Active medical comorbidities: apply once if any of the following are present – cancer, heart failure; systemic lupus erythematosus, inflammatory polyarthropathy or inflammatory bowel disease; nephrotic syndrome; type I or type 2 diabetes mellitus with nephropathy; sickle cell disease.

Table 27 shows the simplified revised Geneva score. Age and heart rate were dichotomised using the thresholds developed for our analysis rather than those in the original rule. We were unable to determine whether or not significant injury in the previous 4 weeks involved a lower limb fracture, so this criterion was considered to be positive for any woman with a significant injury. We were unable to determine whether or not previous cancer was an active malignant condition, so this criterion was considered to be positive for any woman with previous cancer. Unilateral lower limb pain was identified only if it was recorded as free text after 'Presenting features – other' had been ticked. We were unable to determine whether or not clinical signs of DVT specifically involved pain on lower limb palpation, so this criterion was considered to be positive in any woman with clinical signs of DVT recorded.

TABLE 25 The pregnancy-adapted PERC score

Criterion	CRF item and dichotomisation used (when relevant)
Aged ≥ 50 years	Calculated from year of birth, using a threshold of > 35 years
Heart rate of ≥ 100 b.p.m.	Heart rate, dichotomised using > 100 b.p.m. in the first or second trimester or post partum and < 110 b.p.m. in the third trimester
Oxygen saturation on room air of < 95%	Oxygen saturation on room air, dichotomised using a threshold of < 94%
Prior history of DVT/PE	Does the woman have a past history of thrombosis (either in previous pregnancies or when not pregnant)?
Recent trauma or surgery	Either (1) did the woman have surgery in the 4 weeks prior to PE in this pregnancy or (2) did the woman have a significant injury in the 4 weeks prior to PE in this pregnancy?
Haemoptysis	Presenting feature haemoptysis ticked
Exogenous oestrogen	Removed from rule
Unilateral leg swelling	Presenting feature other – lower limb swelling

b.p.m., beats per minute.

TABLE 26 The pregnancy-adapted Wells's criteria

Criterion	CRF item and dichotomisation used (when relevant)
Clinical signs or symptoms of DVT	Either (1) presenting feature other – lower limb pain or (2) clinical signs of DVT?
PE is the most likely diagnosis OR equally likely	What was considered to be the most likely diagnosis after initial clinical assessment? See details in text
Heart rate of > 100 b.p.m.	Heart rate, dichotomised using > 100 b.p.m. in the first or second trimester or post partum and < 110 b.p.m. in the third trimester
Immobilisation for at least 3 days OR surgery in the previous 4 weeks	Either (1) did the woman have surgery in the 4 weeks prior to PE in this pregnancy or (2) period of immobility/bed rest during this pregnancy (≥ 3 days)
Previous objectively diagnosed PE or DVT	Does the woman have a past history of thrombosis (either in previous pregnancies or when not pregnant)?
Haemoptysis	Presenting feature haemoptysis ticked
Malignancy with treatment within 6 months or palliative	Did the woman have any other previous or pre-existing medical problem? Selected if cancer was recorded

b.p.m., beats per minute.

TABLE 27 The pregnancy-adapted simplified revised Geneva score

Criterion	CRF item
Aged > 65 years	Calculated from year of birth, using a threshold of > 35 years
Previous DVT or PE	Does the woman have a past history of thrombosis (either in previous pregnancies or when not pregnant)?
Surgery (under general anaesthesia) or lower limb fracture in the past month	Either (1) did the woman have surgery in the 4 weeks prior to PE in this pregnancy or (2) did the woman have a significant injury in the 4 weeks prior to PE in this pregnancy?
Active malignant condition	Did the woman have any other previous or pre-existing medical problem? Selected if cancer was recorded
Unilateral lower limb pain	Presenting feature other – lower limb pain
Haemoptysis	Presenting feature haemoptysis ticked
Heart rate	Heart rate, dichotomised using > 100 b.p.m. in first or second trimester or post partum and < 110 b.p.m. in the third trimester
Pain on limb palpation	Clinical signs of DVT?

b.p.m., beats per minute.

Appendix 3 Recruitment of women with suspected pulmonary embolism and diagnosed deep-vein thrombosis

Women with suspected PE and diagnosed DVT were prospectively recruited from 11 sites, incorporating 10 emergency departments and seven maternity units. Service delivery for women with suspected PE varied between sites, with some sites managing the patients through the emergency department, other sites managing them through the maternity unit and other sites managing them through either route. *Table 28* shows recruitment by each unit and indicates whether recruitment was led by the emergency department or the maternity team. All of the women diagnosed with DVT were recruited through maternity units, whereas 182 women with suspected PE were recruited through emergency departments and 142 women with suspected PE were recruited through maternity units.

TABLE 28 Recruitment by participating site

Site and unit/department	Cohort			
	Diagnosed DVT	Suspected PE: recruited	Suspected PE: not recruited	Time open to recruitment
Royal Bolton Hospital Emergency Department	0	40	12	16 months, 10 days
Bradford Royal Infirmary Maternity Department	1	35	12	17 months, 15 days
Hull Royal Infirmary Emergency Department	0	15	5	16 months, 10 days
Sheffield Teaching Hospital Emergency Department	0	25	28	18 months, 20 days
Sheffield Teaching Hospital Maternity Department	0	5	8	
Leeds Teaching Hospitals NHS Trust Emergency Department	0	6	0	15 months, 20 days
Leeds Teaching Hospitals NHS Trust Maternity Department	3	42	4	
Manchester Royal Infirmary Emergency Department	0	20	6	15 months, 3 days
Portsmouth Hospitals NHS Trust Emergency Department	0	3	0	12 months, 25 days
Portsmouth Hospitals NHS Trust Maternity Department	3	18	6	
Nottingham University Hospitals NHS Trust Maternity Department (City Hospital)	0	7	0	10 months, 30 days
Nottingham University Hospitals NHS Trust Maternity Department (Queen's Medical Centre)	0	16	0	
Royal Berkshire Hospital Emergency Department	0	16	0	10 months, 26 days
Royal Berkshire Hospital Maternity Department	2	11	3	
Royal London Hospital Emergency Department	0	33	7	13 months, 28 days
Whipps Cross University Hospital Emergency Department	0	8	4	
GSTT Maternity Department	9	24	0	12 months, 17 days

The number of women with suspected PE who were not recruited varied markedly between trusts, with some reporting no patients. This variation is very likely to be attributable to variation in the ability to identify eligible but non-recruited women, so we have not calculated the proportion of eligible women recruited.

Overall, the 11 sites recruited 324 women with suspected PE over 190.52 site months of recruitment. This suggests that any future prospective cohort study of women with suspected PE should be designed on the basis of an estimated recruitment rate of 1.7 per site per month of recruitment (with each site including both emergency department and maternity department recruitment, if both provide diagnostic assessment of suspected PE).

Figure 9 shows the recruitment chart for the study with the actual number of women recruited compared with the predicted number of women recruited. As described in the main text, we deliberately over-recruited to ensure sufficient numbers in the primary analysis that excludes women with clinically diagnosed or ruled-out PE.

The recruitment of women diagnosed with DVT occurred at 5 out of 11 sites; however, 50% of cases of DVT were identified at St Thomas' Hospital. This is likely to represent the local referral pathway for women with VTE in pregnancy for which St Thomas' Hospital is regarded as a specialist centre.

Figure 10 shows the flow of patients with suspected PE and those with diagnosed PE through the study and into the analysis populations.

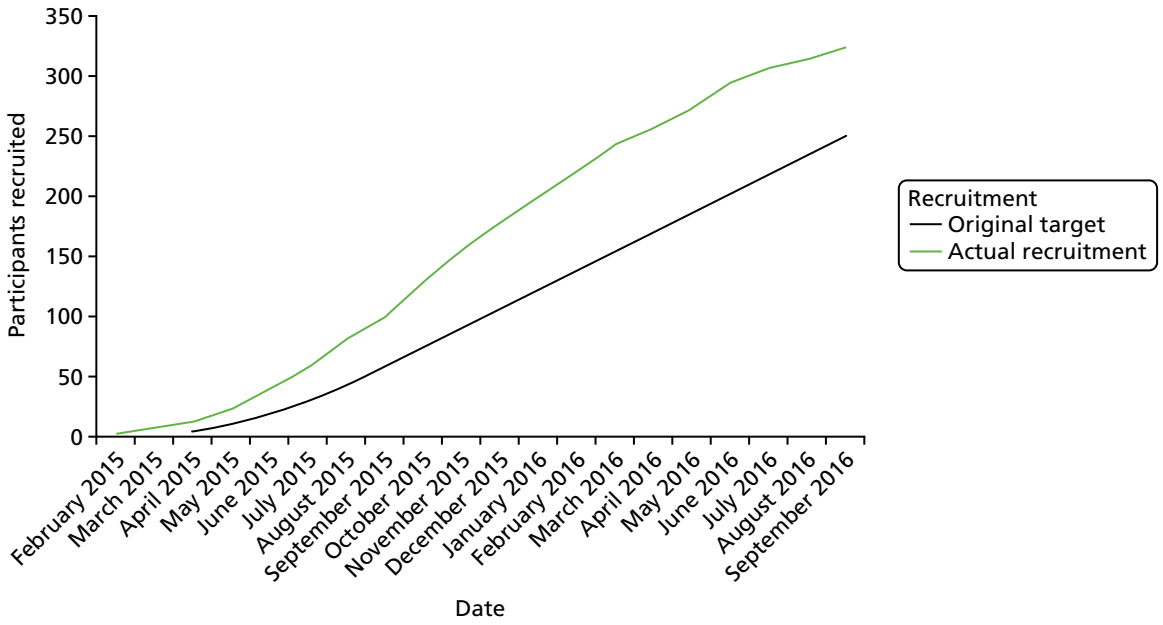


FIGURE 9 Actual and target recruitment for women with a suspected PE.

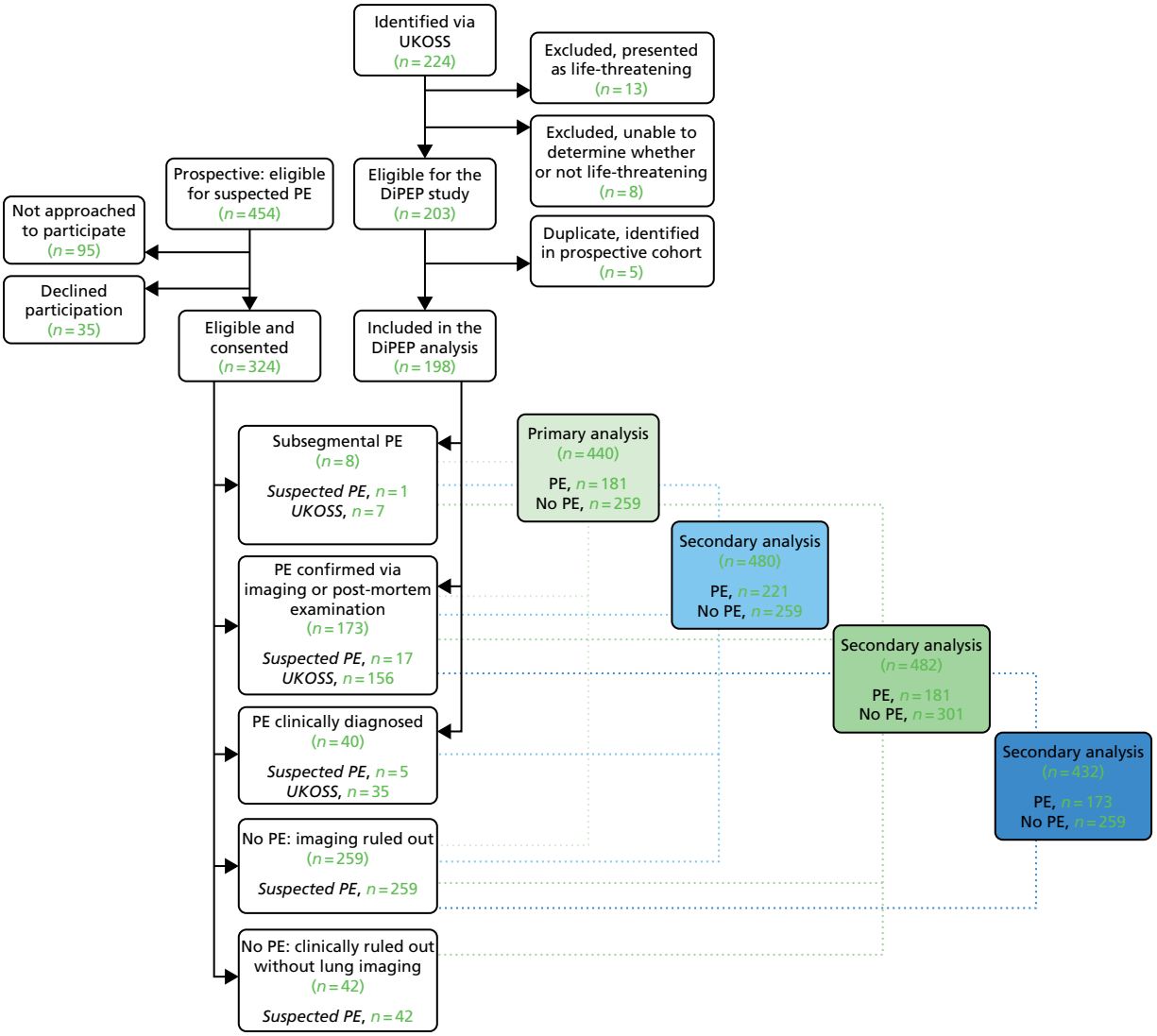


FIGURE 10 Recruitment flow and analysis populations.

Appendix 4 Clinical variables missing data

As described in the methods, we excluded patients from the multivariable analysis and the analysis of CDRs, if any of the following criteria were met:

- More than one of heart rate, respiratory rate and oxygen saturation were missing.
- More than half of the predictors relating to previous medical history were missing.
- More than half of the predictors relating to the current pregnancy were missing.

Table 29 shows the number of missing physiological variables, *Table 30* shows the number of missing variables for previous medical history, *Table 31* shows the number of missing variables for the current pregnancy and *Table 32* shows the number of missing variables for previous pregnancies.

TABLE 29 Number of missing physiological variables by cohort

	Number (% of total) of missing physiological variables (out of 8)							
Cohort	0	1	2	3	4	5	6	7
Diagnosed PE	141 (71.21)	33 (16.67)	6 (3.03)	5 (2.53)	2 (1.01)	1 (0.51)	10 (5.05)	0 (0.00)
Non-recruited	49 (51.58)	38 (40.00)	5 (5.26)	1 (1.05)	0 (0.00)	0 (0.00)	0 (0.00)	2 (2.11)
Suspected PE	286 (88.27)	35 (10.80)	1 (0.31)	1 (0.31)	0 (0.00)	0 (0.00)	1 (0.31)	0 (0.00)

Variables: age, BMI, heart rate, respiratory rate, systolic blood pressure, diastolic blood pressure, temperature and oxygen saturation.

TABLE 30 Number of missing medical history variables by cohort

	Number (% of total) of missing medical history variables (out of 7)				
Cohort	0	1	2	5	6
Diagnosed PE	190 (95.96)	5 (2.53)	2 (1.01)	0 (0.00)	1 (0.51)
Non-recruited	88 (92.63)	4 (4.21)	1 (1.05)	1 (1.05)	1 (1.05)
Suspected PE	323 (99.69)	1 (0.31)	0 (0.00)	0 (0.00)	0 (0.00)

Variables: history of thrombosis in relatives, history of varicose veins, history of intravenous drug use, known thrombophilia, surgery within the previous 4 weeks, injury within the previous 4 weeks, history of thrombosis.

TABLE 31 Number of missing variables for the current pregnancy by cohort

	Number (% of total) of missing variables for the current pregnancy (out of 3)		
Cohort	0	1	2
Diagnosed PE	189 (95.45)	5 (2.53)	4 (2.02)
Non-recruited	90 (94.74)	2 (2.11)	3 (3.16)
Suspected PE	324 (100.00)	0 (0.00)	0 (0.00)

Variables: multiple pregnancy, ≥ 4 hours' travel during pregnancy, ≥ 3 days of immobility.

TABLE 32 Number of missing variables for previous pregnancies by cohort

Cohort	Number (% of total) of missing variables for previous pregnancies (out of 3)		
	0	1	2
Diagnosed PE	157 (79.29)	37 (18.69)	4 (2.02)
Non-recruited	83 (87.37)	10 (10.53)	2 (2.11)
Suspected PE	302 (93.21)	22 (6.79)	0 (0.00)

Variables: previous pregnancies lasting for ≥ 24 weeks' gestation, previous pregnancies lasting for ≤ 24 weeks' gestation, previous pregnancy problems.

Table 33 shows missing data rates for all key variables. Presenting features were recorded using a tick box to indicate when they were present (with the assumption being that they were absent if the box was not ticked), so missing data could not be reported for these variables. Most women did not have D-dimer measurement as part of their usual care, so the missing rate was high for this variable in all groups. Missing data rates were generally very low in the recruited cohort with suspected PE, with only previous pregnancy problems (22/324), temperature (18/324) and D-dimer (207/324) missing in more than 5% of women. Missing data rates were higher in the cohort with diagnosed PE, with employment (13/198), previous pregnancy problems (37/198), all physiological variables (14/198 to 30/198), ECG (10/198), likely diagnosis (19/154) and D-dimer measurement (154/198) missing in > 5% of women.

TABLE 33 Missing data for all key variables by group

Variable	Cohort, *n*		
	Diagnosed PE (*n* = 198)	Non-recruited (*n* = 95)	Suspected PE (*n* = 324)
Year of birth	0	1	0
Ethnicity	3	2	0
Marital status	3	1	1
Employment	13	16	2
Height	5	34	11
Weight	5	25	7
Smoking status	1	4	1
Previous pregnancies lasting for > 24 weeks' gestation	3	2	0
Previous pregnancies lasting for < 24 weeks' gestation	8	6	0
Previous pregnancy problems	37	8	22
Thrombotic event during pregnancy	3	1	0
Receiving thromboprophylaxis	2	4	0
Family history of VTE	1	5	2
History of varicose veins	4	5	1
History of i.v. drug abuse	4	2	0
History of injury in the last 4 weeks	2	4	0

TABLE 33 Missing data for all key variables by group (*continued*)

Variable	Cohort, *n*		
	Diagnosed PE (*n* = 198)	Non-recruited (*n* = 95)	Suspected PE (*n* = 324)
History of thrombophilia	3	2	0
Previous VTE	1	3	0
Other medical problems	1	2	0
History of surgery in the last 4 weeks	1	1	0
Expected date of delivery	0	1	0
Multiple pregnancy	0	3	0
History of long-haul travel	7	4	0
History of immobilisation	6	4	0
Heart rate	14	2	2
Respiratory rate	30	5	9
Oxygen saturation	21	4	3
Systolic BP (mmHg)	15	3	1
Diastolic BP (mmHg)	15	4	1
Temperature	32	10	18
Clinical signs of DVT	2	1	0
ECG	10	1	1
Chest radiograph	9	1	0
Likely diagnosis after clinical assessment	19	1	11
D-dimer measurement	154	66	207
D-dimer normal range	163	56	141

BP, blood pressure; i.v., intravenous.

Appendix 5 Distributions of physiological measures for women with and without pulmonary embolism

Figures 11–17 show the distributions of age and physiological parameters for women with and without PE and those excluded from the analysis.

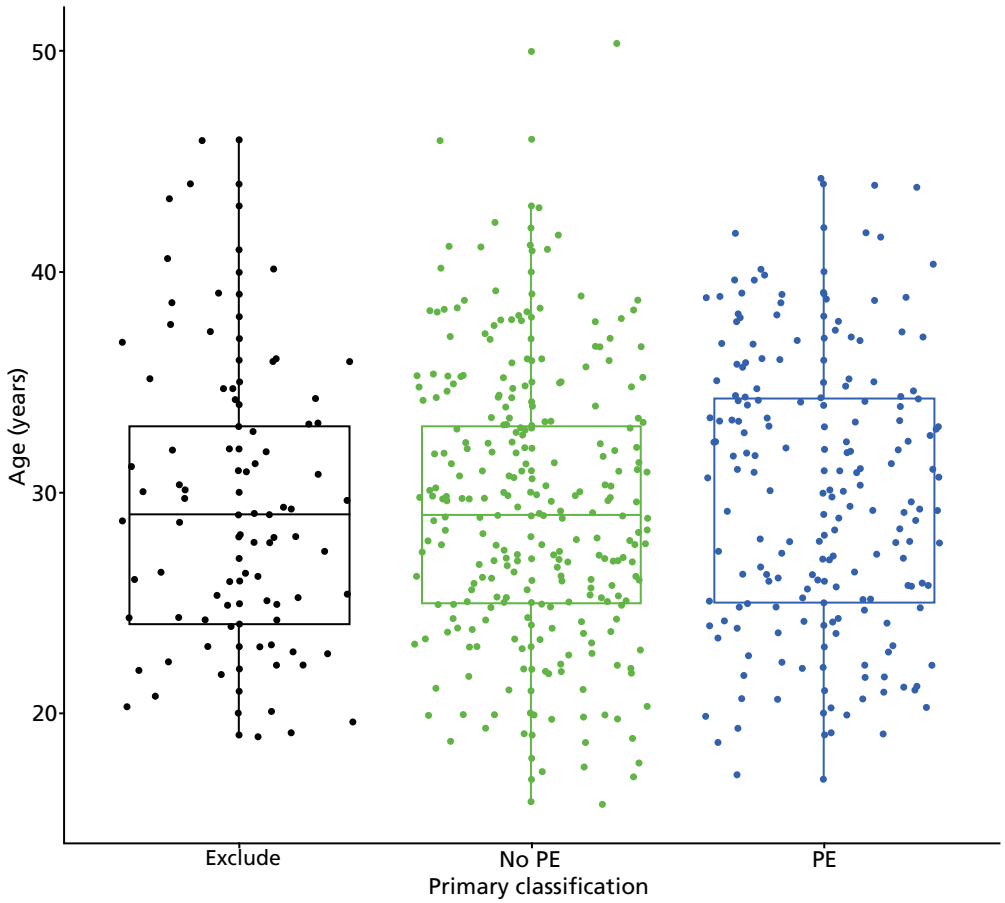


FIGURE 11 Distribution of age by classification in the primary analysis.

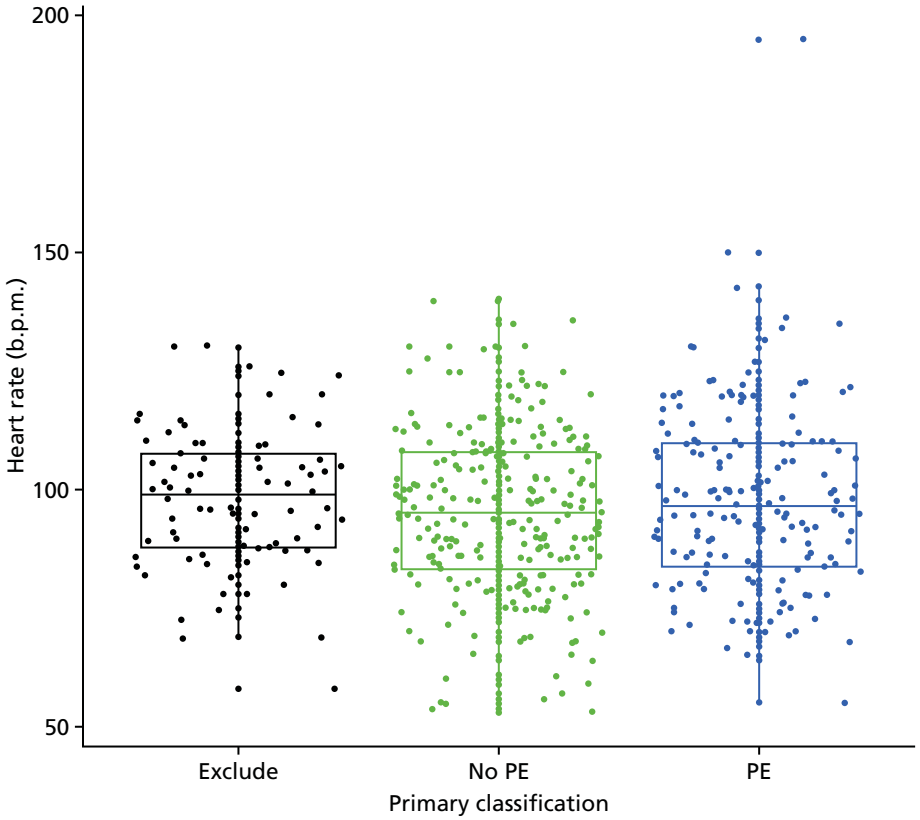


FIGURE 12 Distribution of heart rate by classification in the primary analysis. b.p.m., beats per minute.

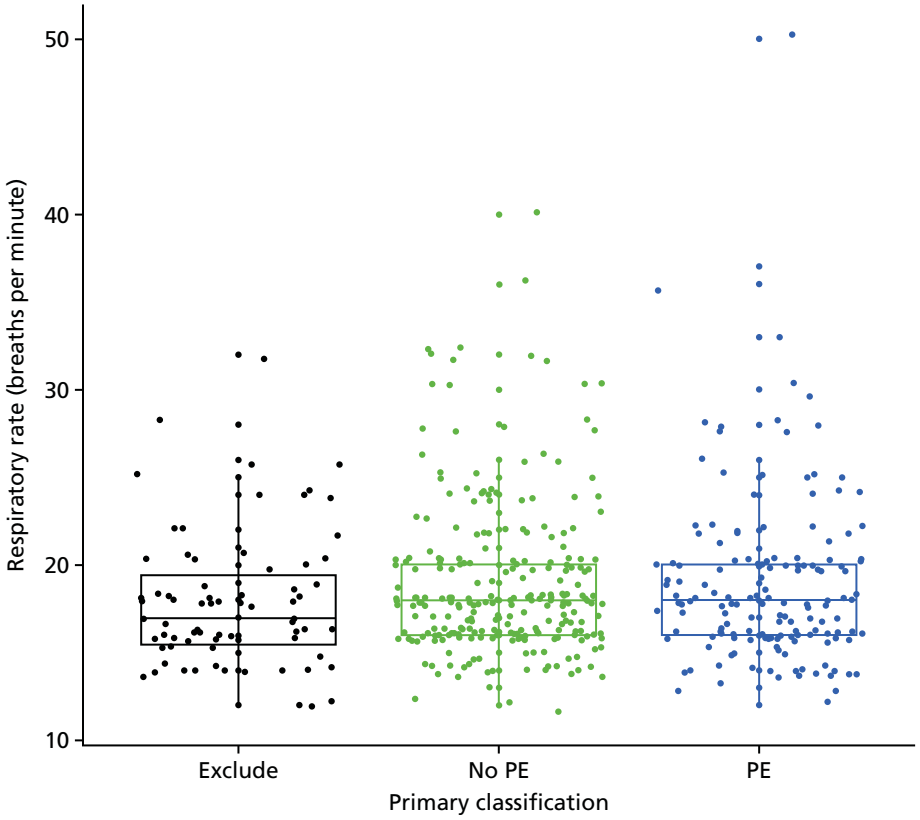


FIGURE 13 Distribution of respiratory rate by classification in the primary analysis.

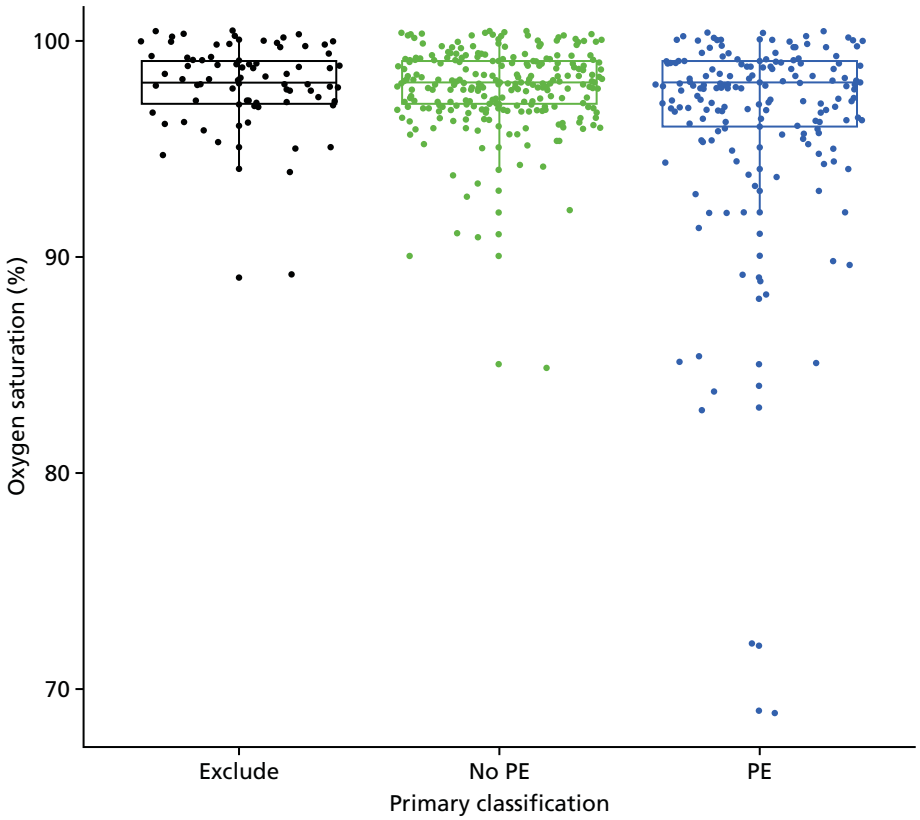


FIGURE 14 Distribution of oxygen saturation by classification in the primary analysis.

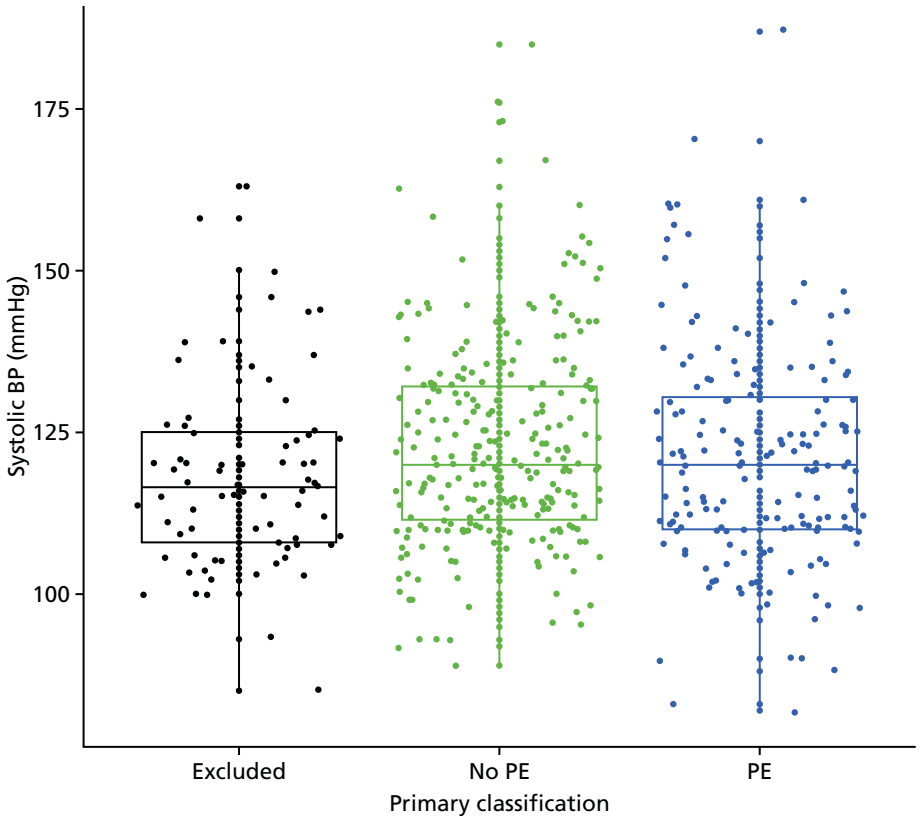


FIGURE 15 Distribution of systolic blood pressure by classification in the primary analysis. BP, blood pressure.

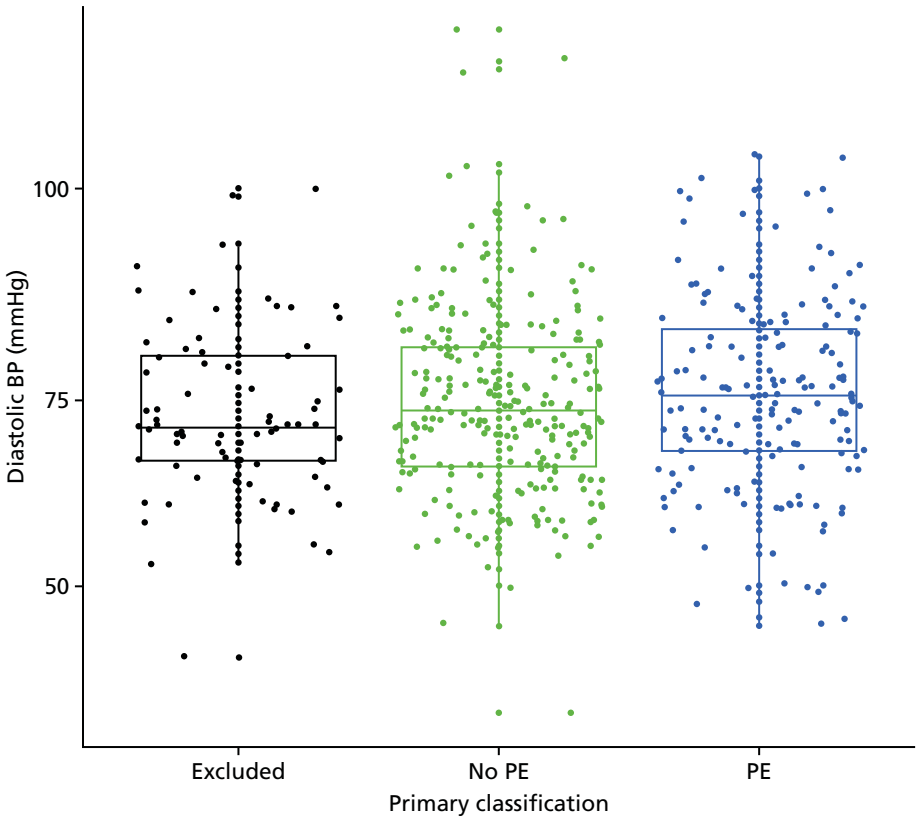


FIGURE 16 Distribution of diastolic blood pressure by classification in the primary analysis. BP, blood pressure.

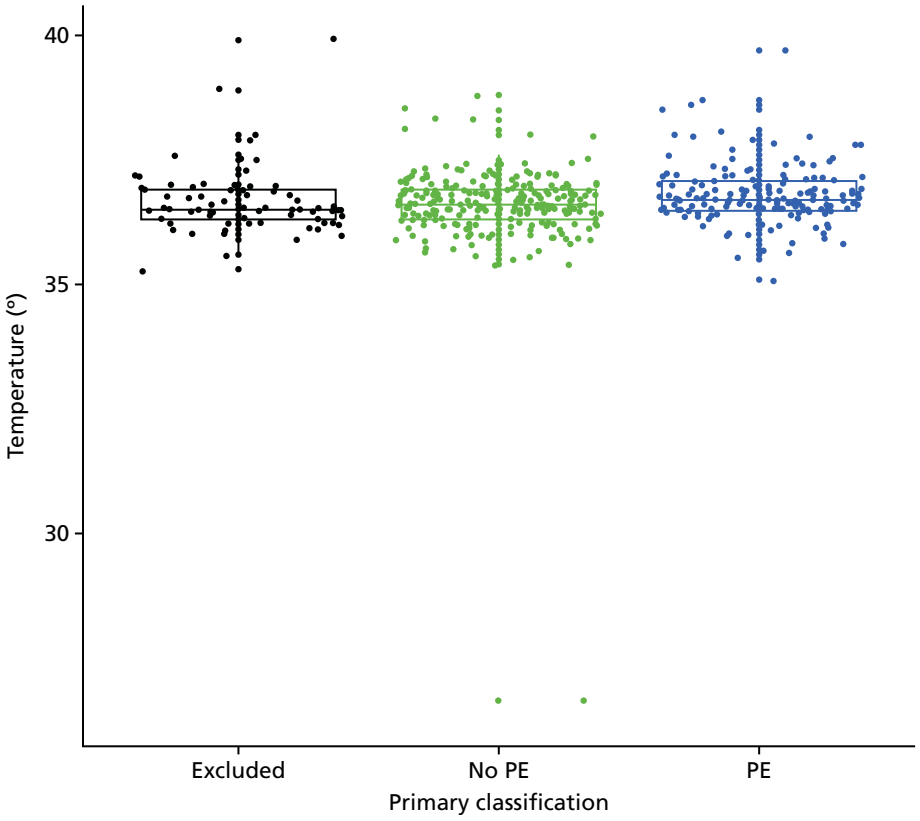


FIGURE 17 Distribution of temperature by classification in the primary analysis.

Appendix 6 Diagnostic performance of the clinical decision rules

Figures 18–24 show the diagnostic performance of each CDR across its range of possible scores.

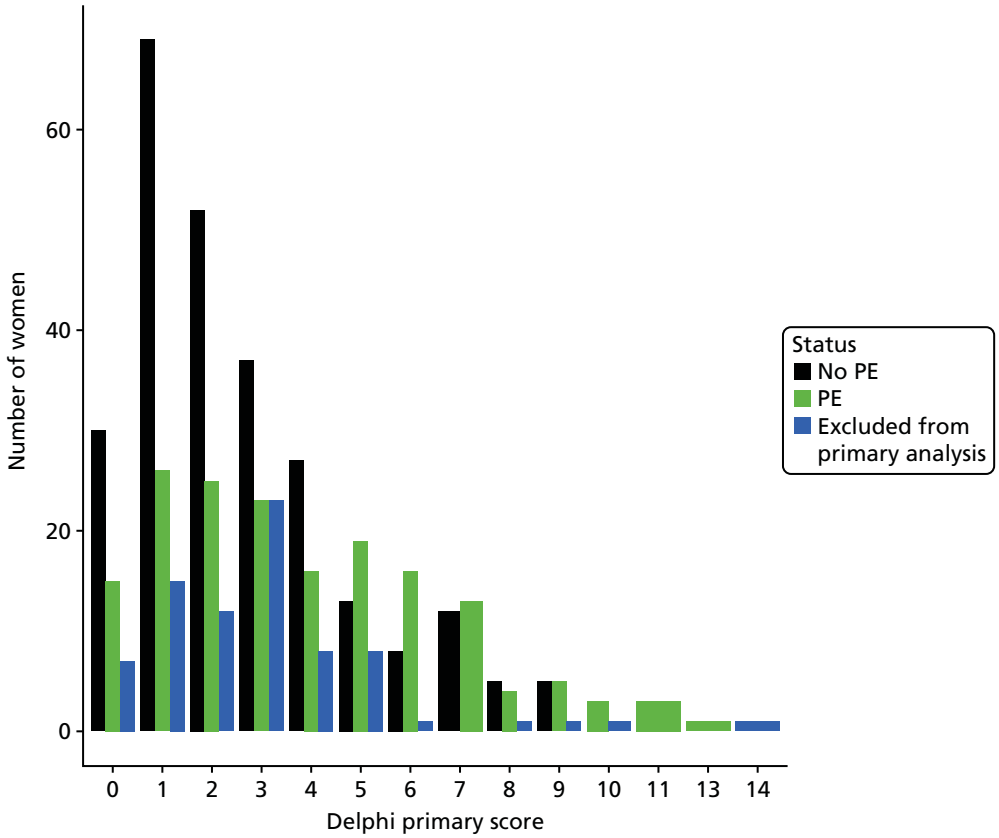


FIGURE 18 Number of women with and without PE and excluded (primary analysis) by the primary consensus rule score.

