HEALTH TECHNOLOGY ASSESSMENT

VOLUME 20 ISSUE 87 NOVEMBER 2016
ISSN 1366-5278

Placental growth factor (alone or in combination with soluble fms-like tyrosine kinase 1) as an aid to the assessment of women with suspected pre-eclampsia: systematic review and economic analysis

Geoff K Frampton, Jeremy Jones, Micah Rose and Liz Payne

National Institute for Health Research

DOI 10.3310/hta20870

Placental growth factor (alone or in combination with soluble fms-like tyrosine kinase 1) as an aid to the assessment of women with suspected pre-eclampsia: systematic review and economic analysis

Geoff K Frampton,* Jeremy Jones, Micah Rose and Liz Payne

Southampton Health Technology Assessments Centre, University of Southampton, Southampton, UK

*Corresponding author

Declared competing interests of authors: none

Published November 2016
DOI: 10.3310/hta20870

This report should be referenced as follows:

Frampton GK, Jones J, Rose M, Payne L. Placental growth factor (alone or in combination with soluble fms-like tyrosine kinase 1) as an aid to the assessment of women with suspected pre-eclampsia: systematic review and economic analysis. *Health Technol Assess* 2016;**20**(87).

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

NICE TAR and DAR

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 4.058

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, The Cochrane Library and the ISI 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: nihredit@southampton.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 commissioned and funded by the HTA programme on behalf of NICE as project number 14/69/05. The protocol was agreed in March 2015. The assessment report began editorial review in September 2015 and was accepted for publication in June 2016. 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.

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

Abstract

Placental growth factor (alone or in combination with soluble fms-like tyrosine kinase 1) as an aid to the assessment of women with suspected pre-eclampsia: systematic review and economic analysis

Geoff K Frampton,* Jeremy Jones, Micah Rose and Liz Payne

Southampton Health Technology Assessments Centre, University of Southampton, Southampton, UK

*Corresponding author G.K.Frampton@soton.ac.uk

Background: Pre-eclampsia (PE) prediction based on blood pressure, presence of protein in the urine, symptoms and laboratory test abnormalities can result in false-positive diagnoses. This may lead to unnecessary antenatal admissions and preterm delivery. Blood tests that measure placental growth factor (PlGF) or the ratio of soluble fms-like tyrosine kinase 1 (sFlt-1) to PlGF could aid prediction of PE if either were added to routine clinical assessment or used as a replacement for proteinuria testing.

Objectives: To evaluate the diagnostic accuracy and cost-effectiveness of PlGF-based tests for patients referred to secondary care with suspected PE in weeks 20–37 of pregnancy.

Design: Systematic reviews and an economic analysis.

Data sources: Bibliographic databases including MEDLINE, EMBASE, Web of Science and The Cochrane Library and Database of Abstracts of Reviews of Effects were searched up to July 2015 for English-language references. Conferences, websites, systematic reviews and confidential company submissions were also accessed.

Review methods: Systematic reviews of test accuracy and economic studies were conducted to inform an economic analysis. Test accuracy studies were required to include women with suspected PE and report quantitatively the accuracy of PlGF-based tests; their risk of bias was assessed using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) criteria. The economic studies review had broad eligibility criteria to capture any types of economic analysis; critical appraisal employed standard checklists consistent with National Institute for Health and Care Excellence criteria. Study selection, critical appraisal and data extraction in both reviews were performed by two reviewers.

Economic analysis: An independent economic analysis was conducted based on a decision tree model, using the best evidence available. The model evaluates costs (2014, GBP) from a NHS and Personal Social Services perspective. Given the short analysis time horizon, no discounting was undertaken.

Results: Four studies were included in the systematic review of test accuracy: two on Alere's Triage® PlGF test (Alere, Inc., San Diego, CA, USA) for predicting PE requiring delivery within a specified time and two on Roche Diagnostics' Elecsys® sFlt-1 to PlGF ratio test (Roche Diagnostics GmbH, Mannheim, Germany) for predicting PE within a specified time. Three studies were included in the systematic review of economic studies, and two confidential company economic analyses were assessed separately. Study heterogeneity precluded meta-analyses of test accuracy or cost-analysis outcomes, so narrative syntheses were conducted to inform the independent economic model. The model predicts that, when supplementing routine clinical assessment for rule-out and rule-in of PE, the two tests would be cost-saving in weeks 20–35 of gestation,

and marginally cost-saving in weeks 35–37, but with minuscule impact on quality of life. Length of neonatal intensive care unit stay was the most influential parameter in sensitivity analyses. All other sensitivity analyses had negligible effects on results.

Limitations: No head-to-head comparisons of the tests were identified. No studies investigated accuracy of PlGF-based tests when used as a replacement for proteinuria testing. Test accuracy studies were found to be at high risk of clinical review bias.

Conclusions: The Triage and Elecsys tests would save money if added to routine clinical assessment for PE. The magnitude of savings is uncertain, but the tests remain cost-saving under worst-case assumptions. Further research is required to clarify how the test results would be interpreted and applied in clinical practice.

Study registration: This study is registered as PROSPERO CRD42015017670.

Funding: The National Institute for Health Research Health Technology Assessment programme.

Contents

List of tables

List of figures

List of boxes

List of abbreviations

ACOG	American Congress of Obstetricians and Gynecologists
ALT	alanine aminotransferase
AUC	area under the curve
BMI	body mass index
CE	Conformité Européenne
CG	clinical guideline
CI	confidence interval
CPCI-S	Conference Proceedings Citation Index – Science
CPEP	Calcium for Pre-eclampsia Prevention
CRD	Centre for Reviews and Dissemination
DARE	Database of Abstracts of Reviews of Effects
EAG	External Assessment Group
EQ-5D	European Quality of Life-5 Dimensions
GDG	Guideline Development Group
HELLP	haemolysis, elevated liver enzymes, low platelet count
HRQoL	health-related quality of life
HYPITAT	Hypertension and Pre-eclampsia Intervention Trial at Term
ISSHP	International Society for the Study of Hypertension in Pregnancy
IUGR	intrauterine growth restriction
NICE	National Institute for Health and Care Excellence
NICU	neonatal intensive care unit
NPV	negative predictive value
PE	pre-eclampsia
PlGF	placental growth factor
PPV	positive predictive value
QALY	quality-adjusted life-year
QUADAS	Quality Assessment of Diagnostic Accuracy Studies
ROC	receiver operator characteristic
SCI-EXPANDED	Science Citation Index Expanded
SF-12	Short Form questionnaire-12 items
SF-36	Short Form questionnaire-36 items
SF-6D	Short Form questionnaire-6 items
sFlt-1	soluble fms-like tyrosine kinase 1
TTO	time trade-off
VEGFR1	vascular endothelial growth factor receptor 1

Note

This monograph is based on the Technology Assessment Report produced for NICE. The full report contained a considerable number of data that were deemed confidential. The full report was used by the Appraisal Committee at NICE in its deliberations. The full report with each piece of confidential data removed and replaced by the statement 'confidential information (or data) removed' is available on the NICE website: www.nice.org.uk.

The present monograph presents as full a version of the report as is possible while retaining readability. Where confidential data have been removed, this is clearly indicated. The influence of the confidential information on the interpretation and conclusions of this diagnostic assessment is discussed in *Chapter 6*, *Discussion*.

Plain English summary

Pre-eclampsia affects some pregnant women, with potentially serious consequences for the mother and/or baby if not identified and treated. However, not all women suspected of having pre-eclampsia develop it. In the NHS, routine pregnancy care involves checking for signs of pre-eclampsia.

Two new tests (Triage and Elecsys) measure proteins in blood which are often abnormal in women with pre-eclampsia. We investigated whether or not these blood tests accurately predict the risk of pre-eclampsia developing in women suspected of having the condition after week 20 of pregnancy. These tests could help to identify women who require hospital admission for further assessment and women who could safely be monitored by their midwife or family doctor, potentially improving care and saving money.

We conducted extensive medical evidence searches, using review methods that minimised the risk of error and bias. The costs and accuracy of the blood tests were used to develop an economic model. This model estimated costs and benefits to predict whether or not the tests would be good value for money to the NHS.

Our results predict that the 'Triage' and 'Elecsys' tests would improve care and save money if used in addition to routine pregnancy care in women with suspected pre-eclampsia in early pregnancy (20–35 weeks) compared with routine pregnancy care alone. However, cost savings for late pregnancy (35–37 weeks) would be small. There is uncertainty around the size of the cost savings, but the tests were cost-saving even when tested in 'worst-case' scenarios. Research recommendations are made to reduce this uncertainty.

Scientific summary

Background

Pre-eclampsia (PE) is a potentially serious condition affecting up to 5% of pregnancies, most frequently after 20 weeks of gestation. If undetected and untreated it may result in serious maternal and neonatal complications. Suspected PE affects health services by necessitating regular monitoring, testing and treatment. Uncertainty around PE prediction increases the economic burden on the NHS as a result of unnecessary antenatal admissions, fetal monitoring and preterm delivery associated with false-positive diagnoses. Women with PE have longer inpatient stays and their neonates require longer neonatal intensive care unit (NICU) stays than babies born to women without PE. Suspected PE may affect pregnant women through hospitalisation, loss of work days or anxiety. The only cure for PE is to deliver the placenta (and, therefore, the baby), so women are monitored until the optimum time for delivery.

In current practice, the presence or absence of hypertension and proteinuria aid diagnosis of PE, but these markers do not accurately identify or exclude disease with poor pregnancy outcome. Blood tests that could potentially predict PE have recently been developed. These measure the levels of two proteins in blood: placental growth factor (PlGF), which occurs in abnormally low levels in women with PE; and soluble fms-like tyrosine kinase 1 (sFlt-1), which occurs in abnormally high levels in women with PE. However, the diagnostic accuracy and cost-effectiveness of these tests are unclear. The tests specified in the National Institute for Health and Care Excellence (NICE) scope, and included in this diagnostic assessment and economic evaluation, are the Triage® (Alere, Inc., San Diego, CA, USA) PlGF test, the DELFIA® Xpress PlGF 1-2-3 test (PerkinElmer, Wallac Oy, Turku, Finland), the Elecsys® sFlt-1 to PlGF ratio test (Roche Diagnostics GmbH, Mannheim, Germany) and the BRAHMS® sFlt-1 Kryptor/BRAHMS PlGF plus Kryptor PE ratio (Thermo Fisher Scientific GmbH, Hennigsdorf, Germany).

Objectives

The aim was to evaluate the accuracy and cost-effectiveness of biomarker tests at identifying PE for women presenting with suspected PE between 20 weeks and 36^{+6} weeks of gestation who have received blood pressure assessment and qualitative (dipstick) proteinuria assessment. Specific objectives were to determine the accuracy and cost-effectiveness of the Triage PlGF test, Elecsys sFlt-1 to PlGF ratio test, DELFIA Xpress PlGF test and BRAHMS Kryptor sFlt-1 to PlGF ratio test for the diagnosis of PE in the second and third trimesters of pregnancy:

- in addition to standard clinical assessment
- as a replacement for quantitative proteinuria tests.

Methods

Systematic review of test accuracy

A systematic review of diagnostic and prognostic accuracy evidence was undertaken following a peer-reviewed protocol. Searches were based on a comprehensive search strategy. Bibliographic databases including MEDLINE, EMBASE, Web of Science and The Cochrane Library and Database of Abstracts of Reviews of Effects were searched for English-language references in March 2015, and these searches were updated in July 2015. Conferences, websites, systematic reviews and confidential company submissions were also obtained, and reference lists of identified relevant documents were checked. Studies were eligible if they included women with suspected PE in weeks 20–37 of pregnancy, and reported accuracy of

at least one of the specified tests for identifying PE quantitatively relative to standard clinical practice. Risks of bias and generalisability of the included studies were assessed using a modified version of the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) instrument. Study selection, data extraction and critical appraisal were each performed by at least two reviewers, with any disagreements resolved through discussion. Data were synthesised narratively, with an option of conducting a preplanned meta-analysis if data were sufficiently homogeneous. An advisory group comprising five independent clinical experts informed the review by providing comments on draft versions of the protocol and final report.