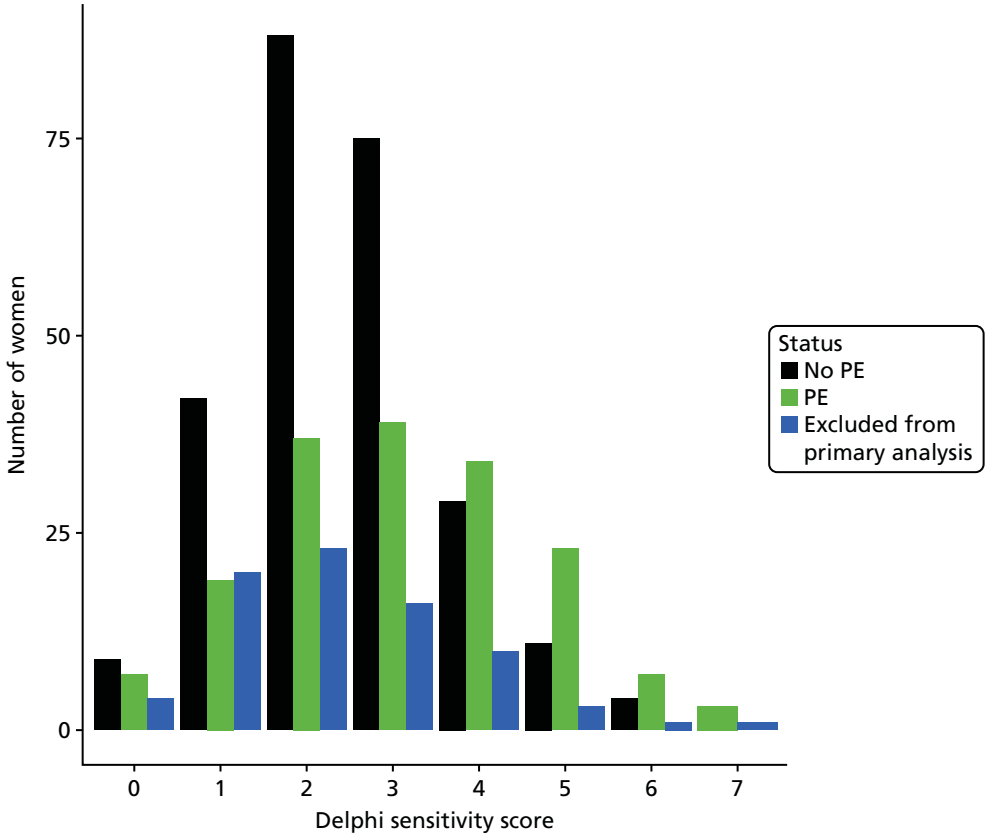


FIGURE 19 Number of women with and without PE and excluded (primary analysis) by the sensitive consensus rule score.

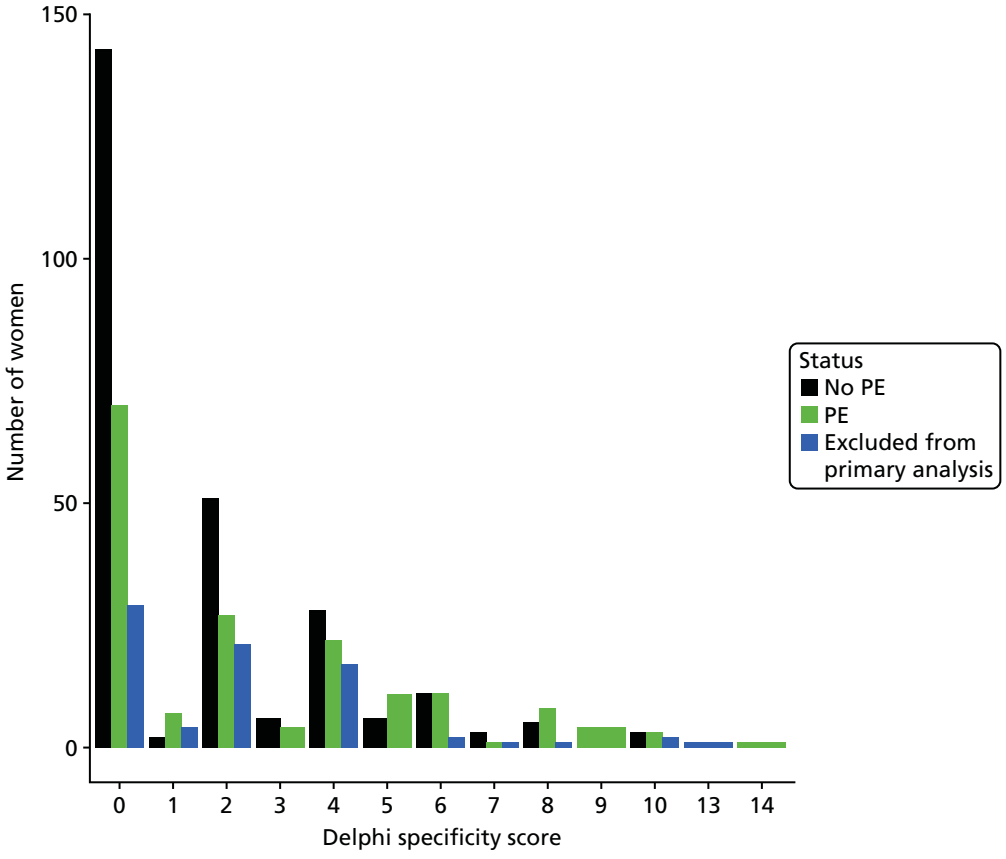


FIGURE 20 Number of women with and without PE and excluded (primary analysis) by the specific consensus rule score.

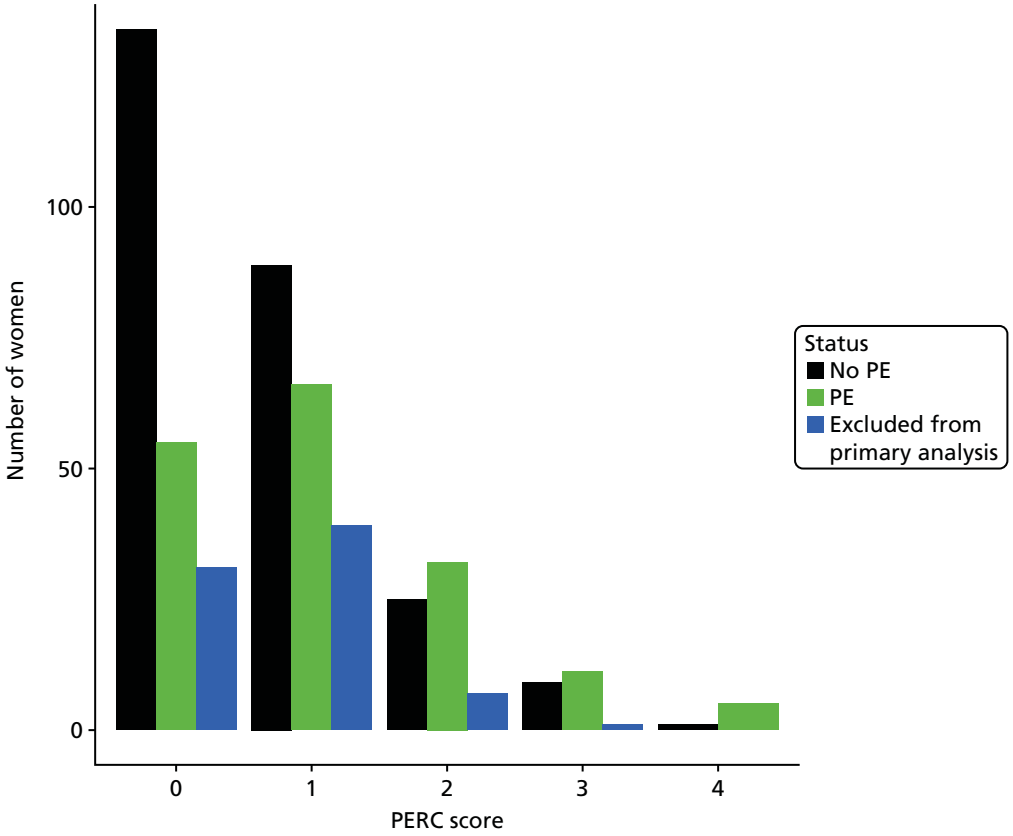


FIGURE 21 Number of women with and without PE and excluded (primary analysis) by the PERC score.

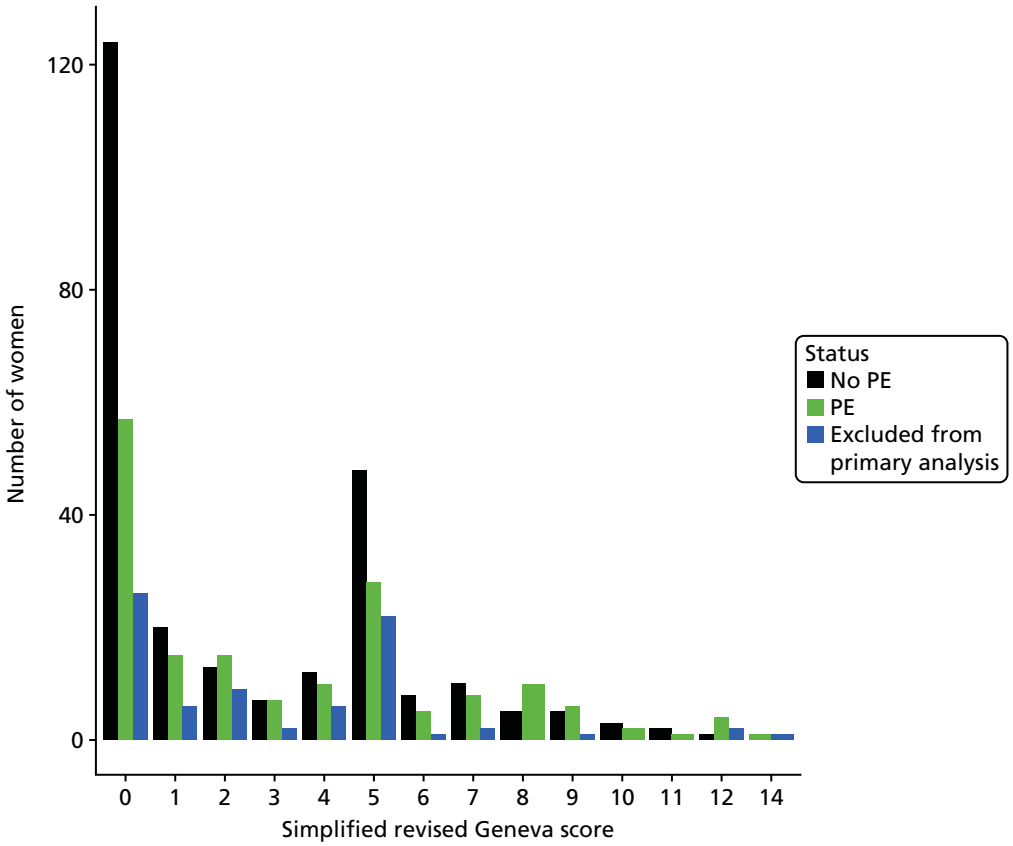


FIGURE 22 Number of women with and without PE and excluded (primary analysis) by the simplified Geneva score.

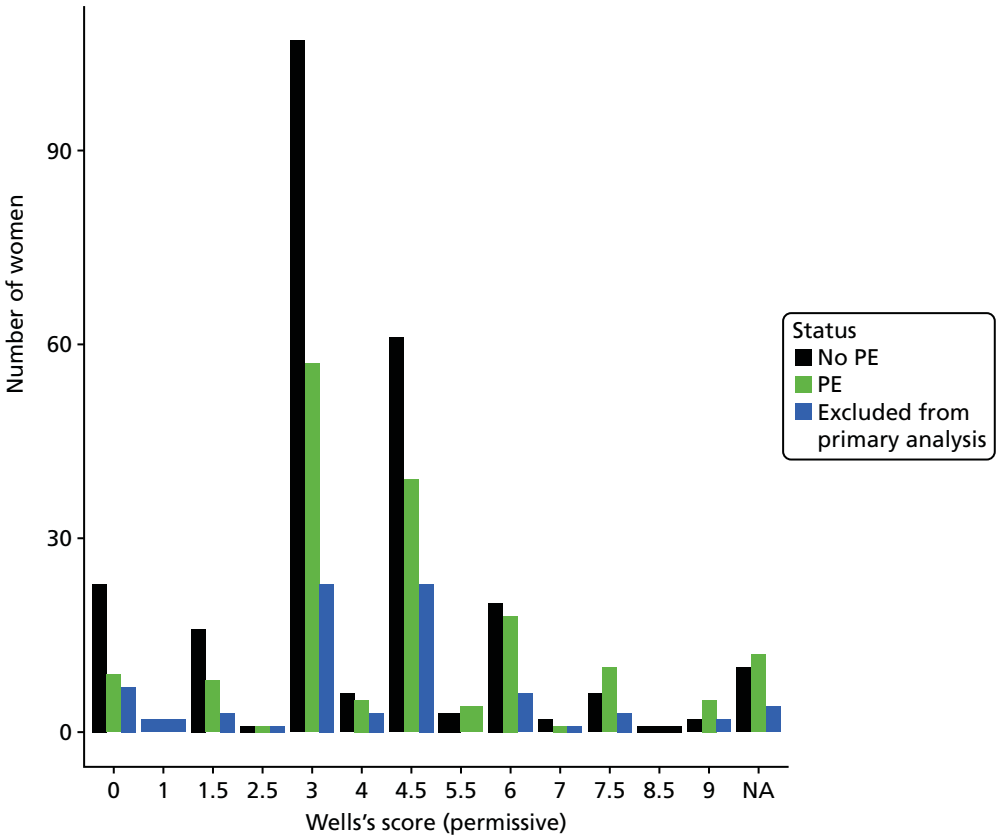


FIGURE 23 Number of women with and without PE and excluded (primary analysis) by Wells's score (permissive application of PE likely). NA, not applicable.

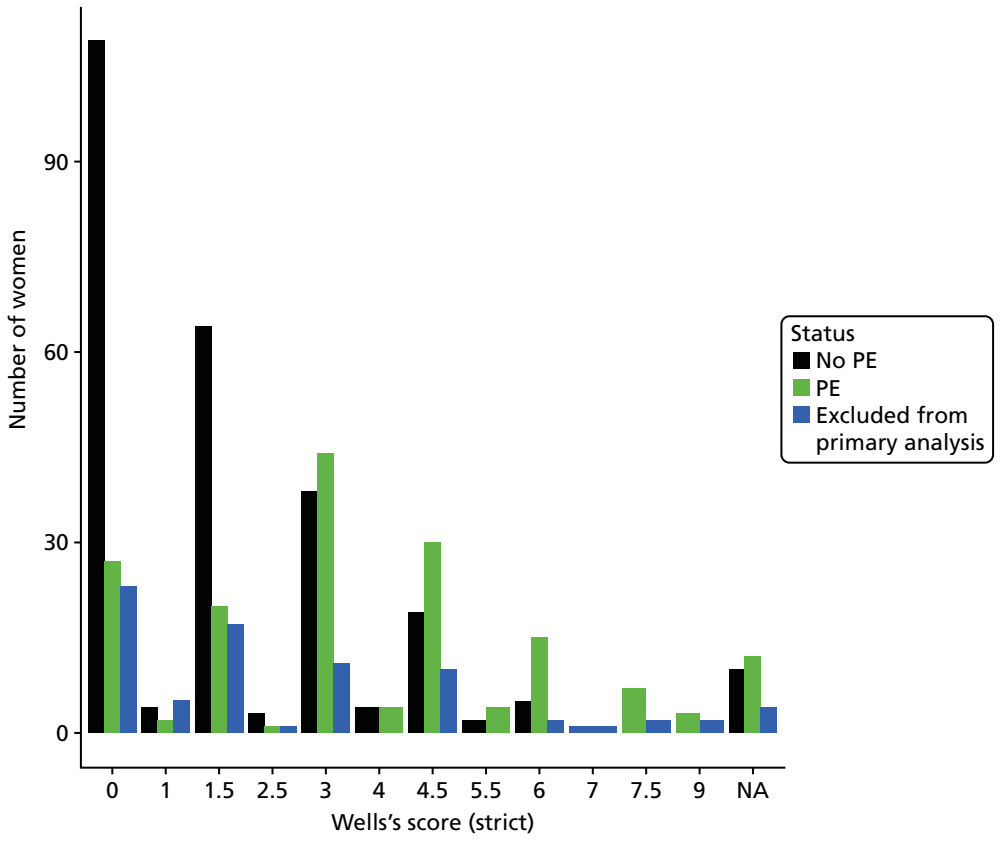


FIGURE 24 Number of women with and without PE and excluded (primary analysis) by Wells's score (strict application of PE likely). NA, not applicable.

Appendix 7 Contributions of the individual elements of clinical decision rules

Tables 34–37 show the odds ratio and *p*-value for each of the individual elements of each CDR.

TABLE 34 Contributions of the individual elements of the consensus-derived CDRs

Criterion	Odds ratio	95% CI	*p*-value
Haemoptysis	1.93	0.83 to 4.61	0.129
Pleuritic chest pain	0.96	0.66 to 1.41	0.842
Previous VTE	1.91	0.95 to 3.92	0.073
Family history of VTE in a first-degree relative	0.71	0.41 to 1.20	0.205
Surgery other than caesarean section	11.1	1.95 to 209	0.025
Significant injury	0.95	0.12 to 5.81	0.959
Obstetric complication[a]	4.12	2.68 to 6.41	< 0.001
Active medical comorbidities[b]	1.02	0.26 to 2.62	0.978
Third trimester	0.69	0.33 to 1.46	0.324
Post partum	1.91	0.89 to 4.19	0.101
Raised BMI of ≥ 30 kg/m^2	1.01	0.68 to 1.52	0.942
Clinical symptoms of DVT	2.19	0.61 to 8.65	0.231
Clinical signs of DVT	1.49	0.81 to 2.77	0.199
Oxygen saturation of < 94% on room air	4.37	2.12 to 9.71	< 0.001
Tachycardia of > 100 b.p.m. (in first or second trimester, or post partum)/tachycardia > 110 b.p.m. (in third trimester)	1.13	0.75 to 1.72	0.556
Increased respiratory rate of > 24 breaths per minute	1.03	0.54 to 1.95	0.919

b.p.m., beats per minute.

a Obstetric complications: apply once if any of the following are present – pre-eclampsia in current pregnancy, assisted reproductive technology/in vitro fertilisation (antenatal only), multiple pregnancy, caesarean section in labour, elective caesarean section, mid-cavity or rotational operative delivery, prolonged labour (> 24 hours), postpartum haemorrhage (> 1 litre or transfusion), preterm birth < 37^{+0} weeks in current pregnancy, stillbirth in current pregnancy, hyperemesis, ovarian hyperstimulation syndrome (first trimester only).

b Active medical comorbidities: apply once if any of the following are present – cancer, heart failure; systemic lupus erythematosus, inflammatory polyarthropathy or inflammatory bowel disease; nephrotic syndrome; type I or type 2 diabetes mellitus with nephropathy; sickle cell disease.

TABLE 35 Contributions of the individual elements of the pregnancy-adapted PERC score

Criterion	Odds ratio	95% CI	*p*-value
Aged > 35 years	1.41	0.86 to 2.31	0.176
Heart rate of > 100 b.p.m. (first or second trimester or post partum) or > 110 b.p.m. in the third trimester	1.13	0.75 to 1.72	0.556
Oxygen saturation of < 94% on room air	4.37	2.12 to 9.71	< 0.001
Prior history of DVT/PE	1.91	0.95 to 3.92	0.073
Recent trauma	0.95	0.12 to 5.81	0.959
Recent surgery	2.72	1.53 to 4.92	0.001
Haemoptysis	1.93	0.83 to 4.61	0.129
Unilateral leg swelling (clinical signs of DVT)	3.08	0.29 to 6.65	0.360

b.p.m., beats per minute.

TABLE 36 Contributions of the individual elements of the pregnancy-adapted Wells's criteria

Criterion	Odds ratio	95% CI	*p*-value
Clinical symptoms of DVT	2.19	0.61 to 8.65	0.231
Clinical signs of DVT	1.49	0.81 to 2.77	0.199
PE is the most likely diagnosis OR equally likely (strict)	9.15	5.70 to 15.0	< 0.001
PE is the most likely diagnosis OR equally likely (permissive)	1.48	0.85 to 2.62	0.174
Heart rate of > 100 b.p.m. (first or second trimester or post partum) or > 110 b.p.m. in the third trimester	1.13	0.75 to 1.72	0.556
Immobilisation for at least 3 days	0.95	0.46 to 1.91	0.887
Surgery in the previous 4 weeks	2.72	1.53 to 4.92	0.001
Previous objectively diagnosed PE or DVT	1.91	0.95 to 3.92	0.073
Haemoptysis	1.93	0.83 to 4.61	0.129
Malignancy with treatment within 6 months or palliative	5.99	0.32 to 112	0.983

b.p.m., beats per minute.

TABLE 37 Contributions of the individual elements of the pregnancy-adapted simplified revised Geneva score

Criterion	Odds ratio	95% CI	*p*-value
Aged > 35 years	1.41	0.86 to 2.31	0.176
Previous DVT or PE	1.91	0.95 to 3.92	0.073
Surgery in the past month	2.72	1.53 to 4.92	0.001
Lower limb fracture in the past month (significant injury)	0.95	0.12 to 5.81	0.959
Active malignant condition	5.99	0.32 to 112.0	0.983
Unilateral lower limb pain	2.19	0.61 to 8.65	0.231
Haemoptysis	1.93	0.83 to 4.61	0.129
Heart rate of > 100 b.p.m. (first or second trimester or post partum) or > 110 b.p.m. in the third trimester	1.13	0.75 to 1.72	0.556
Pain on limb palpation (clinical signs of DVT)	1.49	0.58 to 2.41	0.199

b.p.m., beats per minute.

Appendix 8 Details of the multivariable analysis

The LASSO is a method of automated selection of covariates/predictor variables that maximises the accuracy of the model without inflating the estimated coefficients for each variable. The R package glmnet was used to fit a model using the LASSO. Exactly what predictors are included/used in a model varies, as LASSO regression seeks to select a subset of variables. *Figure 25* shows the change in the coefficients for each predictor variable over iterations of estimation via LASSO. L1 norm is a constraint placed on the analysis for the sum of all coefficients and is one of the unique features of the LASSO.

Cross-validation

Leave-one-out cross-validation has been used internally at each step of the fitting of the LASSO to shrink the point estimate and inform the next iteration. The model should not be overfitted though, and this can be achieved by considering the parameter lambda and either taking the minimum value or the value corresponding to 1 × SE of the point estimate of the mean squared error, which are the dashed lines on *Figure 26*.

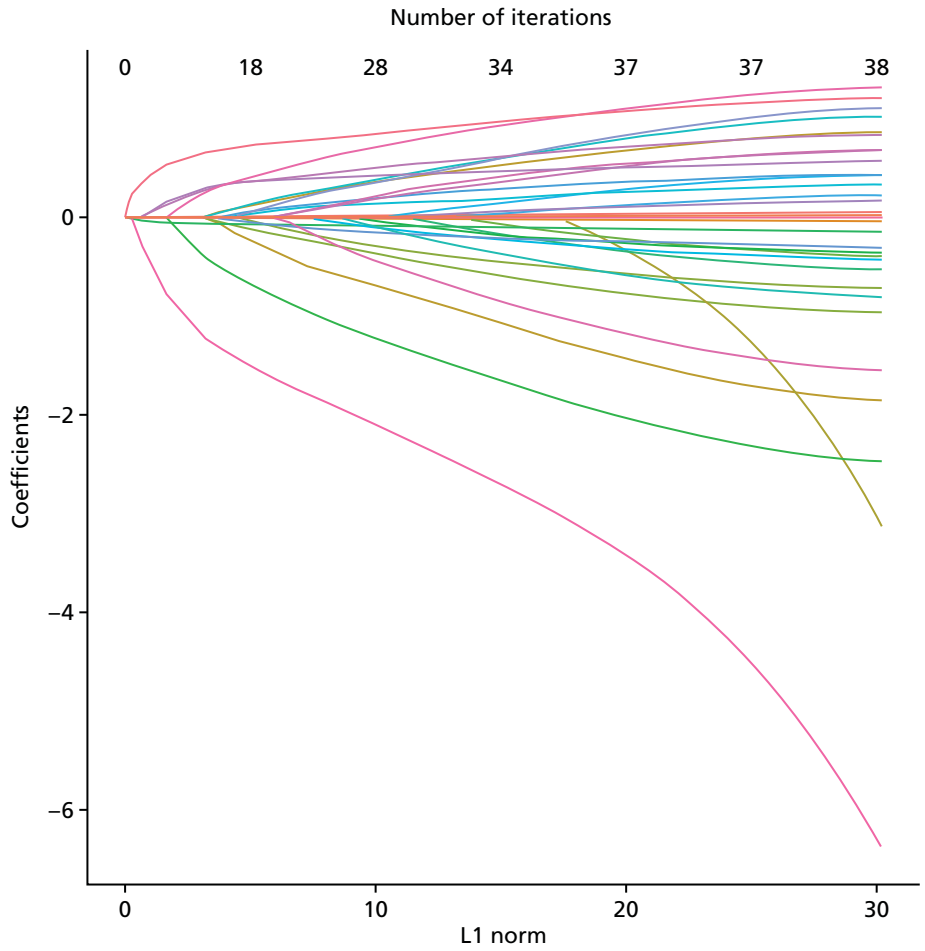


FIGURE 25 Change in the coefficients for each predictor variable over iterations of estimation via LASSO.

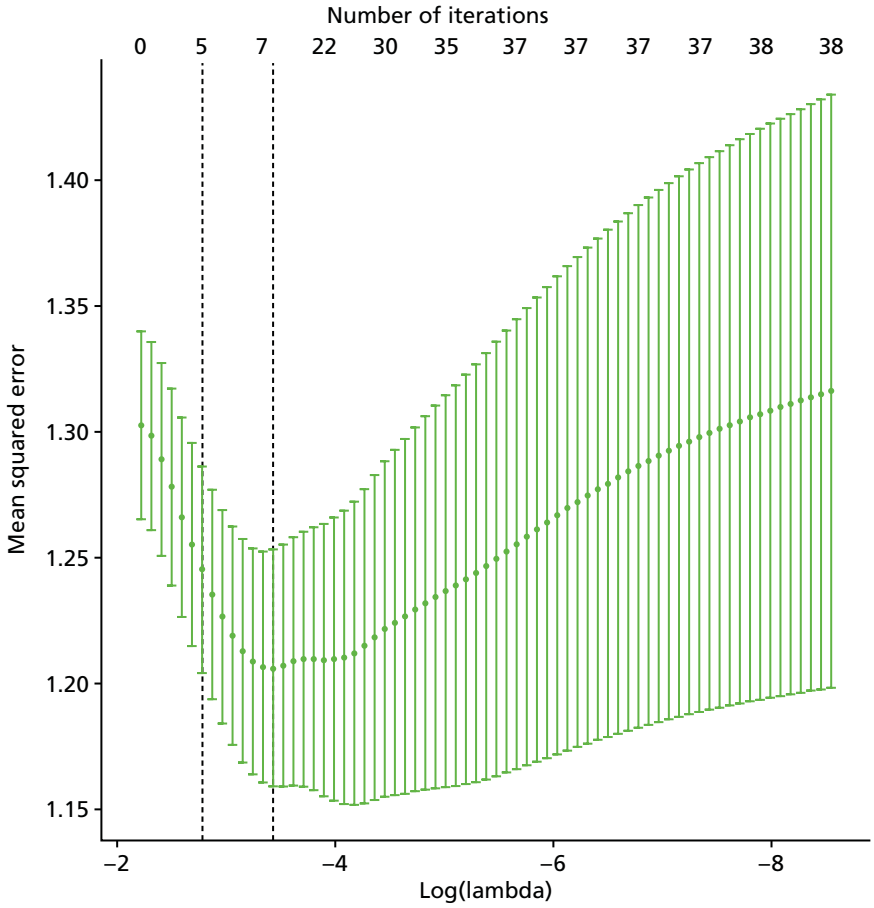


FIGURE 26 Leave-one-out cross-validation.

Receiver operating characteristic curve

Having selected an optimal value for lambda, the predicted probabilities are then obtained and ROC curves are plotted along with the calculated AUC statistic, as shown in *Figure 27*. A cut-off point for probability can be chosen, but in the absence of any choice, a default of $p = 0.5$ has been used in order to calculate the various performance metrics shown in *Table 38*. *Table 39* shows the coefficients for the lambda thresholds.

Trade-off between complexity and predictive value

The question naturally arises as to whether or not a parsimonious model with fewer variables from an earlier stage in the LASSO can be used without losing the predictive ability of the model. It is self-evident that simpler models will have poorer predictive value because they use less information in order to make the prediction but, in order to provide a visual overview and numerical quantification of the trade-off between complexity and predictive value, *all* steps from the LASSO are plotted as ROC curves in *Figure 28* and goodness-of-fit statistics have been calculated for each sequential step.

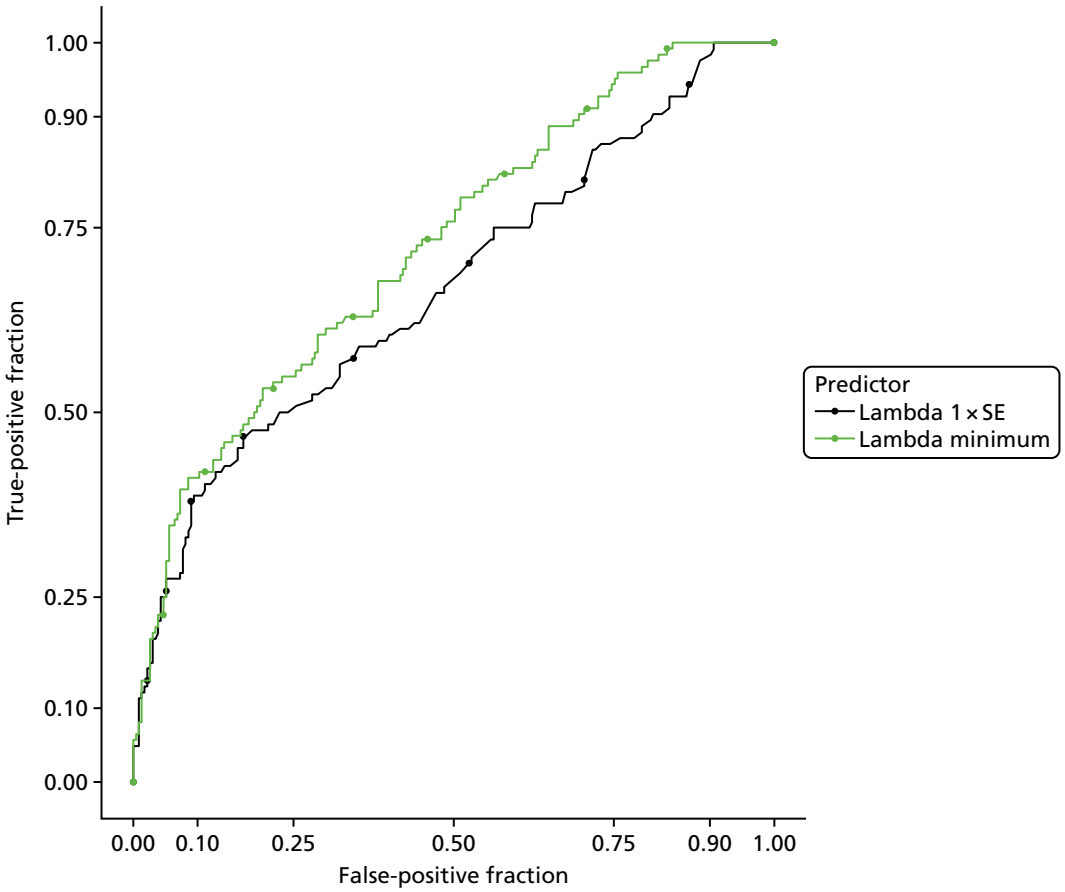


FIGURE 27 Receiver operating characteristic curves for the leave-one-out cross-validation.

TABLE 38 Predictive statistics for the lambda thresholds

Term	True positive	True negative	False positive	False negative	AUC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Lambda 1 × SE	124	18	215	0	0.668 (0.607 to 0.729)	1.000 (0.971 to 1.000)	0.077 (0.046 to 0.119)
Lambda minimum	103	91	142	21	0.724 (0.669 to 0.779)	0.831 (0.753 to 0.892)	0.391 (0.328 to 0.456)

TABLE 39 Coefficients for lambda thresholds

Term	1 × SE	Minimum
(Intercept)	1.915	–3.987
≥ 3 days of immobility/bed rest during pregnancy	0.000	0.000
Age (continuous)	0.000	0.000
BMI (continuous)	0.000	0.000
Cough	0.000	0.000
Diastolic (continuous)	0.000	0.000
dvt.cat	0.000	0.013

continued

TABLE 39 Coefficients for lambda thresholds (*continued*)

Term	1 × SE	Minimum
ECG.PE	0.000	0.000
Family history of thrombosis	0.000	0.000
Haemoptysis	0.000	0.000
Heart rate (continuous)	0.000	0.000
History of i.v. drug use	0.000	0.000
History of thrombosis	0.000	0.256
History of varicose veins	0.000	0.000
Injury	0.000	0.000
Known thrombophilia	0.000	0.000
Long-haul travel during pregnancy	–0.428	–1.225
medical.probs	0.000	0.000
Multiple pregnancy	0.000	–0.402
Non-pleuritic	0.000	0.000
O_2 saturation (continuous)	–0.041	–0.065
Other	0.000	0.000
Other problem with this pregnancy (VTE related)	0.000	0.000
Palpitations	0.000	0.000
Pleuritic	0.000	0.000
Pregnancies < 24 weeks' gestation (continuous)	0.000	0.000
Pregnancies > 24 weeks' gestation (continuous)	0.000	0.000
Problems with this pregnancy (including other)	0.000	0.006
respiratory.rate	0.000	0.000
SOB (exertion)	0.000	0.000
SOB (rest)	0.000	0.000
Smoking (as recorded)	0.000	0.000
Surgery in previous 4 weeks	0.028	0.299
Syncope	0.000	0.000
Systolic (continuous)	0.000	0.000
Temperature (continuous)	0.037	0.273
thromb.event	0.000	0.000
Trimester	0.000	0.000
X-ray: normal.PE	0.413	0.660

i.v., intravenous; SOB, shortness of breath.

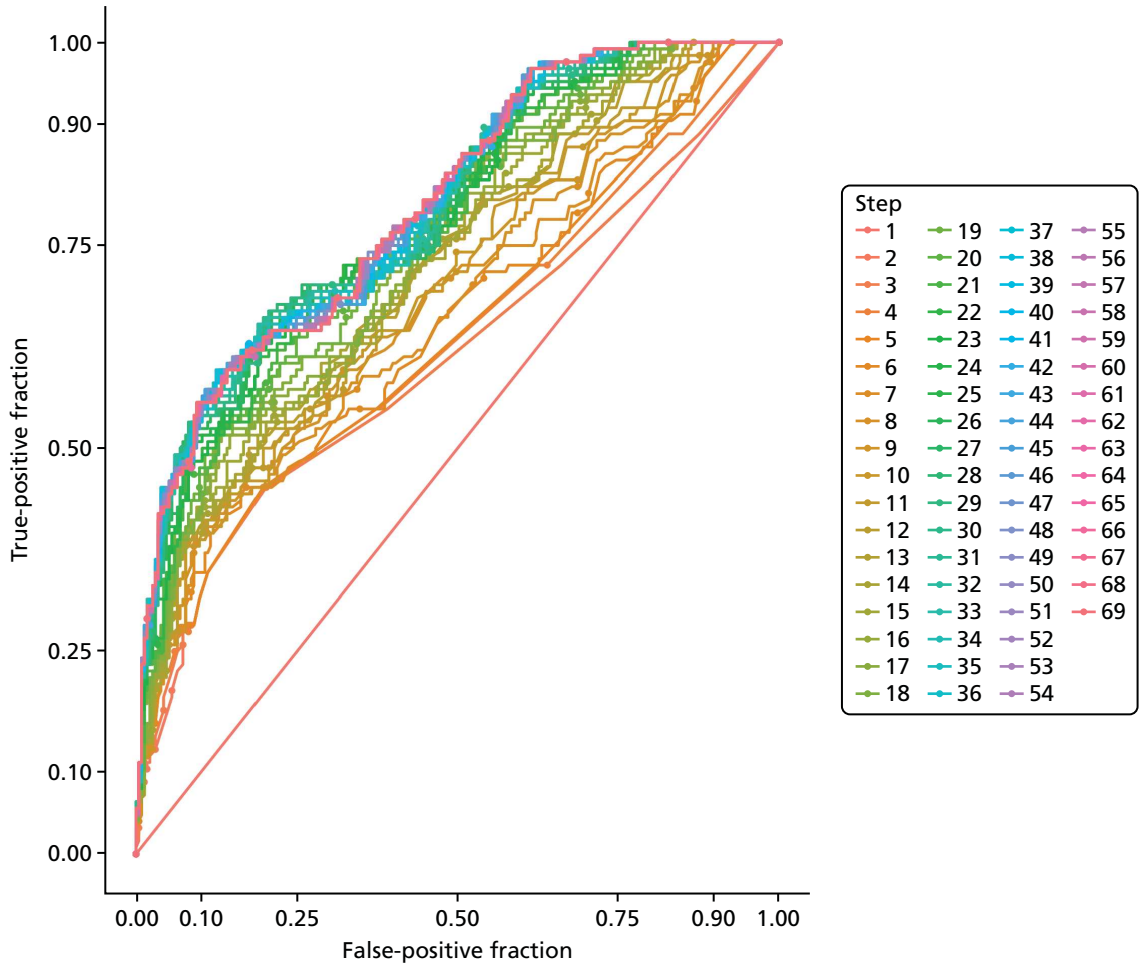


FIGURE 28 Receiver operating characteristic for all steps of the cross-validated LASSO.

Appendix 9 Recursive partitioning

Recursive partitioning is a method of automatically selecting variables and, when continuous, cut-off points within the range of a given variable, which maximise the classification of individuals. The R package rpart was used to fit a model using recursive partitioning. *Table 40* explains the definitions and terms used in recursive partitioning. There are a number of control parameters that are used in running a recursive partitioning model, and these are described in *Table 41*.

Table 42 shows the summary table for the full model. This model, because it was forced to fit a full model that categorised everyone, is overfitted, meaning that its generalisability and application in individuals not in the cohort will be poor. To improve the generalisability of the model, we pruned the tree by selecting a more permissive value for the complexity parameter (cp). It is recommended that a decision tree is pruned using the complexity parameter that corresponds to the minimum cross-validated error (xerror in the table below[48]). The complexity parameter (cp), along with the associated cross-validated error (xerror), is included in *Table 42* and a plot of the two parameters is shown in *Figure 29*. The complexity parameter (cp) at the step/number of splits (0.018) that corresponds to the minimum cross-validated error is 0.018.

TABLE 40 Definitions and terms in recursive partitioning

Term	Definition
Leave-one-out cross-validation	Observations are dropped, one at a time from the data set, the model fitted and the excluded persons outcome predicted. This is repeated for each observation
Overfitting	Trees produced that classify all people are too specific to the data set and will not be useful in predicting outcomes in new patients
Pruning	The process of trimming back an overfitted tree using the complexity parameter
Node	A split in the partitioning tree
Complexity parameter	A metric that quantifies the reduction in error afforded by a given split. As successive splits are made, the reduction in error diminishes
Minimum bucket	A control parameter for fitting trees, which forces each split to have a minimum number of observations classified to each node

TABLE 41 Control parameters used in fitting models with rpart

Control parameter	Value	Explanation
minsplit	4	The minimum number of observations that must exist in a node in order for a split to be attempted
minbucket	2	The mimimum number of observations in any terminal node
cp	–1	The complexity parameter; a negative value ensures that a full model is fitted when everyone is classified

TABLE 42 Summary table for the overfitted full model

Complexity parameter	Splits	Relative error	Cross-validated error	Cross-validated SD
0.118	0	1.000	1.000	0.060
0.053	1	0.882	0.976	0.060
0.047	2	0.828	0.988	0.060
0.024	3	0.781	0.893	0.058
0.022	4	0.757	0.893	0.058
0.018	7	0.692	0.899	0.059
0.012	11	0.621	0.959	0.059
0.009	20	0.509	0.935	0.059
0.008	33	0.337	0.976	0.060
0.006	42	0.249	0.982	0.060
0.003	60	0.130	1.012	0.060
0.000	62	0.124	1.018	0.060
–1.000	78	0.124	1.018	0.060

SD, standard deviation.

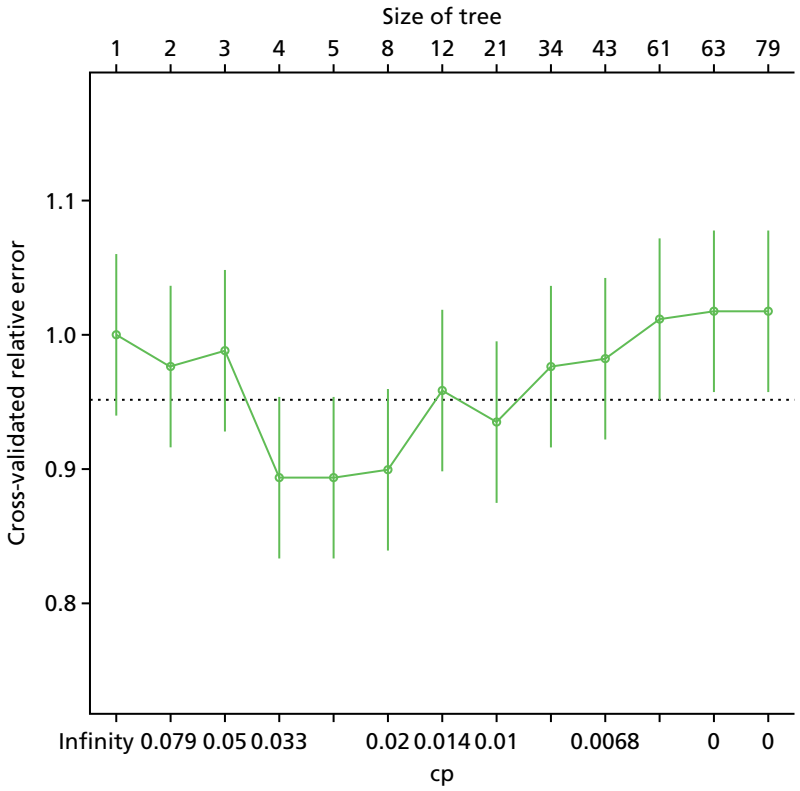


FIGURE 29 Complexity parameter along with the associated cross-validated error.

We now needed to calculate the sensitivity and specificity. There was no single value for either of these metrics, as individuals have a predicted probability of classification in the range of $0 < p < 1$, rather than a binary classification. The trade-off between sensitivity and specificity is shown in the ROC curve (*Figure 30*).

The question arises, however, as to whether or not it is possible to utilise a simpler, more parsimonious tree with fewer splits, but still retain the ability to make useful and accurate predictions. To this end, the ROC curve for each step/split in the recursive partitioning process is plotted in *Figure 31*, and *Table 43* provides the performance statistics for each step. Dichotomising individuals' predicted probability of disease is required in order to calculate the sensitivity and specificity. For now, a cut-off point of $p = 0.05$ has been used, but this is unlikely to be optimal for any of the trees.

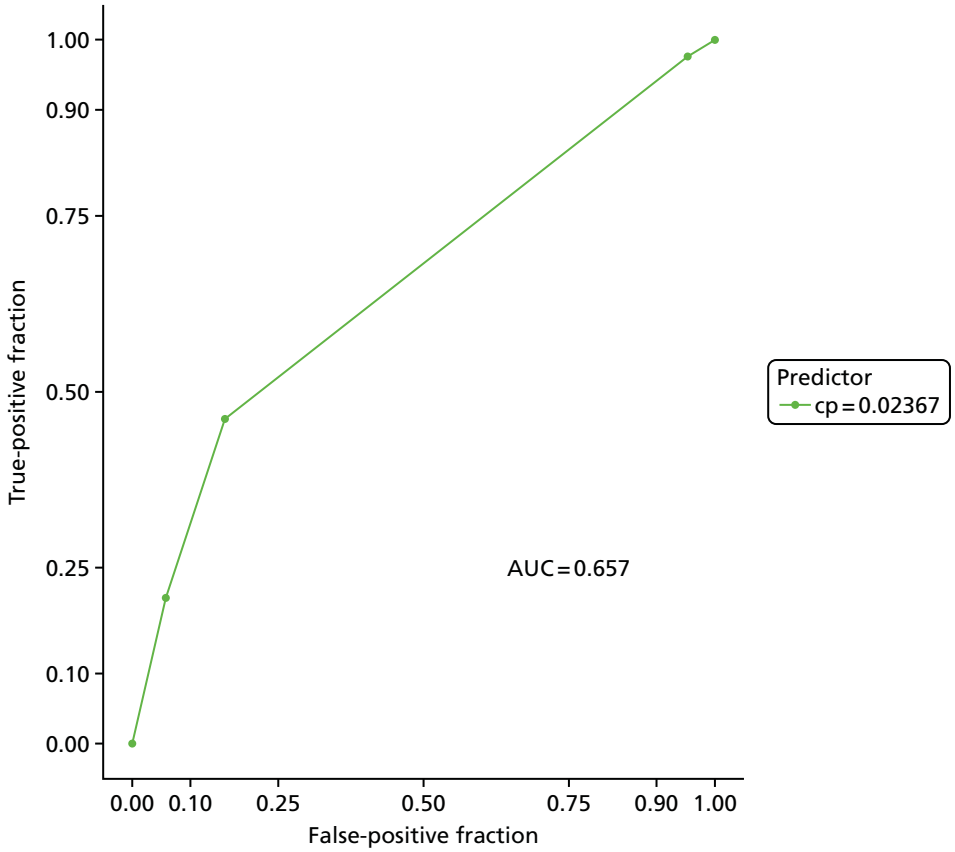


FIGURE 30 Receiver operating characteristic curves for the pruned tree minimising the cross-validation error.

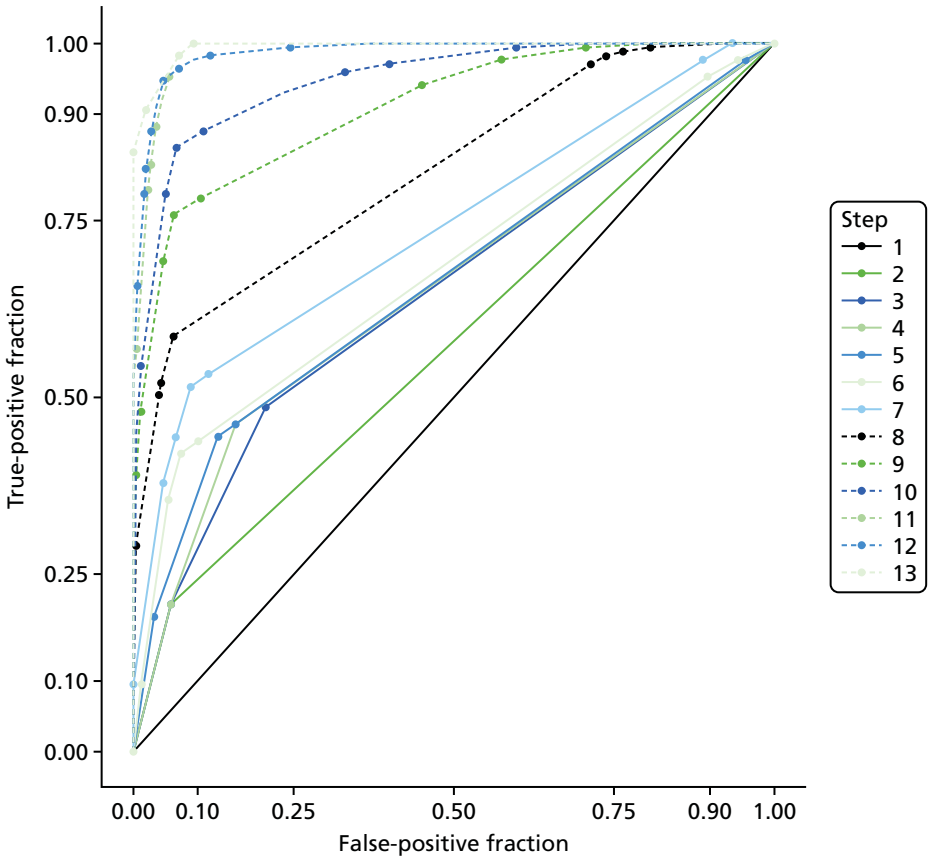


FIGURE 31 Receiver operating characteristic curves for each pruned tree.

TABLE 43 Predictive statistics for the sequential splits of recursive partitioning, using a cut-off point of $p = 0.5$

Term	AUC	AUC lower CI to AUC upper CI	Sensitivity	Sensitivity lower CI to sensitivity upper CI	Specificity	Specificity lower CI to specificity upper CI
1	0.500	0.500 to 0.500	0.000	0.000 to 0.022	1.000	0.986 to 1.000
2	0.574	0.541 to 0.608	0.207	0.149 to 0.276	0.942	0.906 to 0.967
3	0.647	0.601 to 0.693	0.485	0.408 to 0.563	0.795	0.740 to 0.842
4	0.657	0.611 to 0.703	0.462	0.385 to 0.540	0.841	0.791 to 0.883
5	0.664	0.619 to 0.710	0.444	0.368 to 0.522	0.868	0.821 to 0.907
6	0.684	0.639 to 0.729	0.420	0.345 to 0.498	0.926	0.887 to 0.955
7	0.738	0.695 to 0.781	0.515	0.437 to 0.592	0.911	0.869 to 0.943
8	0.819	0.784 to 0.855	0.586	0.508 to 0.661	0.938	0.901 to 0.964
9	0.909	0.882 to 0.937	0.757	0.686 to 0.820	0.938	0.901 to 0.964
10	0.949	0.929 to 0.968	0.852	0.789 to 0.902	0.934	0.897 to 0.961
11	0.985	0.977 to 0.994	0.953	0.909 to 0.979	0.946	0.911 to 0.970
12	0.987	0.979 to 0.995	0.947	0.901 to 0.975	0.953	0.920 to 0.976
13	0.995	0.991 to 0.998	0.982	0.949 to 0.996	0.930	0.892 to 0.958

Appendix 10 Results of the secondary analyses

Table 44 shows the *p*-values for the univariable analysis. There was little difference between the analyses. Some variables were significant predictors ($p < 0.05$) on the secondary analysis that were not significant predictors on the primary analysis, for example giving up smoking during the pregnancy, history of varicose veins and previous VTE, but this simply reflected the *p*-value moving to the other side of the 0.05 threshold, presumably because of a slightly larger sample size.

TABLE 44 Primary and secondary univariable analysis (*p*-values)

Variable	Analysis			
	Primary	Secondary, with clinically diagnosed PE	Secondary, with clinically ruled-out PE	Secondary, with subsegmental PE excluded
Age, years (continuous)	0.179	0.307	0.166	0.187
Aged over 35 years	0.176	0.247	0.177	0.256
BMI, kg/m^2 (continuous)	0.372	0.587	0.256	0.253
BMI of $\geq$ 30 kg/m^2	0.942	0.631	0.775	0.781
Ex-smoker (prior)	0.837	0.901	0.842	0.792
Ex-smoker (during)	0.082	0.082	0.035	0.080
Current smoker	0.279	0.159	0.318	0.354
Pregnancies lasting for < 24 weeks' gestation (continuous)	0.509	0.386	0.679	0.440
One or more pregnancy lasting for < 24 weeks' gestation	0.980	0.937	0.776	0.883
Pregnancies lasting for > 24 weeks' gestation (continuous)	0.017	0.014	0.014	0.017
One or more pregnancy lasting for > 24 weeks' gestation	0.198	0.243	0.217	0.224
Previous pregnancy problems	0.442	0.888	0.252	0.493
Family history of VTE	0.205	0.125	0.324	0.160
History of varicose veins	0.045	0.084	0.033	0.056
History of i.v. drug use	0.800	0.910	0.719	0.775
Known thrombophilia	0.745	0.993	0.934	0.801
Surgery in the previous 4 weeks	0.001	0.003	0.000	0.000
Injury in the previous 4 weeks	0.959	0.845	0.910	0.998
Previous VTE	0.073	0.092	0.025	0.081
Other previous medical problem	0.829	0.989	0.760	0.955
Other previous medical problem (VTE related)	0.941	0.941	0.941	0.941
Second trimester	0.234	0.573	0.064	0.185
Third trimester	0.324	0.635	0.196	0.283
Post partum	0.101	0.105	0.168	0.142
Multiple pregnancy	0.191	0.097	0.299	0.220
Long-haul travel during pregnancy	0.006	0.002	0.011	0.007

continued

TABLE 44 Primary and secondary univariable analysis (*p*-values) (*continued*)

Variable	Analysis			
	Primary	Secondary, with clinically diagnosed PE	Secondary, with clinically ruled-out PE	Secondary, with subsegmental PE excluded
≥ 3 days of immobility/bed rest during pregnancy	0.887	0.988	0.970	0.995
Received thromboprophylaxis	0.000	0.000	0.000	0.000
Previous thrombotic event this pregnancy	0.226	0.085	0.157	0.203
Other problem with this pregnancy	0.006	0.095	0.003	0.006
Other problem with this pregnancy (VTE related)	0.713	0.815	0.903	0.772
Presenting: pleuritic chest pain	0.842	0.774	0.982	0.767
Presenting: non-pleuritic chest pain	0.457	0.391	0.464	0.405
Presenting: SOB (exertion)	0.520	0.895	0.796	0.317
Presenting: SOB (rest)	0.142	0.135	0.104	0.239
Presenting: haemoptysis	0.129	0.109	0.320	0.103
Presenting: cough	0.988	0.913	0.772	0.940
Presenting: syncope	0.218	0.090	0.189	0.291
Presenting: palpitations	0.598	0.424	0.521	0.595
Presenting: other	0.914	0.771	0.888	0.989
Temperature of > 37.5 °C	0.020	0.017	0.022	0.015
Temperature, °C (continuous)	0.003	0.002	0.001	0.003
Diastolic BP of < 50 mmHg	0.221	0.195	0.162	0.202
Diastolic BP (continuous)	0.256	0.503	0.210	0.265
Systolic BP of < 90 mmHg	0.205	0.164	0.162	0.191
Systolic (continuous)	0.322	0.137	0.480	0.350
Oxygen saturation of < 94%	0.000	0.001	0.000	0.000
Oxygen saturation, % (continuous)	0.000	0.001	0.000	0.000
Respiratory rate of > 24 breaths per minute	0.919	0.956	0.722	0.798
Respiratory rate (continuous)	0.948	0.841	0.592	0.869
Heart rate of > 100 b.p.m. (110 b.p.m. in the third trimester)	0.556	0.618	0.615	0.524
Heart rate, b.p.m. (continuous)	0.126	0.063	0.126	0.084
Clinical signs of DVT	0.199	0.104	0.418	0.274
PE-related ECG abnormality	0.579	0.580	0.760	0.631
PE-related chest radiograph abnormality	0.010	0.018	0.007	0.019
Other chest radiograph abnormality	0.001	0.001	0.001	0.001
PE is the most likely diagnosis or equally likely (permissive)	0.156	0.156	0.156	0.156
PE is the most likely diagnosis or equally likely (strict)	0.000	0.000	0.000	0.000

BP, blood pressure; b.p.m., beats per minute; i.v., intravenous; SOB, shortness of breath.

Table 45 shows the AUROC estimates for the CDRs. The results for the secondary analyses were similar to those of the primary analysis.

Table 46 shows the results for D-dimer analysis using the hospital laboratory measurements. The inclusion of women with clinically diagnosed PE and the exclusion of women with subsegmental PE resulted in small changes to sensitivity, whereas the inclusion of women with clinically ruled-out PE resulted in a small change in specificity. None of the estimates in the secondary analysis differed in a meaningful way from the primary analysis.

TABLE 45 Primary and secondary analysis AUC estimates for the CDRs

Decision rule	Analysis			
	Primary	Secondary, with clinically diagnosed PE	Secondary, with clinically ruled-out PE	Secondary, with subsegmental PE excluded
Primary consensus	0.626	0.621	0.626	0.629
Sensitive consensus	0.620	0.599	0.629	0.622
Specific consensus	0.589	0.592	0.582	0.592
PERC score	0.621	0.610	0.619	0.623
Simplified revised Geneva score	0.579	0.575	0.572	0.579
Wells's score (permissive)	0.577	0.580	0.577	0.578
Wells's score (strict)	0.732	0.716	0.728	0.731

TABLE 46 Primary and secondary analysis of D-dimer (hospital) sensitivity and specificity

Threshold	Parameter	Analysis			
		Primary analysis	Secondary, with clinically diagnosed PE	Secondary, with clinically ruled-out PE	Secondary, with subsegmental PE excluded
Standard	Sensitivity	0.884	0.878	0.884	0.878
	95% CI	0.741 to 0.956	0.7445 to 0.945	0.741 to 0.956	0.730 to 0.954
	n/*N*	38/43	43/49	38/43	36/41
	Specificity	0.088	0.088	0.092	0.088
	95% CI	0.047 to 0.156	0.047 to 0.156	0.052 to 0.156	0.047 to 0.156
	n/*N*	11/125	11/125	13/141	11/125
Gestation specific	Sensitivity	0.698	0.694	0.698	0.707
	95% CI	0.537 to 0.823	0.544 to 0.813	0.537 to 0.823	0.543 to 0.833
	n/*N*	30/43	34/49	30/43	29/41
	Specificity	0.328	0.328	0.355	0.328
	95% CI	0.248 to 0.419	0.248 to 0.419	0.277 to 0.440	0.248 to 0.419
	n/*N*	41/125	41/125	50/141	41/125

Tables 47 and *48* show the coefficients for the multivariable models. The 1 × SE models for the secondary analyses included the same five terms as the primary analysis (long-haul travel, oxygen saturation, surgery, temperature and PE-related chest radiograph abnormality), but also included previous VTE when clinically ruled-out PE was included or subsegmental PE was excluded. The models had similar accuracy in the secondary analyses, with AUC estimates of around 0.7. The minimum value models for the secondary analysis included similar terms to the primary analysis, but also included a number of additional terms, especially when clinically diagnosed PE was included. The models had slightly higher accuracy, with AUCs

TABLE 47 Primary and secondary multivariable analysis 1 × SE model

Term or diagnostic parameter	Analysis			
	Primary	Secondary, with clinically diagnosed PE	Secondary, with clinically ruled-out PE	Secondary, with subsegmental PE excluded
(Intercept)	1.915	–1.202	0.523	–0.813
Immobility/bed rest during pregnancy	0.000	0.000	0.000	0.000
Age, years (continuous)	0.000	0.000	0.000	0.000
BMI, kg/m^2 (continuous)	0.000	0.000	0.000	0.000
Cough	0.000	0.000	0.000	0.000
Diastolic BP (continuous)	0.000	0.000	0.000	0.000
Clinical signs of DVT	0.000	0.000	0.000	0.000
PE-related ECG abnormality	0.000	0.000	0.000	0.000
Family history of VTE	0.000	0.000	0.000	0.000
Haemoptysis	0.000	0.000	0.000	0.000
Heart rate, b.p.m. (continuous)	0.000	0.000	0.000	0.000
History of i.v. drug use	0.000	0.000	0.000	0.000
Previous VTE	0.000	0.000	0.015	0.036
History of varicose veins	0.000	0.000	0.000	0.000
Injury	0.000	0.000	0.000	0.000
Known thrombophilia	0.000	0.000	0.000	0.000
Long-haul travel during pregnancy	–0.428	–0.692	–0.432	–0.693
Previous medical problem	0.000	0.000	0.000	0.000
Multiple pregnancy	0.000	0.000	0.000	0.000
Non-pleuritic	0.000	0.000	0.000	0.000
Oxygen saturation, % (continuous)	–0.041	–0.027	–0.055	–0.053
Other presenting complaint	0.000	0.000	0.000	0.000
Other problem with this pregnancy (VTE related)	0.000	0.000	0.000	0.000
Palpitations	0.000	0.000	0.000	0.000
Pleuritic	0.000	0.000	0.000	0.000
Pregnancies lasting for < 24 weeks' gestation (continuous)	0.000	0.000	0.000	0.000
Pregnancies lasting for > 24 weeks' gestation (continuous)	0.000	0.000	0.000	0.000
Problems with this pregnancy	0.000	0.000	0.000	0.000

TABLE 47 Primary and secondary multivariable analysis 1 × SE model (*continued*)

Term or diagnostic parameter	Analysis			
	Primary	Secondary, with clinically diagnosed PE	Secondary, with clinically ruled-out PE	Secondary, with subsegmental PE excluded
Respiratory rate (breaths per minute)	0.000	0.000	0.000	0.000
SOB (exertion)	0.000	0.000	0.000	0.000
SOB (rest)	0.000	0.000	0.000	0.000
Smoking (as recorded)	0.000	0.000	0.000	0.000
Surgery in the previous 4 weeks	0.028	0.000	0.170	0.193
Syncope	0.000	0.000	0.000	0.000
Systolic BP (continuous)	0.000	0.000	0.000	0.000
Temperature, °C (continuous)	0.037	0.099	0.101	0.140
Thrombotic event in this pregnancy	0.000	0.000	0.000	0.000
Trimester	0.000	0.000	0.000	0.000
PE-related chest radiograph abnormality	0.413	0.394	0.525	0.556
AUC (95%CI)	0.668 (0.607 to 0.729)	0.677 (0.621 to 0.732)	0.692 (0.634 to 0.750)	0.708 (0.650 to 0.766)
Sensitivity (95% CI)	1.00 (0.971 to 1.000)	1.000 (0.975 to 1.000)	0.613 (0.521 to 0.699)	0.832 (0.752 to 0.894)
Specificity (95% CI)	0.077 (0.046 to 0.119)	0.077 (0.046 to 0.119)	0.644 (0.584 to 0.702)	0.373 (0.311 to 0.439)

BP, blood pressure; b.p.m., beats per minute; i.v., intravenous; SOB, shortness of breath.

TABLE 48 Primary and secondary multivariable analysis minimum value model

Term or diagnostic parameter	Analysis			
	Primary	Secondary, with clinically diagnosed PE	Secondary, with clinically ruled-out PE	Secondary, with subsegmental PE excluded
(Intercept)	–3.987	–7.262	–5.164	–5.440
Immobility/bed rest during pregnancy	0.000	0.000	0.000	0.000
Age, years (continuous)	0.000	0.000	0.000	0.000
BMI, kg/m^2 (continuous)	0.000	0.000	0.000	0.000
Cough	0.000	0.000	0.000	0.000
Diastolic BP (continuous)	0.000	0.000	0.000	0.000
Clinical signs of DVT	0.013	0.209	0.000	0.000
PE-related ECG abnormality	0.000	–0.453	0.000	0.000
Family history of VTE	0.000	–0.157	0.000	0.000
Haemoptysis	0.000	0.117	0.000	0.061
Heart rate, b.p.m. (continuous)	0.000	0.002	0.000	0.000
History of i.v. drug use	0.000	0.000	0.000	0.000

continued

TABLE 48 Primary and secondary multivariable analysis minimum value model (*continued*)

Term or diagnostic parameter	Analysis			
	Primary	Secondary, with clinically diagnosed PE	Secondary, with clinically ruled-out PE	Secondary, with subsegmental PE excluded
Previous VTE	0.256	0.353	0.421	0.352
History of varicose veins	0.000	–0.108	0.000	0.000
Injury	0.000	0.000	0.000	0.000
Known thrombophilia	0.000	0.000	0.000	0.000
Long-haul travel during pregnancy	–1.225	–1.786	–1.090	–1.248
Previous medical problems	0.000	0.000	0.000	0.000
Multiple pregnancy	–0.402	–0.887	–0.290	–0.429
Non-pleuritic	0.000	0.044	0.000	0.000
Oxygen saturation, % (continuous)	–0.065	–0.056	–0.075	–0.069
Other presenting complaint	0.000	0.000	0.000	0.000
Other problem with this pregnancy (VTE related)	0.000	0.000	0.000	0.000
Palpitations	0.000	0.000	0.000	0.000
Pleuritic	0.000	0.000	0.000	0.000
Pregnancies lasting for < 24 weeks' gestation (continuous)	0.000	0.000	0.000	0.000
Pregnancies lasting for > 24 weeks' gestation (continuous)	0.000	0.000	0.000	0.000
Problems with this pregnancy	0.006	0.000	0.000	0.000
Respiratory rate (breaths per minute)	0.000	0.000	0.000	0.000
SOB (exertion)	0.000	0.000	0.000	0.089
SOB (rest)	0.000	0.000	–0.038	0.000
Smoking (as recorded)	0.000	0.000	0.000	0.000
Surgery in the previous 4 weeks	0.299	0.323	0.379	0.366
Syncope	0.000	0.085	0.000	0.000
Systolic BP (continuous)	0.000	–0.006	0.000	–0.002
Temperature, °C (continuous)	0.273	0.380	0.313	0.318
Thrombotic event in this pregnancy	0.000	0.313	0.000	0.000
Trimester	0.000	0.000	0.000	0.000
PE-related chest radiograph abnormality	0.660	0.664	0.731	0.729
AUC (95%CI)	0.724 (0.669 to 0.779)	0.735 (0.685 to 0.786)	0.731 (0.677 to 0.785)	0.746 (0.692 to 0.800)
Sensitivity (95% CI)	0.831 (0.753 to 0.892)	0.972 (0.931 to 0.992)	0.621 (0.529 to 0.707)	0.748 (0.660 to 0.823)
Specificity (95% CI)	0.391 (0.328 to 0.456)	0.262 (0.207 to 0.323)	0.674 (0.614 to 0.730)	0.545 (0.479 to 0.610)

BP, blood pressure; b.p.m., beats per minute; i.v., intravenous; SOB, shortness of breath.

of up to 0.757. The slightly higher accuracy in the secondary analysis may reflect the fact that the increase in the size of the data set provided more statistical power to develop a more accurate model or this may reflect bias, with the clinical diagnosis or the ruling out of PE being based on variables included in the models.

Accuracy remained short of what would be required from an acceptable model, whereas problems of overfitting and potential bias are likely to be greater than those in the primary analysis.

Table 49 shows the AUROC for each biomarker with a 95% CI. There were no meaningful differences between the analyses.

TABLE 49 Primary and secondary analysis AUROC (95% CI) for each biomarker

Biomarker	Analysis			
	Primary	Secondary, with clinically diagnosed PE	Secondary, with clinically ruled-out PE	Secondary, with subsegmental PE excluded
APTT (minutes)	0.669 (0.570 to 0.768)	0.638 (0.54 to 0.735)	0.681 (0.584 to 0.778)	0.676 (0.576 to 0.777)
BNP level	0.549 (0.453 to 0.645)	0.543 (0.449 to 0.636)	0.549 (0.451 to 0.647)	0.546 (0.448 to 0.645)
CRP level (pg/ml)	0.542 (0.445 to 0.639)	0.557 (0.465 to 0.649)	0.542 (0.447 to 0.636)	0.531 (0.434 to 0.628)
Clauss fibrinogen	0.589 (0.476 to 0.701)	0.559 (0.452 to 0.666)	0.600 (0.489 to 0.712)	0.604 (90.492 to 0.715)
D-dimer (ELISA)	0.668 (0.561 to 0.776)	0.64 (0.537 to 0.743)	0.669 (0.561 to 0.776)	0.679 (0.571 to 0.787)
D-dimer (Innovance)	0.651 (0.545 to 0.758)	0.624 (0.522 to 0.725)	0.655 (0.549 to 0.761)	0.667 (0.562 to 0.772)
MRproANP	0.524 (0.418 to 0.630)	0.523 (0.423 to 0.622)	0.526 (0.422 to 0.631)	0.483 (0.376 to 0.591)
PF 1 + 2	0.562 (0.462 to 0.661)	0.546 (0.453 to 0.638)	0.569 (0.470 to 0.668)	0.556 (0.455 to 0.658)
Plasmin–antiplasmin level	0.639 (0.536 to 0.742)	0.615 (0.518 to 0.712)	0.637 (0.534 to 0.740)	0.633 (0.528 to 0.738)
PT (minutes)	0.613 (0.508 to 0.718)	0.588 (0.488 to 0.688)	0.623 (0.518 to 0.727)	0.623 (0.517 to 0.729)
TG (endogenous potential)	0.559 (0.437 to 0.681)	0.556 (0.448 to 0.665)	0.566 (0.442 to 0.690)	0.566 (0.440 to 0.692)
TG (lag time)	0.702 (0.598 to 0.806)	0.656 (0.551 to 0.761)	0.721 (0.622 to 0.819)	0.704 (0.598 to 0.811)
TG (peak)	0.596 (0.478 to 0.715)	0.569 (0.459 to 0.679)	0.610 (0.492 to 0.729)	0.597 (0.475 to 0.719)
TG (time to peak)	0.655 (0.541 to 0.769)	0.613 (0.503 to 0.723)	0.675 (0.564 to 0.786)	0.657 (0.540 to 0.774)
Tissue factor (pg/ml)	0.531 (0.424 to 0.638)	0.565 (0.464 to 0.666)	0.515 (0.409 to 0.620)	0.518 (0.411 to 0.625)
Troponin level (ng/ml)	0.597 (0.499 to 0.695)	0.559 (0.462 to 0.655)	0.609 (0.514 to 0.704)	0.608 (0.510 to 0.706)

Appendix 11 Details of the biomarker analysis (including women who had received anticoagulation treatment)

Figures 32–47 show, for each biomarker, a box-and-whisker plot comparing biomarker levels for women with DVT, women with PE, women with no PE and women excluded from the primary analysis.

Activated partial thromboplastin time

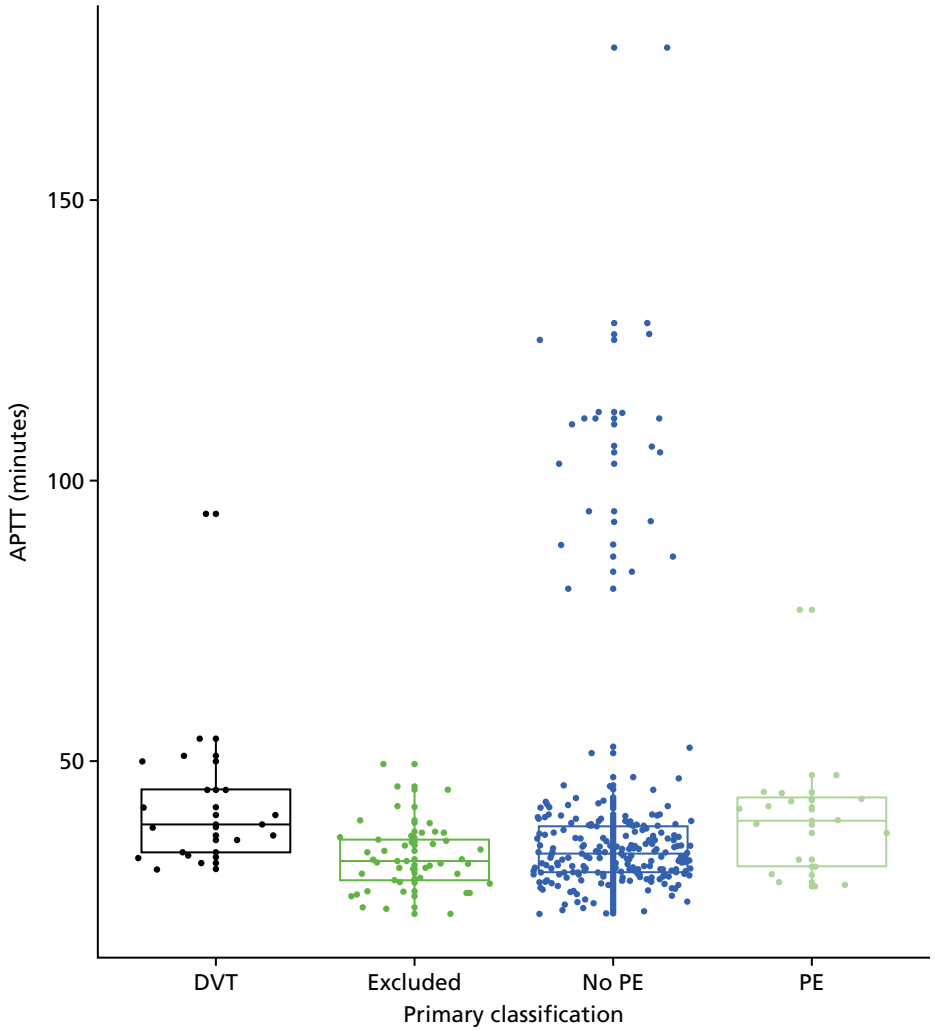


FIGURE 32 Box-and-whisker plot comparing the APTT levels for women with DVT, women with PE, women with no PE and women excluded from the primary analysis.

Clauss fibrinogen

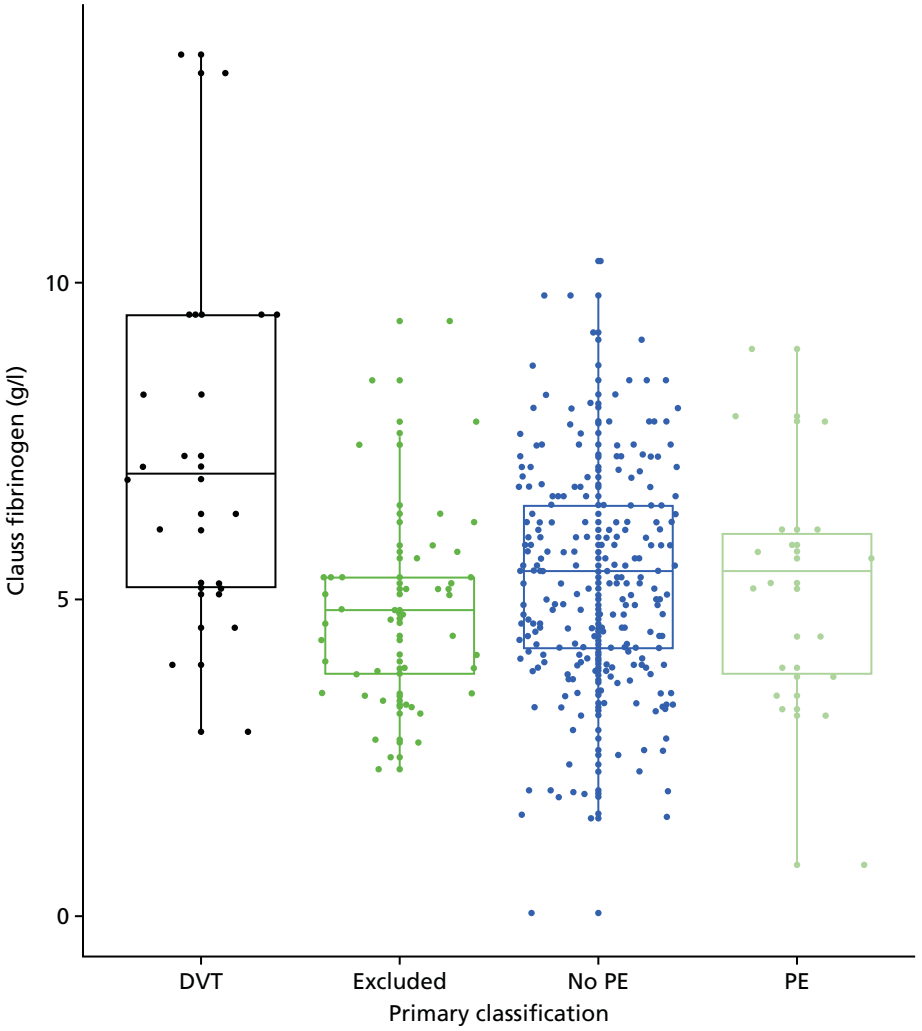


FIGURE 33 Box-and-whisker plot comparing Clauss fibrinogen levels for women with DVT, women with PE, women with no PE and women excluded from the primary analysis.