Systematic review of economic studies

A systematic review of economic studies followed the same process as the review of test accuracy but with modified eligibility criteria for study designs and outcomes. Studies were included if they were full economic evaluations, assessing both costs and consequences, or cost studies for the specified biomarker tests. Outcomes were those consistent with full economic evaluations and cost studies, including intermediate outcomes (budget impact, cost per patient, cost per case of PE correctly managed), or final outcomes [life-years or quality-adjusted life-years (QALYs) gained]. Studies were critically appraised using standard checklists consistent with NICE criteria. Each step of the review was conducted by two health economists, with any disagreements resolved by discussion. Outcomes were synthesised narratively.

Economic evaluation

The External Assessment Group (EAG) developed a de novo decision-analytic model to assess the cost-effectiveness of PlGF tests or sFlt-1 to PlGF ratio tests for the diagnosis of PE when used in addition to standard clinical assessment compared with standard clinical assessment alone. The model was informed by the systematic review of economic studies, confidential company submissions and information provided by clinical experts and the advisory group. Test accuracy parameters and maternal and fetal outcomes were obtained from the systematic review of test accuracy studies, health-related quality of life (HRQoL) was obtained by a systematic search for HRQoL studies, and cost and resource parameters were obtained by targeted searches in relevant sources. The model is a decision tree incorporating the management of clinical symptoms of suspected PE, the timing and mode of delivery, and maternal and neonatal outcomes. Costs (2014, GBP) are evaluated from a NHS and Personal Social Services perspective. Given the analysis time horizon (under 1 year), no discounting was undertaken.

Results

Number and quality of test accuracy studies

Searches yielded 1972 unique bibliographic records, and a further 20 documents were identified through company submissions. After screening these, the systematic review included 12 documents that reported four unique studies: two used the Triage PlGF test and two employed the Elecsys sFlt-1 to PlGF ratio test. One of the included studies on the Triage test, PETRA, was unpublished and confidential when the present report was prepared; this is excluded from the present report, but was available to the EAG and the NICE Diagnostics Assessment Committee.

The three published studies generally rated well on QUADAS criteria, although all three studies had a high risk of clinical review bias. This is because only test results were used to diagnose PE in the primary studies, whereas in clinical practice test results would be interpreted in conjunction with hypertension, proteinuria and/or other signs or symptoms.

Test accuracy outcomes

Meta-analysis of sensitivity and specificity was not feasible because of the heterogeneity of the study populations and outcomes. Test accuracy outcomes differed among studies in terms of the test cut-off points employed, time periods of gestation covered, and time periods following testing to which the outcomes applied. The Triage PlGF test predicts PE requiring delivery within 14 days of testing (i.e. prognosis) for women presenting in weeks 20–35 and in weeks 35–37 of pregnancy, whereas the Elecsys

sFlt-1 to PlGF ratio test is diagnostic, predicting rule-out or rule-in of PE within a specified number of weeks for women presenting at any time in weeks 20–37.

For the Triage PlGF test, data are available for test-positive cut-off points of < 100 pg/ml, < 12 pg/ml and < 5th percentile of PlGF concentration, but the < 12 pg/ml cut-off point had low sensitivity (≤ 63%). The < 100 pg/ml and < 5th percentile cut-off points both had high sensitivity (96%) for identifying women likely to develop PE requiring delivery within 14 days, when presenting with suspected PE up to 35 weeks of gestation. However, sensitivity was lower after 35 weeks of gestation (70% for the < 5th percentile cut-off point). Diagnostic accuracy outcomes for the Elecsys sFlt-1 to PlGF ratio are for three test cut-off points: 23, 38 and 85. However, the majority of data are from one study (PROGNOSIS) that employed the 38-week cut-off point. The PROGNOSIS study outcomes suggest that the Elecsys sFlt-1 to PlGF ratio is appropriate for rule-out of PE within 1 week of testing (sensitivity 85.7%, negative predictive value 99.1%) and for rule-in of PE within 4 weeks of testing (specificity 83.1%), although with a relatively high likelihood of false positives (positive predictive value 38.6%).

Number and quality of economic studies

Three documents were included in the systematic review of economic studies, which reported on three unique studies. These were cost analyses, focusing on potential savings in health sector resources through improved accuracy of diagnosis of PE. None of the three studies formally evaluated maternal or neonatal outcomes (other than admission to intensive care or to a special care baby unit, which were included in the cost analysis). These studies all have limitations, including that none measured health benefits, none adequately described and justified its resource costs, and none reported whether or not its model was validated. Owing to heterogeneity of the study designs and outcomes, meta-analysis was inappropriate and the results were synthesised narratively.

A further two cost studies for the Triage PlGF test and the Elecsys sFlt-1 to PlGF ratio test were provided in confidential company evidence submissions as part of the NICE Diagnostics Assessment process. These are not described in the current report, but were taken into consideration by the EAG when planning the de novo independent economic analysis.

Results of the cost-effectiveness analysis

The EAG cost-effectiveness model predicts that, when supplementing routine clinical assessment for rule-out and rule-in of PE in women with suspected PE, the Triage PlGF test and the Elecsys sFlt-1 to PlGF ratio test would both be cost-saving in weeks 20–35 of gestation and marginally cost-saving in weeks 35–37, but with a minuscule impact on QALYs. Cost differences slightly favour the Triage PlGF test for both gestational periods. The magnitude of savings is uncertain, but the tests remain cost-saving under worst-case assumptions. Length of NICU stay was the most influential parameter in sensitivity analyses. All other sensitivity analyses had negligible effects on results.

Scenario analyses assessing the effects of replacing quantitative proteinuria testing with biomarker testing, and assessing near-patient testing instead of central laboratory testing found negligible impacts on cost-savings for the biomarker tests.

Discussion

Strengths of the evidence synthesis

The current diagnostic assessment was based on a prespecified, peer-reviewed protocol. It included comprehensive literature searches in a wide variety of data sources undertaken by an experienced information specialist. The study selection and data extraction steps were based on standard pilot tested worksheets. Evidence was critically appraised using prespecified and internationally accepted criteria. Study selection, data extraction and critical appraisal were conducted by at least two reviewers to minimise

risks of errors and bias. All excluded full-text documents are listed with the reasons for exclusion. An independent advisory group informed the protocol, economic model and draft report.

Limitations of the test accuracy evidence synthesis

No head-to-head comparisons of relevant biomarker tests were identified. The included evidence base addresses only part of the decision problem, as no relevant studies were found for the DELFIA Xpress PlGF test or the BRAHMS Kryptor sFlt-1 to PlGF ratio test. No relevant studies have investigated the accuracy of PlGF-based tests when used as a replacement for proteinuria testing. Test accuracy studies were at high risk of clinical review bias. Meta-analysis was not possible because the studies employed different outcome measures, test cut-off points and gestational periods. Searches were limited to English-language references; however, we consider it unlikely that this resulted in us missing relevant evidence. The current report does not present information from confidential studies that were available to the EAG and NICE Diagnostics Assessment Committee. However, as explained in *Chapter 6*, the excluded information would not materially affect the conclusions.

Limitations of the economic analysis

Owing to lack of adequate diagnostic effectiveness data, only the Triage PlGF test and the Elecsys sFlt-1 to PlGF ratio test could be analysed. The economic analysis is based on several assumptions because of data insufficiency.

Uncertainties

Although data about clinical outcomes other than those related to diagnosing PE (maternal and fetal morbidity and mortality, emergency admission) were reported in some studies, heterogeneity between studies prevented useful assessment of test effects on these outcomes.

Data are lacking for women presenting before 30 weeks of gestation who are at high risk of adverse maternal and neonatal outcomes. There were also insufficient HRQoL data for women with gestational hypertension and PE. The EAG relied heavily on mapping algorithms from the Short Form questionnaire-36 items to provide European Quality of Life-5 Dimensions (EQ-5D) utility estimates. These appeared to overestimate EQ-5D utility scores compared with those measured directly using EQ-5D. However, as no studies have validated the EQ-5D for use in pregnancy or post-partum periods, we cannot rule out the possibility that EQ-5D might have underestimated HRQoL in these periods.

Data are lacking for long-term maternal and neonatal outcomes in women with gestational hypertension, in the general population of pregnant women who give birth preterm, and in high-risk subgroups of women with previous PE, multiple pregnancies, diabetes mellitus (pre-existing or gestational) or renal or autoimmune conditions.

Conclusions

The PlGF and sFlt-1 to PlGF ratio tests are currently used to predict PE in only a few UK hospitals. However, our results suggest that there would be clinical benefits and cost savings of using the Triage PlGF test or the Elecsys sFlt-1 to PlGF ratio test, when added to standard clinical assessment, for women presenting with suspected PE between 20 and 37 weeks of gestation. Sensitivity analyses indicate that replacing quantitative proteinuria testing with a PlGF test or a sFlt-1 to PlGF ratio test, or conducting the biomarker tests in a near-patient (e.g. antenatal clinic) setting (as opposed to a central laboratory), would have negligible impact on cost-effectiveness. The most appropriate location and type of testing would vary by local needs and local acquisition and maintenance costs for the test equipment. Investment in equipment and training will be required for any of the biomarker tests to be employed in NHS practice. Further information on the DELFIA Xpress PlGF test and the BRAHMS Kryptor sFlt-1 to PlGF ratio test would be helpful to allow adequate evaluation of their potential test accuracy and cost-effectiveness compared with the Triage PlGF test and the Elecsys sFlt-1 to PlGF ratio test.

Research recommendations

Observational research studies are needed to clarify long-term fetal, neonatal and maternal outcomes for women diagnosed with PE and the utilities associated with these.

Pragmatic research studies should clarify how the PlGF test and sFlt-1 to PlGF ratio test influence key decisions in a clinical setting.

Head-to-head comparisons of PlGF-based tests would help to clarify which test(s) could be most cost-saving for the NHS. This would require that the tests employ the same diagnostic or prognostic end points and cover the same periods of gestation. Such studies should be designed so as to minimise bias, pragmatically reflect UK clinical practice, include women with suspected PE between 20 and 30 weeks of gestation (in addition to other gestational age groups) and employ definitions of PE that are consistent with those employed in UK clinical practice.

Study registration

This study is registered as PROSPERO CRD42015017670.

Funding

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

Chapter 1 Background

Description of the health problem

Pre-eclampsia (PE) is a potentially serious complication that occurs in some pregnant women, most frequently during the second half of pregnancy (after 20 weeks of gestation). It is associated with placental dysfunction, whereby blood flow through the placenta is reduced, and is characterised by maternal hypertension and proteinuria, although not all women have both of these manifestations.[1] If PE is undetected and untreated, it may result in complications that include disseminated intravascular coagulation, stroke and organ dysfunction, or can develop into eclampsia, a potentially life-threatening convulsive condition. The only cure for PE is to deliver the placenta (and, therefore, the baby), and this may avoid or remedy complications associated with PE.[2] Women who have hypertension or PE during pregnancy may also have a higher risk of complications from placental abruption (when the placental lining separates from the uterus before delivery).[2] Gestational hypertension (high blood pressure that develops during pregnancy) and PE can also affect the fetus, increasing the risk of intrauterine growth restriction (IUGR) and intrauterine death.[3] PE can also develop in women with chronic hypertension before pregnancy, and in such cases is known as superimposed PE.[4] PE is frequently asymptomatic and, in such cases, may be detected only through routine antenatal testing. Symptoms of PE can include neurological symptoms (headache or visual disturbances), epigastric or right upper-quadrant pain,[5] oedema (swelling of the hands, face or feet) and oliguria (low output of urine).[6] Although most cases of PE are mild and cause no problems, the condition can worsen and be serious for both mother and baby.[7] However, PE before week 34 of pregnancy is less common but, when it occurs, is often more severe.[8]

Epidemiology

Pre-eclampsia affects up to 5% of pregnancies, and severe PE occurs in about 1–2% of pregnancies.[7] In 2012–13, 12,356 pregnant women were admitted to hospital in England for PE, and 294 for eclampsia.[9] Maternal deaths attributable to PE have fallen,[10] and in the UK and Ireland , in 2010–12, only nine deaths directly attributable to PE or eclampsia were recorded (0.38 per 100,000), although deaths from related conditions also occurred, including two caused by placental abruption (0.49 per 100,000).[11] According to Action on Pre-eclampsia, fetal mortality is much higher, and around 1000 babies die each year as a result of PE, mostly because of complications associated with early delivery.[12]

Definitions of pre-eclampsia and related conditions

There is no international consensus on the criteria by which to diagnose PE and related conditions, although criteria provided by different organisations overlap, as shown in *Table 1*. The criteria for diagnosing PE that are relevant to the current diagnostic assessment are those provided by the National Institute for Health and Care Excellence (NICE),[13] the American Congress of Obstetricians and Gynecologists (ACOG)[14] and the International Society for the Study of Hypertension in Pregnancy (ISSHP),[15] as these are the criteria referred to in studies of the diagnostic accuracy of placental growth factor (PlGF) tests and soluble fms-like tyrosine kinase 1 (sFlt-1) to PlGF ratio tests.