Prothrombin time

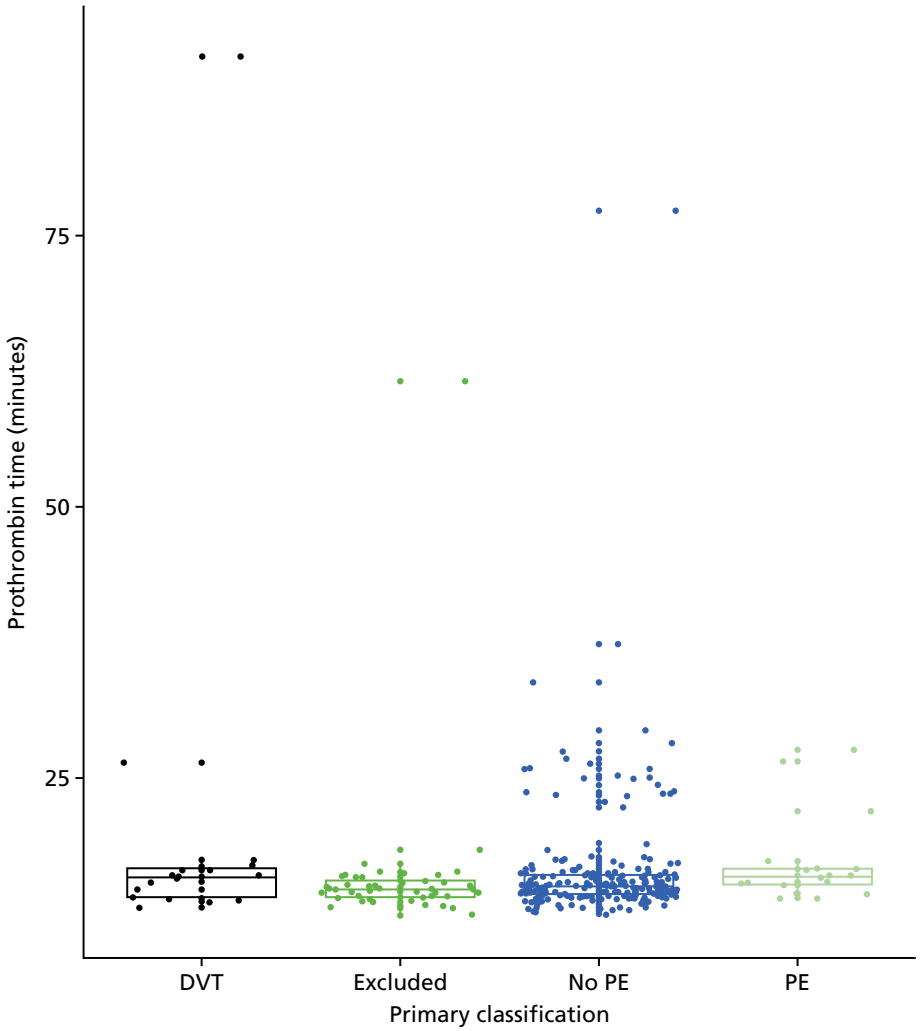


FIGURE 34 Box-and-whisker plot comparing PT levels for women with DVT, women with PE, women with no PE and women excluded from the primary analysis.

D-dimer (Innovance) levels

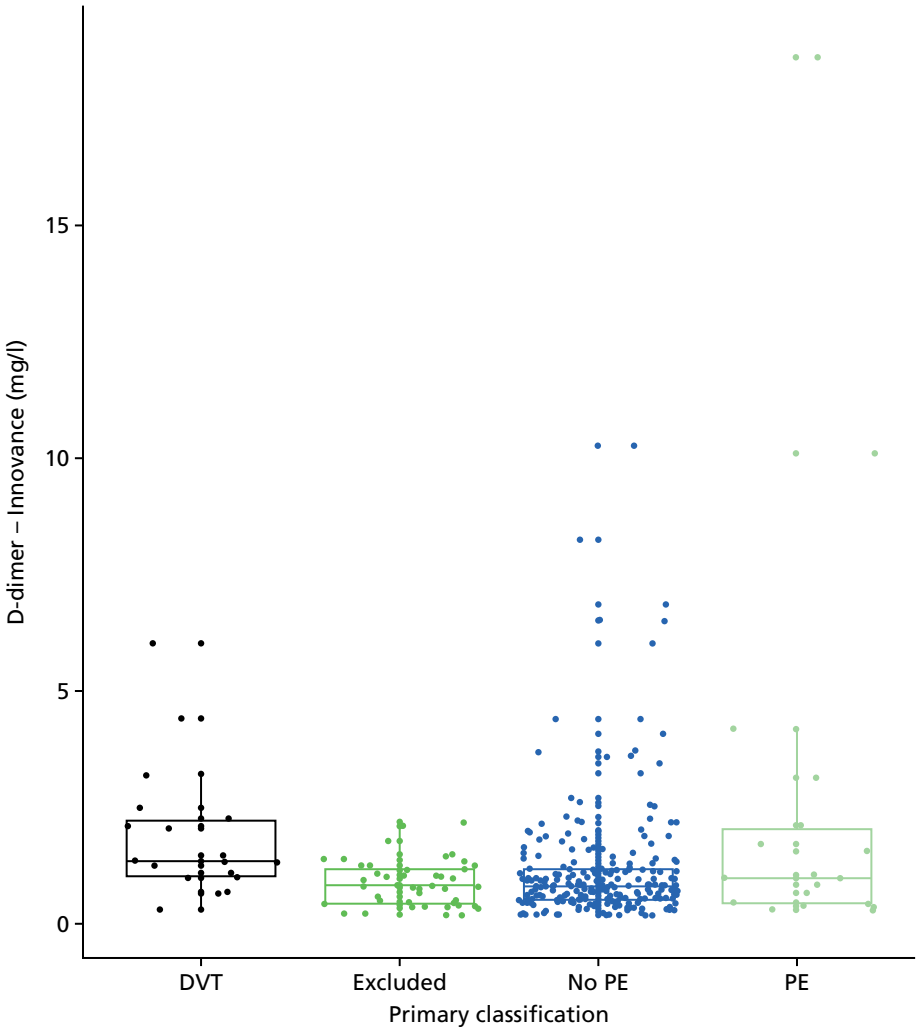


FIGURE 35 Box-and-whisker plot comparing D-dimer (Innovance) levels for women with DVT, PE, no PE and those excluded from the primary analysis.

D-dimer (ELISA) levels

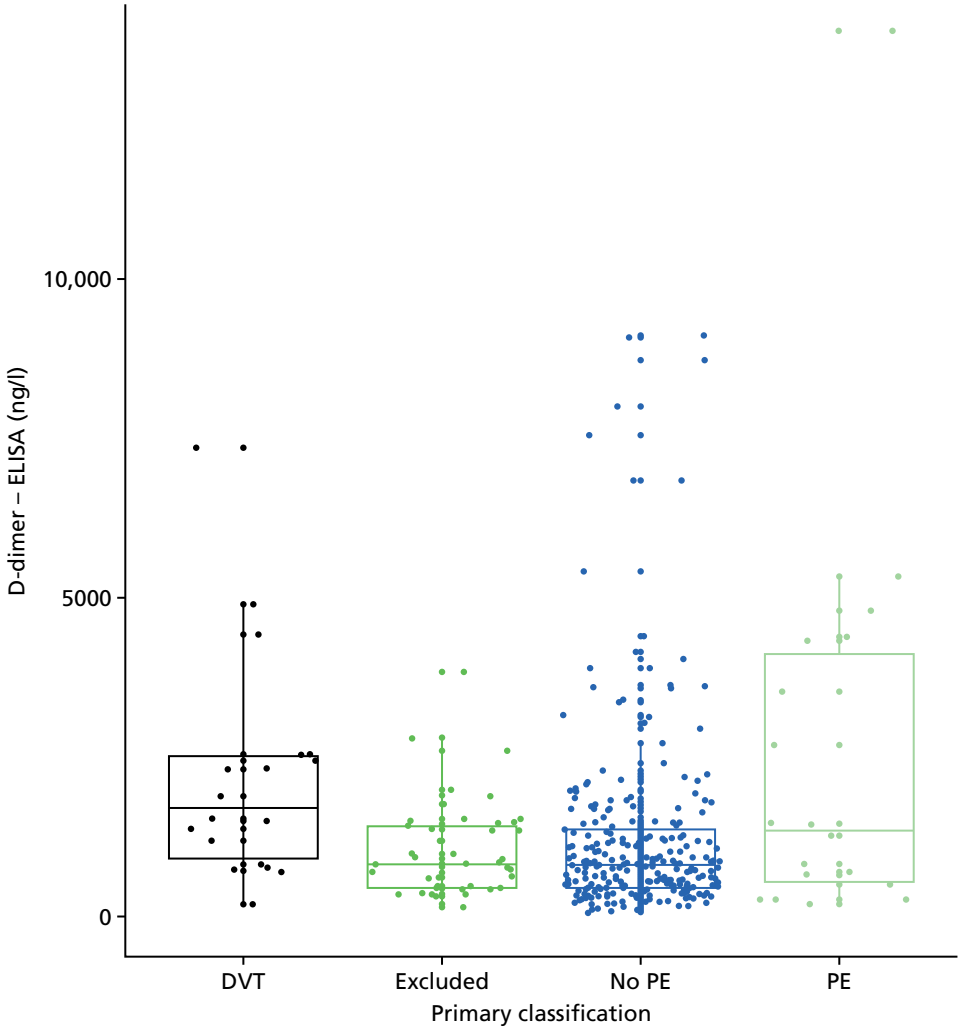


FIGURE 36 Box-and-whisker plot comparing D-dimer (ELISA) levels for women with DVT, women with PE, women with no PE and women excluded from the primary analysis.

Thrombin generation (lag time)

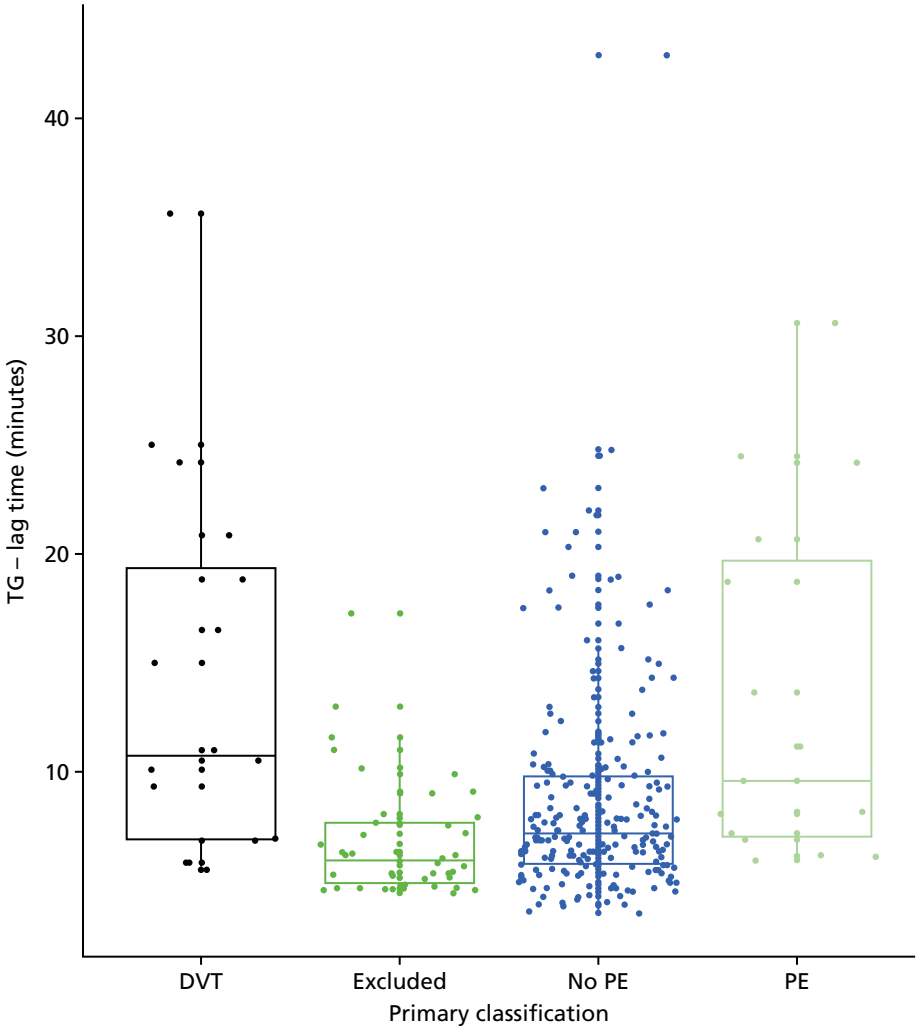


FIGURE 37 Box-and-whisker plot comparing TG (lag time) levels for women with DVT, women with PE, women with no PE and women excluded from the primary analysis.

Thrombin generation (endogenous potential)

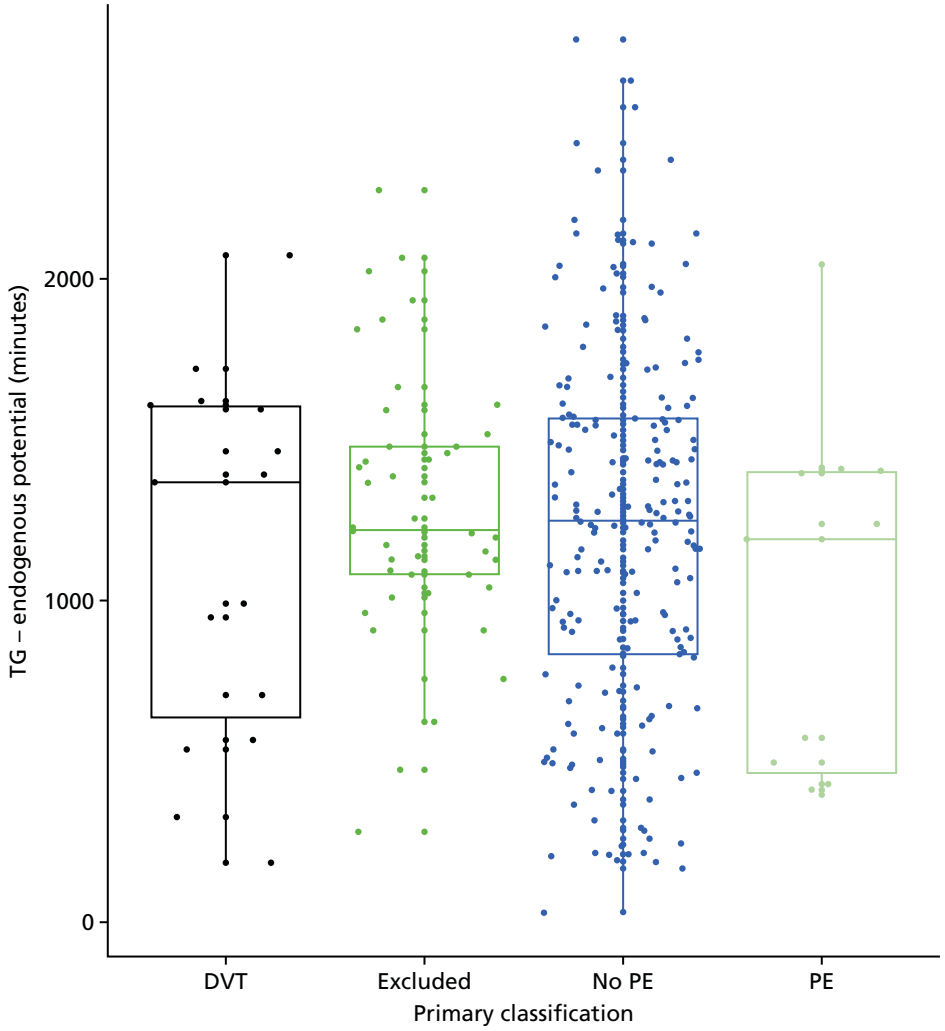


FIGURE 38 Box-and-whisker plot comparing TG (endogenous potential) levels for women with DVT, women with PE, women with no PE and women excluded from the primary analysis.

Thrombin generation (peak)

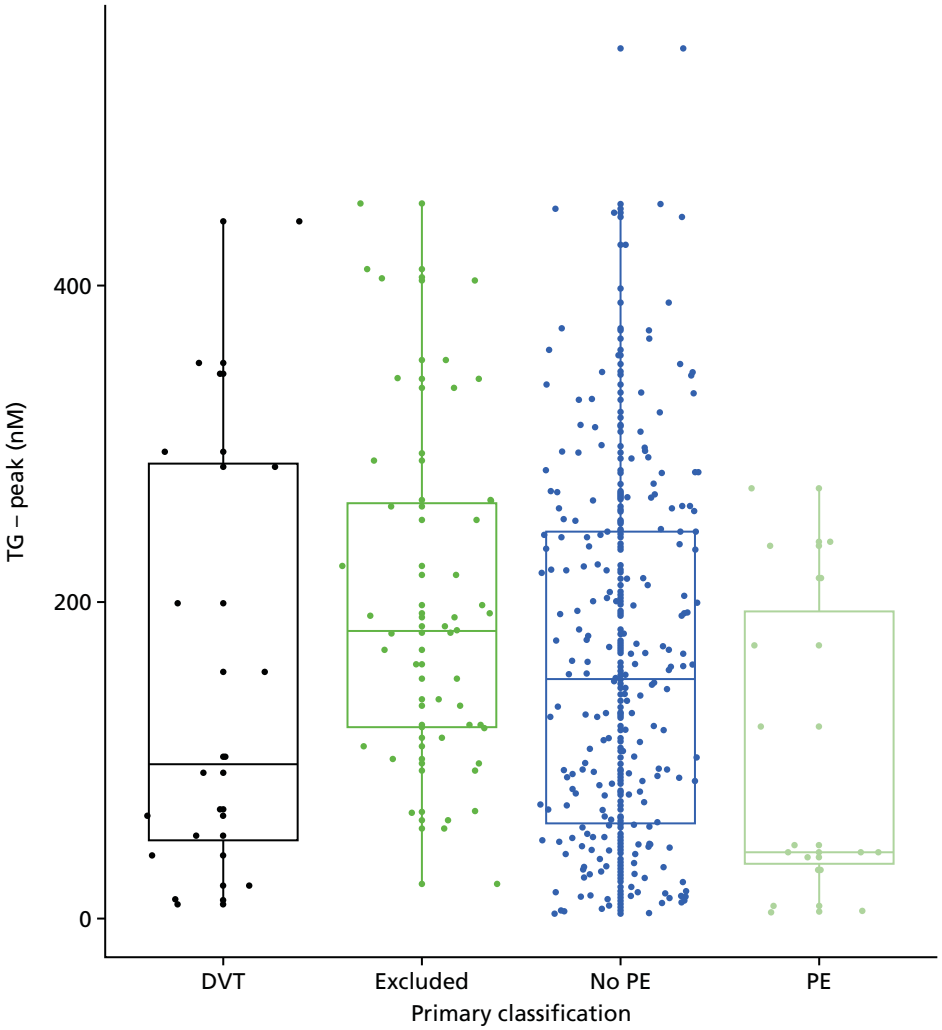


FIGURE 39 Box-and-whisker plot comparing TG (peak) levels for women with DVT, women with PE, women with no PE and women excluded from the primary analysis.

Thrombin generation (time to peak)

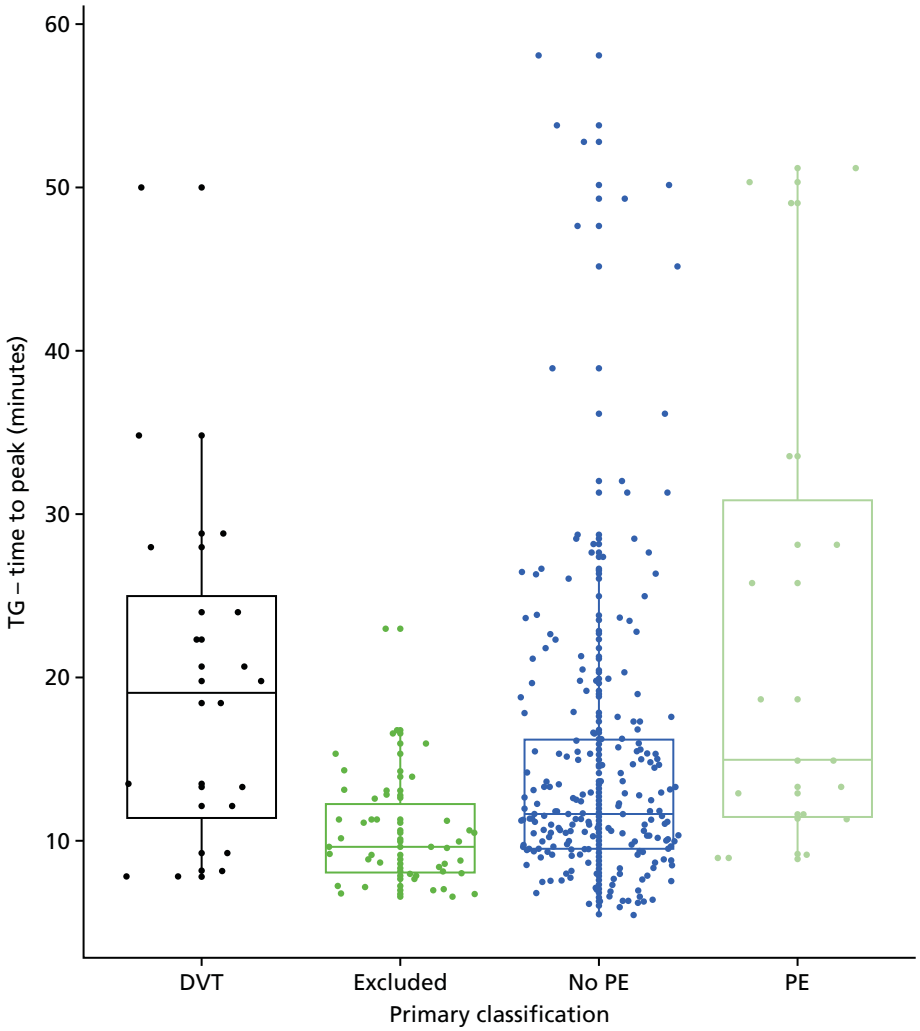


FIGURE 40 Box-and-whisker plot comparing TG (time to peak) levels for women with DVT, women with PE, women with no PE and women excluded from the primary analysis.

Plasmin–antiplasmin

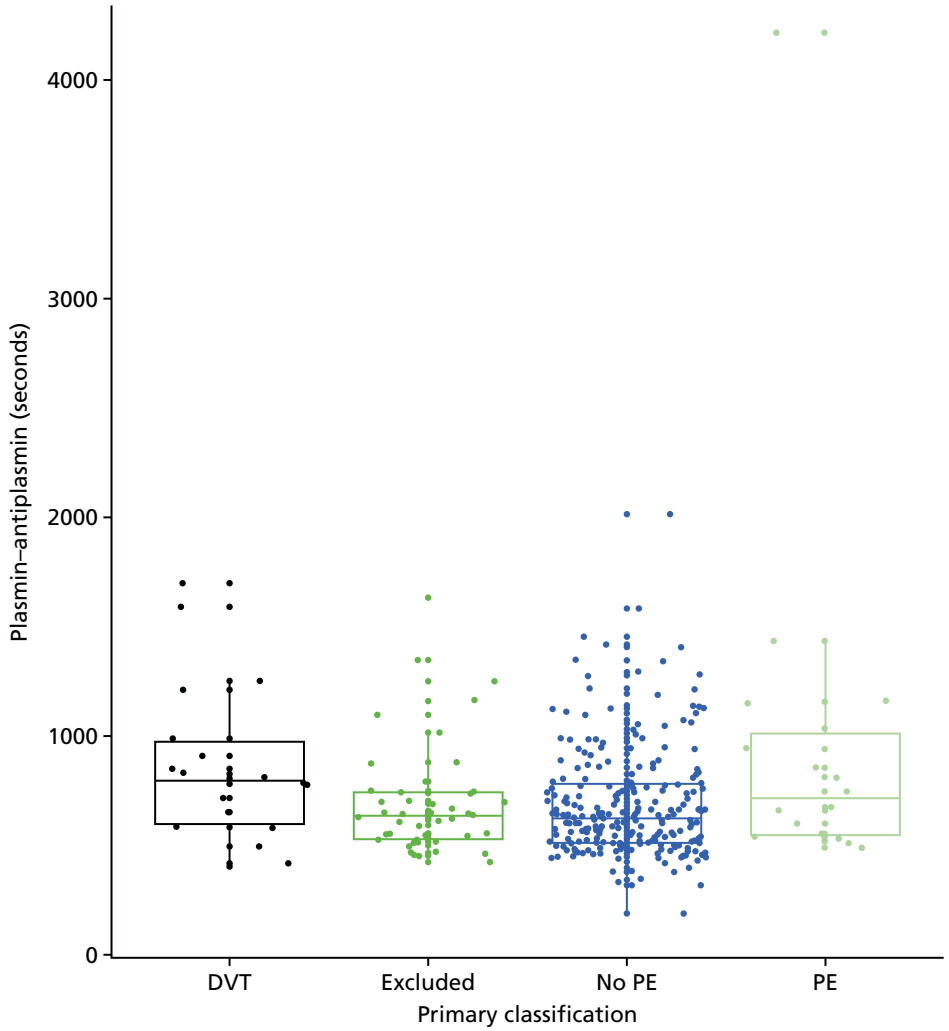


FIGURE 41 Box-and-whisker plot comparing the plasmin–antiplasmin levels for women with DVT, women with PE, women with no PE and women excluded from the primary analysis.

Mid-regional pro-atrial natriuretic peptide

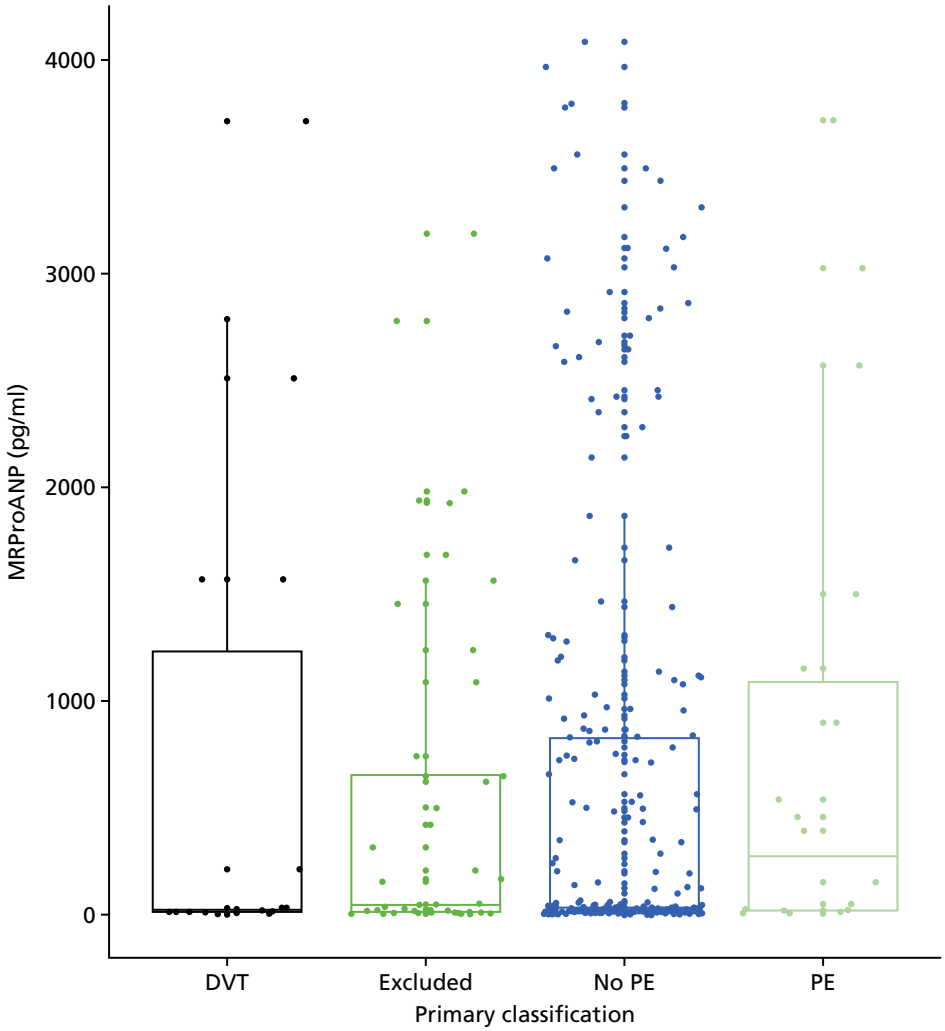


FIGURE 42 Box-and-whisker plot comparing MRProANP levels for women with DVT, women with PE, women with no PE and women excluded from the primary analysis.

B-type natriuretic peptide

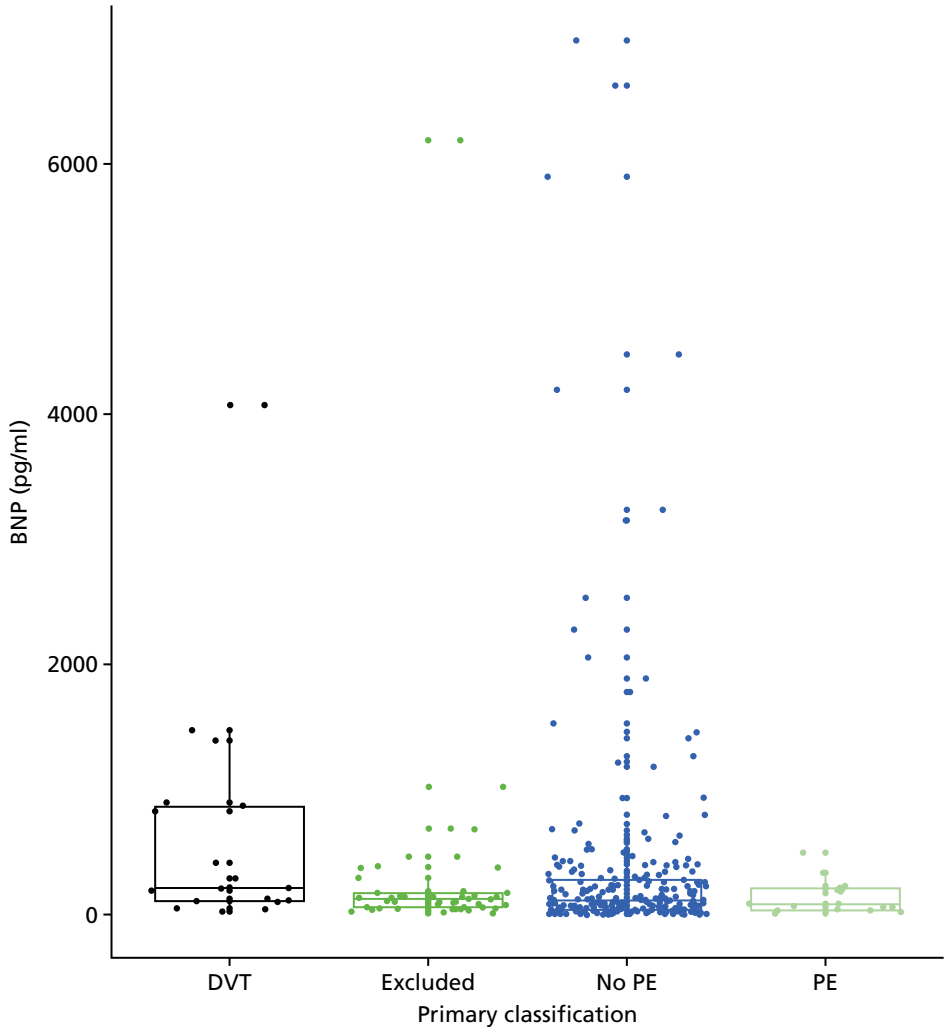


FIGURE 43 Box-and-whisker plot comparing BNP levels for women with DVT, women with PE, women with no PE and women excluded from the primary analysis.

Tissue factor

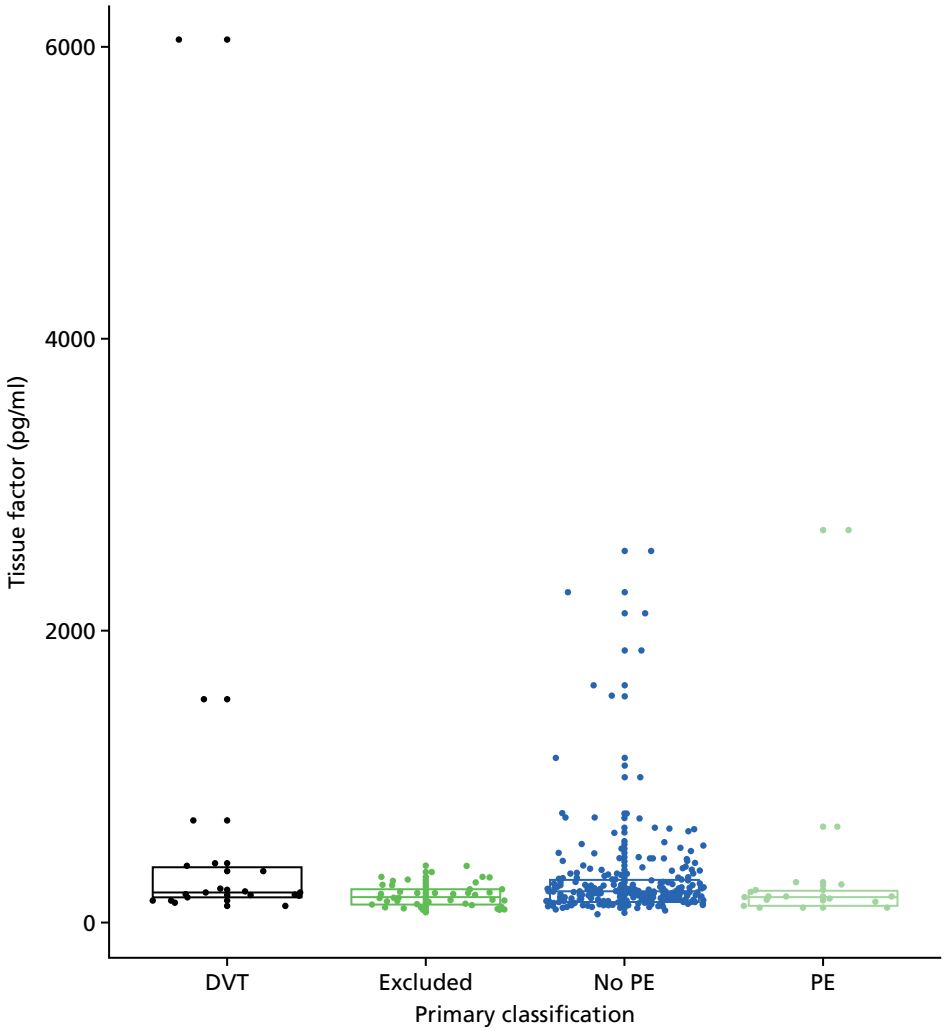


FIGURE 44 Box-and-whisker plot comparing tissue factor levels (pg/ml) for women with DVT, women with PE, women with no PE and women excluded from the primary analysis.

PF 1 + 2

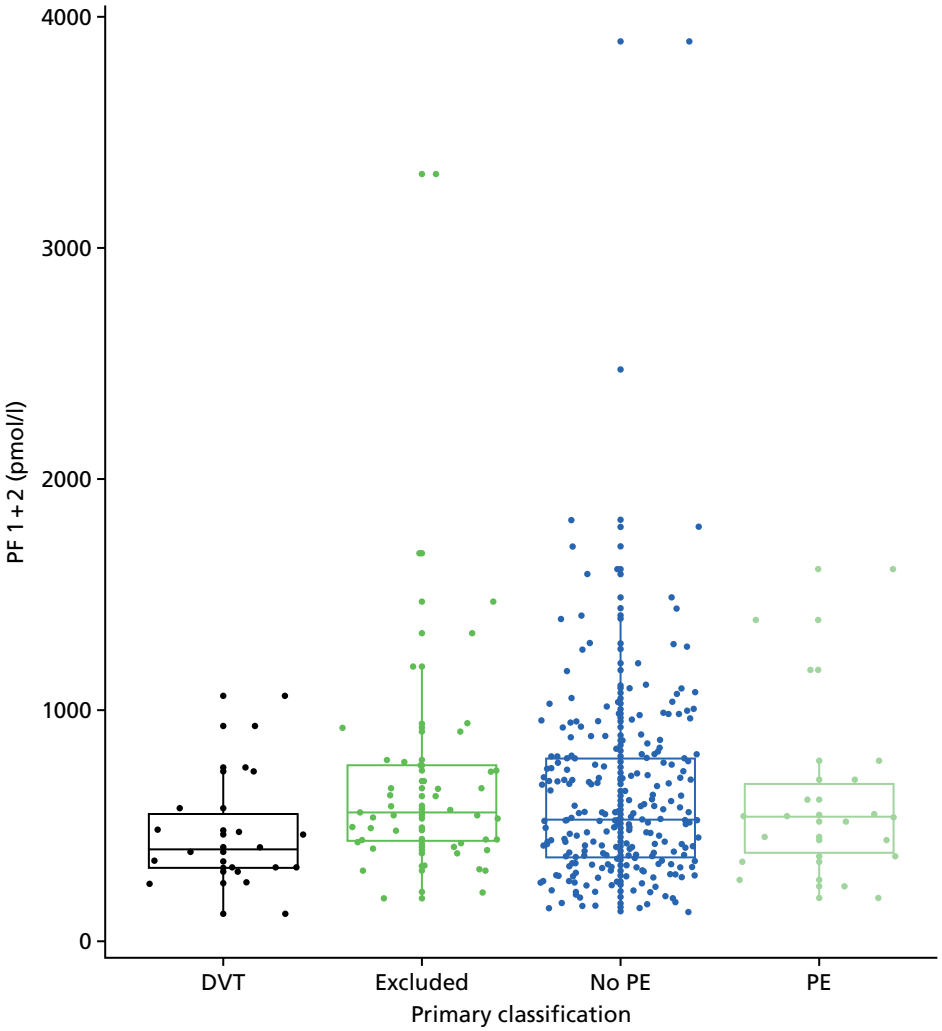


FIGURE 45 Box-and-whisker plot comparing PF 1 + 2 levels (pmol/l) for women with DVT, women with PE, women with no PE and women excluded from the primary analysis.