Pre-eclampsia is defined by NICE as gestational hypertension accompanied by gestational proteinuria (i.e. hypertension and proteinuria occurring during pregnancy) after week 20 of pregnancy.[13] The presence of either hypertension or proteinuria alone during pregnancy can also indicate a risk of developing PE.[13] PE is classified as early onset if it occurs before week 34 of pregnancy or as late onset if it occurs after week 34.[4]

Hypertension in pregnancy is defined by NICE, ACOG and ISSHP as a systolic blood pressure of at least 140 mmHg and/or a diastolic blood pressure of at least 90 mmHg, and they all define proteinuria in

TABLE 1 Criteria for diagnosing PE and related conditions

Condition	Organisation		
	ACOG (2002 bulletin)[14]	NICE[13]	ISSHP[15]
Gestational hypertension	Elevated systolic blood pressure (≥ 140 mmHg) or diastolic blood pressure ≥ 90 mmHg without proteinuria after gestational week 20, with blood pressure returning to normal post partum	New hypertension (systolic blood pressure ≥ 140 mmHg or diastolic blood pressure ≥ 90 mmHg) presenting after 20 weeks without significant proteinuria	Elevated systolic blood pressure (≥ 140 mmHg) or diastolic blood pressure (≥ 90 mmHg) after gestational week 20 or 'after mid-pregnancy' (ISSHP does not specify which it adopts, but notes that these definitions are 'very similar')
Proteinuria	Presence of ≥ 0.3 g of protein in a 24-hour urine specimen. This finding usually correlates with a dipstick finding of 1+ or greater, but should be confirmed using a random urine dipstick evaluation and a 24-hour or 'timed' collection	Diagnose significant proteinuria if the urinary protein-to-creatinine ratio is > 30 mg/mmol or a validated 24-hour urine collection result shows > 300 mg of protein	Presence of > 300 mg/day of urinary protein by urine spot dipstick confirmed by 24-hour or 'timed' quantitative measure if possible
PE	Hypertension and proteinuria that may be associated with other myriad signs and symptoms, such as oedema, visual disturbances, headache, epigastric pain or laboratory abnormalities indicating HELLP syndrome. An updated ACOG definition published in 2013[16] allows for new onset of hypertension without new-onset proteinuria if there are also new-onset signs or features of the syndrome	New hypertension presenting after 20 weeks with significant proteinuria (see above)	Incorporates PE and eclampsia Two different definitions For research: new-onset hypertension after gestation week 20 plus proteinuria For clinical practice: new hypertension after gestational week 20 plus new onset of one or more of: • proteinuria ≥ 300 mg/day or spot urine protein-to-creatinine ratio ≥ 30 mg/mmol • renal insufficiency • liver disease (raised transaminases and/or severe right upper quadrant pain) • neurological problems (eclampsia, severe headaches with hyperreflexia, persistent visual disturbances) • haematological disturbances (thrombocytopenia, disseminated intravascular coagulation, haemolysis) • fetal growth restriction
HELLP syndrome	Haemolysis, elevated liver enzymes, low platelet count	Haemolysis, elevated liver enzymes and low platelet count	Incorporated in the definition of PE
Superimposed PE	New-onset proteinuria in a woman who has hypertension before 20 weeks of gestation, or a sudden increase in proteinuria if already present in early gestation, or a sudden increase in hypertension or the development of HELLP syndrome	NICE does not explicitly define superimposed PE, but acknowledges that PE can arise in pregnant women with chronic hypertension (defined as hypertension that is present at the booking visit or before 20 weeks or if the woman is	New signs and/or symptoms of PE after gestational week 20 in a woman with chronic hypertension (hypertension pre-conception or in the first half of pregnancy, either essential if no underlying cause, or secondary)

TABLE 1 Criteria for diagnosing PE and related conditions (*continued*)

Condition	Organisation		
	ACOG (2002 bulletin)[14]	NICE[13]	ISSHP[15]
	Women with chronic hypertension who develop headache, scotomata, or epigastric pain may also have superimposed PE	already taking antihypertensive medication when referred to maternity services. It can be primary or secondary)	
Severe PE	Elevated systolic blood pressure (≥ 160 mmHg) or diastolic blood pressure (≥ 110 mmHg) on two occasions at least 6 hours apart while patient is on bed rest; proteinuria of ≥ 5 g in a 24-hour urine specimen or ≥ 3+ on two random urine samples collected at least 4 hours apart; oliguria of less than 500 ml in 24 hours; cerebral or visual disturbances; pulmonary oedema or cyanosis; epigastric or upper right quadrant pain; impaired liver function; thrombocytopenia; fetal growth restriction	PE with severe hypertension and/or with symptoms, and/or biochemical and/or haematological impairment. (Severe hypertension is a diastolic blood pressure of ≥ 110 mmHg, systolic blood pressure of ≥ 160 mmHg)	Incorporated in the definition of PE
Eclampsia	New-onset grand mal seizures in a woman with PE (although other causes of seizures include bleeding arteriovenous malformation, ruptured aneurysm or idiopathic seizure disorder)	A convulsive condition associated with PE	Incorporated in the definition of PE

HELLP, haemolysis, elevated liver enzymes, low platelet count.

pregnancy as 300 mg or more of protein in a 24-hour urine collection (which equates to 30 mg/dl or 30 mg/mmol)[15] (see *Table 1*).

Some researchers have developed amended versions of the ACOG definition of PE. For example, the PETRA study[6] (included in this diagnostic assessment, see *Chapter 4*) used an 'expanded' definition of PE, to address the limitations of the 2002 ACOG definition, in that the diagnostic criteria were considered too rigid and give insufficient recognition to the progressive nature and varied rate of development of PE.[16]

Another condition associated with PE is haemolysis, elevated liver enzymes, low platelet count (HELLP) syndrome. Experts disagree about whether or not this is a variant of PE or a separate syndrome[17] (see *Table 1*), but HELLP syndrome is diagnosed in pregnant women with signs of PE or even eclampsia, combined with laboratory findings of haemolysis (breakdown of red blood cells), elevated liver function test results and low platelet count, hence the acronym HELLP.

Women with gestational hypertension without proteinuria occasionally develop severe PE, eclampsia, HELLP syndrome, disseminated intravascular coagulation, acute renal or hepatic failure, or placental abruption.[18] Some women with isolated gestational proteinuria may also later develop hypertension and PE,[19] and in these cases the PE may be more severe than in women who present with both hypertension and proteinuria.[20]

Impact of pre-eclampsia

Impact on pregnant women and babies

Hypertension in pregnancy carries risks for mother and baby and increases a woman's lifetime risk of hypertension, PE in subsequent pregnancies,[21] ischaemic heart disease, stroke, type 2 diabetes mellitus and venous thromboembolism.[10,22] Negative consequences of PE for the baby include fetal growth restriction and preterm birth,[3] which can lead to complications including intracranial haemorrhage, nutritional compromise, necrotising enterocolitis and breathing difficulties (neonatal respiratory distress syndrome)[7] and necessitate a stay in a neonatal intensive care unit (NICU). Decisions about when to deliver the baby when mothers have PE involve a balance between the best outcomes for both the mother and the baby.[23] Before 34 weeks of gestation, clinicians would aim to prolong the pregnancy long enough for the fetus to develop as much as possible before birth. Some babies die because of complications related to early delivery, and a few are stillborn.[7] Babies born early, or small for gestational age, may experience preschool developmental delays[24] or be at increased risk of adult disease.[25] However, the baby may be delivered early if there is a risk that the mother may develop severe PE, HELLP syndrome, disseminated intravascular coagulation, acute renal failure, hepatic failure, placental abruption or eclampsia.

Suspected PE may have a negative impact on pregnant women if it involves hospitalisation, loss of work days and/or anxiety, and previously pre-eclamptic women, particularly those with severe PE, have reported poorer quality of life than women with normotensive pregnancies.[26,27] PE can be stressful for both parents, owing to worry about the condition of the unborn baby and the risk of morbidity and mortality as a result of preterm birth.[28] Having a condition that can deteriorate rapidly, being kept in hospital for monitoring, uncertainty about what will happen and undergoing emergency caesarean section can also cause women to experience fear, anxiety, loss of control over their situation and anxiety about future pregnancies.[29] Women's partners and friends can be affected because of fear of losing the mother or the baby.[29] Evidence is mixed but, generally, PE or HELLP syndrome is associated with increased prevalence or severity of depression and post-traumatic stress disorder.[30] In one study[31] post-partum depression was identified in 23% of women with mild PE and in 44% of women with severe PE, although the higher incidence among women who had suffered severe PE was attributed to infants' admission to a neonatal unit or perinatal death, rather than to the severity of disease alone.

Significance for the NHS

Pregnant women need to be monitored during routine antenatal care, and should be given advice about the action that they need to take if they experience symptoms that may indicate PE.[13,32] If proteinuria is identified on a dipstick ('qualitative') test, a spot urinary protein-to-creatinine ratio or 24-hour urine collection is needed to quantify the level of proteinuria. Twenty-four-hour urine collection may necessitate an overnight stay in hospital, refrigeration of the urine during collection and laboratory-based analysis. Women suspected of having PE should be referred to a specialist and admitted to hospital for both maternal and fetal monitoring; if not admitted to hospital, women would need ongoing regular monitoring in case signs of PE develop. The uncertainty around PE prediction increases the economic burden on the NHS by increasing false-positive diagnoses, antenatal admissions, fetal monitoring and preterm delivery.[33] A 2011 study of women who had PE in a previous pregnancy found that there were longer maternal inpatient stays (12.97 days for women with PE in their current pregnancy compared with 5.06 days for women without PE in their current pregnancy) and more inpatient days in a neonatal or special care baby unit (14.7 days for babies whose mothers had PE in their current pregnancy compared with 1.35 days for babies of women without PE).[34]

Care pathway

The NICE care pathway for women at risk of or with PE is shown in *Figure 1*. PE may progress unpredictably, within hours or over weeks,[23] so women are assessed at antenatal appointments, and women with one high-risk factor (hypertensive disease during a previous pregnancy, chronic kidney