Troponin

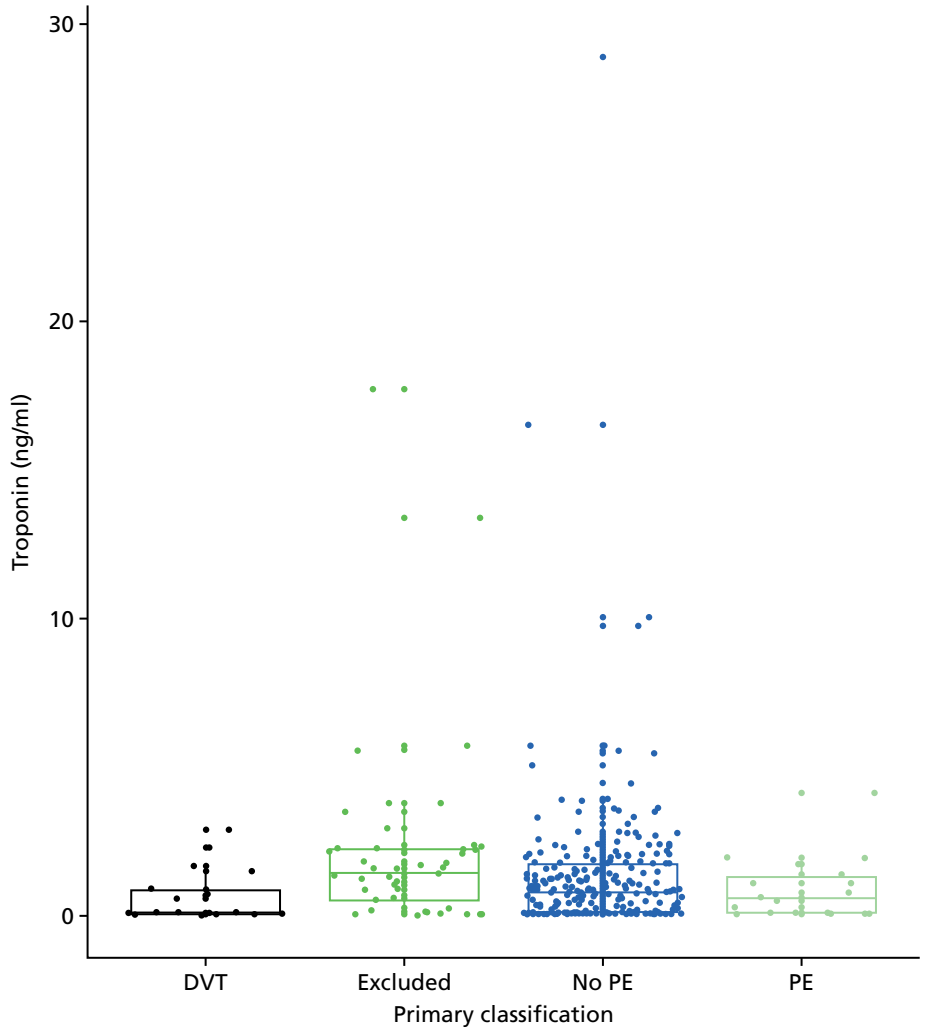


FIGURE 46 Box-and-whisker plot comparing troponin levels (ng/ml) for women with DVT, women with PE, women with no PE and women excluded from the primary analysis.

C-reactive protein

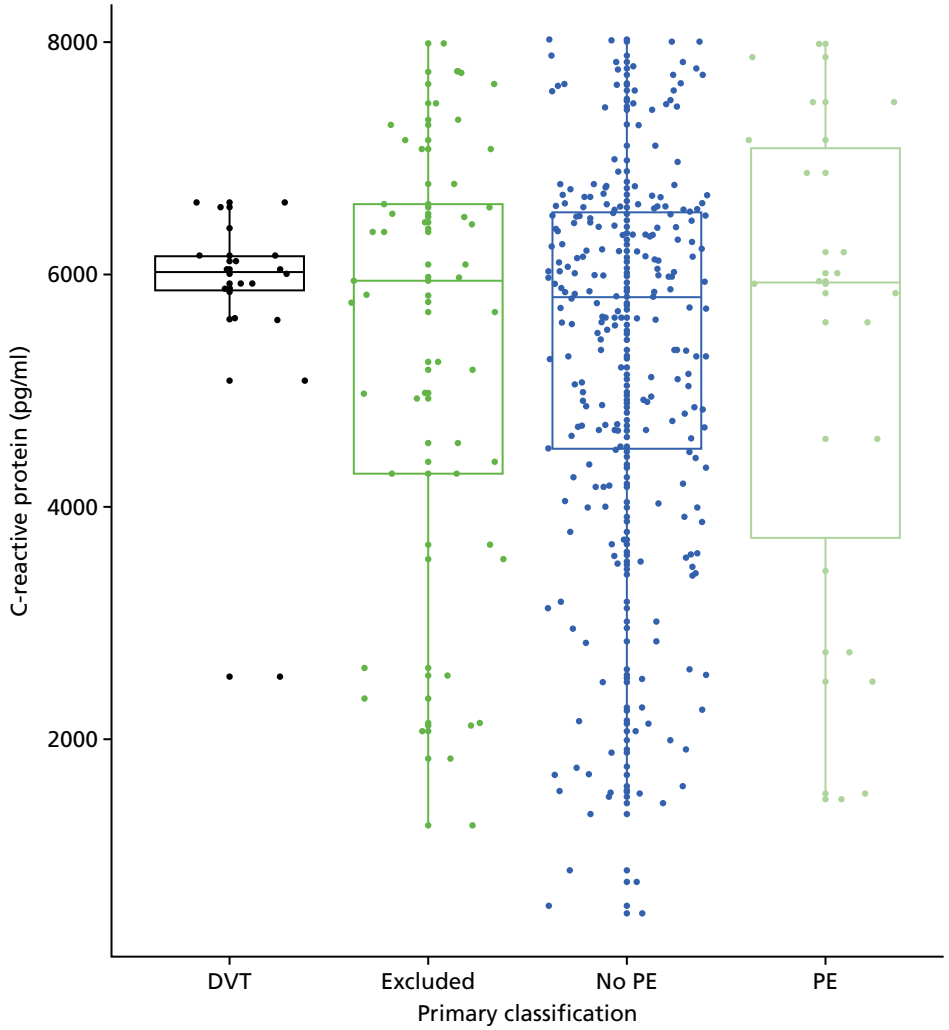


FIGURE 47 Box-and-whisker plot comparing CRP levels for women with DVT, women with PE, women with no PE and women excluded from the primary analysis.

Details of the biomarker analysis (excluding women who had received anticoagulation treatment)

We repeated the primary biomarker analysis, having excluded 240 out of 328 women who had received anticoagulation treatment prior to blood sampling. The analysis involved only 66 women, of whom four women had VTE. *Table 50* compares the mean biomarker levels between women with and without VTE in the primary analysis. The differences observed in the main analysis disappeared or even reversed when women receiving anticoagulation treatment were removed, but this probably reflects the small numbers available for this analysis. There were no significant differences in biomarker levels between the two groups.

Figures 48–50 show the ROC curves for the D-dimer biomarkers, the apothrombin, PF 1 + 2, prothrombin and TG biomarkers and the other biomarkers.

TABLE 50 Mean (standard deviation) biomarker levels for the patient groups with those having received anticoagulation treatment excluded

	Mean biomarker level (SD) in women with		
Biomarker	No VTE ($n = 62$)	VTE ($n = 4$)	*p*-value
APTT (minutes)	33.4 (16.67)	33.4 (6.57)	0.993
PT (minutes)	14.8 (2.108)	14.2 (0.772)	0.610
Clauss fibrinogen	5.41 (1.81)	6.61 (2.61)	0.219
D-dimer (ELISA)	1114 (848)	832 (667)	0.517
D-dimer (Innovance)	1.126 (0.826)	0.797 (0.420)	0.432
TG (lag time)	6.20 (1.646)	6.98 (0.919)	0.354
TG (endogenous potential)	1501 (389)	1575 (351)	0.711
TG (time to peak)	10.03 (2.57)	10.31 (1.40)	0.823
TG (peak)	235 (100.3)	248 (71.0)	0.798
Plasmin–antiplasmin level	678 (205)	821 (276)	0.204
BNP level	256 (586.31)	29 (5.47)	0.205
MRproANP	478 (904)	1371 (1358)	0.095
Tissue factor (pg/ml)	222 (157.8)	164 (37.4)	0.428
PF 1 + 2 (pmol/l)	711 (386)	373 (161)	0.095
Troponin level (ng/ml)	1.03 (1.24)	2.12 (1.65)	0.122
CRP level (pg/ml)	5410 (1596)	5884 (1734)	0.564

SD, standard deviation.

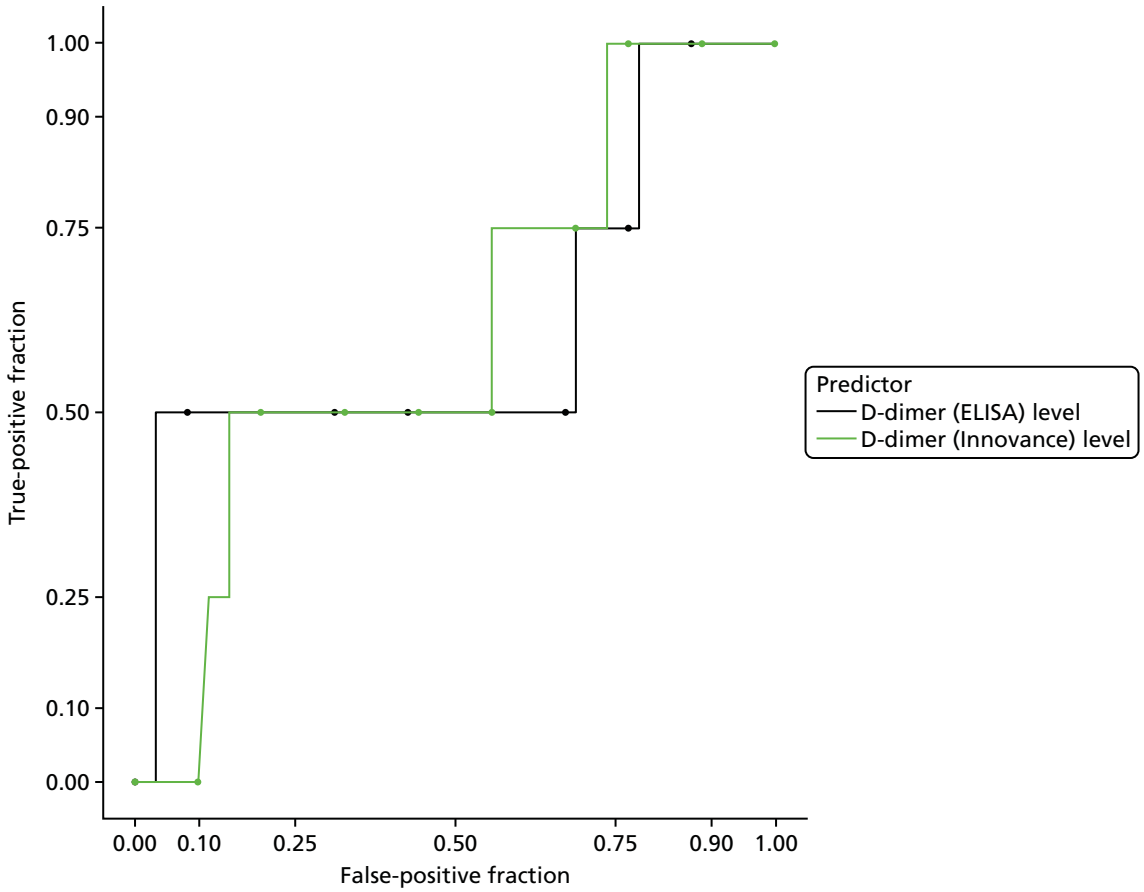


FIGURE 48 Receiver operating characteristic curves for the D-dimer level biomarkers.

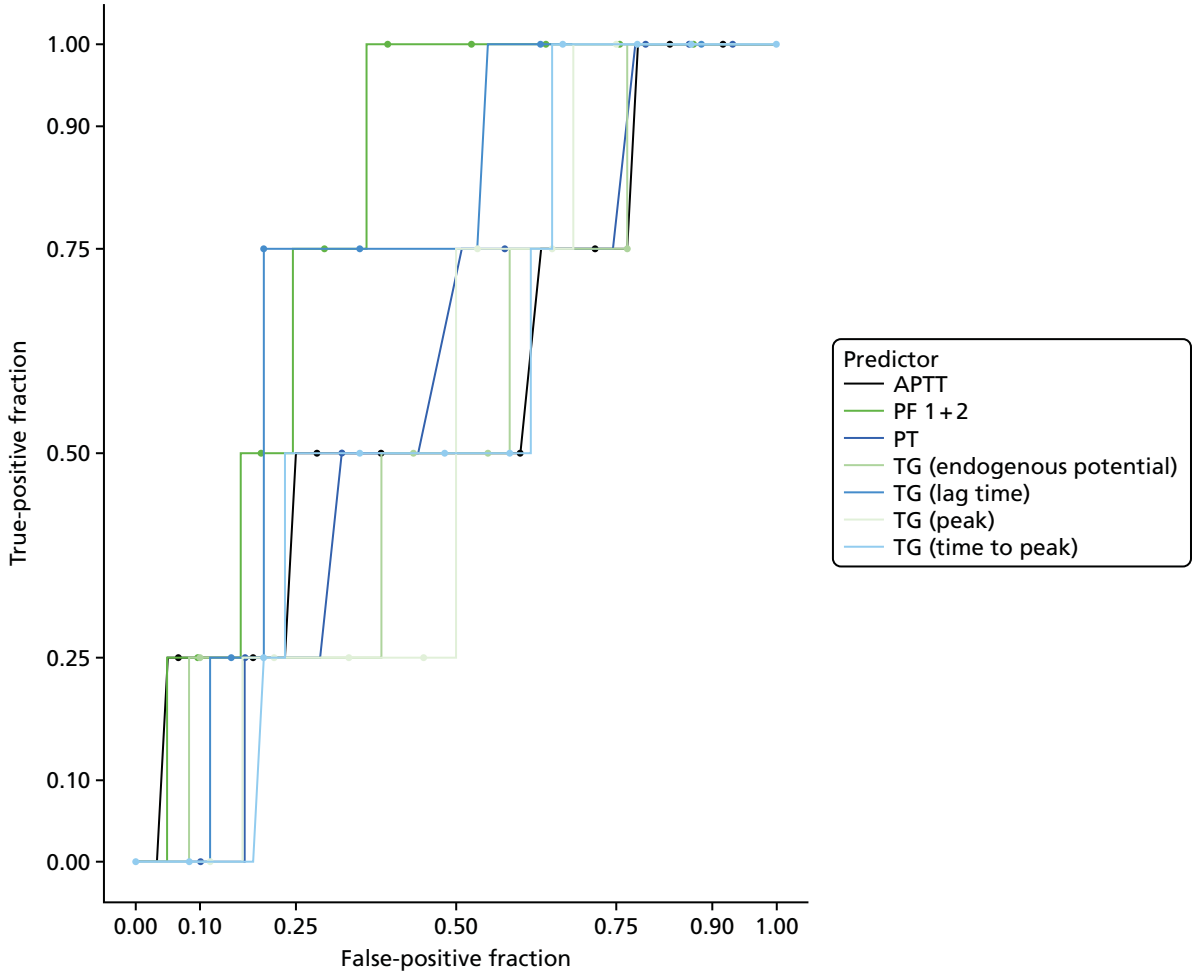


FIGURE 49 Receiver operating characteristic curves for apothrombin, PF 1 + 2, prothrombin and TG biomarkers (excluding women who had received anticoagulation treatment).

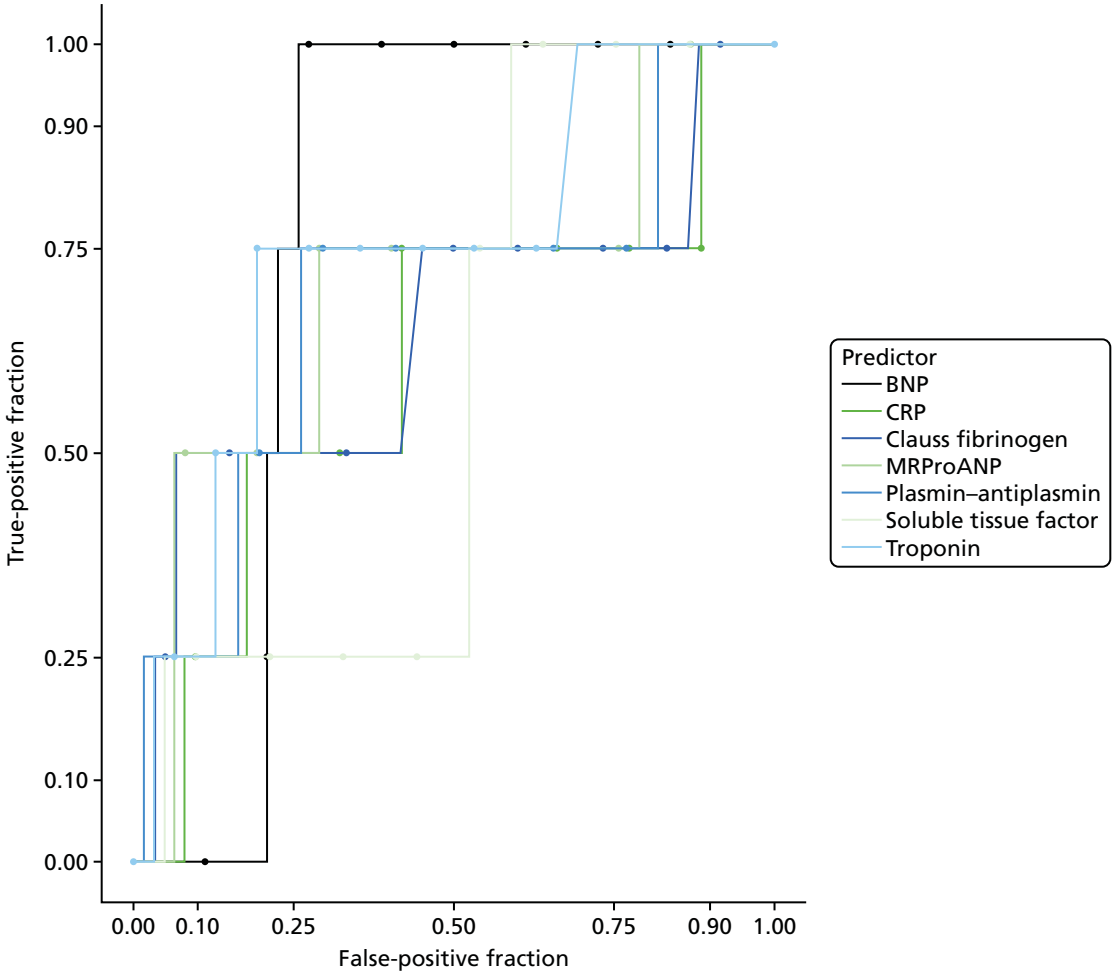


FIGURE 50 Receiver operating characteristic curves for other biomarkers (excluding women who had received anticoagulation treatment).

Table 51 reports the AUROC for the continuous biomarkers and diagnostic parameters for the biomarkers at the predefined threshold for positivity and the threshold that optimised sensitivity (> 95%) at the expense of specificity. The analysis was limited by small numbers, especially of women with VTE. BNP, PF 1 + 2, TG (lag time) and troponin may have some potential to rule out VTE with acceptable sensitivity, but the CIs are wide and the estimates would need to be validated in a larger cohort of women with VTE.

TABLE 51 Area under the receiver operating characteristic value, sensitivity and specificity for each biomarker, excluding women who had received anticoagulation treatment

		At predefined threshold (95% CI)		At threshold with optimal sensitivity (95% CI)	
Biomarker	AUC (95% CI)	Sensitivity	Specificity	Sensitivity	Specificity
APTT (minutes)	0.581 (0.244 to 0.919)	0.00 (0.000 to 0.602)	0.967 (0.885 to 0.996)	1 (0.398 to 1)	0.217 (0.121 to 0.342)
BNP level	0.774 (0.670 to 0.878)	0.00 (0.000 to 0.602)	0.935 (0.843 to 0.982)	1 (0.398 to 1)	0.742 (0.615 to 0.845)
CRP level (pg/ml)	0.609 (0.250 to 0.968)	1.00 (0.398 to 1.000)	0.097 (0.036 to 0.199)	1 (0.398 to 1)	0.113 (0.047 to 0.219)
Clauss fibrinogen	0.648 (0.259 to 1.000)	0.75 (0.194 to 0.994)	0.250 (0.147 to 0.379)	1 (0.398 to 1)	0.117 (0.048 to 0.226)
D-dimer (ELISA)	0.615 (0.210 to 1.000)	0.50 (0.068 to 0.932)	0.148 (0.070 to 0.262)	1 (0.398 to 1)	0.213 (0.119 to 0.337)
D-dimer (Innovance)	0.613 (0.299 to 0.926)	0.25 (0.006 to 0.806)	0.672 (0.540 to 0.787)	1 (0.398 to 1)	0.262 (0.158 to 0.391)
MRproANP	0.698 (0.357 to 1.000)	0.50 (0.068 to 0.932)	0.823 (0.705 to 0.908)	1 (0.398 to 1)	0.210 (0.117 to 0.332)
PF 1 + 2 (pmol/l)	0.795 (0.644 to 0.947)	0.00 (0.000 to 0.602)	0.918 (0.819 to 0.973)	1 (0.398 to 1)	0.639 (0.506 to 0.758)
Plasmin–antiplasmin level	0.684 (0.335 to 1.000)	0.50 (0.068 to 0.932)	0.770 (0.645 to 0.868)	1 (0.398 to 1)	0.180 (0.094 to 0.300)
PT (minutes)	0.572 (0.306 to 0.838)	0.00 (0.000 to 0.602)	0.831 (0.710 to 0.916)	1 (0.398 to 1)	0.220 (0.123 to 0.347)
TG (lag time)	0.735 (0.531 to 0.940)	1.00 (0.398 to 1.000)	0.00 (0.000 to 0.060)	1 (0.398 to 1)	0.450 (0.321 to 0.584)
TG (endogenous potential)	0.454 (0.155 to 0.753)	0.50 (0.068 to 0.932)	0.525 (0.393 to 0.654)	1 (0.398 to 1)	0.233 (0.134 to 0.360)
TG (peak)	0.462 (0.229 to 0.696)	0.00 (0.000 to 0.602)	0.852 (0.738 to 0.930)	1 (0.398 to 1)	0.317 (0.203 to 0.450)
TG (time to peak)	0.577 (0.320 to 0.834)	1.00 (0.398 to 1.000)	0.213 (0.119 to 0.337)	1 (0.398 to 1)	0.350 (0.231 to 0.484)
Tissue factor (pg/ml)	0.422 (0.159 to 0.686)	0.00 (0.000 to 0.602)	0.885 (0.778 to 0.953)	1 (0.398 to 1)	0 (0.286 to 0.543)
Troponin level (ng/ml)	0.742 (0.453 to 1.000)	0.25 (0.006 to 0.806)	0.903 (0.801 to 0.964)	1 (0.398 to 1)	0.306 (0.196 to 0.437)

Appendix 12 Results of the health economic literature search

Databases

Date searched: 25 August 2016.

MEDLINE and MEDLINE In-Process & Other Non-Indexed Citations

Search terms

1. Pregnancy
2. Pulmonary Embolism/di [Diagnosis]
3. Pulmonary Embolism/ra [Radiography]
4. Pulmonary Embolism/ri [Radionuclide Imaging]
5. 2 or 3 or 4
6. 1 and 5
7. 6 and Scottish intercollegiate guidelines network (SIGN) economic search filters, available from www.sign.ac.uk/search-filters.html

NHS Economic Evaluation Database

Search terms

1. Pulmonary embolism [exp]

Selection criteria

Table 52 shows the selection criteria for the literature search.

TABLE 52 The inclusion criteria used to assess the economic studies

Study characteristic	Criteria
Design	Cost–consequences analysis, cost-effectiveness analysis, cost–benefit analysis, cost–utility analysis, cost studies
Population	Patients aged ≥ 18 years
	For pregnant or postpartum women
	Mixed populations included, if the data can be extracted on pregnant or postpartum women
Intervention	CDR to send the woman for scanning
Outcomes	Cost-effectiveness, cost estimates, quality-of-life estimates

Results of the literature review

Figure 51 shows the PRISMA flow diagram.

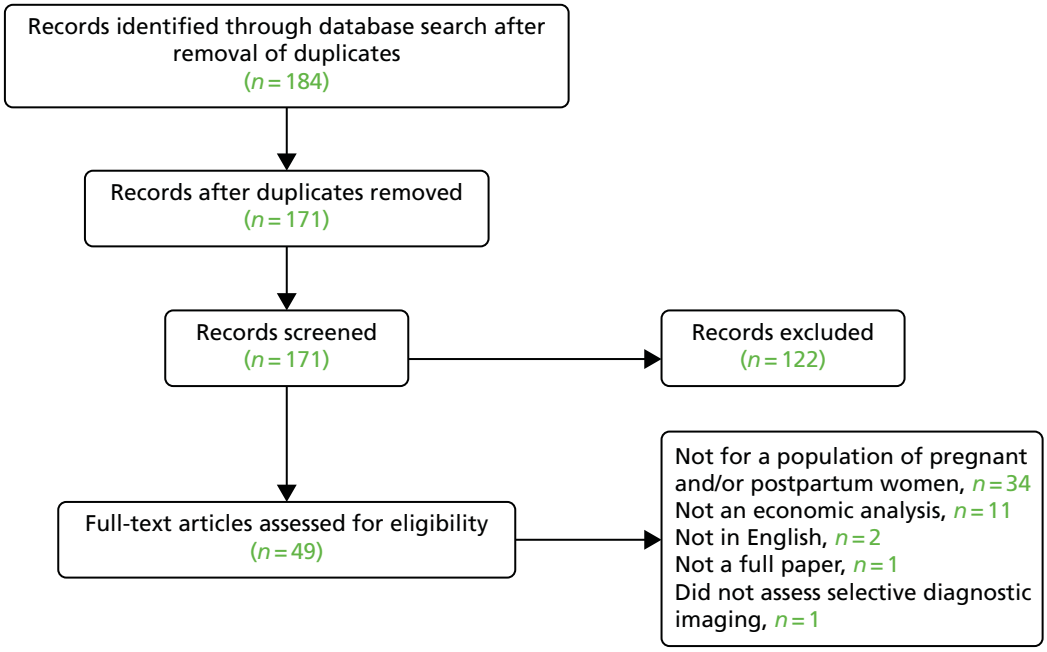


FIGURE 51 The PRISMA flow diagram of the economic model results.

Appendix 13 The optimal number of bootstraps in the decision-analysis model

Bootstrapping procedure

Within each bootstrap, the model was run once in the PE cohort and once in the no PE cohort. A total of 157 patients in the PE cohort and 248 patients in the no PE cohort were sampled with replacement, with both of these analyses drawing the patient data from the observed patient characteristics in the DiPEP case-control study and/or the UKOSS cases.

Determining the number of model bootstraps to use

An estimator of unbiased costs and QALYs was obtained by running the model once for each patient with PE and each patient with no PE. It should be noted that these estimators may not be completely unbiased estimators of the population mean cost and QALY outcomes, as random numbers determine the scanning methodology used (CTPA or VQ SPECT) and the risk of death from PE. As a result, there may be some random noise in these estimators. From this analysis, it was determined that 'scan all' would be the cost-effective strategy at £20,000 per QALY gained (*Table 53*).

TABLE 53 The results of the analysis used to determine the unbiased estimator of costs and QALYs for each strategy in the population of women with suspected PE

			Incremental				
Strategy	Costs (£)	QALYs	Costs (£)	QALYs	ICER	NMB (£)	iNMB vs. the next most cost-effective option
No scanning, treat none	1953	19.9164	–	–	Dominated by scan all	396,374	–
Wells's score (strict)	1673	20.0889	–	–	Dominated by scan all	400,105	–
Delphi score (specific)	1729	20.0935	–	–	Dominated by scan all	400,142	–
Geneva score	1673	20.1227	–	–	Dominated by scan all	400,781	–
Wells's score (permissive)	1644	20.1461	–	–	Dominated by scan all	401,278	–
Delphi score (primary)	1579	20.2056	–	–	Dominated by scan all	402,532	–
PERC score	1546	20.2265	–	–	Dominated by scan all	402,983	–
No scanning, treat all	2404	20.2686	–	–	Dominated by scan all	402,968	–
Delphi score (sensitive)	1409	20.3511	–	–	Dominated by scan all	405,614	–
Scan all	1367	20.3705	–586	0.4541	Dominant	406,042	£428

NMB, net monetary benefit; iNMB, incremental net monetary benefit; –, no data.

The model was then run deterministically with 200 bootstraps, with the results from each bootstrap being recorded. The stability of the model results with respect to the number of bootstraps was assessed using the iNMB of 'scan all' versus the most cost-effective treatment option out of the remaining nine strategies at a MAICER of £20,000 per QALY gained. To determine the optimal number of bootstraps, 200 bootstraps were conducted, with the costs and QALYs of each strategy being recorded in each bootstrap.

The average incremental net monetary benefit of 'scan all' versus the next most cost-effective strategy at £20,000 per QALY with respect to the number of bootstraps conducted was calculated. The 95% CIs were constructed around this measure using the normal approximation, percentile and bias-corrected methods. The percentile *t* method to calculate the 95% CI was not used, as this method requires a further inner loop in the bootstrapping procedure and, as such, was deemed to be too computationally expensive to conduct.

Results

All methods showed that the 95% CI of the iNMB of 'scan all' versus the next most cost-effective option at £20,000 per QALY gained does not cross zero when more than 100 bootstrap samples are taken (*Figure 52*). Therefore, there is a < 2.5% probability that the true iNMB of 'scan all' versus the next most cost-effective option is < 0 (i.e. another strategy is cost-effective at £20,000 per QALY gained). Therefore, 100 bootstraps per analysis was determined to be sufficient for running the model to determine the cost-effectiveness of the different strategies.

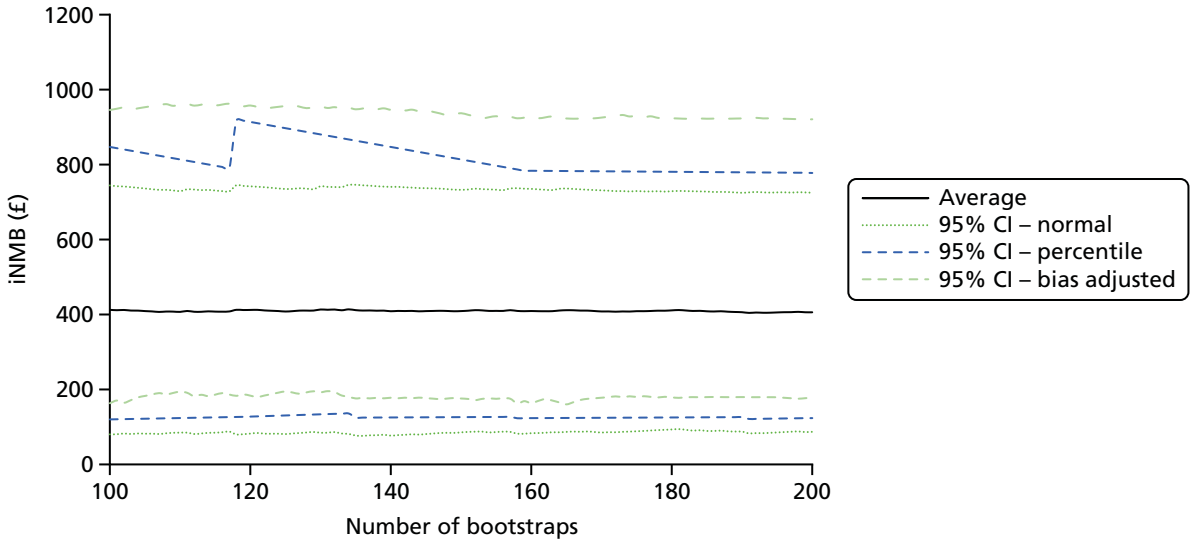


FIGURE 52 The average incremental net monetary benefit of 'scan all' compared with the next most cost-effective option with respect to the number of bootstraps. iNMB, incremental net monetary benefit.

Appendix 14 The clinical parameters used in the decision-analysis model

Table 54 shows the clinical parameters used in the decision-analysis model.

TABLE 54 The clinical parameters used in the decision-analysis model

Parameter	Mean	Distribution	95% CI	Alpha/*n*	Beta	Source	Notes
Initial PE							
Probability that a woman had PE	6.5%	Beta	3.9% to 9.7%	18	259	Suspected PE data set	
Probability of death from PE	See *Appendix 16*					Expert elicitation	See the beta regression results, *Appendix 16*
Harms of withholding anticoagulation treatment							
Odds ratio of fatal PE	12.517	Normal on log-odds	0.636 to 246.384			Barrit and Jordan[70]	The distributions were constrained on the grounds of clinical plausibility, so that odds ratios did not go below 1
Odds ratio of a recurrent VTE or death	16.667	Normal on log-odds	1.818 to 152.770			Barrit and Jordan[70]	
Probability of recurrent VTEs							
Recurrent PE with 3 months of anticoagulation treatment	2.95%	Beta	2.5% to 3.6%	103	3319	Carrier *et al.*[66]	See *Table 4*
Recurrent VTE with 3 months of anticoagulation treatment	3.62%	Beta	3.0% to 4.2%	123	3299	Carrier *et al.*[66]	See *Table 4*
Recurrent PE with 6 months of anticoagulation treatment	4.83%	Beta	3.6% to 5.3%	92	2001	Carrier *et al.*[66]	See *Table 4*
Recurrent VTE with 6 months of anticoagulation treatment	2.17%	Beta	4.0% to 5.9%	103	1990	Carrier *et al.*[66]	See *Table 4*
Case fatality rates for recurrent VTEs							
3 months of anticoagulation treatment	29.97%	Beta	22.3% to 38.4%	37	86	Carrier *et al.*[66]	See *Table 4*
6 months of anticoagulation treatment	31.19%	Beta	13.2% to 28.6%	21	82	Carrier *et al.*[66]	See *Table 4*
Bleeding from anticoagulation							
Major bleeding event with 3 months of anticoagulation treatment	1.95%	Beta	1.4% to 2.3%	62	3360	Carrier *et al.*[66]	See *Table 6*
Major bleeding event with 6 months of anticoagulation treatment	2.36%	Beta	1.5% to 2.8%	44	2049	Carrier *et al.*[66]	See *Table 6*

Parameter	Mean	Distribution	95% CI	Alpha/*n*	Beta	Source	Notes
Fatal major bleeding event with 3 months of anticoagulation treatment	0.25%	Beta	0.1% to 0.4%	7	3415	Carrier *et al.*[66]	See *Table 6*. This was for the whole population, not just those who had a bleed
Fatal major bleeding event with 6 months of anticoagulation treatment	0.62%	Beta	0.3% to 1.0%	13	2080	Carrier *et al.*[66]	
Split of non-fatal major bleeds							
All gastrointestinal bleeds	35.9%	Dirichlet		499		Ensor *et al.*[67]	
All intracranial bleeds	18.1%	Dirichlet		245		Ensor *et al.*[67]	
All other bleeds	45.9%	Dirichlet		622		Ensor *et al.*[67]	
Probability that an intracranial bleed is fatal	39.03%	Beta	26.5% to 38.2%	79	166	Ensor *et al.*[67]	
Probability that a gastrointestinal bleed is fatal	19.3%	Beta	15.2% to 22.0%	92	407	Ensor *et al.*[67]	
Probability that another major bleed is fatal	9.5%	Beta	8.2% to 13.0%	65	557	Ensor *et al.*[67]	
Standardised mortality ratios after an intracranial bleed							
Year 1	4.50	Normal on the log-scale	2.496 to 8.115			Fogelholm *et al.*[75]	The log-SE was assumed to be 20% of the log-mean
Years 2–6	2.20	Normal on the log-scale	1.615 to 2.997			Fogelholm *et al.*[75]	
Chronic thromboembolic pulmonary hypertension							
Probability of CTEPH	N/A					Assumption	Assumed to be the same as the predicted risk of death from PE
Probability that CTEPH is surgically treated	59.5%	Beta	55.8% to 63.2%	404	275	Delcroix *et al.*[72]	
Life expectancy after surgically treated CTEPH – log-normal mean	5.081	Multivariate normal	3.956 to 6.206			See *Appendix 17*	See *Appendix 17*, for the variance–covariance matrices and SEs for these parameters
Life expectancy after surgically treated CTEPH – log-normal SD	3.343	Multivariate normal	2.646 to 4.224			See *Appendix 17*	
Life expectancy after medically treated CTEPH – exponential rate	0.1168	Normal	0.0950 to 0.1436			See *Appendix 17*	

continued

TABLE 54 The clinical parameters used in the decision-analysis model (*continued*)

Parameter	Mean	Distribution	95% CI	Alpha/*n*	Beta	Source	Notes
Cancer risks to the woman							
Lifetime attributable risk to the mother of breast cancer as a result of scanning at the age of 25 years	0.1%	Fixed				Hurwitz *et al.*[98]	No information on the uncertainty in these parameters was presented
Lifetime attributable risk to the mother of breast cancer as a result of scanning at the age of 55 years	0.02%	Fixed				Hurwitz *et al.*[98]	
Lifetime attributable risk to the mother of lung cancer as a result of scanning at the age of 25 years	0.12%	Fixed				Hurwitz *et al.*[98]	
Lifetime attributable risk to the mother of lung cancer as a result of scanning at the age of 55 years	0.09%	Fixed				Hurwitz *et al.*[98]	
Reduction in the risk of breast cancer as a result of using a VQ SPECT rather than a CTPA scan	97%	Fixed				RCOG[2]	
Cancer risks to the fetus							
Risk of developing childhood cancer from a CTPA scan	55/10,000,000	Fixed	10/10,000,000 to 100/10,000,000			Wall *et al.*[61]	Table on page 8 of Wall *et al.*[61] The values are upper and lower bounds rather than 95% CIs

N/A, not applicable; SD, standard deviation.

Appendix 15 The results of the beta regressions conducted on the data collected in the expert elicitation exercise

Results of the base-case analysis

The results of the beta regression conducted in the primary statistical analysis population and fitted to the average expert elicitation answers for all four experts are provided in *Figure 53*. The exponential of the mean effect coefficients in beta regression is equivalent to the odds ratio. For the dispersion coefficients, negative numbers indicate a higher mean variance and positive numbers indicate a lower mean variance. Although the dispersion parameter is not clinically meaningful, it is incredibly useful for economic analyses as it allows the variance term to change with the fitted value predicted by the mean effect parameters. This allows the heterogeneity in each patient's outcome to be incorporated into the economic model by

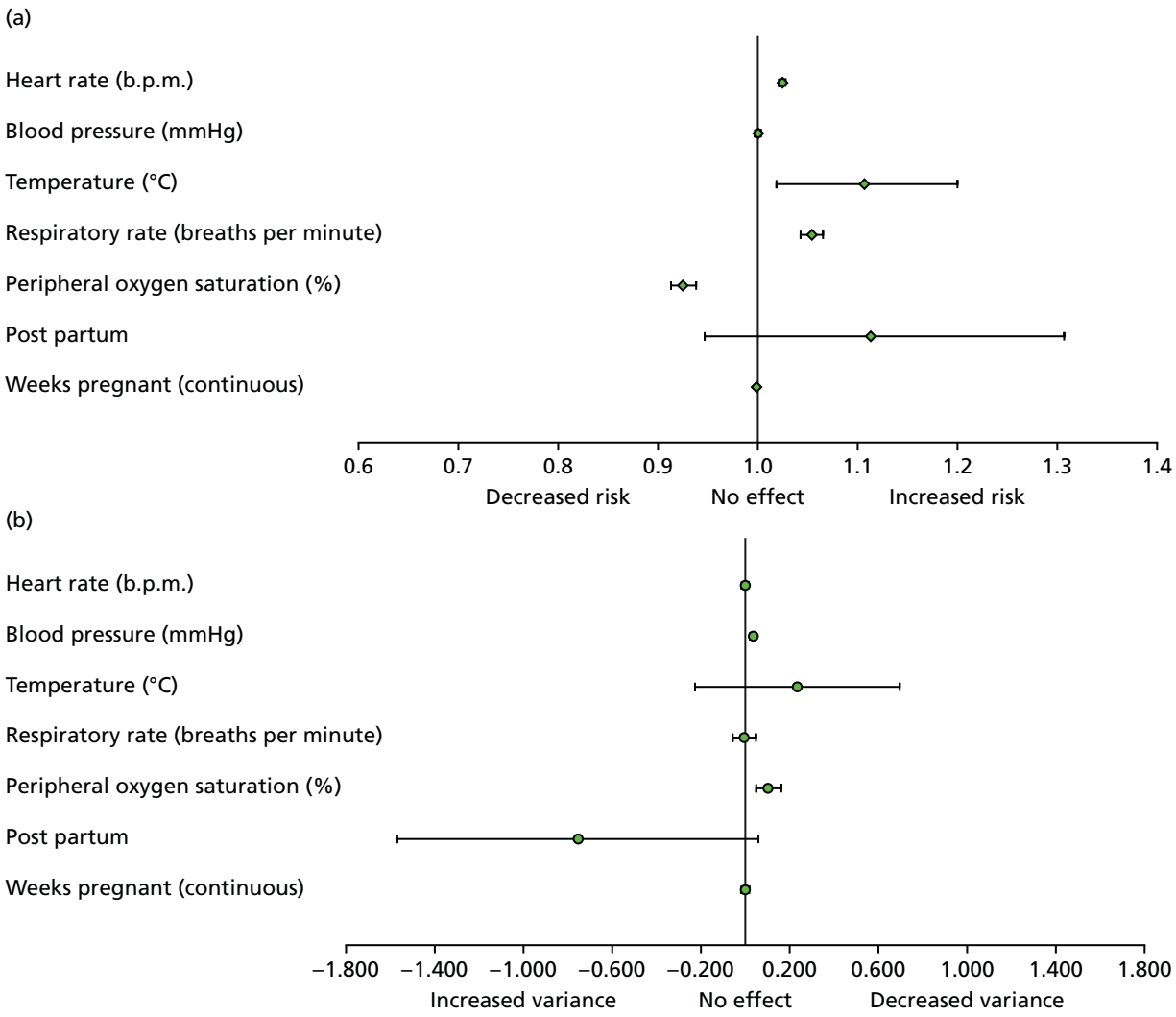


FIGURE 53 The results of the beta regression fitted to the average probability, from all four experts, of 30-day mortality for women with PE in the primary statistical analysis population with 95% CIs for (a) the effect of model coefficients on the odds ratio; and (b) the effect of model coefficients on the dispersion parameter. b.p.m., beats per minute.

estimating the predicted mean effect and the predicted variance in the mean effect, and sampling the modelled outcome from a beta distribution.

Part (a) of *Figure 53* presents the results for the odds ratio for death within 30 days associated with the fitted covariates. It was found that a unit increase in heart rate, systolic blood pressure, temperature or respiratory rate led to a statistically significant increase in the 30-day mortality for a pregnant woman with PE at the 5% level. Similarly, a higher peripheral oxygen saturation led to a statistically significant decreased risk of 30-day mortality for a pregnant woman with PE at the 5% level. The duration of the pregnancy in weeks (continuous) and whether or not the woman was post partum were not statistically significant predictors.

Part (b) of *Figure 53* presents the results for the impact on the included variables on the dispersion parameter. It was found that an increased blood pressure and peripheral oxygen saturation predicted a statistically significant decreased variance in the 30-day mortality for a pregnant woman with PE (holding all other factors constant) at the 5% level. No other factors had a statistically significant effect.

Tables 55–63 provide the results of the beta regressions for all modelled scenario analyses. The variance–covariance matrix for the base-case analysis is also provided in *Table 56*, as this was used to parameterise the beta regression coefficients as multivariate normal within the model PSA.

TABLE 55 The beta regression coefficients for the base-case analysis (average of all experts' answers, primary statistical population)

Parameter	Mean	SE	95% CI
Logit (mean effect)			
Intercept	–3.815	1.68	–7.108 to –0.521
Heart rate (b.p.m.)	0.024	0.00	0.021 to 0.027
Blood pressure (mmHg)	0.000	0.00	–0.002 to 0.003
Temperature (°C)	0.100	0.04	0.018 to 0.183
Respiratory rate (breaths per minute)	0.053	0.01	0.042 to 0.063
Peripheral oxygen saturation (%)	–0.077	0.01	–0.091 to –0.064
Post partum	0.107	0.08	–0.054 to 0.268
Weeks pregnant (continuous)	–0.001	0.00	–0.002 to 0.001
Ln (dispersion)			
Intercept	–15.382	9.04	–33.100 to 2.335
Heart rate (b.p.m.)	–0.004	0.01	–0.019 to 0.012
Blood pressure (mmHg)	0.034	0.01	0.019 to 0.050
Temperature (°C)	0.233	0.24	–0.228 to 0.694
Respiratory rate (breaths per minute)	–0.007	0.03	–0.060 to 0.046
Peripheral oxygen saturation (%)	0.103	0.03	0.048 to 0.158
Post partum	–0.756	0.42	–1.571 to 0.059
Weeks pregnant (continuous)	–0.001	0.00	–0.011 to 0.008

b.p.m., beats per minute; Ln, natural logarithm.

TABLE 56 The variance–covariance matrix associated with the base-case economic analysis

Parameter	(Intercept)	Heart.Rate	SBP	Temp	Resp_rate	Perip_ox_sat	Post_partum	Weeks_pregnant_preg_interact	(phi)_(Intercept)	(phi)_Heart.Rate	(phi)_SBP	(phi)_Temp	(phi)_Resp_rate	(phi)_Perip_ox_sat	(phi)_Post_partum	(phi)_Weeks_pregnant_preg_interact
(Intercept)	$2.82^{x10}+00$	$3.98^{x10}-04$	$-1.49^{x10}-04$	$-6.34^{x10}-02$	$2.46^{x10}-04$	$-5.23^{x10}-03$	$1.08^{x10}-02$	$1.20^{x10}-05$	$-2.43^{x10}+00$	$-3.57^{x10}-04$	$8.23^{x10}-05$	$5.55^{x10}-02$	$3.59^{x10}-04$	$4.13^{x10}-03$	$-8.36^{x10}-03$	$-1.57^{x10}-05$
Heart.Rate	$3.98^{x10}-04$	$2.23^{x10}-06$	$1.19^{x10}-07$	$-1.80^{x10}-05$	$-2.26^{x10}-06$	$6.40^{x10}-07$	$1.48^{x10}-05$	$4.37^{x10}-08$	$-3.57^{x10}-04$	$-2.10^{x10}-06$	$-1.31^{x10}-07$	$1.61^{x10}-05$	$2.21^{x10}-06$	$-4.34^{x10}-07$	$-1.35^{x10}-05$	$-4.05^{x10}-08$
SBP	$-1.49^{x10}-04$	$1.19^{x10}-07$	$1.72^{x10}-06$	$-2.15^{x10}-06$	$1.52^{x10}-06$	$-4.08^{x10}-07$	$-1.97^{x10}-06$	$7.02^{x10}-08$	$1.16^{x10}-04$	$-1.33^{x10}-07$	$-1.65^{x10}-06$	$2.80^{x10}-06$	$-1.61^{x10}-06$	$4.47^{x10}-07$	$1.54^{x10}-06$	$-6.70^{x10}-08$
Temp	$-6.34^{x10}-02$	$-1.80^{x10}-05$	$-2.15^{x10}-06$	$1.76^{x10}-03$	$-3.39^{x10}-05$	$1.39^{x10}-05$	$-4.05^{x10}-04$	$-6.74^{x10}-07$	$5.54^{x10}-02$	$1.61^{x10}-05$	$3.97^{x10}-06$	$-1.54^{x10}-03$	$1.54^{x10}-05$	$-1.12^{x10}-05$	$3.19^{x10}-04$	$9.41^{x10}-07$
Resp_rate	$2.46^{x10}-04$	$-2.26^{x10}-06$	$1.52^{x10}-06$	$-3.39^{x10}-05$	$3.00^{x10}-05$	$5.01^{x10}-06$	$-9.26^{x10}-05$	$-5.83^{x10}-07$	$3.10^{x10}-04$	$2.22^{x10}-06$	$-1.76^{x10}-06$	$1.51^{x10}-05$	$-2.57^{x10}-05$	$-4.08^{x10}-06$	$1.01^{x10}-04$	$5.18^{x10}-07$
Perip_ox_sat	$-5.23^{x10}-03$	$6.40^{x10}-07$	$-4.08^{x10}-07$	$1.39^{x10}-05$	$5.01^{x10}-06$	$4.77^{x10}-05$	$1.67^{x10}-05$	$-3.17^{x10}-07$	$4.15^{x10}-03$	$-4.50^{x10}-07$	$3.73^{x10}-07$	$-1.09^{x10}-05$	$-4.86^{x10}-06$	$-3.78^{x10}-05$	$-1.22^{x10}-05$	$2.43^{x10}-07$
Post_partum	$1.08^{x10}-02$	$1.48^{x10}-05$	$-1.97^{x10}-06$	$-4.05^{x10}-04$	$-9.26^{x10}-05$	$1.67^{x10}-05$	$6.75^{x10}-03$	$4.27^{x10}-05$	$-8.35^{x10}-03$	$-1.36^{x10}-05$	$1.29^{x10}-06$	$3.40^{x10}-04$	$9.86^{x10}-05$	$-1.95^{x10}-05$	$-6.48^{x10}-03$	$-4.14^{x10}-05$
Weeks_pregnant_preg_interact	$1.20^{x10}-05$	$4.37^{x10}-08$	$7.02^{x10}-08$	$-6.74^{x10}-07$	$-5.83^{x10}-07$	$-3.17^{x10}-07$	$4.27^{x10}-05$	$7.45^{x10}-07$	$-1.92^{x10}-05$	$-4.00^{x10}-08$	$-6.25^{x10}-08$	$1.14^{x10}-06$	$4.88^{x10}-07$	$2.04^{x10}-07$	$-4.16^{x10}-05$	$-7.16^{x10}-07$
(phi)_(Intercept)	$-2.43^{x10}+00$	$-3.57^{x10}-04$	$1.16^{x10}-04$	$5.54^{x10}-02$	$3.10^{x10}-04$	$4.15^{x10}-03$	$-8.35^{x10}-03$	$-1.92^{x10}-05$	$8.17^{x10}+01$	$5.64^{x10}-03$	$-4.66^{x10}-03$	$-2.00^{x10}+00$	$-1.97^{x10}-02$	$-8.07^{x10}-02$	$3.52^{x10}-01$	$-8.09^{x10}-04$
(phi)_Heart.Rate	$-3.57^{x10}-04$	$-2.10^{x10}-06$	$-1.33^{x10}-07$	$1.61^{x10}-05$	$2.22^{x10}-06$	$-4.50^{x10}-07$	$-1.36^{x10}-05$	$-4.00^{x10}-08$	$5.64^{x10}-03$	$6.24^{x10}-05$	$-2.77^{x10}-06$	$-3.28^{x10}-04$	$-8.21^{x10}-05$	$2.34^{x10}-05$	$4.64^{x10}-04$	$-1.50^{x10}-06$
(phi)_SBP	$8.23^{x10}-05$	$-1.31^{x10}-07$	$-1.65^{x10}-06$	$3.97^{x10}-06$	$-1.76^{x10}-06$	$3.73^{x10}-07$	$1.29^{x10}-06$	$-6.25^{x10}-08$	$-4.66^{x10}-03$	$-2.77^{x10}-06$	$6.29^{x10}-05$	$-4.85^{x10}-05$	$3.52^{x10}-05$	$-1.42^{x10}-05$	$-4.19^{x10}-04$	$-5.82^{x10}-07$
(phi)_Temp	$5.55^{x10}-02$	$1.61^{x10}-05$	$2.80^{x10}-06$	$-1.54^{x10}-03$	$1.51^{x10}-05$	$-1.09^{x10}-05$	$3.40^{x10}-04$	$1.14^{x10}-06$	$-2.00^{x10}+00$	$-3.28^{x10}-04$	$-4.85^{x10}-05$	$5.53^{x10}-02$	$-1.00^{x10}-04$	$5.22^{x10}-05$	$-1.45^{x10}-02$	$-8.72^{x10}-06$
(phi)_Resp_rate	$3.59^{x10}-04$	$2.21^{x10}-06$	$-1.61^{x10}-06$	$1.54^{x10}-05$	$-2.57^{x10}-05$	$-4.86^{x10}-06$	$9.86^{x10}-05$	$4.88^{x10}-07$	$-1.97^{x10}-02$	$-8.21^{x10}-05$	$3.52^{x10}-05$	$-1.00^{x10}-04$	$7.38^{x10}-04$	$1.43^{x10}-04$	$-2.15^{x10}-03$	$-6.76^{x10}-06$
(phi)_Perip_ox_sat	$4.13^{x10}-03$	$-4.34^{x10}-07$	$4.47^{x10}-07$	$-1.12^{x10}-05$	$-4.08^{x10}-06$	$-3.78^{x10}-05$	$-1.95^{x10}-05$	$2.04^{x10}-07$	$-8.07^{x10}-02$	$2.34^{x10}-05$	$-1.42^{x10}-05$	$5.22^{x10}-05$	$1.43^{x10}-04$	$7.80^{x10}-04$	$1.35^{x10}-03$	$1.78^{x10}-06$
(phi)_Post_partum	$-8.36^{x10}-03$	$-1.35^{x10}-05$	$1.54^{x10}-06$	$3.19^{x10}-04$	$1.01^{x10}-04$	$-1.22^{x10}-05$	$-6.48^{x10}-03$	$-4.16^{x10}-05$	$3.52^{x10}-01$	$4.64^{x10}-04$	$-4.19^{x10}-04$	$-1.45^{x10}-02$	$-2.15^{x10}-03$	$1.35^{x10}-03$	$1.73^{x10}-01$	$1.32^{x10}-03$
(phi)_Weeks_pregnant_preg_interact	$-1.57^{x10}-05$	$-4.05^{x10}-08$	$-6.70^{x10}-08$	$9.41^{x10}-07$	$5.18^{x10}-07$	$2.43^{x10}-07$	$-4.14^{x10}-05$	$-7.16^{x10}-07$	$-8.09^{x10}-04$	$-1.50^{x10}-06$	$-5.82^{x10}-07$	$-8.72^{x10}-06$	$-6.76^{x10}-06$	$1.78^{x10}-06$	$1.32^{x10}-03$	$2.28^{x10}-05$

SBP, systolic blood pressure.

TABLE 57 The beta regression coefficients for the base-case analysis (average of all experts' answers, secondary statistical population)

Parameter	Mean	SE	95% CI
Logit (mean effect)			
Intercept	–4.35744	1.425821	–7.152 to –1.56288
Heart rate (b.p.m.)	0.024	0.00	0.020921 to 0.026877
Blood pressure (mmHg)	–0.002	0.00	–0.00407 to 0.001067
Temperature (°C)	0.130	0.04	0.059289 to 0.200863
Respiratory rate (breaths per minute)	0.049	0.01	0.038068 to 0.059609
Peripheral oxygen saturation (%)	–0.079	0.01	–0.09166 to –0.06675
Post partum	0.077	0.08	–0.07849 to 0.23275
Weeks pregnant (continuous)	–0.002	0.00	–0.00327 to 4.78E – 05
Ln (dispersion)			
Intercept	–10.324	7.73	–25.4684 to 4.819999
Heart rate (b.p.m.)	–0.001	0.01	–0.01518 to 0.013829
Blood pressure (mmHg)	0.024	0.01	0.009269 to 0.0384
Temperature (°C)	0.166	0.20	–0.22148 to 0.552588
Respiratory rate (breaths per minute)	–0.022	0.03	–0.07306 to 0.02831
Peripheral oxygen saturation (%)	0.090	0.03	0.036247 to 0.144063
Post partum	–0.808	0.39	–1.57152 to –0.04532
Weeks pregnant (continuous)	–0.004	0.00	–0.01257 to 0.004924

b.p.m., beats per minute.

TABLE 58 The beta regression coefficients for the base-case analysis (average of all experts' answers, tertiary statistical population)

Parameter	Mean	SE	95% CI
Logit (mean effect)			
Intercept	–3.81468	1.68042	–7.10825 to –0.52112
Heart rate (b.p.m.)	0.024	0.00	0.021198 to 0.027051
Blood pressure (mmHg)	0.000	0.00	–0.00217 to 0.002972
Temperature (°C)	0.100	0.04	0.018308 to 0.182639
Respiratory rate (breaths per minute)	0.053	0.01	0.041905 to 0.063361
Peripheral oxygen saturation (%)	–0.077	0.01	–0.0908 to –0.06373
Post partum	0.107	0.08	–0.05427 to 0.267697
Weeks pregnant (continuous)	–0.001	0.00	–0.00238 to 0.001002
Ln (dispersion)			
Intercept	–15.382	9.04	–33.0996 to 2.335226
Heart rate (b.p.m.)	–0.004	0.01	–0.01909 to 0.011877
Blood pressure (mmHg)	0.034	0.01	0.018624 to 0.049721
Temperature (°C)	0.233	0.24	–0.22815 to 0.693645
Respiratory rate (breaths per minute)	–0.007	0.03	–0.06021 to 0.046304
Peripheral oxygen saturation (%)	0.103	0.03	0.04805 to 0.157502
Post partum	–0.756	0.42	–1.57078 to 0.058617
Weeks pregnant (continuous)	–0.001	0.00	–0.01067 to 0.008039

b.p.m., beats per minute.

TABLE 59 The beta regression coefficients for the base-case analysis (average of all experts' answers, quaternary statistical population)

Parameter	Mean	SE	95% CI
Logit (mean effect)			
Intercept	–3.59	1.677087	–6.87703 to –0.30297
Heart rate (b.p.m.)	0.024	0.00	0.021011 to 0.026792
Blood pressure (mmHg)	0.001	0.00	–0.00136 to 0.003529
Temperature (°C)	0.097	0.04	0.013497 to 0.180258
Respiratory rate (breaths per minute)	0.052	0.01	0.041777 to 0.063114
Peripheral oxygen saturation (%)	–0.079	0.01	–0.09238 to –0.06563
Post partum	0.155	0.08	–0.00156 to 0.312137
Weeks pregnant (continuous)	–0.001	0.00	–0.00219 to 0.001128
Ln (dispersion)			
Intercept	–14.672	9.08	–32.4684 to 3.123832
Heart rate (b.p.m.)	–0.003	0.01	–0.01838 to 0.013273
Blood pressure (mmHg)	0.039	0.01	0.022945 to 0.054807
Temperature (°C)	0.192	0.24	–0.27241 to 0.6572
Respiratory rate (breaths per minute)	–0.009	0.03	–0.06267 to 0.044416
Peripheral oxygen saturation (%)	0.104	0.03	0.049105 to 0.159394
Post partum	–0.650	0.43	–1.48727 to 0.187474
Weeks pregnant (continuous)	–0.001	0.00	–0.01052 to 0.008347

b.p.m., beats per minute.

TABLE 60 The beta regression coefficients for the base-case analysis (expert one only, primary statistical population)

Parameter	Mean	SE	95% CI
Logit (mean effect)			
Intercept	–10.6046	3.027756	–16.5389 to –4.67033
Heart rate (b.p.m.)	0.034	0.00	0.02926 to 0.039054
Blood pressure (mmHg)	0.004	0.00	0.000921 to 0.007971
Temperature (°C)	0.273	0.07	0.126813 to 0.419116
Respiratory rate (breaths per minute)	0.081	0.01	0.070187 to 0.092472
Peripheral oxygen saturation (%)	–0.093	0.01	–0.11333 to –0.07261
Post partum	0.186	0.14	–0.09588 to 0.467419
Weeks pregnant (continuous)	–0.001	0.00	–0.00421 to 0.001827
Ln (dispersion)			
Intercept	1.458	8.94	–16.0548 to 18.9707
Heart rate (b.p.m.)	–0.056	0.01	–0.07094 to –0.04027
Blood pressure (mmHg)	0.036	0.01	0.020354 to 0.050788
Temperature (°C)	–0.181	0.23	–0.63721 to 0.274615
Respiratory rate (breaths per minute)	0.065	0.03	0.012839 to 0.116601
Peripheral oxygen saturation (%)	0.112	0.03	0.059088 to 0.164267
Post partum	–1.246	0.41	–2.0579 to –0.43364
Weeks pregnant (continuous)	–0.001	0.00	–0.01 to 0.008805

b.p.m., beats per minute.

TABLE 61 The beta regression coefficients for the base-case analysis (expert two only, primary statistical population)

Parameter	Mean	SE	95% CI
Logit (mean effect)			
Intercept	–5.32288	1.940926	–9.12702 to –1.51873
Heart rate (b.p.m.)	0.012	0.00	0.0093 to 0.014984
Blood pressure (mmHg)	0.001	0.00	–0.00048 to 0.003333
Temperature (°C)	0.176	0.05	0.079966 to 0.273029
Respiratory rate (breaths per minute)	0.047	0.01	0.036686 to 0.057529
Peripheral oxygen saturation (%)	–0.080	0.01	–0.095 to –0.06561
Post partum	0.076	0.08	–0.08993 to 0.240977
Weeks pregnant (continuous)	0.000	0.00	–0.00191 to 0.001482
Ln (dispersion)			
Intercept	–4.394	8.98	–22.0034 to 13.21459
Heart rate (b.p.m.)	–0.033	0.01	–0.04806 to –0.01708
Blood pressure (mmHg)	0.028	0.01	0.012448 to 0.043463
Temperature (°C)	–0.006	0.23	–0.46311 to 0.451116
Respiratory rate (breaths per minute)	–0.025	0.03	–0.07724 to 0.027719
Peripheral oxygen saturation (%)	0.119	0.03	0.064695 to 0.173743
Post partum	–0.375	0.42	–1.19287 to 0.443607
Weeks pregnant (continuous)	0.007	0.00	–0.00207 to 0.016705

b.p.m., beats per minute.

TABLE 62 The beta regression coefficients for the base-case analysis (expert three only, primary statistical population)

Parameter	Mean	SE	95% CI
Logit (mean effect)			
Intercept	4.400131	0.963907	2.510908 to 6.289355
Heart rate (b.p.m.)	0.011	0.00	0.008667 to 0.013769
Blood pressure (mmHg)	–0.001	0.00	–0.00436 to 0.001715
Temperature (°C)	–0.049	0.03	–0.10065 to 0.003562
Respiratory rate (breaths per minute)	0.016	0.00	0.009221 to 0.023241
Peripheral oxygen saturation (%)	–0.075	0.00	–0.07879 to –0.07122
Post partum	–0.018	0.09	–0.1918 to 0.155907
Weeks pregnant (continuous)	–0.001	0.00	–0.00238 to 0.001343
Ln (dispersion)			
Intercept	–7.329	9.00	–24.9632 to 10.30504
Heart rate (b.p.m.)	0.032	0.01	0.01627 to 0.04678
Blood pressure (mmHg)	0.010	0.01	–0.00567 to 0.025051
Temperature (°C)	0.311	0.23	–0.14807 to 0.769327
Respiratory rate (breaths per minute)	0.062	0.03	0.009499 to 0.11547
Peripheral oxygen saturation (%)	–0.044	0.03	–0.09868 to 0.010815
Post partum	–0.048	0.41	–0.85734 to 0.760827
Weeks pregnant (continuous)	0.004	0.00	–0.00518 to 0.013551

b.p.m., beats per minute.

TABLE 63 The beta regression coefficients for the base-case analysis (expert four only, primary statistical population)

Parameter	Mean	SE	95% CI
Logit (mean effect)			
Intercept	2.593387	2.010734	–1.34758 to 6.534354
Heart rate (b.p.m.)	0.023	0.00	0.018729 to 0.027486
Blood pressure (mmHg)	–0.009	0.00	–0.01316 to –0.00442
Temperature (°C)	0.058	0.05	–0.03968 to 0.156479
Respiratory rate (breaths per minute)	0.015	0.01	0.000676 to 0.029638
Peripheral oxygen saturation (%)	–0.097	0.01	–0.11525 to –0.07895
Post partum	0.337	0.10	0.137803 to 0.537004
Weeks pregnant (continuous)	0.000	0.00	–0.00226 to 0.002564
Ln (dispersion)			
Intercept	–10.013	8.96	–27.5773 to 7.552107
Heart rate (beats per minute)	0.016	0.01	0.000644 to 0.0314
Blood pressure (mmHg)	–0.018	0.01	–0.03271 to –0.00234
Temperature (°C)	0.147	0.23	–0.31141 to 0.605584
Respiratory rate (breaths per minute)	–0.068	0.03	–0.11987 to –0.01596
Peripheral oxygen saturation (%)	0.122	0.03	0.069916 to 0.174771
Post partum	–0.235	0.42	–1.05028 to 0.581006
Weeks pregnant (continuous)	–0.005	0.00	–0.01465 to 0.004101

b.p.m., beats per minute.

Appendix 16 The survival curves fitted to the Kaplan–Meier curves presented in Delcroix *et al.*[72]

Figure 54 indicates that the surgically treated group had a relatively smooth hazard function, whereas there was a discontinuity in the hazard function for the medically treated group at around 3 years.

Figure 55 indicates that an exponential curve may be appropriate for the medically treated group, but not for the surgically treated group. This is because the medically treated group approximated a straight line at 45 degrees, whereas the surgically treated group line did not.

Figure 56 indicates that the log-logistic curve may be appropriate for either of the groups, as the lines are approximately straight.

Figure 57 indicates that the log-normal curve may be appropriate for either of the groups, as the lines are approximately straight.

Figure 58 shows the estimated survivor functions for people whose CTEPH was surgically treated over a 20-year horizon. *Figure 59* shows the corresponding data for those people whose CTEPH was medically treated.

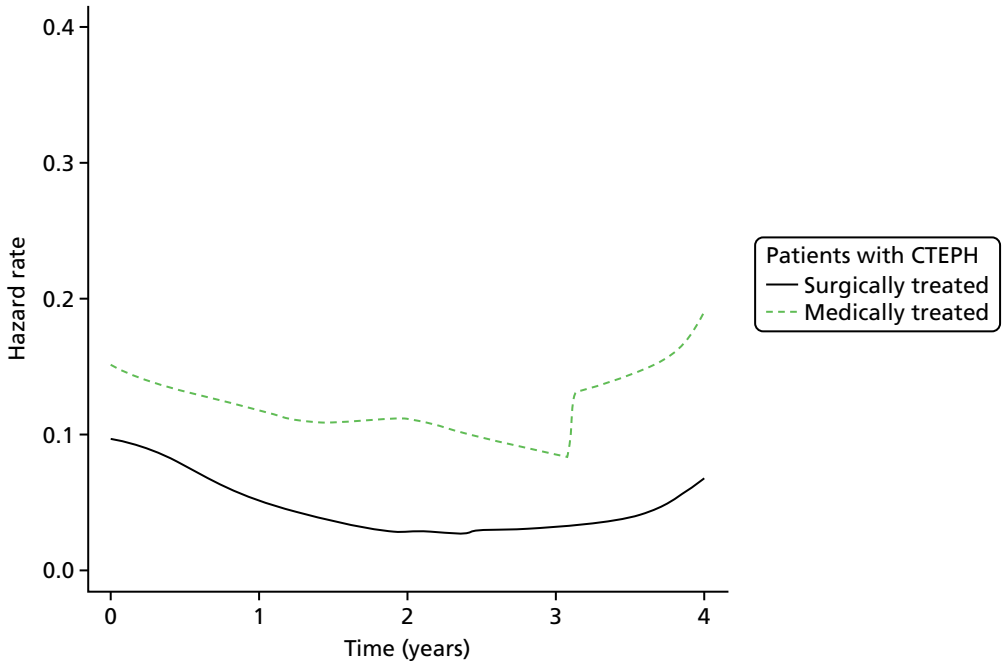


FIGURE 54 The empirical hazard plot for the surgically and non-surgically treated CTEPH patients.

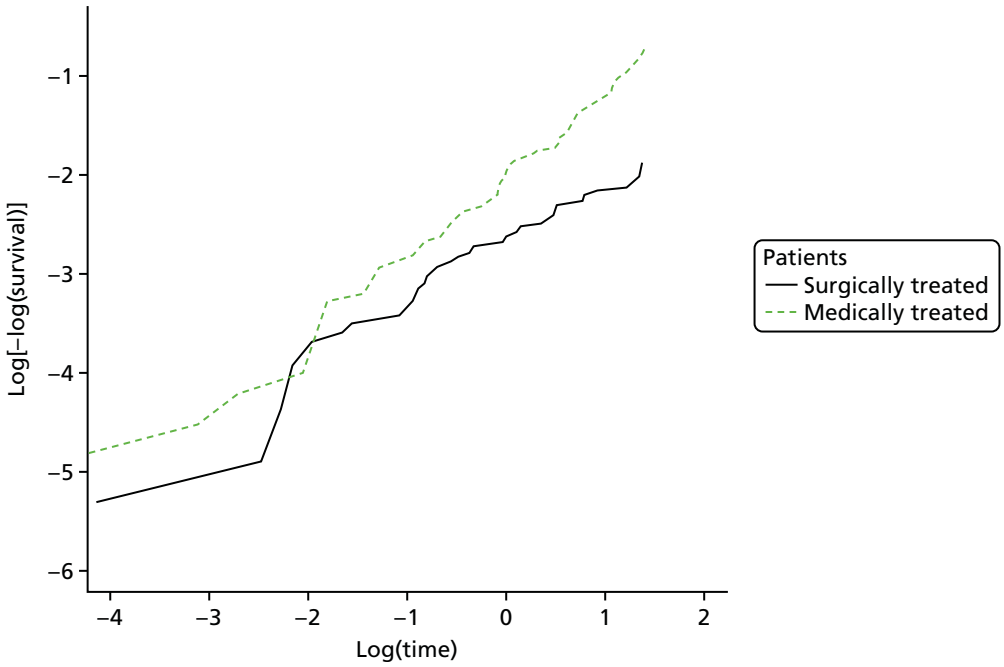


FIGURE 55 The plot of the log-cumulative hazard vs. the log-time to assess the suitability of the Weibull and exponential parametric distributions.

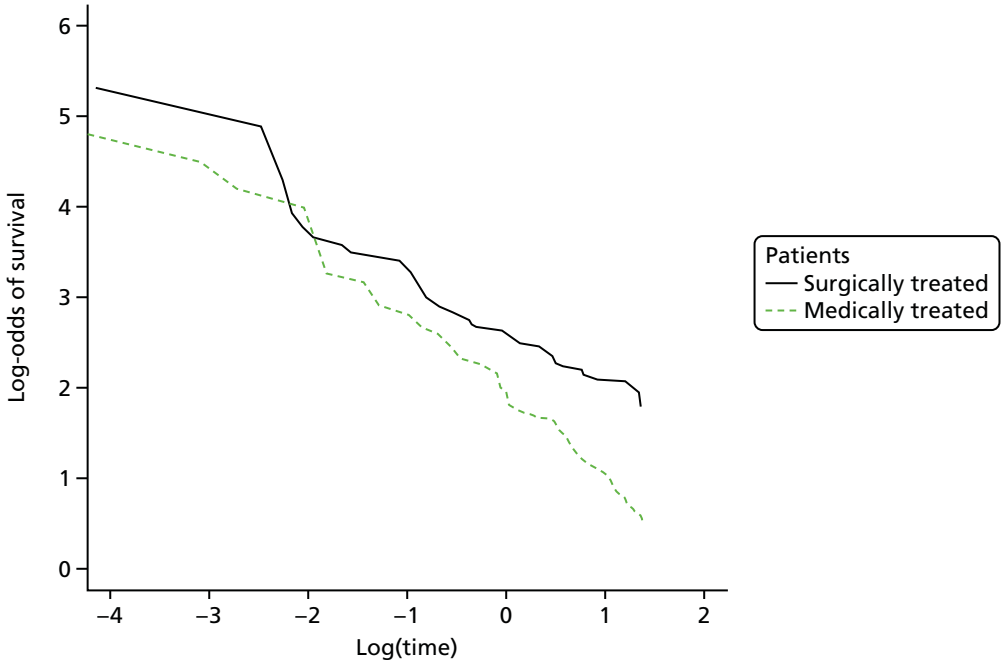


FIGURE 56 The plot of the log-odds of survival vs. the log-time to assess the suitability of a log-logistic distribution.

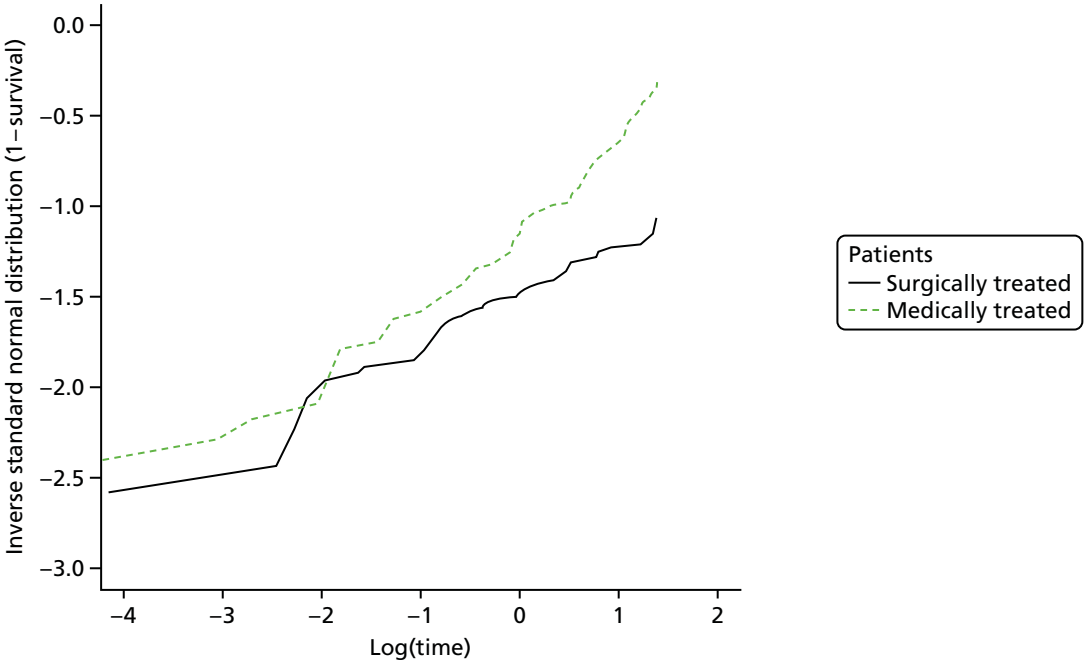


FIGURE 57 The plot of the inverse standard normal distribution vs. the log-time to assess the suitability of a log-normal distribution.

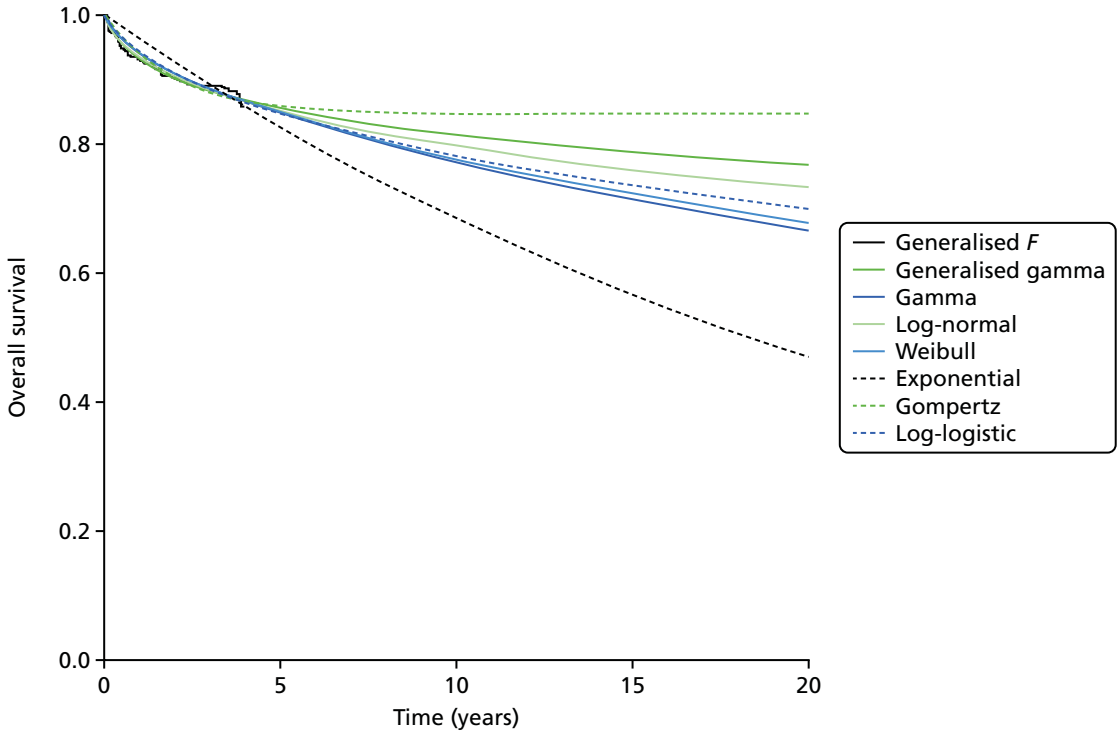


FIGURE 58 The fit of the parametric survival curves to the reconstructed Kaplan–Meier data for people with CTEPH who were surgically treated.