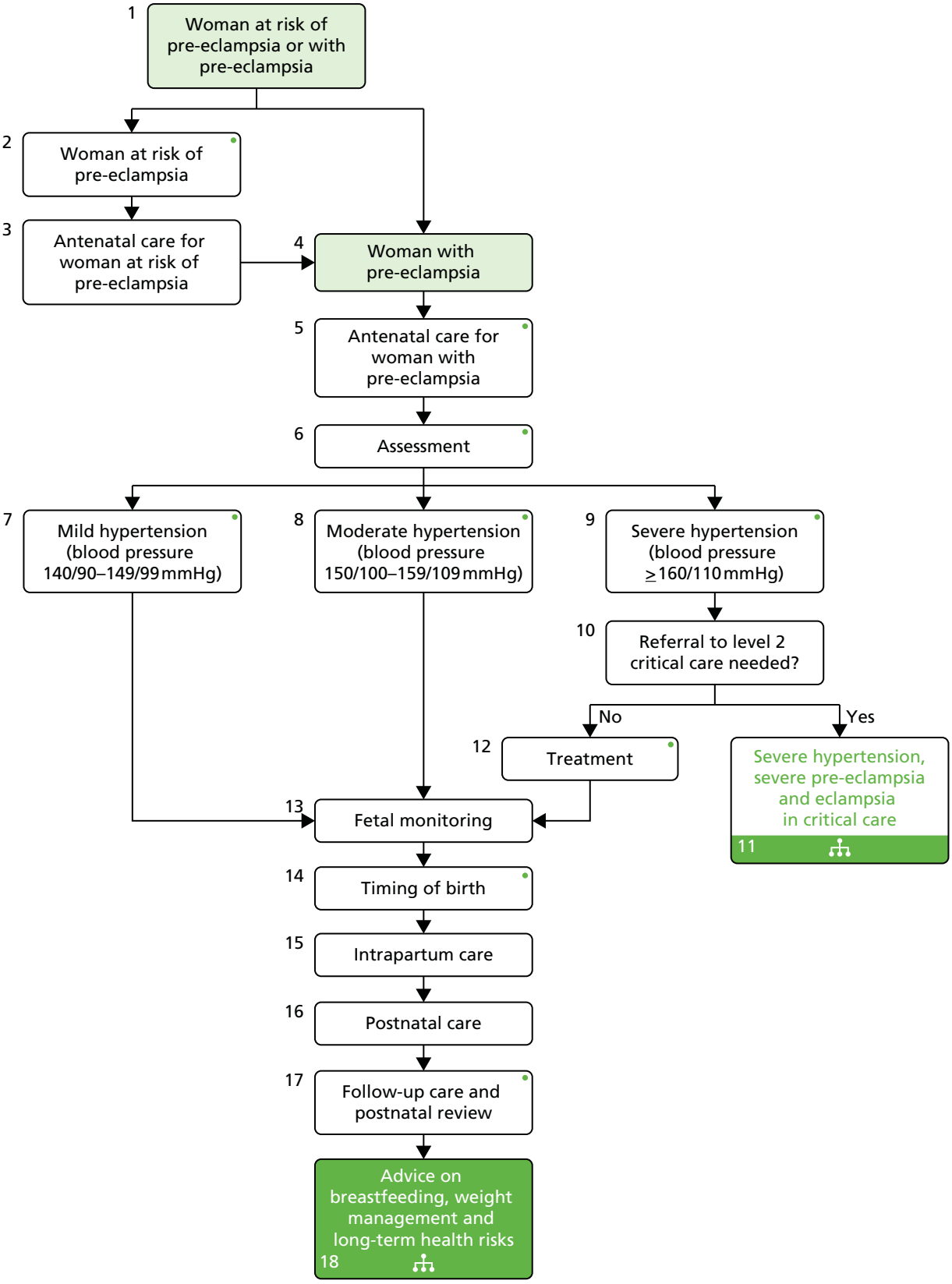

FIGURE 1 Overview of hypertension in pregnancy. National Institute for Health and Care Excellence (2016). Adapted from *NICE Pathway: Hypertension in Pregnancy*. Available from https://pathways.nice.org.uk/pathways/hypertension-in-pregnancy.[35] Reproduced with permission from NICE. The material was accurate at the time of going to press.

disease, autoimmune disease such as systemic lupus erythematosus or antiphospholipid syndrome, type 1 or type 2 diabetes mellitus, chronic hypertension) or more than one moderate-risk factor for PE [age over 40 years, first pregnancy, pregnancy interval over 10 years, family history of PE, previous history of PE, body mass index (BMI) of ≥ 30 kg/m^2, pre-existing hypertension, multiple pregnancy] may be advised to take 75 mg of aspirin daily from 12 weeks of gestation until the birth of the baby.[13] NICE clinical guideline (CG) 107 recommends immediate hospital referral for assessment of mother and fetus if PE is suspected to determine whether or not PE is an appropriate diagnosis.[36]

NICE recommends that, once PE is diagnosed, women are assessed at each consultation by a suitably trained health-care professional and offered an integrated package of care that includes hospital admission, testing and treatment relating to the severity of hypertension.[13] Conservative management in hospital (or the community) continues until 34 weeks, unless there is clinical and test evidence of severe hypertension or potential harm to the baby.[13,32] NICE CG107[13] recommends management according to blood pressure thresholds (*Tables 2* and *3*). Antihypertensive drugs (labetalol, methyldopa or nifedipine) are given, with a target systolic blood pressure of 150 mmHg.[37] PE can be cured only by delivering the baby, so women are monitored until delivery is optimal for both mother and baby. This is usually around 37–38 weeks of pregnancy, but may be earlier in more severe cases.[7] For women with PE and mild or moderate hypertension, delivery is offered between 34 weeks and 36^{+6} weeks of gestation, depending on maternal and fetal condition, risk factors and availability of neonatal intensive care.[13] Delivery within 24–48 hours is recommended for women with PE and mild or moderate hypertension after 37 weeks of gestation.[13]

During hospitalisation for PE, ultrasonography is carried out to monitor fetal growth and well-being (by blood flow measurements in the umbilical cord).[7] Cardiotocography is used to measure the baby's heart rate to detect any signs of compromise.[7]

Diagnosis of gestational hypertension and pre-eclampsia

Diagnosing PE is challenging because symptoms and signs are highly variable: women can be asymptomatic despite severe disease and the disease can progress over several weeks before diagnosis is confirmed.[1] Assessment therefore begins during routine antenatal appointments, when blood pressure is measured, urinalysis for protein is carried out and risk factors for PE are assessed.[32] Women with risk

TABLE 2 Ongoing testing for patients with gestational hypertension

Action	Degree of hypertension		
	Mild hypertension (140/90–149/99 mmHg)	Moderate hypertension (150/100–159/109 mmHg)	Severe hypertension (160/110 mmHg or higher)
Admit to hospital	No	No	Yes (until blood pressure is 159/109 mmHg or lower)
Measure blood pressure	Not more than once a week	At least twice a week	At least four times a day
Test for proteinuria	At each visit using automated reagent strip reading device or urinary protein-to-creatinine ratio	At each visit using automated reagent strip reading device or urinary protein-to-creatinine ratio	Daily using automated reagent strip reading device or urinary protein-to-creatinine ratio
Blood tests	Only those for routine antenatal care	Test kidney function, electrolytes, full blood count, transaminases, bilirubin Do not carry out further blood tests if no proteinuria at subsequent visits	Test at presentation and then monitor weekly: kidney function, electrolytes, full blood count, transaminases, bilirubin

Birth before 37 weeks should not be offered to women with gestational hypertension whose blood pressure is lower than 160/110 mmHg with or without antihypertensive treatment. National Institute for Health and Clinical Excellence (2010). Adapted from *CG107: Hypertension in Pregnancy: Diagnosis and Management*.[13] Available from www.nice.org.uk/Guidance/cg107.[13] Reproduced with permission from NICE. The material was accurate at the time of going to press.

TABLE 3 Further testing after diagnosis of PE

Action	Degree of hypertension		
	Mild hypertension (140/90–149/99 mmHg)	Moderate hypertension (150/100–159/109 mmHg)	Severe hypertension (160/110 mmHg or higher)
Admit to hospital	Yes	Yes	Yes
Measure blood pressure	At least four times a day	At least four times a day	More than four times a day, depending on clinical circumstances
Test for proteinuria	Do not repeat quantification of proteinuria	Do not repeat quantification of proteinuria	Do not repeat quantification of proteinuria
Blood tests	Monitor using the following tests twice a week: kidney function, electrolytes, full blood count, transaminases, bilirubin	Monitor using the following tests three times a week: kidney function, electrolytes, full blood count, transaminases, bilirubin	Monitor using the following tests three times a week: kidney function, electrolytes, full blood count, transaminases, bilirubin

National Institute for Health and Clinical Excellence (2010). Adapted from *CG107: Hypertension in Pregnancy: Diagnosis and Management*.[13] Available from www.nice.org.uk/Guidance/cg107.[13] Reproduced with permission from NICE. The material was accurate at the time of going to press.

factors may then undergo more frequent blood pressure monitoring, and surveillance is increased for those with significant hypertension (diastolic pressure of 90–110 mmHg) and/or proteinuria (1+ on urinalysis reagent strips).[32] Women are also advised to seek health-care advice if they experience symptoms of PE, including severe headache, vision problems, pain just below the ribs, vomiting or sudden swelling of the face, hands or feet.[13] NICE offers guidance on further tests to monitor proteinuria and identify HELLP syndrome, including kidney function, electrolytes, full blood count, transaminases and bilirubin, and on the frequency of tests to be carried out at antenatal visits or in hospital, depending on the severity of hypertension[13] (see *Table 2*).

Clinical experts advising the current diagnostic assessment suggested that most women with suspected PE present with gestational hypertension, while a minority present with other signs and symptoms, of which the most common is proteinuria. Women who have proteinuria without hypertension may also therefore have more regular blood pressure assessment. Clinical experts also suggested that around 20% of pregnant women presenting with new gestational hypertension and 30–50% of pregnant women presenting with quantitatively measured proteinuria will have PE.

Guidance is offered by NICE on the measures and frequency of testing that should be followed once PE is diagnosed (see *Table 3*), including blood pressure, proteinuria and indicators of HELLP syndrome.[13]

Fetal monitoring can provide additional information in cases of suspected PE. Such monitoring includes ultrasound fetal growth measurement, amniotic fluid volume assessment and umbilical artery blood flow measured by Doppler velocimetry[13] to aid identification of IUGR. In women with chronic hypertension or high risk of PE these tests are carried out between 28 and 30 weeks and again between 32 and 34 weeks, while in women with gestational hypertension tests are carried out before 34 weeks, and they may be used in conservative management of severe gestational hypertension or PE.[13] Cardiotocography is carried out at diagnosis of severe gestational hypertension or PE, or in mild or moderate hypertension if fetal activity is abnormal.[13]

The presence or absence of the current diagnostic and prognostic markers of PE (hypertension and proteinuria) does not accurately identify or exclude disease with poor pregnancy outcome. More accurate diagnosis and prediction of PE is needed to inform clinicians' decisions about optimal management to improve outcomes for mothers and babies,[38] and may reduce costs by reducing unnecessary hospitalisations and procedures due to uncertain diagnosis.[39]

Description of the diagnostic technologies under assessment

Tests that enable an earlier and more accurate prediction of the risk of PE may enable those at low risk to remain in the community setting. Tests have been developed that measure the levels of two proteins in blood, each of which can be abnormal in women with PE. The first, PlGF, promotes the development of new blood vessels (which is important for a healthy placenta) and is found in abnormally low levels in women with PE. The second, sFlt-1, blocks the effect of PlGF and occurs in abnormally high levels in women with PE. The tests measure the blood level either of PlGF or the ratio of sFlt-1 to PlGF, and are intended for use in conjunction with clinical judgement and other existing diagnostic tests to aid the diagnosis of PE. These tests may provide earlier and more accurate prediction of the risk of PE in pregnant women who have signs and symptoms suggestive of the condition.

The tests specified in the NICE scope, and included in this health diagnostic assessment, are the Alere Triage® PlGF test (Alere, Inc., San Diego, CA, USA), the PerkinElmer DELFIA® Xpress PlGF 1-2-3 test (PerkinElmer, Wallac Oy, Turku, Finland), the Elecsys® immunoassay measuring the sFlt-1 to PlGF ratio (Roche Diagnostics Limited, Burgess Hill, UK) and the BRAHMS® sFlt-1 Kryptor/BRAHMS PlGF plus Kryptor PE ratio test (Thermo Fisher Scientific GmbH, Hennigsdorf, Germany).

Triage PlGF test (Alere)

The Triage PlGF test is a Conformité Européenne (CE)-marked single-use fluorescence immunoassay device that is used in conjunction with the Alere Triage MeterPro (Alere Inc., San Diego, CA, USA) point-of-care analyser for the quantitative determination of PlGF in blood plasma samples.[40] The test is intended for use in conjunction with clinical judgement and other existing diagnostic tests, to aid the diagnosis of PE and to assess the level of risk for delivery arising from PE within 14 days of testing.[40] Each Triage PlGF test device contains mouse monoclonal antibodies against PlGF, fluorescent dye and stabilisers. Prior to use of the test device, a blood sample is centrifuged for around 3 minutes to obtain an ethylenediaminetetraacetic acid (used to prevent clotting)-anticoagulated plasma specimen. A 250-µl sample of plasma is then added to the Triage PlGF test device's sample port where it reacts with fluorescent antibody conjugates and flows through the test device, via capillary action, to a measurement zone in which complexes of the fluorescent antibody conjugates are captured.[41] The test device is inserted into the Triage MeterPro analyser, which measures levels of fluorescence from the antibody–conjugate complexes. It has been reported that the test has a limit of detection of 9 pg/ml and a measurable range of 12–3000 pg/ml.[42] The test turnaround time is reported as approximately 15 minutes.[41]

The Triage PlGF test is recommended for use in pregnant women with a gestational age of between 20 weeks and 34 weeks^{+6} days. The test cut-off points, derived from a population with suspected PE[5] and recommended by the company, are shown in *Table 4*.

DELFIA Xpress PlGF test (PerkinElmer)

The DELFIA Xpress PlGF test is a CE-marked, solid-phase, two-site fluoroimmunometric sandwich assay for the quantitative determination of PlGF in serum samples. The test is intended as an aid to the diagnosis of

TABLE 4 Recommended cut-off points for the Triage PlGF test

Test cut-off point	Classification	Interpretation
PlGF < 12 pg/ml	Test positive – highly abnormal	Highly abnormal and suggestive of patients with severe placental dysfunction and at an increased risk for preterm delivery
PlGF ≥ 12 pg/ml and < 100 pg/ml	Test positive – abnormal	Abnormal and suggestive of patients with placental dysfunction and at an increased risk for preterm delivery
PlGF ≥ 100 pg/ml	Test negative – normal	Normal and suggestive of patients without placental dysfunction and unlikely to progress to delivery within 14 days of the test

PE during the second and third trimesters of pregnancy, is used in conjunction with clinical assessment and is a laboratory-based rather than a near-patient test.[43] The assay includes both immobilised and europium-labelled monoclonal antibodies, which bind to PlGF molecules present in the sample to form PlGF–monoclonal antibody complexes. The resulting europium fluorescence from each sample is proportional to the concentration of PlGF. The assay has a limit of detection of 1.9 pg/ml (measuring range 1.9–4000 pg/ml) and a limit of quantitation of 3.3 pg/ml. The assay is compatible with the PerkinElmer 6000 DELFIA Xpress random access analyser (PerkinElmer, Wallac Oy, Turku, Finland). The company advises that cut-off values for PlGF measurements obtained during the second trimester are highly dependent on gestational day and should be established by individual laboratories.[43] In the third trimester the company advises that, in addition to laboratory calculated cut-off values based on the gestational day, a fixed cut-off point of 184 pg/ml can be used. Levels of PlGF lower than 184 pg/ml indicate an elevated probability of PE developing.[43] Cut-off values were calculated in a case–control study that included samples from women with PE but without chronic hypertension.[44]

Elecsys sFlt-1 to PlGF ratio test (Roche Diagnostics)

The Elecsys immunoassay measures the relative amounts of sFlt-1 to PlGF [also known as vascular endothelial growth factor receptor 1 (VEGFR1)] in serum samples from women with suspected PE.[45] The ratio is formed by combining the results from two CE-marked sandwich electrochemiluminescence immunoassays (the Elecsys PlGF and Elecsys sFlt-1 assays), which are compatible with both the Elecsys and the Cobas® e automated clinical chemistry analysers (Roche Diagnostics Limited, Burgess Hill, UK).[46,47] The sFlt-1 to PlGF ratio is calculated and reported to the user alongside the individual assay values by the laboratory information system. The Elecsys immunoassay sFlt-1 to PlGF ratio is intended for use in aiding the diagnosis of PE in conjunction with clinical judgement and other diagnostic tests. In addition, the ratio may be used as an aid to predict PE, eclampsia and HELLP syndrome in the short term.[45] The Elecsys sFlt-1 assay has a limit of detection of 10 pg/ml (measuring range 10–85,000 pg/ml) and a limit of quantitation of 15 pg/ml. The Elecsys PlGF assay has a limit of detection of 3 pg/ml (measuring range 3–10,000 pg/ml) and a limit of quantitation of 10 pg/ml.[42] The time required to measure the sFlt-1 to PlGF ratio is 18 minutes.[48]

The Elecsys immunoassay sFlt-1 to PlGF ratio may be used for testing pregnant women with suspected PE from a gestational age of 20 weeks up until the time of delivery. The test cut-off points previously recommended by the company, derived from case–control studies of patients with PE or normal pregnancy outcome, are shown in *Table 5*.[49–51]

After further research and clinical consensus, these recommendations have been updated to include a more confident rule-in and rule-out of PE with a sFlt-1 to PlGF ratio cut-off point of 38, so that clinicians have greater certainty when making decisions about patient management (*Table 6*).

TABLE 5 Recommended cut-off points for the Elecsys sFlt-1 to PlGF ratio test

Test role	Gestation period	sFlt-1 to PlGF ratio cut-off point	
Aid in diagnosis of PE	Week 20^{+0} to week 33^{+6}	Rule out	< 33
		Rule in	> 85
	Week 34^{+0} to delivery	Rule out	< 33
		Rule in	> 110
Short-term prediction of PE	Week 24^{+0} to week 36^{+6}	Rule out[a]	≤ 38
		Rule in[b]	> 38

a Rule out PE for 1 week.[51]
b Rule in PE within 4 weeks.[51]

TABLE 6 Updated recommended cut-off points for the Elecsys sFlt-1 to PlGF ratio test

sFlt-1 to PlGF ratio cut-off point and interpretation	Early-onset: weeks 20 to 33^{+6} of gestation	Late-onset: week 34 to end of pregnancy
≥ 85: diagnosis – rule in PE	Specificity: 99.5% Sensitivity: 88.0%	Not applicable
≥ 110: diagnosis – rule in PE	Not applicable	Specificity: 95.5% Sensitivity: 58.2%
≥ 38: prediction – rule in PE within next 4 weeks	PPV: 38.6%	
< 38: prediction – rule out PE for the next week	NPV: 99.1%	

NPV, negative predictive value; PPV, positive predictive value.
Data are from the Elecsys sFlt-1 test and PlGF test product inserts.[46,50]

BRAHMS sFlt-1 Kryptor to BRAHMS PlGF plus Kryptor ratio test (Thermo Fisher Scientific)

The BRAHMS Kryptor sFlt-1 to PlGF ratio test is formed by combining the results from two automated immunofluorescent sandwich assays, the BRAHMS sFlt-1 Kryptor and BRAHMS PlGF plus Kryptor assays. The assays are indicated for the quantitative determination of sFlt-1 and PlGF in serum samples and are compatible with the BRAHMS Kryptor compact plus analyser (Thermo Fisher Scientific GmbH, Hennigsdorf, Germany). The assays are intended to be run simultaneously, with the analyser reporting both the concentrations for each assay and the sFlt-1 to PlGF ratio to the user. The BRAHMS Kryptor sFlt-1 to PlGF ratio test is intended to be used in conjunction with clinical assessment to aid the diagnosis of PE.[53] The total durations of the assays are 9 minutes (sFlt-1) and 29 minutes (PlGF).[48]

The BRAHMS sFlt-1 Kryptor assay has a limit of detection of 22 pg/ml (measuring range 22–90,000 pg/ml) and a limit of quantification of 34 pg/ml.[48] The BRAHMS PlGF plus Kryptor assay has a limit of detection of 3.6 pg/ml (measuring range 3.6–7000 pg/ml) and a limit of quantification of 6.9 pg/ml.[48] Reference ranges for each of the assays for singleton pregnancies with a normal outcome are provided in the product inserts,[54,55] and the company recommends that individual laboratories should validate these ranges or establish their own reference ranges prior to use. The company suggests a cut-off point of 85, based on a study of PE cases and controls that included women with singleton pregnancies and normal pregnancy outcome.[56]

Important subgroups

Four subgroups of women are potentially relevant to this diagnostic assessment, subject to data being available. These are women with chronic hypertension, pre-existing or gestational diabetes mellitus, renal conditions and/or an autoimmune condition. These subgroups are important because these conditions are associated with higher risk of PE.[13] PE may be difficult to diagnose in pregnant women with underlying chronic kidney disease or lupus,[57] and these conditions may be associated with more severe PE[58] or worse maternal, perinatal and neonatal outcomes.[59]

Current use of the diagnostic technologies in the NHS

The Triage PlGF test is used in three hospitals in England.[60] The other three tests are not currently used in the UK.[43,45,53]

Expected costs associated with the diagnostic technologies

Costs of the diagnostic technologies are presented in detail in *Chapter 5*. The cost of the Triage PlGF test includes £1000 for a cassette of 25 tests, £1400 for an Alere Triage MeterPro, £50 for Alere Triage PlGF Control L1 (Alere Inc., San Diego, CA, USA), £50 for Alere Triage PlGF Control L2 (Alere Inc., San Diego, CA, USA), £960 for an additional box of cassettes for quality control (24 per annum required), £800 for a Hettich EBA 20 centrifuge (Hettich Lab Technology North America, Beverly, MA, USA) and £259 for an annual service charge (payable from year 2).[60]

The cost of a Roche Diagnostics Elecsys sFlt-1 reagent kit (Roche Diagnostics Limited, Burgess Hill, UK) or a Roche Diagnostics Elecsys PlGF reagent kit (Roche Diagnostics Limited, Burgess Hill, UK) is £2861.47 per 100 tests, and the list price per Elecsys immunoassay sFlt-1 to PlGF ratio is £57.23.[61] Instruments and service costs are subject to local contracts and specifications of the service that extend over multiple assays. Costs for calibration and controls may apply based on usage and local protocols. List prices may be subject to local discounts. There are typically no upfront instrument or service costs for the customer as most agreements are on a 'reagent rental' basis for which only consumables are charged. A significant number of laboratories operate under managed services contracts that are value-added tax exempt.[61]

For the PerkinElmer test, the price of one test is €40 (£29.40) for < 1000 tests per year and €25 (£18.40) for ≥ 1000 tests per year. The price includes the cost of DELFIA Xpress instrument, instrument service and PlGF 1-2-3 kit.[43] For the BRAHMS Kryptor sFlt-1 to PlGF ratio test, costs were provided by the company as confidential information[53] and are not reported here.

Chapter 2 Definition of the decision problem

Decision problem

Tests that measure the concentration of PlGF or the sFlt-1 to PlGF ratio could have the potential to aid clinicians in diagnosing PE during the second half of pregnancy. However, the diagnostic accuracy, and the clinical effectiveness and cost-effectiveness of these tests is unclear. A systematic review and an economic evaluation are needed to answer the following questions:

- What is the clinical effectiveness and cost-effectiveness of the Triage PlGF test, Elecsys immunoassay sFlt-1 to PlGF ratio, DELFIA Xpress PlGF test and BRAHMS Kryptor sFlt-1 to PlGF ratio test in addition to clinical assessment for the diagnosis of PE in the second and third trimesters of pregnancy?
- What is the clinical effectiveness and cost-effectiveness of the Triage PlGF test, Elecsys immunoassay sFlt-1 to PlGF ratio, DELFIA Xpress PlGF test and BRAHMS Kryptor sFlt-1 to PlGF ratio test as a replacement for quantitative proteinuria tests in the diagnosis of PE in the second and third trimesters of pregnancy?

Population

The population of relevance to the decision problem is pregnant women, between gestation week 20 and gestation week 36^{+6} who, on the basis of screening tests and clinical symptoms (hypertension plus other signs or symptoms that may include proteinuria, haematological abnormalities, frontal headache, severe pain just below the ribs, vision problems, vomiting and/or severe swelling of the face or hands), are suspected of having PE.

There are four potential subgroups of women for this decision problem: those with chronic hypertension; those with pre-existing or gestational diabetes mellitus; those with renal conditions; and those with an autoimmune condition.

Index tests

According to the paradigm of diagnostic test accuracy assessment, the index test is the new test that is to be evaluated for its diagnostic accuracy when compared against an existing gold standard test (the reference standard). Four index tests are eligible for inclusion in the current diagnostic assessment. These are:

1. Triage PlGF test
2. Elecsys immunoassay sFlt-1 to PlGF ratio
3. DELFIA Xpress PlGF test
4. BRAHMS Kryptor sFlt-1 to PlGF ratio test.

As stated in the NICE scope, each of these tests could be employed:

- in conjunction with standard clinical assessment of suspected PE (i.e. as an add-on test to assessments of hypertension, proteinuria, and other clinical criteria for suspecting PE)
- in conjunction with standard clinical assessment excluding quantitative determination of proteinuria.

For assessing diagnostic accuracy these index tests should be compared against the reference standard, that is, standard clinical assessment of suspected PE. Assessment of the relative diagnostic accuracy of any pairwise comparisons among the four index tests is permissible, subject to the availability of relevant evidence (i.e. any direct head-to-head comparisons among pairs of index tests and/or indirect comparisons).

Reference standard

The gold standard reference for diagnosis of PE is clinical assessment, guided by a combination of the following clinical information:

- maternal hypertension (categorised as mild, moderate or severe)
- quantitative proteinuria test
- clinical symptoms suggestive of PE (e.g. headache, oedema or visual disturbances)
- fetal growth restriction.

Maternal hypertension and/or proteinuria with or without clinical symptoms may be sufficient to diagnose PE, or they may also occur in combination with fetal growth restriction and/or signs of biochemical or haematological impairment.