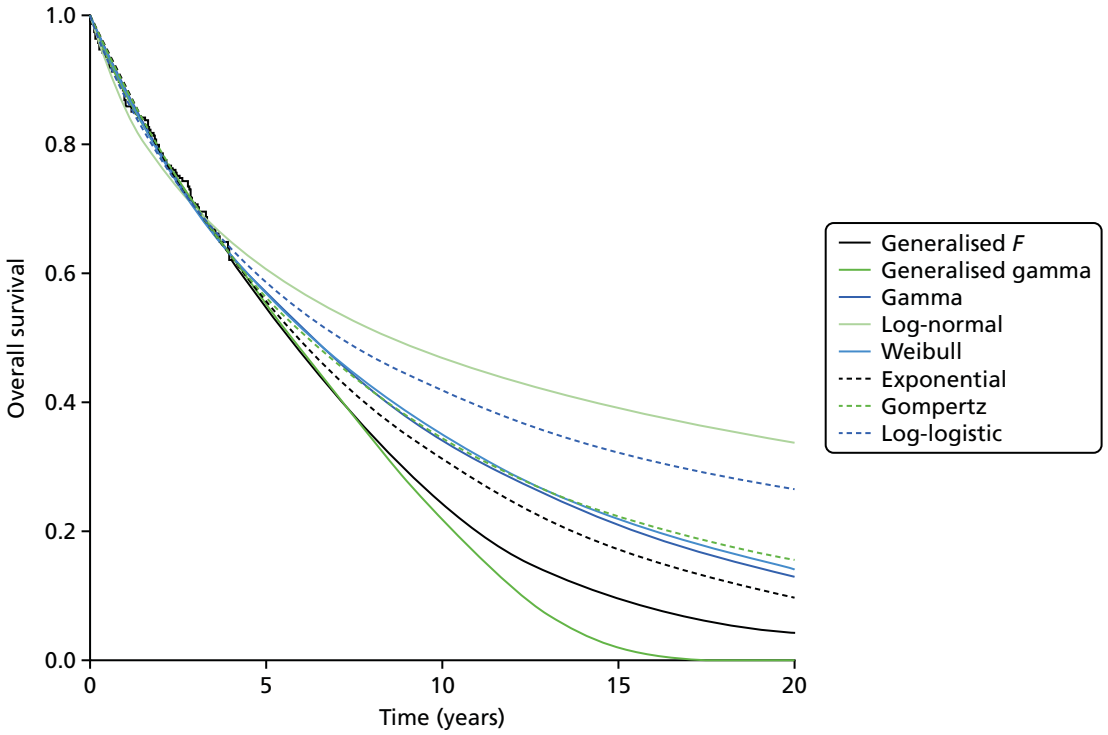


FIGURE 59 The fit of the parametric survival curves to the reconstructed Kaplan–Meier data for people with CTEPH who were medically treated.

Table 64 shows the Akaike information criterion (AIC) and the BIC for each of the parametric survival curves fitted to the reconstructed Kaplan–Meier data obtained from Delcroix *et al.*[72] The curve with the lowest BIC (and AIC) was considered to be the base-case survival curve used in the economic model. Each of the other candidate curves was considered to be eligible for inclusion in the scenario analyses unless there was very strong evidence against the curve compared with the best-fitting curve. This excluded the generalised *F* and exponential curves for the surgically treated group and the generalised *F* and log-normal for the medically treated group. The results of the fitted curves and the associated variance–covariance

TABLE 64 The AIC and the BIC for each of the modelled survival curves

	Surgically treated			Medically treated		
Parametric surviver curve	AIC	BIC	Evidence against[a]	AIC	BIC	Evidence against[a]
Generalised *F*	433.0	449.0	Very strong	572.6	587.1	Very strong
Generalised gamma	431.0	443.0	Positive	570.6	581.4	Strong
Gamma	432.5	440.5	Positive	569.3	576.6	Positive
Log-normal	429.9	437.9	Base case	576.9	584.2	Very strong
Log-logistic	431.9	439.9	Positive	571.0	578.2	Strong
Gompertz	435.2	443.2	Positive	570.4	577.7	Positive
Weibull	432.3	440.3	Positive	569.5	576.7	Positive
Exponential	446.4	450.4	Very strong	568.5	572.2	Base case

a Evidence against is based on the recommendations of Kass *et al.*,[99] which state that differences in BICs of 0–6 indicate that there is positive evidence that the lowest BIC is the best fit, BICs of 6–10 indicate that there is strong evidence that the lowest BIC is the best fit and BICs of > 10 indicate that there is decisive evidence that the lowest BIC is the best fit.

matrices for the log-normal curve for the surgically treated CTEPH patients and the exponential curve for the medically treated CTEPH patients are given in *Tables 65–67*.

The results of the other parametric curves are provided in *Tables 68–78*.

TABLE 65 The result of the fitted log-normal curve in the surgically treated population

Parameter	Mean	SE	95% CI
Mean log	5.081	0.574	3.956 to 6.206
SD log	3.343	0.399	2.646 to 4.224

SD, standard deviation.

TABLE 66 The variance covariance matrix for the fitted log-normal curve in the surgically treated population

Parameter	Mean log	SD log
Mean log	0.01770766	–0.05571957
SD log	–0.05571957	0.23093510

SD, standard deviation.

TABLE 67 The results of the fitted exponential curve in the medically treated population

Parameter	Mean	SE	95% CI
Lambda	0.1168	0.0123	0.0950 to 0.1436

TABLE 68 The result of the fitted generalised gamma curve in the surgically treated population

Parameter	Mean	SE	95% CI
Mu	4.57	1.41	1.80 to 7.33
Sigma	5.01	1.44	2.85 to 8.82
Q	–1.21	1.44	–4.03 to 1.60

TABLE 69 The result of the fitted gamma curve in the surgically treated population

Parameter	Mean	SE	95% CI
Shape	0.59128	0.08370	0.4480 to 0.78034
Rate	0.00709	0.00417	0.00224 to 0.02245

TABLE 70 The result of the fitted log-logistic curve in the surgically treated population

Parameter	Mean	SE	95% CI
Shape	0.6281	0.0836	0.4839 to 0.8152
Scale	76.5809	36.8015	29.8586 to 196.4134

TABLE 71 The result of the fitted Gompertz curve in the surgically treated population

Parameter	Mean	SE	95% CI
Shape	–0.4812	0.1385	–0.7527 to –0.2097
Rate	0.0805	0.0183	0.0515 to 0.1257

TABLE 72 The result of the fitted Weibull curve in the surgically treated population

Parameter	Mean	SE	95% CI
Shape	0.6073	0.0823	0.4655 to 0.7921
Scale	96.1513	48.4245	35.83 to 258.0

TABLE 73 The result of the fitted generalised gamma curve in the medically treated population

Parameter	Mean	SE	95% CI
Mu	2.34	0.677	1.02 to 3.67
Sigma	0.435	1.37	9.23e – 04 to 2.05e + 02[a]
Q	2.80	8.80	–14.4 to 20.0

a Scientific notation is used to denote the number(s), as these are either very large or very close to zero.

TABLE 74 The result of the fitted gamma curve in the medically treated population

Parameter	Mean	SE	95% CI
Shape	0.8829	0.1023	(0.7036 to 1.1079)
Rate	0.0926	0.0232	(0.0567 to 0.1514)

TABLE 75 The result of the fitted log-logistic curve in the medically treated population

Parameter	Mean	SE	95% CI
Shape	0.9905	0.0948	0.8212 to 1.1948
Scale	7.1227	1.1133	5.2433 to 9.6758

TABLE 76 The result of the fitted Gompertz curve in the medically treated population

Parameter	Mean	SE	95% CI
Shape	–0.0289	0.0970	–0.2190 to 0.1613
Rate	0.1227	0.0238	0.0838 to 0.1794

TABLE 77 The result of the fitted Weibull curve in the medically treated population

Parameter	Mean	SE	95% CI
Shape	0.9053	0.0892	0.7463 to 1.0982
Scale	9.4806	1.4976	6.9563 to 12.9210

TABLE 78 The results of the fitted exponential curve in the medically treated population

Parameter	Mean	SE	95% CI
Lambda	0.1168	0.0123	0.0950 to 0.1436

Appendix 17 The utility parameters used in the decision-analysis model

Table 79 shows the utility parameters used in the decision-analysis model.

TABLE 79 The utility parameters used in the decision-analysis model

Parameter	Mean	SE	Distribution used	95% CI/IQR		Alpha	Beta	Source	Notes
Baseline utility									
Constant	0.95086		Multivariate t, df = 26678 (n – 1)					Ara and Brazier[76]	Variance-covariance matrix is available from http://eprints.whiterose.ac.uk/10880/1/HEDSDP 09-12.pdf (accessed 24 July 2018)
Male	0.02121								
Age	–0.00026								
Age^2	–0.00003								
Utility multiplier for a person with PE									
No PE or DVT or bleeding	0.96	0.011	Beta	0.82	1	10.658	0.953	Locadia *et al.*[77]	SE was calculated from an IQR[a]
PE	0.63	0.033	Beta	0.36	0.86	77.478	0.660		
Duration of the multiplier for PE	4 weeks		Fixed					Clinical input	
Utility multiplier for a person with DVT									
Utility of a person with DVT	0.84	0.033	Beta	0.64	0.98	219.537	41.817	Locadia *et al.*[77]	SE was calculated from an IQR[a]
Duration of the multiplier for DVT	4 weeks		Fixed					Clinical input	
Utility multiplier for a person with a gastrointestinal bleeding event									
Utility for a person with a gastrointestinal bleed	0.65	0.025	Beta	0.49	0.86	130.898	0.674	Locadia *et al.*[77]	SE was calculated from an IQR[a]
Duration of the multiplier for a gastrointestinal bleed	4 weeks		Fixed					Clinical input	

Parameter	Mean	SE	Distribution used	95% CI/IQR		Alpha	Beta	Source	Notes
Utility multiplier for a person with CTEPH									
Utility of a person without CTEPH	0.795	0.021	Beta	0.771	0.831	183.2	155.4	Ara and Brazier[78]	On the basis of clinical input, it was assumed that other heart problems would be a similar condition to CTEPH in terms of quality of life
Utility of a person with CTEPH	0.672	0.012	Beta	0.649	0.694	524.7	0.69		
Duration of the multiplier for CTEPH	Permanent							Clinical input	
Utility multiplier for a person with an intracranial haemorrhage									
Utility of a person without an intracranial haemorrhage	0.828	0.012	Beta	0.804	0.851	785.607	163.194	Ara and Brazier[78]	The effects of an intracranial haemorrhage were assumed to be equivalent to those of a stroke
Utility of a person with an intracranial haemorrhage	0.541	0.027	Beta	0.488	0.593	183.177	155.413		
Duration of the multiplier for an intracranial haemorrhage	Permanent							Clinical input	
Utility multiplier for a person with cancer									
Utility of a person without cancer	0.697	0.020	Beta	0.657	0.736	352.718	153.333	Ara and Brazier[78]	
Utility of a person with cancer	0.795	0.021	Beta	0.754	0.836	295.290	76.144		
Duration of the multiplier for cancer	Permanent							Clinical input	

df, degrees of freedom; IQR, interquartile range.

a The SE was calculated from the interquartile ranges using the following two formulae:
- standard deviation = (interquartile range high – interquartile rangelow)/(2 × 0.6475).
- SE = standard deviation/$(n - 1)^{0.5}$.

Appendix 18 The discounted costs and quality-adjusted life-year losses associated with radiation-induced cancers in the mother by the age at which they were scanned

Table 80 shows the discounted costs and QALY losses associated with radiation-induced cancers, when they present over a lifetime, in the mother by the age of the mother when she was scanned. *Table 81* shows the discounted costs and QALY losses associated with radiation-induced cancers, when they present within 15 years, in the mother by the age at which they were scanned. All childhood cancers were assumed to present within 12 years of the initial scan. The associated QALY loss per surviving fetus was –0.00004 QALYs per scan and the associated cost was £0.57 per scan.

TABLE 80 The discounted costs and QALY losses associated with radiation-induced cancers, when they present over a lifetime, in the mother by the age at which they were scanned

Age (years) at which the mother was scanned	Cancer							
	Breast				Lung			
	QALYs		Cost (£)		QALYs		Cost (£)	
	CT scan	VQ SPECT	CT scan	VQ SPECT	CT scan	VQ SPECT	CT scan	VQ SPECT
16	–0.00140	–0.00004	4.56	0.14	–0.00141	–0.00141	3.08	3.08
17	–0.00142	–0.00004	4.62	0.14	–0.00144	–0.00144	3.16	3.16
18	–0.00143	–0.00004	4.67	0.14	–0.00148	–0.00148	3.24	3.24
19	–0.00145	–0.00004	4.72	0.14	–0.00152	–0.00152	3.33	3.33
20	–0.00146	–0.00004	4.76	0.14	–0.00156	–0.00156	3.41	3.41
21	–0.00147	–0.00004	4.81	0.14	–0.00160	–0.00160	3.50	3.50
22	–0.00148	–0.00004	4.85	0.15	–0.00164	–0.00164	3.59	3.59
23	–0.00150	–0.00004	4.88	0.15	–0.00168	–0.00168	3.68	3.68
24	–0.00151	–0.00005	4.92	0.15	–0.00172	–0.00172	3.78	3.78
25	–0.00151	–0.00005	4.95	0.15	–0.00177	–0.00177	3.87	3.87
26	–0.00152	–0.00005	4.97	0.15	–0.00181	–0.00181	3.97	3.97
27	–0.00152	–0.00005	4.98	0.15	–0.00186	–0.00186	4.07	4.07
28	–0.00152	–0.00005	4.99	0.15	–0.00190	–0.00190	4.17	4.17
29	–0.00152	–0.00005	5.00	0.15	–0.00195	–0.00195	4.28	4.28
30	–0.00152	–0.00005	5.00	0.15	–0.00200	–0.00200	4.39	4.39
31	–0.00150	–0.00005	4.98	0.15	–0.00204	–0.00204	4.49	4.49
32	–0.00149	–0.00004	4.96	0.15	–0.00209	–0.00209	4.61	4.61
33	–0.00148	–0.00004	4.94	0.15	–0.00214	–0.00214	4.72	4.72
34	–0.00146	–0.00004	4.90	0.15	–0.00219	–0.00219	4.83	4.83
35	–0.00144	–0.00004	4.87	0.15	–0.00225	–0.00225	4.95	4.95

continued

TABLE 80 The discounted costs and QALY losses associated with radiation-induced cancers, when they present over a lifetime, in the mother by the age at which they were scanned (*continued*)

Age (years) at which the mother was scanned	Cancer							
	Breast				Lung			
	QALYs		Cost (£)		QALYs		Cost (£)	
	CT scan	VQ SPECT	CT scan	VQ SPECT	CT scan	VQ SPECT	CT scan	VQ SPECT
36	–0.00141	–0.00004	4.80	0.14	–0.00230	–0.00230	5.07	5.07
37	–0.00138	–0.00004	4.73	0.14	–0.00235	–0.00235	5.19	5.19
38	–0.00135	–0.00004	4.66	0.14	–0.00240	–0.00240	5.31	5.31
39	–0.00132	–0.00004	4.58	0.14	–0.00245	–0.00245	5.44	5.44
40	–0.00129	–0.00004	4.49	0.13	–0.00251	–0.00251	5.57	5.57
41	–0.00123	–0.00004	4.36	0.13	–0.00256	–0.00256	5.69	5.69
42	–0.00118	–0.00004	4.23	0.13	–0.00260	–0.00260	5.81	5.81
43	–0.00112	–0.00003	4.09	0.12	–0.00265	–0.00265	5.94	5.94
44	–0.00107	–0.00003	3.94	0.12	–0.00270	–0.00270	6.07	6.07
45	–0.00101	–0.00003	3.79	0.11	–0.00276	–0.00276	6.20	6.20
46	–0.00094	–0.00003	3.59	0.11	–0.00279	–0.00279	6.32	6.32
47	–0.00087	–0.00003	3.39	0.10	–0.00283	–0.00283	6.45	6.45
48	–0.00079	–0.00002	3.18	0.10	–0.00287	–0.00287	6.57	6.57
49	–0.00072	–0.00002	2.96	0.09	–0.00291	–0.00291	6.70	6.70
50	–0.00065	–0.00002	2.74	0.08	–0.00296	–0.00296	6.83	6.83

TABLE 81 The discounted costs and QALY losses associated with radiation-induced cancers, when they present within 15 years, in the mother by the age at which they were scanned

Age (years) at which the mother was scanned	Cancer							
	Breast				Lung			
	QALYs		Cost (£)		QALYs		Cost (£)	
	CT scan	VQ SPECT	CT scan	VQ SPECT	CT scan	VQ SPECT	CT scan	VQ SPECT
16	–0.00771	–0.00023	14.99	0.45	–0.01498	–0.01498	14.79	14.79
17	–0.00749	–0.00022	14.70	0.44	–0.01485	–0.01485	14.67	14.67
18	–0.00732	–0.00022	14.50	0.43	–0.01473	–0.01473	14.55	14.55
19	–0.00718	–0.00022	14.32	0.43	–0.01460	–0.01460	14.42	14.42
20	–0.00705	–0.00021	14.17	0.43	–0.01447	–0.01447	14.30	14.30
21	–0.00674	–0.00020	13.74	0.41	–0.01435	–0.01435	14.18	14.18
22	–0.00651	–0.00020	13.43	0.40	–0.01422	–0.01422	14.05	14.05
23	–0.00632	–0.00019	13.19	0.40	–0.01410	–0.01410	13.93	13.93
24	–0.00616	–0.00018	12.97	0.39	–0.01397	–0.01397	13.80	13.80
25	–0.00601	–0.00018	12.77	0.38	–0.01385	–0.01385	13.68	13.68

TABLE 81 The discounted costs and QALY losses associated with radiation-induced cancers, when they present within 15 years, in the mother by the age at which they were scanned (*continued*)

Age (years) at which the mother was scanned	Cancer							
	Breast				Lung			
	QALYs		Cost (£)		QALYs		Cost (£)	
	CT scan	VQ SPECT	CT scan	VQ SPECT	CT scan	VQ SPECT	CT scan	VQ SPECT
26	–0.00567	–0.00017	12.27	0.37	–0.01372	–0.01372	13.56	13.56
27	–0.00541	–0.00016	11.89	0.36	–0.01360	–0.01360	13.43	13.43
28	–0.00519	–0.00016	11.58	0.35	–0.01347	–0.01347	13.31	13.31
29	–0.00500	–0.00015	11.30	0.34	–0.01335	–0.01335	13.19	13.19
30	–0.00483	–0.00015	11.05	0.33	–0.01322	–0.01322	13.06	13.06
31	–0.00454	–0.00014	10.59	0.32	–0.01310	–0.01310	12.94	12.94
32	–0.00430	–0.00013	10.22	0.31	–0.01297	–0.01297	12.81	12.81
33	–0.00410	–0.00012	9.89	0.30	–0.01285	–0.01285	12.69	12.69
34	–0.00391	–0.00012	9.59	0.29	–0.01272	–0.01272	12.57	12.57
35	–0.00374	–0.00011	9.31	0.28	–0.01260	–0.01260	12.44	12.44
36	–0.00354	–0.00011	8.97	0.27	–0.01247	–0.01247	12.32	12.32
37	–0.00336	–0.00010	8.65	0.26	–0.01235	–0.01235	12.20	12.20
38	–0.00318	–0.00010	8.34	0.25	–0.01222	–0.01222	12.07	12.07
39	–0.00302	–0.00009	8.03	0.24	–0.01210	–0.01210	11.95	11.95
40	–0.00286	–0.00009	7.74	0.23	–0.01197	–0.01197	11.83	11.83
41	–0.00269	–0.00008	7.40	0.22	–0.01185	–0.01185	11.70	11.70
42	–0.00252	–0.00008	7.07	0.21	–0.01172	–0.01172	11.58	11.58
43	–0.00236	–0.00007	6.74	0.20	–0.01160	–0.01160	11.45	11.45
44	–0.00220	–0.00007	6.41	0.19	–0.01147	–0.01147	11.33	11.33
45	–0.00205	–0.00006	6.07	0.18	–0.01135	–0.01135	11.21	11.21
46	–0.00186	–0.00006	5.66	0.17	–0.01122	–0.01122	11.08	11.08
47	–0.00168	–0.00005	5.26	0.16	–0.01109	–0.01109	10.96	10.96
48	–0.00151	–0.00005	4.86	0.15	–0.01097	–0.01097	10.84	10.84
49	–0.00135	–0.00004	4.46	0.13	–0.01084	–0.01084	10.71	10.71
50	–0.00119	–0.00004	4.06	0.12	–0.01072	–0.01072	10.59	10.59

Appendix 19 The cost parameters used in the decision-analysis model

Table 82 shows the cost parameters used in the decision-analysis model.

TABLE 82 The cost parameters used in the decision-analysis model

Parameter	Mean	SE	Distribution used	95% CI	Alpha/*n*	Beta	Source	Notes
Decision rule costs								
Cost per hour of registrar time	£42	–	Fixed	–	–	–	PSSRU (2015–16)[87]	–
Time taken to apply the DR	5 minutes	–	Fixed	–	–	–	Clinical input	–
Cost of a decision rule	£3.50	–	Fixed	–	–	–	–	–
Scanning costs								
CTPA scan	£130	4.9	Normal	£121 to £140	–	–	*NHS Reference Costs 2015 to 2016*[90]	Direct access imaging CT scan of one area, post-contrast only
VQ scan (2009–10 prices)	£253	–	Gamma	£142 to £395	15.15	16.69	NICE guideline CG144[33]	p. 543
Cost of modelled events								
Cost of treating PE	£4778	224.7	Normal	£4337 to £5218	–	–	*NHS Reference Costs 2015 to 16*[90]	Currency codes DZ09J–DZ09K[a]
Cost of treating a DVT	£2612	68.6	Normal	£2478 to £2747	–	–	*NHS Reference Costs 2015 to 16*[90]	Currency codes YQ51A–YQ51E[a]
Cost of treating a gastrointestinal bleed	£2201	65.0	Normal	£2074 to £2328	–	–	*NHS Reference Costs 2015 to 16*[90]	Currency codes FZ38G–FZ38P[a]
Cost of treating an intracranial haemorrhage – year 1	£11,707	2341	Gamma	£7576 to £16,722	25.00	468.27	Luengo-Fernandez *et al.*[91]	Assumed SE = 20% of the mean
Cost of treating an intracranial haemorrhage – ongoing	£1686	337	Gamma	£1091 to £2409	25.00	67.45	Luengo-Fernandez *et al.*[91]	Assumed SE = 20% of the mean
Cost of CTEPH surgery	£6558	–	Normal	£3987 to £9129	–	–	NICE[100]	–
Ongoing quarterly cost of CTEPH	£15,968	–	Normal	£9709 to £22,227	–	–	NICE[100]	–
Lifetime cost of breast cancer	£13,241	591	Normal	£12,108 to £14,426	–	–	Hall *et al.*[92]	–
Lifetime cost of childhood cancer	£126,273	5103	Normal	£116,271 to £136,277	–	–	Van Listenburg *et al.*[94]	–
Lifetime cost of stage 1 NSCLC	£16,408	3282	Normal	£9976 to £22,840	–	–	Incisive Health[93]	Assumed SE = 20% of the mean

Parameter	Mean	SE	Distribution used	95% CI	Alpha/*n*	Beta	Source	Notes
Lifetime cost of stage 2 NSCLC	£19,113	3823	Normal	£11,621 to £26,606	–	–	Incisive Health[93]	Assumed SE = 20% of the mean
Lifetime cost of stage 3 NSCLC	£21,454	4291	Normal	£13,044 to £29,863	–	–	Incisive Health[93]	Assumed SE = 20% of the mean
Lifetime cost of stage 4 NSCLC	£13,371	2674	Normal	£8192 to £18,612			Incisive Health[93]	Assumed SE = 20% of the mean
Proportion of women diagnosed with stage 1 NSCLC	16.9%	–	Dirichlet	–	2557	–	Cancer Research UK[101]	The stage distribution was multiplied by the number of women diagnosed with NSCLC in England in 2013
Proportion of women diagnosed with stage 2 NSCLC	8.2%	–	–	–	1245	–	Cancer Research UK[101]	
Proportion of women diagnosed with stage 3 NSCLC	21.5%	–	–	–	3264	–	Cancer Research UK[101]	–
Proportion of women diagnosed with stage 4 NSCLC	53.4%	–	–	–	8092	–	Cancer Research UK[101]	–
Cost of anticoagulation								
Proportion on enoxaparin (Clexane®, Sanofi)	51.2%	–	Fixed	–	–	–	DiPEP cohort	–
Proportion on dalteparin (Fragmin®, Pfizer)	31.4%	–	Fixed	–	–	–	DiPEP cohort	–
Proportion on tinzaparin (Innohep®, Leo Pharma)	17.4%	–	Fixed	–	–	–	DiPEP cohort	–
Dose-dependent drug cost	Varies	–	Fixed	–	–	–	BNF[102]	It was assumed that the woman would receive the dose of the drug closest to her calculated therapeutic dose

NSCLC, non-small cell lung cancer; –, no data.
a Non-elective inpatient stays and associated excess bed-days.

Appendix 20 The stability of the model results with respect to the number of probabilistic sensitivity analysis runs

Figures 60 and *61* show the stability of the average QALYs and that of the average costs, respectively, with regard to the number of PSA runs.

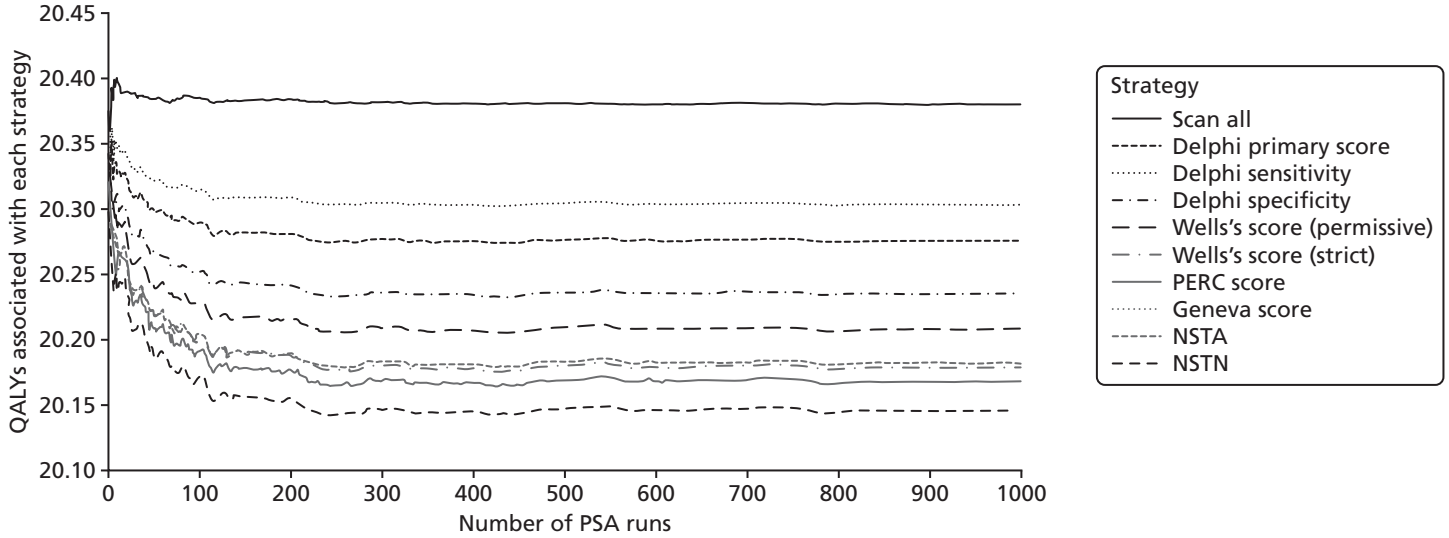


FIGURE 60 The stability of the average QALYs for each strategy with regard to the number of PSA runs. NSTA, no scanning, treat all; NSTN, no scanning, treat none.

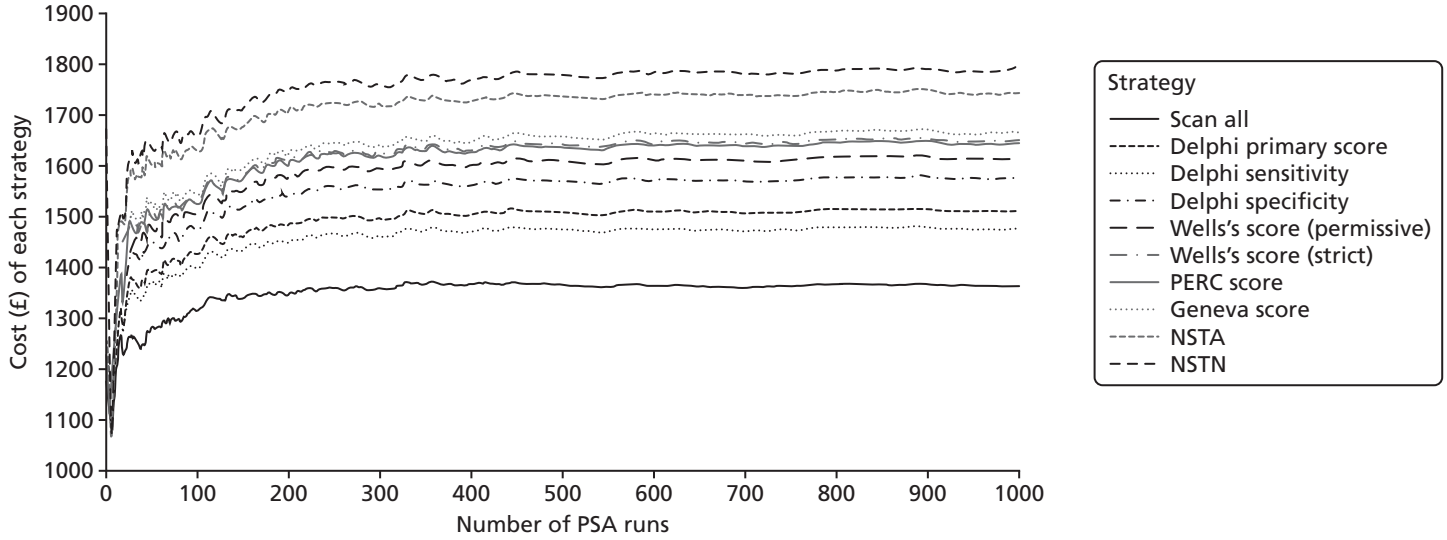


FIGURE 61 The stability of the average costs for each strategy with regard to the number of PSA runs.

Appendix 21 Detailed results of the decision-analysis modelling

Figure 62 shows the full cost-effectiveness acceptability curve associated with the base-case PSA.

Tables 83–113 show the detailed results of the scenario analysis conducted in the decision-analysis modelling.

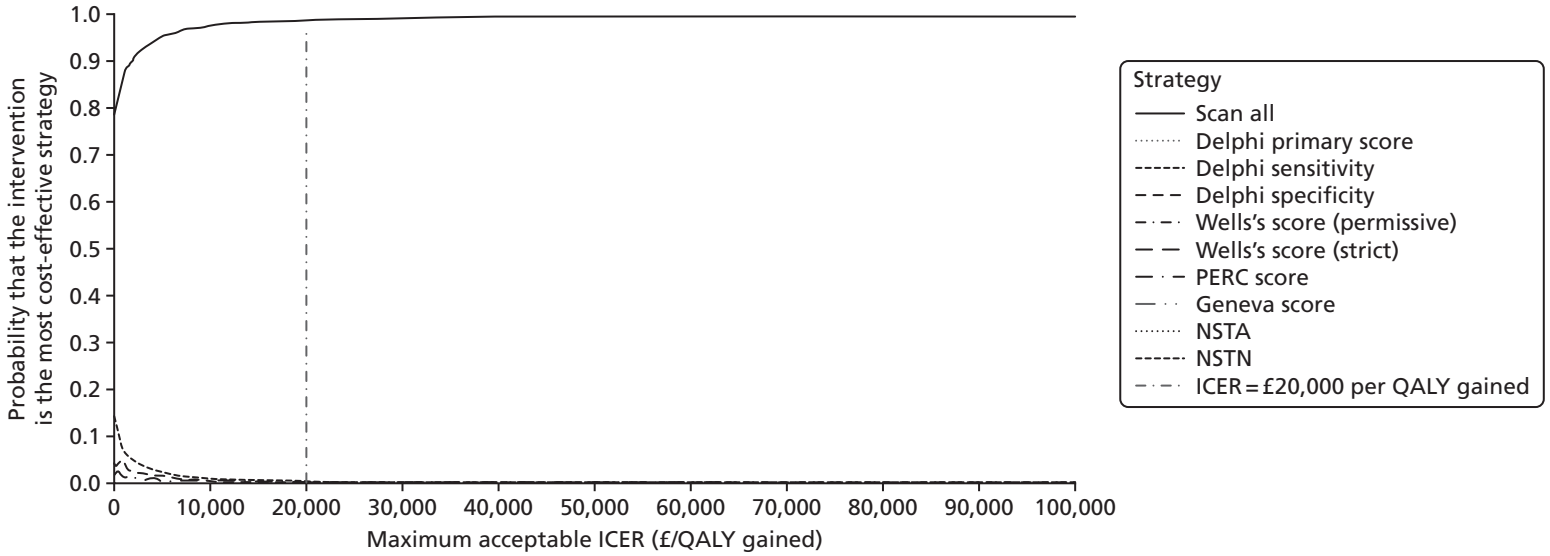


FIGURE 62 The cost-effectiveness acceptability curve for the base-case decision-analysis modelling. NSTA, no scanning, treat all; NSTN, no scanning, treat none.

TABLE 83 The results of imaging tests lead to an imperfect diagnosis of PE

	Costs (£)									Incremental		
Strategy	DR	Scans	Drugs	VTE	Bleeds	Induced cancers	CTEPH	Total	QALYs	Costs (£)	QALYs	ICER (£/QALY gained)
No scan, treat none	0	0	15	57	1	0	2751	2824	19.8537	–	–	Dominated
Delphi specificity score	4	48	92	154	7	1	2047	2352	20.0370	–	–	Dominated
Wells's score (strict)	4	27	84	155	5	1	2013	2288	20.0396	–	–	Dominated
Geneva score	4	81	124	172	9	2	1881	2273	20.0721	–	–	Dominated
Wells's score (permissive)	4	86	129	184	10	3	1791	2207	20.0953	–	–	Dominated
Delphi primary score	4	94	145	216	11	3	1547	2020	20.1606	–	–	Dominated
PERC score	4	107	157	230	12	3	1435	1949	20.1851	–	–	Dominated
No scan, treat all	0	0	1258	322	141	0	641	2363	20.2948	–	–	Dominated
Delphi sensitivity score	4	215	246	303	21	7	852	1647	20.3225	–	–	Dominated
Scan all	0	222	258	312	22	7	772	1593	20.3400	–1231	0.4863	Dominant

DR, decision rule; –, no data.

TABLE 84 The risk of CTEPH is obtained from Klok *et al.*[69]

	Costs (£)									Incremental		
Strategy	DR	Scans	Drugs	VTE	Bleeds	Induced cancers	CTEPH	Total	QALYs	Costs (£)	QALYs	ICER (£/QALY gained)
No scan, treat none	0	0	20	75	2	0	551	648	19.9253	–	–	–
Wells's score (strict)	4	27	75	167	4	1	371	650	20.1077	1	0.1824	£7
Delphi specificity score	4	49	72	167	4	1	380	677	20.1088	–	–	Extendedly dominated
Geneva score	4	83	86	184	5	3	342	705	20.1415	–	–	Dominated
Wells's score (permissive)	4	86	90	196	5	3	319	703	20.1654	–	–	Extendedly dominated
Delphi primary score	4	95	105	229	6	3	263	704	20.2289	–	–	Extendedly dominated
PERC score	4	108	109	243	7	3	235	709	20.2541	60	0.1464	£408
No, scan treat all	0	0	1254	323	122	0	75	1775	20.3221	–	–	Dominated
Delphi sensitivity score	4	215	143	314	9	7	96	787	20.3880	–	–	Extendedly dominated
Scan all	0	223	150	323	9	7	75	788	20.4061	78	0.1520	£516

DR, decision rule; –, no incremental comparison of costs or QALY, was conducted by this option, as it was either dominated, extendedly dominated or was the least effective strategy in terms of QALYs accrued.

TABLE 85 Chronic thromboembolic pulmonary hypertension is not modifiable by anticoagulation

	Costs (£)									Incremental		
Strategy	DR	Scans	Drugs	VTE	Bleeds	Induced cancers	CTEPH	Total	QALYs	Costs (£)	QALYs	ICER (£/QALY gained)
No scan, treat none	0	0	21	76	2	0	425	524	19.9315	–	–	–
Wells's score (strict)	4	27	74	166	4	1	484	760	20.1008	236	0.1693	£1392
Delphi specificity score	4	48	72	167	4	1	500	797	20.1058	–	–	Extendedly dominated
Geneva score	4	83	85	182	5	3	498	859	20.1339	–	–	Extendedly dominated
Wells's score (permissive)	4	88	90	195	5	3	508	892	20.1578	–	–	Extendedly dominated
Delphi primary score	4	94	104	228	6	3	538	976	20.2179	–	–	Extendedly dominated
PERC score	4	109	109	243	7	3	546	1020	20.2436	260	0.1428	£1822
No scan, treat all	0	0	1261	323	122	0	583	2289	20.3053	–	–	Dominated
Delphi sensitivity score	4	215	143	313	9	7	580	1270	20.3711	–	–	Extendedly dominated
Scan all	0	223	150	323	9	7	583	1295	20.3895	275	0.1459	£1888

DR, decision rule; –, no incremental comparison of costs or QALYs was conducted for this option, as it was either dominated, extendedly dominated or was the least effective strategy in terms of QALYs accrued.

TABLE 86 One hundred per cent of patients with CTEPH were surgically treated

	Costs (£)									Incremental		
Strategy	DR	Scans	Drugs	VTE	Bleeds	Induced cancers	CTEPH	Total	QALYs	Costs (£)	QALYs	ICER (£/QALY gained)
No scan, treat none	0	0	15	56	1	0	2756	2829	19.8978	–	–	Dominated
Delphi specificity score	4	49	70	155	4	1	2002	2284	20.0810	–	–	Dominated
Wells's score (strict)	4	27	73	156	4	1	1969	2234	20.0816	–	–	Dominated
Geneva score	4	82	84	173	5	3	1831	2181	20.1149	–	–	Dominated
Wells's score (permissive)	4	87	88	186	5	3	1740	2112	20.1377	–	–	Dominated
Delphi primary score	4	95	103	221	6	3	1475	1908	20.2027	–	–	Dominated
PERC score	4	109	108	236	6	3	1358	1825	20.2272	–	–	Dominated
No scan, treat all	0	0	1264	322	122	0	667	2375	20.2961	–	–	Dominated
Delphi sensitivity score	4	214	143	312	9	7	757	1445	20.3620	–	–	Dominated
Scan all	0	222	150	322	9	7	667	1378	20.3806	–1451	0.4828	Dominant

DR, decision rule; –, no incremental comparison of costs or QALYs was conducted for this option, as it was either dominated, extendedly dominated or was the least effective strategy in terms of QALYs accrued.

TABLE 87 A Weibull curve is used to estimate the life expectancy of women with a surgically treated CTEPH

	Costs (£)									Incremental		
Strategy	DR	Scans	Drugs	VTE	Bleeds	Induced cancers	CTEPH	Total	QALYs	Costs (£)	QALYs	ICER (£/QALY gained)
No scan, treat none	0	0	15	57	1	0	2752	2825	19.8553	–	–	Dominated
Delphi specificity score	4	49	70	155	4	2	1988	2270	20.0556	–	–	Dominated
Wells's score (strict)	4	27	72	156	4	1	1955	2219	20.0577	–	–	Dominated
Geneva score	4	82	83	173	5	3	1822	2170	20.0926	–	–	Dominated
Wells's score (permissive)	4	87	89	188	5	3	1708	2083	20.1225	–	–	Dominated
Delphi primary score	4	95	103	222	6	3	1450	1883	20.1912	–	–	Dominated
PERC score	4	108	108	236	6	3	1342	1807	20.2156	–	–	Dominated
No scan, treat all	0	0	1262	322	122	0	638	2345	20.3008	–	–	Dominated
Delphi sensitivity score	4	214	144	312	9	7	725	1414	20.3654	–	–	Dominated
Scan all	0	222	150	322	9	7	638	1349	20.3852	–1476	0.5299	Dominant

DR, decision rule; –, no incremental comparison of costs or QALYs was conducted for this option, as it was either dominated, extendedly dominated or was the least effective strategy in terms of QALYs accrued.