Outcomes

Diagnostic and prognostic test accuracy

The key test accuracy outcome measures are sensitivity and specificity, positive and negative likelihood ratios, positive predictive values (PPVs) and negative predictive values (NPVs), and the area under the receiver operator characteristic (ROC) curve. The working definitions of each of these outcomes are as follows:[62]

- Sensitivity: the rate of correct identification of people with the condition of interest. It is also known as the true-positive rate. A high sensitivity implies that a negative result rules out a condition.
- Specificity: also known as the true-negative rate, it indicates the rate of correct identification of people without the condition. A test with high specificity implies that a positive result confirms the condition.
- Likelihood ratios: a positive likelihood ratio is the ratio of the true-positive rate to the false-positive rate and is expressed as sensitivity/(100 – specificity), whereas a negative likelihood ratio is the ratio of the false-negative rate to the true-negative rate, expressed as (100 – sensitivity)/specificity. The positive likelihood ratio describes how many times more likely positive index test results are in women with PE than in those without the condition, and should be > 1 for the test to be informative. The negative likelihood ratio describes how many times more likely negative index test results are in women with PE than in those without the condition, and should be < 1 for the test to be informative.
- PPVs and NPVs: PPV is the probability of the condition of interest among people with a positive test result. NPV is the probability of not having the condition among people with a negative test result.
- Area under the ROC curve: this is derived from a plot of sensitivity (*y*-axis) against (1 – specificity) (*x*-axis), which shows the trade-off between the rates of true positives and false positives at different test thresholds (cut-off points) for determining a positive result. The area under the curve (AUC) summarises the entire ROC curve and represents the average value of sensitivity for all possible values of specificity. It can also be interpreted as the probability that a randomly chosen woman with PE is more likely to be diagnosed with PE than a randomly chosen woman without PE.

Other intermediate outcomes

- Time to test result.
- Test failure rate.
- Time to diagnosis.
- Proportion of women diagnosed with PE.
- Time to onset of PE and/or eclampsia.
- Proportion of women returned to less intensive follow-up.
- Length of inpatient hospital stay.
- Time to delivery.

Clinical outcomes

- Maternal morbidity and mortality.
- Fetal morbidity and mortality.
- Emergency admission for hypertensive disease.
- Health-related quality of life (HRQoL), including anxiety.

Outcomes for economic analysis

- Costs, including costs of the tests, training, hospitalisation, and birth with or without complications (considered from a NHS and Personal Social Services perspective).
- Cost-effectiveness of tests, expressed in terms of incremental cost per quality-adjusted life-years (QALYs).

Overall aims and objectives of the assessment

The aim of this diagnostic assessment is to evaluate the clinical effectiveness and cost-effectiveness of tests that could aid in the triage of women presenting with suspected PE between 20 weeks and 36^{+6} weeks of pregnancy who have received blood pressure assessment and qualitative (dipstick) assessment of proteinuria, by predicting whether or not PE can be ruled in or ruled out within specified time periods. Specific objectives are to determine, through a systematic review and economic evaluation, the clinical effectiveness and cost-effectiveness of the Triage PlGF test, Elecsys immunoassay measuring the sFlt-1 to PlGF ratio, DELFIA Xpress PlGF test and BRAHMS sFlt-1 to PlGF ratio:

- in addition to clinical assessment for the diagnosis of PE in the second and third trimesters of pregnancy
- as a replacement for quantitative proteinuria tests in the diagnosis of PE in the second and third trimesters of pregnancy.

The scope of these objectives is as defined by the eligibility criteria (population, intervention, comparators and outcomes) specified below (see *Chapter 3*, *Inclusion and exclusion criteria*).

Chapter 3 Methods for reviewing test accuracy

A review of the evidence for test accuracy was undertaken systematically following the general principles outlined in the Centre for Reviews and Dissemination (CRD)'s guidance; *Systematic Reviews: CRD's Guidance for Undertaking Reviews in Health Care*,[63] the *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy*[62,64] and the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA),[65] taking into consideration specific aspects of methodology that are relevant to the synthesis of evidence of test accuracy. All methods are based on those specified in the peer-reviewed project protocol.[66]

The project was informed by an advisory group of five independent clinical experts (see *Acknowledgements*). This included two obstetricians, one neonatologist, one midwife and one methodologist. The advisory group provided comments on draft versions of the protocol and the final report.

This report contains references to confidential information provided as part of the NICE appraisal process. Confidential information has been removed from the report, and this is clearly marked where it applies in the text, tables and figures. The influence of the excluded confidential information on the interpretation and conclusions is considered in *Chapter 6*, *Discussion*.

Identification of studies

A comprehensive search strategy for studies on the diagnostic accuracy of the four index tests was developed, tested and refined by an experienced information scientist (see *Appendix 1*). The search strategy aimed to identify studies on the diagnosis of PE, based on the prespecified inclusion and exclusion criteria (see *Inclusion and exclusion criteria*).

The following evidence sources were searched:

- General health and biomedical databases: MEDLINE (via Ovid); PREMEDLINE In-Process & Other Non-Indexed Citations; EMBASE; The Cochrane Library; Web of Science; Science Citation Index Expanded (SCI-EXPANDED); Conference Proceedings Citation Index – Science (CPCI-S); Database of Abstracts of Reviews of Effects (DARE) (accessed via CRD); and the CRD's Canadian and International Health Technology Assessment databases.
- Relevant conferences including those of the American Society of Hypertension; British Hypertension Society; British Maternal and Fetal Medicine Society; European Society of Hypertension; ISSHP; and the International Society for Prenatal Diagnosis International Conference on Prenatal Diagnosis and Therapy.
- Internet pages of relevant institutions and other organisations including those of the Royal College of Obstetricians and Gynaecologists; ACOG; International Society of Perinatal Obstetricians; Society for Maternal Fetal Medicine; Action on Pre-Eclampsia; Pre-Eclampsia Foundation; National Childbirth Trust; Cochrane Pregnancy and Childbirth Group; and Tommy's (funds research into pregnancy problems and provides parents with information).
- Grey literature and research in progress: UK Clinical Research Network Portfolio Database; the World Health Organization's International Clinical Trials Registry Platform; International Standard Randomised Controlled Trial Number (controlled and other trials); Clinical Trials.gov; and the UK Clinical Trials Gateway.

All databases were searched from 2000 (clinical experts advised that this was an appropriate start date, given that the technologies under comparison are relatively new) to March 2015, with searches updated in July 2015. Systematic reviews were retrieved only to check their reference lists for potentially relevant primary research studies.

All searches were limited to the English language. This was a pragmatic decision made when developing the review protocol because the current diagnostic assessment is specifically focused on clinical practice in England and Wales. Non-English-language studies would be unlikely to be generalisable to the current clinical setting because the management of women suspected of having PE varies by country.

Searches for evidence relevant to the economic evaluation were based on the search strategy and evidence sources as reported here for the review of test performance, with modifications when appropriate. These searches included economic evaluations, costs, resources and HRQoL (see *Chapter 5*, *Systematic review of economic studies*).

Inclusion and exclusion criteria

The eligibility criteria for the systematic review of test performance are:

- Study design: this review included primary diagnostic research, but was not limited to particular study designs. Instead, issues of methodological rigour relating to study design (specifically, risk of bias and applicability of the study findings) were evaluated during formal quality assessment (see *Critical appraisal*).
- Population: women with suspected PE between 20 weeks and 36^{+6} weeks of pregnancy who have received blood pressure assessment and qualitative assessment of proteinuria. In the present review, 'suspected PE' means that blood has been taken for one of the index tests, but a formal diagnosis of PE has not yet been made.
- Index tests: Triage PlGF, Elecsys immunoassay measuring the sFlt-1 to PlGF ratio, DELFIA Xpress PlGF and BRAHMS Kryptor sFlt-1 to PlGF ratio test, in conjunction with standard clinical assessment, or in conjunction with standard clinical assessment excluding quantitative determination of proteinuria.
- Reference standard: clinical assessment guided by maternal hypertension, proteinuria, symptoms suggestive of PE, and ultrasound fetal growth measurements. A combination of maternal hypertension and/or proteinuria, with or without clinical symptoms, may be sufficient to diagnose PE, or they may also occur in combination with fetal growth restriction and/or signs of biochemical or haematological impairment.
- Test performance outcomes: diagnostic and prognostic test accuracy (sensitivity, specificity, prevalence and related outcome measures) for PE.
- Intermediate measures: time to test result, test failure rate, time to diagnosis, proportion of women diagnosed with PE, time to onset of PE and/or eclampsia, proportion of women returned to less intensive follow-up, length of inpatient hospital stay and time to delivery.

Study selection

Studies were selected for inclusion through a two-stage process using the predefined and explicit criteria specified above (see *Inclusion and exclusion criteria*).

First, all titles and abstracts identified in the searches were screened independently by two reviewers to identify bibliographic records that met the inclusion criteria, using a standard study selection worksheet (see *Appendix 2*). The worksheet was pilot tested on 60 titles and abstracts by three reviewers (in three pairwise combinations of the reviewers with 20 abstracts each) to identify any ways that the worksheet could be improved to minimise errors. Only minor adjustments to the selection worksheet were deemed necessary and these were made before applying the worksheet to all the identified titles and abstracts.

Second, full-text articles were retrieved for those bibliographic records judged to be relevant or unclear at the title and abstract screening stage. If a study was reported in more than one article, all articles relating to the study were grouped together for assessment. Eligibility of each study was then assessed using the same study selection worksheet as applied to titles and abstracts by one reviewer and the decision was checked by a second reviewer.

At each step of the selection process, any disagreements were resolved by discussion among the two reviewers or, if necessary, by involving a third reviewer.

Data extraction

Data extraction was undertaken by one reviewer and checked by a second reviewer using a predesigned form. Any disagreements between reviewers were resolved by consensus or, if necessary, arbitration by a third reviewer. When test accuracy statistics were not reported in the primary studies, these were calculated, when possible, by the reviewers. To ensure ease of use, and to minimise errors, the data extraction form was first pilot tested on one of the included studies.[5] In the pilot test, two reviewers discussed whether or not any errors or disagreements identified could be solved by adjusting the wording or layout of the form. A final version of the form was then agreed and this was applied to all studies included in the review.

Critical appraisal

The methodological rigour of studies reporting diagnostic accuracy was assessed using the Cochrane adaptation[67] of the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) tool[68] (which focuses on methodological rigour rather than quality of reporting). For each of the included studies, judgements on study rigour were made by one reviewer using the QUADAS criteria and were checked by a second reviewer. Any disagreements between reviewers were resolved by consensus or, if necessary, by arbitration by a third reviewer. The QUADAS tool[68] asks 11 questions about the characteristics of the primary studies. These questions aim to identify potential threats to the validity of the study findings, and reflect 10 different types of bias that can be present in studies of test accuracy (*Table 7*).

TABLE 7 Types of bias possible in studies of the accuracy of biomarkers for PE

QUADAS question	Type of bias	Explanation
1	Spectrum	The study population is not representative of those who will receive the index test (PlGF or sFlt-1 to PlGF ratio) in clinical practice
2	Verification	The reference standard does not accurately diagnose PE (i.e. does not reflect usual clinical assessment)
3	Disease progression	The time interval between the index (biomarker) test and reference standard (usual clinical assessment) is long enough that the two tests may not have measured the same disease state
4, 5	Differential verification	Diagnosis is inaccurate because not all patients receive the same reference standard
6	Incorporation	The index (biomarker) test is not independent of the reference standard (e.g. it may be one of several tests used as the reference standard)
7	Diagnostic review	The index (biomarker) test result influences interpretation of the reference standard result
8	Test review	The reference standard result influences interpretation of the index (biomarker) test result
9	Clinical review	The information used when interpreting the index (biomarker) test does not reflect that likely to be available in clinical practice
10	Test classification	Incorrect inclusion or exclusion from the analysis of index test results classified as uninterpretable, intermediate or indeterminate may systematically influence sensitivity or specificity
11	Attrition	Exclusion of patients or test results from analysis may systematically influence sensitivity or specificity if the reason for exclusion is linked to test performance, or if criteria for permitting exclusions differ between biomarker tests, especially if the magnitude of attrition is unbalanced across the test methods

In addition to the QUADAS assessment, the External Assessment Group (EAG) considered the generalisability of the studies, that is, their probable relevance to clinical practice, based on the reported eligibility criteria and population characteristics for each study.

Data synthesis

Diagnostic and prognostic test accuracy outcomes, that is, the sensitivity and specificity of tests for predicting PE or PE-related delivery, were synthesised through a structured narrative review with tabulation of results. We planned (subject to the availability and suitability of the primary data) to conduct one or more meta-analyses of data on test sensitivity and specificity, in order to improve the precision of any estimates of test accuracy. The appropriateness of meta-analysis was determined by critical appraisal of the primary studies during the critical appraisal step (see *Critical appraisal*) together with consideration of the clinical heterogeneity of the studies (i.e. heterogeneity of the study populations and their generalisability to UK clinical practice). To account for correlation between sensitivity and specificity, and their dependence on the prevalence of PE, the planned pooling of sensitivity and specificity outcomes was based on appropriate hierarchical random-effects models [using statistical software such as WinBUGS (MRC Biostatistics Unit, Cambridge, UK) or R (The R Foundation for Statistical Computing, Vienna, Austria)].

The data synthesis approach followed good practice, as recommended by *Systematic Reviews: CRD's Guidance for Undertaking Reviews in Health Care* (see *Chapter 2*, *Systematic reviews of clinical tests*),[63] the *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy*[67,68] and the NICE *Diagnostics Assessment Programme Manual*.[69]

For all other outcomes that inform the economic model, the method of data synthesis is described in *Chapter 5*.

Chapter 4 Assessment of diagnostic test performance

This chapter presents the quantity of research available, including the number of studies, their designs, participant characteristics, and the characteristics of the index tests and reference standards that they compared (see *Quantity and quality of research available*). The included studies are critically appraised, including their risks of bias (see *Critical appraisal of the included studies*), then the assessment of diagnostic accuracy, which takes into consideration the available evidence on diagnostic outcomes as well as any threats to validity highlighted in the preceding sections, is presented (see *Assessment of test accuracy*).

Quantity and quality of research available

A variety of documents was provided, via NICE, by the four companies marketing the index tests (Alere, Roche Diagnostics, PerkinElmer and Thermo Fisher Scientific). As mentioned in the methods (see *Chapter 3*, *Identification of studies*), these documents, which included published and unpublished manuscripts, conference posters, test manuals, product inserts and formal company submissions, were included at the full-text screening stage to check for their relevance to the systematic review of test accuracy.

The selection of evidence for the systematic review of test accuracy is summarised in *Figure 2*. Searches yielded 1972 unique bibliographic records, which were screened using the study selection worksheet

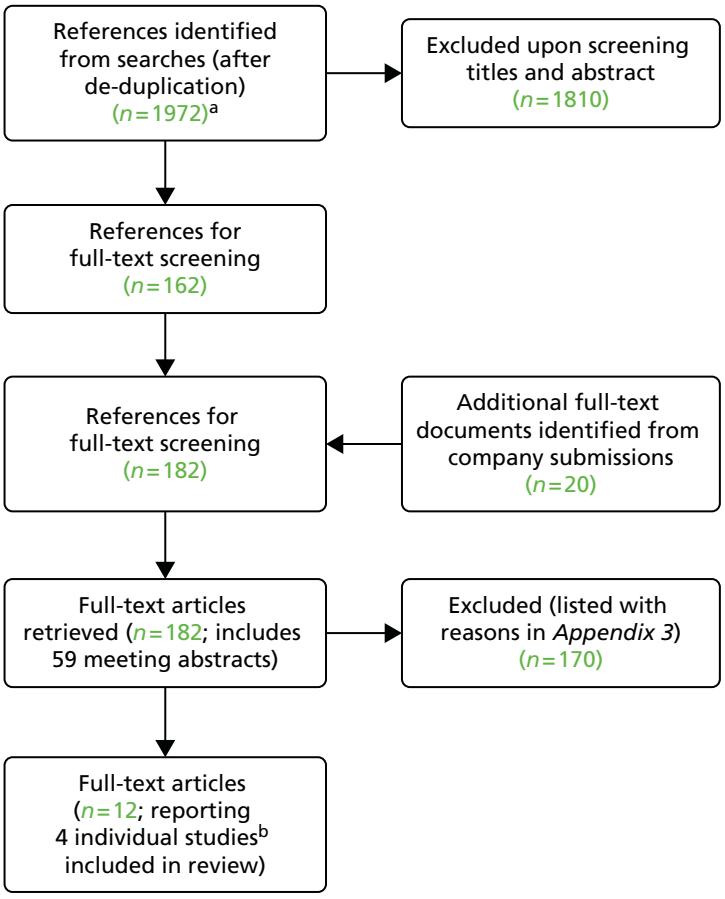

FIGURE 2 Flow chart for the identification of test accuracy studies. a, Of which 113 studies were identified in updated searches, July 2015; b, one study was confidential (see *Characteristics of the included studies*).

(see *Appendix 2*). Of these, 1810 records were excluded on title and/or abstract because they did not meet all of the inclusion criteria. Full-text versions of bibliographic records were obtained for 162 records that met all inclusion criteria or provided insufficient information to make a judgement on eligibility. An additional 20 documents were identified from submissions provided by NICE from the companies that market the PlGF and sFlt-1/PlGF tests. Screening of these 182 full-text documents revealed that 170 were not relevant, primarily because they did not include a 'suspected PE' group. Reasons for excluding the full-text documents are listed in *Appendix 3*. The remaining 12 full-text documents (including full versions of five meeting abstracts, i.e. the most comprehensive versions of the abstracts available from the meeting proceedings) met all the eligibility criteria and are included in the current systematic review.

Characteristics of the included studies

The 12 full-text documents included in the systematic review of test accuracy reported on four unique primary research studies. Two of these studies used the Triage PlGF test and two studies used the Elecsys sFlt-1 to PlGF ratio for predicting PE. None of the studies that met the systematic review inclusion criteria had employed the PerkinElmer DELFIA Xpress PlGF test or the BRAHMS Kryptor sFlt-1 to PlGF ratio for assessment of women with suspected PE.

Studies on the Triage PlGF test included in the review are:

- The PETRA study, reported in an unpublished, confidential company submission to NICE provided by Alere (Sibai[6]). Confidential data from PETRA are not presented in the current report; however, they were available for consideration by the NICE Diagnostics Assessment Committee and the EAG.
- The PELICAN study, reported by Chappell *et al.* in an academic journal paper,[5] three meeting abstracts,[70–72] and the Triage PlGF test product insert.[40]

Studies on the Elecsys sFlt-1 to PlGF ratio test included in the review are:

- the PROGNOSIS study, reported in a company submission to NICE provided by Roche Diagnostics,[61] an unpublished academic manuscript[73] that was subsequently published by Roche Diagnostics (Zeisler *et al.*[51]), a meeting abstract (Zeisler *et al.*[74]) and the Elecsys sFlt-1 to PlGF ratio test product insert[46]
- a study by Álvarez-Fernández *et al.* reported in an academic journal paper[75] and a meeting abstract[76] (the EAG contacted the authors of this study to confirm that it met the eligibility criteria).

For brevity, studies are cited in this report by their primary reference only, unless the reference is intentionally to a different document.

General characteristics of the study participants

The three published studies varied in their designs and locations (*Table 8*).[5,51,75] Both the PELICAN study on the Triage PlGF test[5] and the PROGNOSIS study on the Elecsys sFlt-1 to PlGF ratio test[51] included women in the UK, although in the PROGNOSIS study only 1 of the 30 study centres was in the UK, with 76 UK participants represented among 1050 total participants. The remaining study by Álvarez-Fernández *et al.*[75] was not conducted in the UK. Two of the studies were of a single-cohort design, whereas the PROGNOSIS study[51] had two cohorts: a 'development' cohort to derive a cut-off value-based prediction model for the sFlt-1 to PlGF ratio ($n = 500$) and a 'validation' cohort to test the model ($n = 550$). Two of the studies recruited patients prospectively[5,51] and one retrospectively.[75] The major studies, PELICAN[5] and PROGNOSIS,[51] had relatively large sample sizes for those gestational age groups that are relevant to the current diagnostic assessment, analysing 424–1050 patients. All the studies covered the majority of the period of pregnancy specified in the NICE scope (20^{+0} to 36^{+6} weeks, i.e. the second and third trimesters), although only the PELICAN study[5] covered the whole of this period. However, only the PROGNOSIS study (Zeisler *et al.*[51]) reported how gestational age was determined (calculated from the last menstrual period or first sonography date and recorded at visit 1). Groups of women who presented with fetuses at gestational ages exceeding those specified in the NICE scope (i.e. after 36^{+6} weeks) were reported in two studies;[5,75] these groups are not considered in the current review.

TABLE 8 Overview of the included published studies

Study[a]	Location	Design	Recruitment	Total women analysed	Timing of tests during gestation
Triage PlGF test					
Chappell *et al.*;[5] PELICAN study	UK and Ireland	Prospective single cohort	From seven centres; not reported whether or not consecutive selection. Recruited, $n = 649$; analysed, $n = 625$	424	20^{+0}–34^{+6} weeks ($n = 287$) and 35^{+0}–36^{+6} weeks ($n = 137$); women presenting at 37^{+0}–40^{+6} weeks excluded (outside of scope)
Elecsys sFlt-1 to PlGF ratio test					
Zeisler *et al.*;[51] PROGNOSIS study	14 countries, including UK	Prospective, two cohorts (development, validation)	From 30 centres (1 in UK); not reported whether or not consecutive selection. Enrolled, $n = 1273$; analysed, $n = 1050$	1050	24^{+0}–36^{+6} weeks
Álvarez-Fernández *et al.*[75]	Spain	Retrospective single cohort	From one centre; not reported whether or not consecutive selection. Enrolled, $n = 281$; analysed, $n = 257$	62	20–34 weeks ($n = 62$); women presenting at 34–41 weeks of gestation excluded (outside of scope)

a Excluding the confidential PETRA study (Sibai[6]). Data from PETRA were available for consideration by the EAG and the NICE Diagnostics Assessment Committee.

Key characteristics of the participants in the three published studies are shown in *Table 9*. Women were aged in their early thirties. Median BMI ranged from 24.9 kg/m^2 in the PROGNOSIS study[51] to 31.2 kg/m^2 in the study by Álvarez-Fernández *et al.*[75] Two studies reported the proportion of women who were nulliparous, and this ranged from 43% in the PELICAN study[5] to 84% in the study by Álvarez-Fernández *et al.*[75] One study analysed only singleton pregnancies (PROGNOSIS[51]), while the proportion of singleton pregnancies in the remaining studies ranged from 74% to 93% in the study by Álvarez-Fernández *et al.*[75] and up to 90% to 96% in the PELICAN study.[5] Ethnicity of the participants was reported only in the PELICAN and PROGNOSIS studies, and all three studies were dominated by white or Caucasian participants. None of the published studies reported the socioeconomic status of their participants (i.e. their educational achievement, employment status, income or dependency on any community services). When reported, in only two studies, the majority of participants (> 70%) were not smokers.

Prognostic characteristics of the study participants

Population characteristics relevant to the prognosis of PE were reported inconsistently in the published studies and, therefore, it is difficult to present a comparison of the prognostic characteristics across the studies. Many characteristics of potential prognostic relevance were only reported by one or two studies (*Table 10*).