TABLE 88 A Gompertz curve is used to estimate the life expectancy of women with a surgically treated CTEPH

	Costs (£)									Incremental		
Strategy	DR	Scans	Drugs	VTE	Bleeds	Induced cancers	CTEPH	Total	QALYs	Costs (£)	QALYs	ICER (£/QALY gained)
No scan, treat none	0	0	15	57	1	0	2749	2822	19.8706	–	–	Dominated by scan all
Delphi specificity	4	49	71	157	4	2	1970	2258	20.0717	–	–	Dominated by scan all
Wells's score (strict)	4	27	74	158	4	1	1941	2209	20.0718	–	–	Dominated by scan all
Geneva score	4	82	84	174	5	3	1812	2163	20.1051	–	–	Dominated by scan all
Wells's score (permissive)	4	87	89	188	5	3	1709	2085	20.1320	–	–	Dominated by scan all
Delphi primary score	4	94	104	222	6	3	1454	1887	20.1985	–	–	Dominated by scan all
PERC score	4	107	108	237	6	3	1338	1805	20.2246	–	–	Dominated by scan all
No scan, treat all	0	0	1262	322	122	0	647	2353	20.3031	–	–	Dominated by scan all
Delphi sensitivity score	4	215	143	312	9	7	737	1426	20.3678	–	–	Dominated by scan all
Scan all	0	223	150	322	9	7	647	1359	20.3878	–1463	0.5172	Dominant

DR, decision rule; –, no incremental comparison of costs or QALYs was conducted for this option, as it was either dominated, extendedly dominated or was the least effective strategy in terms of QALYs accrued.

TABLE 89 A log-logistic curve is used to estimate the life expectancy of women with surgically treated CTEPH

	Costs (£)									Incremental		
Strategy	DR	Scans	Drugs	VTE	Bleeds	Induced cancers	CTEPH	Total	QALYs	Costs (£)	QALYs	ICER (£/QALY gained)
No scan, treat none	0	0	15	56	1	0	2755	2828	19.7306	–	–	Dominated by scan all
Delphi specificity score	4	49	70	155	4	2	1968	2251	19.9622	–	–	Dominated by scan all
Wells's score (strict)	4	27	74	157	4	1	1929	2196	19.9671	–	–	Dominated by scan all
Geneva score	4	81	84	174	5	3	1794	2143	20.0074	–	–	Dominated by scan all
Wells's score (permissive)	4	87	89	187	5	3	1691	2065	20.0376	–	–	Dominated by scan all
Delphi primary score	4	94	104	222	6	3	1418	1851	20.1186	–	–	Dominated by scan all
PERC score	4	108	109	238	7	3	1295	1763	20.1503	–	–	Dominated by scan all
No scan, treat all	0	0	1257	322	122	0	592	2293	20.2606	–	–	Dominated by scan all
Delphi sensitivity score	4	214	144	313	9	7	679	1369	20.3218	–	–	Dominated by scan all
Scan all	0	222	151	322	9	7	592	1303	20.3446	–1525	0.6140	Dominant

DR, decision rule; –, no incremental comparison of costs or QALYs was conducted for this option, as it was either dominated, extendedly dominated or was the least effective strategy in terms of QALYs accrued.

TABLE 90 A gamma curve is used to estimate the life expectancy of women with surgically treated CTEPH

	Costs (£)									Incremental		
Strategy	DR	Scans	Drugs	VTE	Bleeds	Induced cancers	CTEPH	Total	QALYs	Costs (£)	QALYs	ICER (£/QALY gained)
No scan, treat none	0	0	15	56	1	0	2753	2826	19.8469	–	–	Dominated by scan all
Delphi specificity score	4	48	70	155	4	2	1992	2274	20.0489	–	–	Dominated by scan all
Wells's score (strict)	4	27	73	156	4	1	1958	2223	20.0502	–	–	Dominated by scan all
Geneva score	4	81	84	174	5	3	1814	2165	20.0877	–	–	Dominated by scan all
Wells's score (permissive)	4	86	89	188	5	4	1712	2088	20.1150	–	–	Dominated by scan all
Delphi primary score	4	93	104	221	6	4	1460	1892	20.1826	–	–	Dominated by scan all
PERC score	4	109	109	238	7	5	1333	1803	20.2119	–	–	Dominated by scan all
No scan, treat all	0	0	1259	322	122	0	643	2347	20.2952	–	–	Dominated by scan all
Delphi sensitivity score	4	214	144	312	9	9	731	1423	20.3585	–	–	Dominated by scan all
Scan all	0	222	151	322	9	10	643	1358	20.3783	–£1468	0.5314	Dominant

DR, decision rule; –, no incremental comparison of costs or QALYs was conducted for this option, as it was either dominated, extendedly dominated or was the least effective strategy in terms of QALYs accrued.

TABLE 91 A generalised gamma curve is used to estimate the life expectancy of women with surgically treated CTEPH

	Costs (£)									Incremental		
Strategy	DR	Scans	Drugs	VTE	Bleeds	Induced cancers	CTEPH	Total	QALYs	Costs (£)	QALYs	ICER (£/QALY gained)
No scan, treat none	0	0	15	57	1	0	2750	2823	19.8659	–	–	Dominated by scan all
Wells's score (strict)	4	27	73	156	4	1	1959	2224	20.0644	–	–	Dominated by scan all
Delphi specificity score	4	49	71	157	4	2	1972	2258	20.0680	–	–	Dominated by scan all
Geneva score	4	82	84	173	5	3	1820	2170	20.1001	–	–	Dominated by scan all
Wells's score (permissive)	4	87	89	186	5	3	1722	2095	20.1262	–	–	Dominated by scan all
Delphi primary score	4	95	104	222	6	3	1453	1887	20.1967	–	–	Dominated by scan all
PERC score	4	109	108	237	6	3	1337	1805	20.2226	–	–	Dominated by scan all
No scan, treat all	0	0	1258	322	122	0	643	2346	20.3034	–	–	Dominated by scan all
Delphi sensitivity score	4	214	144	312	9	7	731	1420	20.3682	–	–	Dominated by scan all
Scan all	0	223	151	322	9	7	643	1355	20.3879	–1468	0.5220	Dominant

DR, decision rule; –, no incremental comparison of costs or QALYs was conducted for this option, as it was either dominated, extendedly dominated or was the least effective strategy in terms of QALYs accrued.

TABLE 92 A Weibull curve is used to estimate the life expectancy of women with medically treated CTEPH

	Costs (£)									Incremental		
Strategy	DR	Scans	Drugs	VTE	Bleeds	Induced cancers	CTEPH	Total	QALYs	Costs (£)	QALYs	ICER (£/QALY gained)
No scan, treat none	0	0	15	56	1	0	2750	2823	19.8551	–	–	Dominated
Wells's score (strict)	4	28	72	156	4	1	1963	2227	20.0530	–	–	Dominated
Delphi specificity score	4	50	69	156	4	2	1981	2266	20.0557	–	–	Dominated
Geneva score	4	84	82	173	5	3	1823	2173	20.0891	–	–	Dominated
Wells's score (permissive)	4	88	87	186	5	3	1729	2101	20.1143	–	–	Dominated
Delphi primary score	4	95	102	221	6	3	1464	1895	20.1841	–	–	Dominated
PERC score	4	109	108	238	7	3	1331	1800	20.2140	–	–	Dominated
No scan, treat all	0	0	1262	322	122	0	644	2351	20.2940	–	–	Dominated
Delphi sensitivity score	4	215	142	312	9	7	736	1424	20.3579	–	–	Dominated
Scan all	0	223	149	322	9	7	644	1355	20.3784	–1468	0.5233	Dominant

DR, decision rule; –, no incremental comparison of costs or QALYs was conducted for this option, as it was either dominated, extendedly dominated or was the least effective strategy in terms of QALYs accrued.

TABLE 93 A Gompertz curve is used to estimate the life expectancy of women with medically treated CTEPH

	Costs (£)									Incremental		
Strategy	DR	Scans	Drugs	VTE	Bleeds	Induced cancers	CTEPH	Total	QALYs	Costs (£)	QALYs	ICER (£/QALY gained)
No scan, treat none	0	0	15	57	1	0	2748	2822	19.8576	–	–	Dominated
Delphi specificity score	4	48	68	153	4	1	2005	2283	20.0518	–	–	Dominated
Wells's score (strict)	4	27	71	154	4	1	1971	2232	20.0533	–	–	Dominated
Geneva score	4	80	82	172	5	3	1832	2177	20.0892	–	–	Dominated
Wells's score (permissive)	4	86	87	185	5	3	1734	2102	20.1150	–	–	Dominated
Delphi primary score	4	93	101	219	6	3	1482	1907	20.1816	–	–	Dominated
PERC score	4	106	106	235	6	3	1355	1816	20.2101	–	–	Dominated
No scan, treat all	0	0	1255	322	122	0	643	2343	20.2961	–	–	Dominated
Delphi sensitivity score	4	214	143	312	9	7	733	1421	20.3599	–	–	Dominated
Scan all	0	222	150	322	9	7	643	1353	20.3802	–1468	0.5226	Dominant

DR, decision rule; –, no incremental comparison of costs or QALYs was conducted for this option, as it was either dominated, extendedly dominated or was the least effective strategy in terms of QALYs accrued.

TABLE 94 A log-logistic curve is used to estimate the life expectancy of women with medically treated CTEPH

	Costs (£)									Incremental		
Strategy	DR	Scans	Drugs	VTE	Bleeds	Induced cancers	CTEPH	Total	QALYs	Costs (£)	QALYs	ICER (£/QALY gained)
No scan, treat none	0	0	15	56	1	0	2752	2825	19.8582	–	–	Dominated
Delphi specificity score	4	48	70	155	4	1	1990	2272	20.0552	–	–	Dominated
Wells's score (strict)	4	27	74	158	4	1	1947	2214	20.0588	–	–	Dominated
Geneva score	4	81	84	174	5	3	1815	2165	20.0926	–	–	Dominated
Wells's score (permissive)	4	87	89	189	5	3	1704	2080	20.1215	–	–	Dominated
Delphi primary score	4	93	104	223	6	3	1453	1885	20.1876	–	–	Dominated
PERC score	4	107	108	237	6	3	1342	1808	20.2119	–	–	Dominated
No scan, treat all	0	0	1253	322	122	0	648	2346	20.2927	–	–	Dominated
Delphi sensitivity score	4	214	143	313	9	7	733	1422	20.3576	–	–	Dominated
Scan all	0	222	150	322	9	7	648	1358	20.3766	–1467	0.5184	Dominant

DR, decision rule; –, no incremental comparison of costs or QALYs was conducted for this option, as it was either dominated, extendedly dominated or was the least effective strategy in terms of QALYs accrued.

TABLE 95 A gamma curve is used to estimate the life expectancy of women with medically treated CTEPH

	Costs (£)									Incremental		
Strategy	DR	Scans	Drugs	VTE	Bleeds	Induced cancers	CTEPH	Total	QALYs	Costs (£)	QALYs	ICER (£/QALY gained)
No scan, treat none	0	0	15	56	1	0	2755	2828	19.9090	–	–	Dominated
Wells's score (strict)	4	28	72	155	4	1	1976	2239	20.0968	–	–	Dominated
Delphi specificity score	4	49	70	155	4	1	1995	2279	20.0992	–	–	Dominated
Geneva score	4	82	83	173	5	3	1836	2185	20.1318	–	–	Dominated
Wells's score (permissive)	4	87	89	188	5	3	1726	2101	20.1597	–	–	Dominated
Delphi primary score	4	95	105	224	6	3	1451	1888	20.2284	–	–	Dominated
PERC score	4	109	109	237	6	3	1352	1819	20.2496	–	–	Dominated
No scan, treat all	0	0	1264	322	122	0	665	2374	20.3211	–	–	Dominated
Delphi sensitivity score	4	215	144	312	9	7	751	1442	20.3874	–	–	Dominated
Scan all	0	223	151	322	9	7	665	1378	20.4058	–1451	0.4968	Dominant

DR, decision rule; –, no incremental comparison of costs or QALYs was conducted for this option, as it was either dominated, extendedly dominated or was the least effective strategy in terms of QALYs accrued.

TABLE 96 A generalised gamma curve is used to estimate the life expectancy of women with medically treated CTEPH

	Costs (£)									Incremental		
Strategy	DR	Scans	Drugs	VTE	Bleeds	Induced cancers	CTEPH	Total	QALYs	Costs (£)	QALYs	ICER (£/QALY gained)
No scan, treat none	0	0	15	56	1	0	2755	2828	19.9090	–	–	Dominated by scan all
Wells's score (strict)	4	28	72	155	4	1	1976	2239	20.0968	–	–	Dominated by scan all
Delphi specificity	4	49	70	155	4	1	1995	2279	20.0992	–	–	Dominated by scan all
Geneva score	4	82	83	173	5	3	1836	2185	20.1318	–	–	Dominated by scan all
Wells's score (permissive)	4	87	89	188	5	3	1726	2101	20.1597	–	–	Dominated by scan all
Delphi primary score	4	95	105	224	6	3	1451	1888	20.2284	–	–	Dominated by scan all
PERC score	4	109	109	237	6	3	1352	1819	20.2496	–	–	Dominated by scan all
No scan, treat all	0	0	1264	322	122	0	665	2374	20.3211	–	–	Dominated by scan all
Delphi sensitivity	4	215	144	312	9	7	751	1442	20.3874	–	–	Dominated by scan all
Scan all	0	223	151	322	9	7	665	1378	20.4058	–1451	0.4968	Dominant

DR, decision rule; –, no incremental comparison of costs or QALYs was conducted for this option, as it was either dominated, extendedly dominated or was the least effective strategy in terms of QALYs accrued.

TABLE 97 There is no risk of death following a recurrent VTE

	Costs (£)									Incremental		
Strategy	DR	Scans	Drugs	VTE	Bleeds	Induced cancers	CTEPH	Total	QALYs	Costs (£)	QALYs	ICER (£/QALY gained)
No scan, treat none	0	0	15	57	1	0	2756	2829	19.8549	–	–	Dominated
Wells's score (strict)	4	26	73	155	4	1	1970	2233	20.0531	–	–	Dominated
Delphi specificity score	4	48	71	155	4	1	1990	2274	20.0553	–	–	Dominated
Geneva score	4	82	85	175	5	3	1812	2164	20.0946	–	–	Dominated
Wells's score (permissive)	4	87	90	187	5	3	1719	2094	20.1193	–	–	Dominated
Delphi primary score	4	94	104	222	6	3	1461	1894	20.1872	–	–	Dominated
PERC score	4	107	110	237	7	3	1338	1805	20.2153	–	–	Dominated
No scan, treat all	0	0	1260	322	122	0	645	2350	20.2967	–	–	Dominated
Delphi sensitivity score	4	215	145	313	9	7	728	1420	20.3622	–	–	Dominated
Scan all	0	223	152	322	9	7	645	1358	20.3809	–1471	0.5260	Dominant

DR, decision rule; –, no incremental comparison of costs or QALYs was conducted for this option, as it was either dominated, extendedly dominated or was the least effective strategy in terms of QALYs accrued.

TABLE 98 There is no anticoagulation treatment cost following a recurrent VTE

	Costs (£)									Incremental		
Strategy	DR	Scans	Drugs	VTE	Bleeds	Induced cancers	CTEPH	Total	QALYs	Costs (£)	QALYs	ICER (£/QALY gained)
No scan, treat none	0	0	15	56	1	0	2756	2828	19.8481	–	–	Dominated
Wells's score (strict)	4	27	41	156	4	1	1961	2193	20.0478	–	–	Dominated
Delphi specificity score	4	48	40	157	4	1	1971	2226	20.0525	–	–	Dominated
Geneva score	4	82	46	175	5	3	1809	2123	20.0876	–	–	Dominated
Wells's score (permissive)	4	87	49	189	5	3	1702	2038	20.1162	–	–	Dominated
Delphi primary score	4	93	55	224	6	3	1439	1824	20.1854	–	–	Dominated
PERC score	4	108	57	238	7	3	1332	1750	20.2092	–	–	Dominated
No scan, treat all	0	0	1181	322	122	0	645	2271	20.2902	–	–	Dominated
Delphi sensitivity	4	214	73	313	9	7	731	1350	20.3550	–	–	Dominated
Scan all	0	222	76	322	9	7	645	1282	20.3742	–1547	0.5261	Dominant

DR, decision rule; –, no incremental comparison of costs or QALYs was conducted for this option, as it was either dominated, extendedly dominated or was the least effective strategy in terms of QALYs accrued.

TABLE 99 The expert elicitation exercise was conducted for women with PE as defined in the secondary statistical population

	Costs (£)									Incremental		
Strategy	DR	Scans	Drugs	VTE	Bleeds	Induced cancers	CTEPH	Total	QALYs	Costs (£)	QALYs	ICER (£/QALY gained)
No scan, treat none	0	0	16	59	1	0	2657	2734	19.8820	–	–	No scan, treat none
Wells's score (strict)	4	27	73	158	4	1	1883	2150	20.0772	–	–	Wells's score (strict)
Delphi specificity score	4	49	71	159	4	2	1898	2187	20.0806	–	–	Delphi specificity score
Geneva score	4	82	84	176	5	3	1738	2092	20.1144	–	–	Geneva score
Wells's score (permissive)	4	87	89	189	5	3	1649	2025	20.1382	–	–	Wells's score (permissive)
Delphi primary score	4	95	103	224	6	3	1393	1828	20.2051	–	–	Delphi primary score
PERC score	4	108	108	239	7	3	1276	1745	20.2314	–	–	PERC score
No scan, treat all	0	0	1259	322	122	0	608	2312	20.3056	–	–	No scan, treat all
Delphi sensitivity	4	214	143	312	9	7	695	1384	20.3704	–	–	Delphi sensitivity
Scan all	0	222	150	322	9	7	608	1319	20.3898	–1415	0.5078	Scan all

DR, decision rule; –, no incremental comparison of costs or QALYs was conducted for this option, as it was either dominated, extendedly dominated or was the least effective strategy in terms of QALYs accrued.

TABLE 100 The expert elicitation exercise on the risk of mortality from PE was conducted for women with PE as defined in the tertiary statistical population

	Costs (£)									Incremental		
Strategy	DR	Scans	Drugs	VTE	Bleeds	Induced cancers	CTEPH	Total	QALYs	Costs (£)	QALYs	ICER (£/QALY gained)
No scan, treat none	0	0	15	57	1	0	2750	2823	19.8673	–	–	Dominated by scan all
Wells's score (strict)	4	28	73	156	4	1	1953	2219	20.0675	–	–	Dominated by scan all
Delphi specificity score	4	49	71	156	4	2	1976	2262	20.0691	–	–	Dominated by scan all
Geneva score	4	83	85	174	5	3	1811	2163	20.1043	–	–	Dominated by scan all
Wells's score (permissive)	4	87	90	188	5	3	1702	2079	20.1329	–	–	Dominated by scan all
Delphi primary score	4	95	105	223	6	3	1445	1879	20.2012	–	–	Dominated by scan all
PERC score	4	109	109	238	7	3	1328	1797	20.2276	–	–	Dominated by scan all
No scan, treat all	0	0	1254	322	122	0	641	2340	20.3083	–	–	Dominated by scan all
Delphi sensitivity	4	214	144	312	9	7	727	1417	20.3727	–	–	Dominated by scan all
Scan all	0	222	151	322	9	7	641	1352	20.3922	–1471	0.5249	Dominant

DR, decision rule; –, no incremental comparison of costs or QALYs was conducted for this option, as it was either dominated, extendedly dominated or was the least effective strategy in terms of QALYs accrued.

TABLE 101 The expert elicitation exercise on the risk of mortality from PE was conducted for women with PE as defined in the quaternary statistical population

	Costs (£)									Incremental		
Strategy	DR	Scans	Drugs	VTE	Bleeds	Induced cancers	CTEPH	Total	QALYs	Costs (£)	QALYs	ICER (£/QALY gained)
No scan, treat none	0	0	15	56	1	0	2776	2848	19.8447	–	–	Dominated
Delphi specificity score	4	49	70	155	4	2	2004	2287	20.0474	–	–	Dominated
Wells's score (strict)	4	28	73	157	4	1	1968	2235	20.0491	–	–	Dominated
Geneva score	4	81	84	174	5	3	1831	2181	20.0852	–	–	Dominated
Wells's score (permissive)	4	86	89	188	5	3	1727	2101	20.1130	–	–	Dominated
Delphi primary score	4	93	103	221	6	3	1476	1906	20.1798	–	–	Dominated
PERC score	4	107	108	236	6	3	1360	1824	20.2063	–	–	Dominated
No scan, treat all	0	0	1264	322	122	0	656	2364	20.2910	–	–	Dominated
Delphi sensitivity score	4	214	144	313	9	7	742	1431	20.3563	–	–	Dominated
Scan all	0	222	150	322	9	7	656	1367	20.3757	–1481	0.5310	Dominant

DR, decision rule; –, no incremental comparison of costs or QALYs was conducted for this option, as it was either dominated, extendedly dominated or was the least effective strategy in terms of QALYs accrued.

TABLE 102 The expert elicitation exercise on the risk of mortality from PE was conducted for expert one's answers with PE as defined in the primary statistical population

	Costs (£)									Incremental		
Strategy	DR	Scans	Drugs	VTE	Bleeds	Induced cancers	CTEPH	Total	QALYs	Costs (£)	QALYs	ICER (£/QALY gained)
No scan, treat none	0	0	10	36	1	0	3031	3077	19.6578	–	–	Dominated
Wells's score (strict)	4	27	69	143	4	1	2310	2557	19.9018	–	–	Dominated
Delphi specificity score	4	49	67	143	4	2	2383	2651	19.9021	–	–	Dominated
Geneva score	4	82	81	161	4	3	2220	2555	19.9471	–	–	Dominated
Wells's score (permissive)	4	87	86	176	5	3	2115	2475	19.9799	–	–	Dominated
Delphi primary score	4	94	101	213	6	3	1923	2344	20.0618	–	–	Dominated
PERC score	4	108	106	230	6	3	1790	2247	20.0969	–	–	Dominated
No scan, treat all	0	0	1256	321	122	0	1109	2808	20.2194	–	–	Dominated
Delphi sensitivity score	4	215	143	311	8	7	1192	1880	20.2796	–	–	Dominated
Scan all	0	223	150	321	9	7	1109	1819	20.3035	–1258	0.6457	Dominant

DR, decision rule; –, no incremental comparison of costs or QALYs was conducted for this option, as it was either dominated, extendedly dominated or was the least effective strategy in terms of QALYs accrued.

TABLE 103 The expert elicitation exercise on the risk of mortality from PE was conducted for expert two's answers with PE as defined the primary statistical population

	Costs (£)									Incremental		
Strategy	DR	Scans	Drugs	VTE	Bleeds	Induced cancers	CTEPH	Total	QALYs	Costs (£)	QALYs	ICER (£/QALY gained)
No scan, treat none	0	0	18	65	1	0	2494	2578	19.9162	–	–	Dominated
Wells's score (strict)	4	26	74	162	4	1	1751	2022	20.0964	–	–	Dominated
Delphi specificity score	4	48	71	162	4	2	1763	2053	20.1001	–	–	Dominated
Geneva score	4	80	85	179	5	3	1608	1964	20.1311	–	–	Dominated
Wells's score (permissive)	4	85	90	191	5	3	1524	1901	20.1528	–	–	Dominated
Delphi primary score	4	93	104	225	6	3	1272	1707	20.2168	–	–	Dominated
PERC score	4	107	110	242	7	3	1147	1619	20.2439	–	–	Dominated
No scan, treat all	0	0	1260	323	122	0	522	2227	20.3031	–	–	Dominated
Delphi sensitivity score	4	215	144	313	9	7	601	1291	20.3704	–	–	Dominated
Scan all	0	223	151	323	9	7	522	1233	20.3875	–1345	0.4712	Dominant

DR, decision rule; –, no incremental comparison of costs or QALYs was conducted for this option, as it was either dominated, extendedly dominated or was the least effective strategy in terms of QALYs accrued.

TABLE 104 The expert elicitation exercise on the risk of mortality from PE was conducted for expert 3's answers with PE as defined in the primary statistical population

	Costs (£)									Incremental		
Strategy	DR	Scans	Drugs	VTE	Bleeds	Induced cancers	CTEPH	Total	QALYs	Costs (£)	QALYs	ICER (£/QALY gained)
No scan, treat none	0	0	15	55	1	0	2807	2878	19.8382	–	–	Dominated
Wells's score (strict)	4	27	72	155	4	1	2005	2268	20.0422	–	–	Dominated
Delphi specificity score	4	49	70	155	4	2	2031	2314	20.0448	–	–	Dominated
Geneva score	4	81	83	173	5	3	1860	2208	20.0795	–	–	Dominated
Wells's score (permissive)	4	86	88	186	5	3	1761	2132	20.1065	–	–	Dominated
Delphi primary score	4	94	103	223	6	3	1487	1919	20.1786	–	–	Dominated
PERC score	4	108	108	237	6	3	1373	1839	20.2052	–	–	Dominated
No scan, treat all	0	0	1256	322	122	0	667	2368	20.2900	–	–	Dominated
Delphi sensitivity score	4	215	143	312	9	7	758	1447	20.3536	–	–	Dominated
Scan all	0	223	150	322	9	7	667	1378	20.3737	–1500	0.5356	Dominant

DR, decision rule; –, no incremental comparison of costs or QALYs was conducted for this option, as it was either dominated, extendedly dominated or was the least effective strategy in terms of QALYs accrued.

TABLE 105 The expert elicitation exercise on the risk of mortality from PE was conducted for expert 4's answers with PE as defined in the primary statistical population

	Costs (£)									Incremental		
Strategy	DR	Scans	Drugs	VTE	Bleeds	Induced cancers	CTEPH	Total	QALYs	Costs (£)	QALYs	ICER (£/QALY gained)
No scan, treat none	0	0	12	43	1	0	2795	2851	19.7291	–	–	Dominated
Delphi specificity score	4	49	70	149	4	2	2180	2456	19.9563	–	–	Dominated
Wells's score (strict)	4	28	72	148	4	1	2127	2383	19.9572	–	–	Dominated
Geneva score	4	83	84	168	4	3	2014	2359	19.9992	–	–	Dominated
Wells's score (permissive)	4	86	89	181	5	3	1941	2307	20.0291	–	–	Dominated
Delphi primary score	4	95	104	218	6	3	1740	2170	20.1036	–	–	Dominated
PERC score	4	109	109	233	6	3	1644	2108	20.1335	–	–	Dominated
No scan, treat all	0	0	1257	321	122	0	1023	2723	20.2418	–	–	Dominated
Delphi sensitivity	4	215	144	310	8	7	1107	1795	20.3013	–	–	Dominated
Scan all	0	223	152	321	9	7	1023	1734	20.3256	–1117	0.5965	Dominant

DR, decision rule; –, no incremental comparison of costs or QALYs was conducted for this option, as it was either dominated, extendedly dominated or was the least effective strategy in terms of QALYs accrued.

TABLE 106 The risk of PE-related mortality is taken from UKOSS

	Costs (£)									Incremental		
Strategy	DR	Scans	Drugs	VTE	Bleeds	Induced cancers	CTEPH	Total	QALYs	Costs (£)	QALYs	ICER (£/QALY gained)
No scan, treat none	0	0	19	68	2	0	2446	2535	19.9641	–	–	Dominated
Delphi specificity score	4	49	72	163	4	2	1693	1987	20.1326	–	–	Dominated
Wells's score (strict)	4	27	75	165	4	1	1676	1951	20.1368	–	–	Dominated
Geneva score	4	82	86	181	5	3	1547	1907	20.1656	–	–	Dominated
Wells's score (permissive)	4	88	91	194	5	3	1450	1833	20.1879	–	–	Dominated
Delphi primary score	4	95	105	227	6	3	1195	1635	20.2449	–	–	Dominated
PERC score	4	108	110	241	7	3	1087	1559	20.2678	–	–	Dominated
No scan, treat all	0	0	1254	323	123	0	439	2139	20.3293	–	–	Dominated
Delphi sensitivity score	4	214	144	314	9	7	518	1209	20.3967	–	–	Dominated
Scan all	0	223	151	323	9	7	439	1151	20.4134	–1383	0.4492	Dominant

DR, decision rule; –, no incremental comparison of costs or QALYs was conducted for this option, as it was either dominated, extendedly dominated or was the least effective strategy in terms of QALYs accrued.

TABLE 107 There is a 12.5 kg reduction in the weight of pregnant women who are more than 20 weeks pregnant for the purpose of calculating their anticoagulation dose

	Costs (£)									Incremental		
Strategy	DR	Scans	Drugs	VTE	Bleeds	Induced cancers	CTEPH	Total	QALYs	Costs (£)	QALYs	ICER (£/QALY gained)
No scan, treat none	0	0	13	56	1	0	2751	2822	19.8482	–	–	Dominated
Wells's score (strict)	4	27	66	156	4	1	1960	2217	20.0481	–	–	Dominated
Delphi specificity score	4	49	64	156	4	2	1976	2255	20.0511	–	–	Dominated
Geneva score	4	82	76	174	5	3	1816	2158	20.0857	–	–	Dominated
Wells's score (permissive)	4	86	79	186	5	3	1722	2085	20.1109	–	–	Dominated
Delphi primary score	4	94	92	221	6	3	1462	1882	20.1791	–	–	Dominated
PERC score	4	108	97	237	6	3	1337	1792	20.2078	–	–	Dominated
No scan, treat all	0	0	1123	322	122	0	645	2212	20.2898	–	–	Dominated
Delphi sensitivity score	4	214	126	313	9	7	731	1403	20.3547	–	–	Dominated
Scan all	0	222	133	322	9	7	645	1338	20.3740	–1484	0.5258	Dominant

DR, decision rule; –, no incremental comparison of costs or QALYs was conducted for this option, as it was either dominated, extendedly dominated or was the least effective strategy in terms of QALYs accrued.

TABLE 108 The cost of CTEPH surgery is £24,000

	Costs (£)									Incremental		
Strategy	DR	Scans	Drugs	VTE	Bleeds	Induced cancers	CTEPH	Total	QALYs	Costs (£)	QALYs	ICER (£/QALY gained)
No scan, treat none	0	0	15	57	1	0	2751	2824	19.8488	–	–	Dominated
Wells's score (strict)	4	27	72	155	4	1	1968	2230	20.0455	–	–	Dominated
Delphi specificity score	4	49	69	155	4	1	1990	2272	20.0475	–	–	Dominated
Geneva score	4	82	83	173	5	3	1823	2172	20.0837	–	–	Dominated
Wells's score (permissive)	4	87	88	186	5	3	1724	2097	20.1097	–	–	Dominated
Delphi primary score	4	95	103	221	6	3	1462	1893	20.1792	–	–	Dominated
PERC score	4	109	109	237	6	3	1338	1805	20.2070	–	–	Dominated
No scan, treat all	0	0	1262	322	122	0	644	2351	20.2891	–	–	Dominated
Delphi sensitivity score	4	216	143	312	9	7	734	1424	20.3532	–	–	Dominated
Scan all	0	224	150	322	9	7	644	1357	20.3734	–1466	0.5246	Dominant

DR, decision rule; –, no incremental comparison of costs or QALYs was conducted for this option, as it was either dominated, extendedly dominated or was the least effective strategy in terms of QALYs accrued.

TABLE 109 The cost of CTEPH management is from Schweikert *et al.*[97]

	Costs (£)									Incremental		
Strategy	DR	Scans	Drugs	VTE	Bleeds	Induced cancers	CTEPH	Total	QALYs	Costs (£)	QALYs	ICER (£/QALY gained)
No scan, treat none	0	0	15	56	1	0	2750	2823	19.8645	–	–	Dominated
Wells's score (strict)	4	28	72	156	4	1	2008	2273	20.0642	–	–	Dominated
Delphi specificity score	4	49	70	155	4	2	2036	2319	20.0643	–	–	Dominated
Geneva score	4	81	83	174	5	2	1876	2224	20.1011	–	–	Dominated
Wells's score (permissive)	4	87	88	187	5	3	1782	2156	20.1277	–	–	Dominated
Delphi primary score	4	94	102	220	6	3	1551	1980	20.1927	–	–	Dominated
PERC score	4	107	108	237	6	3	1431	1896	20.2222	–	–	Dominated
No scan, treat all	0	0	1258	322	122	0	775	2478	20.3049	–	–	Dominated
Delphi sensitivity	4	215	143	312	9	7	861	1549	20.3685	–	–	Dominated
Scan all	0	223	150	322	9	7	775	1486	20.3892	–1337	0.5247	Dominant

DR, decision rule; –, no incremental comparison of costs or QALYs was conducted for this option, as it was either dominated, extendedly dominated or was the least effective strategy in terms of QALYs accrued.

TABLE 110 The cost of CTEPH surgery is £24,000 and the cost of CTEPH management is from Schweikert *et al.*[97]

	Costs (£)									Incremental		
Strategy	DR	Scans	Drugs	VTE	Bleeds	Induced cancers	CTEPH	Total	QALYs	Costs (£)	QALYs	ICER (£/QALY gained)
No scan, treat none	0	0	15	56	1	0	2754	2827	19.8417	–	–	Dominated
Delphi specificity score	4	49	69	154	4	1	2048	2329	20.0401	–	–	Dominated
Wells's score (strict)	4	27	72	156	4	1	2014	2277	20.0417	–	–	Dominated
Geneva score	4	82	83	173	5	3	1885	2233	20.0781	–	–	Dominated
Wells's score (permissive)	4	87	88	187	5	3	1789	2162	20.1051	–	–	Dominated
Delphi primary score	4	94	102	221	6	3	1549	1979	20.1733	–	–	Dominated
PERC score	4	108	107	236	6	3	1438	1903	20.1998	–	–	Dominated
No scan, treat all	0	0	1257	322	122	0	780	2482	20.2841	–	–	Dominated
Delphi sensitivity score	4	214	142	311	9	7	870	1556	20.3465	–	–	Dominated
Scan all	0	222	150	322	9	7	780	1490	20.3681	–1337	0.5264	Dominant

DR, decision rule; –, no incremental comparison of costs or QALYs was conducted for this option, as it was either dominated, extendedly dominated or was the least effective strategy in terms of QALYs accrued.

TABLE 111 Women are not at risk from bleeding, recurrent VTE or CTEPH and the risk of death is from the expert elicitation

	Costs (£)									Incremental		
Strategy	DR	Scans	Drugs	VTE	Bleeds	Induced cancers	CTEPH	Total	QALYs	Costs (£)	QALYs	ICER (£/QALY gained)
No scan, treat none	0	0	0	0	0	0	0	0	20.0380	–	–	–
Delphi specificity score	4	49	60	113	0	2	233	461	20.1799	–	–	Dominated
Wells's score (strict)	4	28	65	118	0	1	224	440	20.1818	440	0.1437	3060
Geneva score	4	83	76	135	0	3	262	562	20.2045	–	–	Extendedly dominated
Wells's score (permissive)	4	87	82	152	0	3	295	623	20.2256	–	–	Extendedly dominated
Delphi primary score	4	95	98	191	0	3	377	768	20.2736	–	–	Extendedly dominated
PERC score	4	109	105	212	0	3	411	843	20.2952	75	0.0216	3490
Delphi sensitivity score	4	215	143	299	0	7	553	1220	20.3932	–	–	Extendedly dominated
Scan all	0	222	151	310	0	7	573	1263	20.4065	420	0.1113	3775
No scan, treat all	0	0	1257	310	0	0	573	2140	20.4091	877	0.0026	337,261

DR, decision rule; –, no incremental comparison of costs or QALYs was conducted for this option, as it was either dominated, extendedly dominated or was the least effective strategy in terms of QALYs accrued.

TABLE 112 Women are not at risk from bleeding, recurrent VTE or CTEPH and the risk of death is from the UKOSS patients

	Costs (£)									Incremental		
Strategy	DR	Scans	Drugs	VTE	Bleeds	Induced cancers	CTEPH	Total	QALYs	Costs (£)	QALYs	ICER (£/QALY gained)
No scan, treat none	0	0	0	0	0	0	0	0	20.1504	–	–	–
Delphi specificity score	4	48	62	115	0	1	143	373	20.2587	–	–	Dominated
Wells's score (strict)	4	27	66	118	0	1	146	362	20.2622	362	0.1118	3237
Geneva score	4	81	78	137	0	2	170	471	20.2799	–	–	Extendedly dominated
Wells's score (permissive)	4	86	84	152	0	2	188	516	20.2940	–	–	Extendedly dominated
Delphi primary score	4	93	100	193	0	2	239	630	20.3310	–	–	Extendedly dominated
PERC score	4	107	106	210	0	3	260	689	20.3462	327	0.0840	3897
Delphi sensitivity score	4	214	145	299	0	5	370	1038	20.4303	–	–	Extendedly dominated
Scan all	0	222	153	310	0	5	384	1075	20.4408	385	0.0946	4072
No scan, treat all	0	0	1259	310	0	0	384	1953	20.4427	879	0.0019	469,304

DR, decision rule; –, no incremental comparison of costs or QALYs was conducted for this option, as it was either dominated, extendedly dominated or was the least effective strategy in terms of QALYs accrued.

TABLE 113 All scanning-induced cancers present within 15 years

	Costs (£)									Incremental		
Strategy	DR	Scans	Drugs	VTE	Bleeds	Induced cancers	CTEPH	Total	QALYs	Costs (£)	QALYs	ICER (£/QALY gained)
No scan, treat none	0	0	15	56	1	0	2751	2824	19.8474	–	–	Dominated
Delphi specificity score	4	28	72	157	4	2	1949	2216	20.0468	–	–	Dominated
Wells's score (strict)	4	49	70	156	4	4	1976	2263	20.0469	–	–	Dominated
Geneva score	4	81	84	175	5	7	1805	2160	20.0817	–	–	Dominated
Wells's score (permissive)	4	86	88	188	5	7	1709	2087	20.1070	–	–	Dominated
Delphi primary score	4	95	103	223	6	8	1448	1886	20.1762	–	–	Dominated
PERC score	4	108	108	238	7	9	1330	1803	20.2021	–	–	Dominated
Delphi sensitivity score	0	0	1257	322	122	0	644	2346	20.2887	–	–	Dominated
Scan all	4	215	143	312	9	17	733	1432	20.3403	–	–	Dominated
No scan, treat all	0	223	150	322	9	18	644	1366	20.3600	–1458	0.5126	Dominant

DR, decision rule; –, no incremental comparison of costs or QALYs was conducted for this option, as it was either dominated, extendedly dominated or was the least effective strategy in terms of QALYs accrued.

EME
HS&DR
HTA
PGfAR
PHR

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

Published by the NIHR Journals Library