New-onset proteinuria was reported in two studies. The proportion of women with new-onset proteinuria was highest in the PELICAN study[5] (56–62%) and lowest in the PROGNOSIS study[51] (30–44%). Álvarez-Fernández *et al.*[75] reported that 73% of the PE group and 3% of those without PE diagnosis had proteinuria, but did not state whether or not this was new onset. Although all three studies reported hypertension, only Álvarez-Fernández *et al.*[75] reported the proportion of patients who had chronic hypertension (8–9%). Cases of new-onset hypertension were reported only in the PELICAN study[5] (54–67%), although the PROGNOSIS study[51] did report cases of elevated blood pressure, without specifying whether or not this meant hypertension; elevated blood pressure was found to be more common in women who developed PE (42–53%) than in those without PE (18–40%). Worsening of

TABLE 9 Overview of participants' characteristics in the included published studies

	Studies on the Triage PlGF test[a]		Studies on the Elecsys test		
	PELICAN study: Chappell *et al.*[5]		PROGNOSIS study: Roche Diagnostics; Zeisler *et al.*[51]		[b]Álvarez-Fernández *et al.*[75]
Population characteristic	Weeks 20–34^{+6}	Weeks 35–36^{+6}	Weeks 24–33^{+6}	Weeks 34–36^{+6}	Weeks 20–41
Age (years), median (IQR)	31.9 (27.0–35.9)	32.4 (27.5–35.4)	PE group: 32 (29–38) No-PE group: 32 (27–36)	PE group: 31 (27–36) No-PE group: 30 (27–35)	PE group: 35 (30–37) No-PE group: 33 (31–37)
BMI (kg/m^2), median (IQR)	28.6 (24.2–33.6)	28.6 (24.4–32.7)	PE group: 24.9 (21.5–30.0)[c] No-PE group: 26.7 (22.6–31.5)[c]	PE group: 26.2 (22.5–30.7)[c] No-PE group: 25.9 (22.2–31.2)[c]	PE group: 31.2 (27.6–35.6) No-PE group: 30.5 (27.6–34.4)
Nulliparous, rounded %	43 [57][d]	44	Not reported	Not reported	PE group: 84 No-PE group: 77
Singleton pregnancy, rounded %	96	90	100% of women in the primary analysis	100% of women in the primary analysis	PE group: 74[e] No-PE group: 93[e]
Ethnicity, rounded %	White: 65 Black: 24 Asian: 7 Other: 4	White: 64 Black: 20 Asian: 9 Other: 7	PE group: • Caucasian: 86 • Asian: 6 • Black: 5 • Other: 3 No-PE group: • Caucasian: 81 • Asian: 5 • Black: 6 • Other: 8	PE group: • Caucasian: 77 • Asian: 9 • Black: 9 • Other: 5 No-PE group: • Caucasian: 85 • Asian: 4 • Black: 3 • Other: 8	Not reported

	Studies on the Triage PlGF test[a]		Studies on the Elecsys test		
	PELICAN study: Chappell *et al.*[5]		PROGNOSIS study: Roche Diagnostics; Zeisler *et al.*[51]		[b]Álvarez-Fernández *et al.*[75]
Population characteristic	Weeks 20–34^{+6}	Weeks 35–36^{+6}	Weeks 24–33^{+6}	Weeks 34–36^{+6}	Weeks 20–41
Socioeconomic status	Not reported	Not reported	Not reported	Not reported	Not reported
Smoking status, rounded %	Never smoked: 73	Never smoked: 76	Past smoker: • PE group: 21 • No-PE group: 21 Current smoker: • PE group: 17 • No-PE group: 16	Past smoker: • PE group: 20 • No-PE group: 20 Current smoker: • PE group: 8 • No-PE group: 14	Not reported

IQR, interquartile range.
a Excluding the confidential PETRA study (Sibai[6]). Data from PETRA were available for consideration by the EAG and the NICE Diagnostics Assessment Committee.
b Baseline data are reported for gestation weeks 20–41 (timing partly outside the NICE scope).
c Pre-pregnancy BMI.
d Value in product insert (in square brackets) exceeds the range reported in the publication.
e Not reported but deduced by reviewer from % of multiple gestations reported.

TABLE 10 Baseline prognostic characteristics of the included published studies

	Studies on the Triage PlGF test[a]		Studies on the Elecsys test		
	PELICAN study: Chappell *et al.*[5]		PROGNOSIS study: Roche Diagnostics; Zeisler *et al.*[51]		[b]Álvarez-Fernández *et al.*[75]
Population characteristic	Weeks 20–34^{+6}	Weeks 35–36^{+6}	Weeks 24–33^{+6}	Weeks 34–36^{+6}	Weeks 20–41
New-onset or worsening proteinuria, rounded %	New onset, dipstick: 56	New onset, dipstick: 62	New onset: • PE group: 39 • No-PE group: 38 Worsening: • PE group: 3 • No-PE group: 1	New onset: • PE group: 44 • No-PE group: 30 Worsening: • PE group: 1 • No-PE group: 2	Not specified as new-onset (missing data $n = 11/257$): • PE group: 73 • No-PE group: 3
Chronic hypertension, rounded %	Not reported	Not reported	Not reported	Not reported	PE group: 8 No-PE group: 9
New-onset hypertension, rounded %	54	67	New-onset elevated BP: • PE group: 42 • No-PE group: 18	New-onset elevated BP: • PE group: 53 • No-PE group: 40	Not reported
Worsening hypertension, rounded %	20	15	• PE group: 24 • No-PE group: 14	• PE group: 11 • No-PE group: 12	Not reported
Persistent epigastric/RUQ pain, rounded %	Persistent epigastric/ RUQ pain: 6	Persistent epigastric/ RUQ pain: 6	Epigastric pain: • PE group: 6 • No-PE group: 7	Epigastric pain: • PE group: 8 • No-PE group: 10	Not reported
Headaches, rounded %	29	32	• PE group: 32 • No-PE group: 28	• PE group: 35 • No-PE group: 32	Not reported
Abnormal fetal growth, rounded %	Suspected fetal growth restriction: 9	Suspected fetal growth restriction: 3	IUGR: • PE group: 23 • No-PE group: 15	IUGR: • PE group: 6 • No-PE group: 16	IUGR: • PE group: 14 • No-PE group: 3

BP, blood pressure; RUQ, right upper quadrant.
a Excluding the confidential PETRA study (Sibai[6]). Data from PETRA were available for consideration by the EAG and the NICE Diagnostics Assessment Committee.
b Baseline data are reported for gestation weeks 20–41 (timing partly outside the NICE scope).

existing hypertension was also reported only in the PELICAN[5] (15–20%) and PROGNOSIS[51] (11–24% in women who developed PE; 12–14% in those who did not) studies. The same two studies reported the proportions of women who had epigastric or upper right quadrant pain or headache, although they differed in whether or not the epigastric pain was stated as being persistent. Overall, the proportion of women with any type of epigastric or upper right quadrant pain in these studies was in the range of 6–10% and the proportion of patients with headache was in the range 28–35%. All three published studies reported the proportion of women who had (suspected) abnormal fetal growth, which was described as suspected fetal growth restriction (small for gestational age, < 10th customised birthweight percentile) in the PELICAN study[5] (3–9%) and as IUGR in the PROGNOSIS study[51] (6–23%) and the study by Álvarez-Fernández *et al.*[75] (3–14%). However, the PROGNOSIS study[51] and Álvarez-Fernández *et al.*[75] defined IUGR differently. The PROGNOSIS study[51] defined IUGR as estimated fetal weight or abdominal circumference < 5th percentile, presence of pathological process that inhibits expression of normal intrinsic growth potential on at least one occasion after gestational week 22, amniotic fluid index < 10th percentile or pulsatility index > 95th percentile, serial ultrasonography growth curve anomalies or serial growth curve anomalies based on local manual measurement. Álvarez-Fernández *et al.*[75] defined IUGR as fetal weight < 3rd percentile for gestational age or between the 3rd and 10th percentiles with abnormal uterine arteries or altered blood flow on ultrasound.

Four potential subgroups of women were identified by NICE for this decision problem: those with chronic hypertension, gestational diabetes mellitus, renal impairment and autoimmune conditions. Although some of the included primary studies collected data on these baseline characteristics, there were no relevant subgroup analyses that could contribute to the present diagnostic assessment.

The published studies included in the review differed in the amount of detail they provided about why PE was suspected (*Table 11*). The most detailed information was presented for the PROGNOSIS study.[51]

TABLE 11 Reasons for suspecting PE in the included published studies

Study	Reasons for suspecting PE
PELICAN study: Chappell *et al.*[5]	Stated that reasons included (so may not have been limited to) headache, visual disturbances, epigastric or right upper quadrant pain, hypertension, dipstick proteinuria or suspected fetal growth restriction
PROGNOSIS study: Roche Diagnostics; Zeisler *et al.*[51] (reported in study protocol: Hund *et al.*[77])	New-onset raised BP (did not need to be defined hypertension ≥ 140 mmHg systolic and/or ≥ 90 mmHg diastolic), aggravation of pre-existing hypertension, new-onset proteinuria (did not need to be defined proteinuria – any protein in urine was sufficient), aggravation of pre-existing proteinuria, or one or more other reason from the following: • epigastric pain • excessive oedema/severe swelling (face, hands, feet) • headache • visual disturbance • sudden weight gain (> 1 kg/week in third trimester) • low platelets • elevated liver transaminases • (suspected) IUGR[a] • abnormal uterine perfusion detected by Doppler sonography, with mean pulsatility index > 95th percentile in the second trimester • bilateral uterine artery notching
Álvarez-Fernández *et al.*[75]	High BP, proteinuria, abnormal uterine artery Doppler (definition not reported), headache not responding to analgesic, visual symptoms (blurry vision or flashing lights), and/or severe oedema affecting hands, feet or face

BP, blood pressure.
a IUGR defined as estimated fetal weight or abdominal circumference < 5th percentile, presence of pathological process demonstrated on at least one occasion after gestational week 22 by amniotic fluid index (< 10th percentile) or pulsatility index > 95th percentile, serial ultrasonography growth curve anomalies, serial growth curve anomalies (manual measurement).

Critical appraisal of the included studies

Risks of bias

The results of the EAG's critical appraisal of the included studies using QUADAS criteria are shown in *Table 12*. At the time of preparing the present report, the PETRA study (Sibai[6]) was confidential. The EAG critically appraised the PETRA study and the results of the appraisal were made available to the NICE Diagnostics Assessment Committee. However, owing to the data confidentiality, the EAG's critical appraisal of the PETRA study is not included in *Table 12* or discussed below. The impact of the PETRA study on the overall conclusions of the current diagnostic assessment is considered in *Chapter 6*, *Discussion*.

'Yes' answers to QUADAS questions 1 to 9 imply low risk of bias for each of the types of bias being assessed (for explanations of the bias types addressed by each question see *Table 7*). 'Yes' answers to QUADAS questions 10 and 11 reflect adequacy of reporting and require further information (see the paragraphs following) in order for the risks of bias to be assessed. Explanations for the judgements reached by the EAG are provided in the data extraction form for each study (see *Appendix 4* for an example data extraction form for review of test accuracy).

TABLE 12 Overview of QUADAS assessments

	Test		
	Triage PlGF[a]	Elecsys	
QUADAS question	PELICAN study: Chappell *et al.*[5]	PROGNOSIS study: Roche Diagnostics;[61] Zeisler *et al.*[51]	Álvarez-Fernández *et al.*[75]
1. Was the spectrum of patients representative of the patients who will receive the test in practice?	Yes	Yes	Yes
2. Is the reference standard likely to classify the target condition correctly?	Yes	Yes	Yes
3. Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes	Yes	Yes
4. Did the whole sample, or a random selection of the sample, receive verification using the intended reference standard?	Yes	Yes	Yes
5. Did patients receive the same reference standard irrespective of the index test result?	Yes	Yes	Yes
6. Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes	Yes	Yes
7. Were the reference standard results interpreted without knowledge of the results of the index test?	Yes	Yes	Yes
8. Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	Yes	Yes
9. Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	No	No	No
10. Were uninterpretable/intermediate test results reported?	Yes	Yes	No
11. Were withdrawals from the study explained?	Yes	Yes	Yes

a Excluding results of the EAG's critical appraisal of the confidential PETRA study (Sibai[6]). Data were available for consideration by the NICE Diagnostics Assessment Committee.

As shown in *Table 12*, all three of the published studies were rated favourably on QUADAS questions 1 to 7, suggesting that the studies would have low risk of spectrum bias, verification bias, disease progression bias, differential verification bias, incorporation bias or diagnostic review bias.

QUADAS question 8 assesses whether or not studies could be at risk of test review bias, that is, when the reference standard result influences interpretation of the index (biomarker) test result. The EAG judged all three studies to be at low risk of this type of bias.

QUADAS question 9 assesses whether or not studies could be at risk of clinical review bias, that is, when the information used when interpreting the index (biomarker) test does not reflect that which would be available in clinical practice. As the PlGF and sFlt-1 to PlGF ratio tests are intended (in this diagnostic assessment) to be add-on tests to standard clinical assessment, it is implicit that in order to make a diagnosis they would be interpreted in conjunction with other clinical information. The EAG judged that in all three studies the interpretation of the biomarker test would not reflect clinical practice, as diagnosis of PE was based solely on whether test values were above or below the cut-off point; in clinical practice such test values would probably be interpreted in conjunction with information about hypertension, proteinuria and/or other clinical signs or symptoms.

QUADAS question 10 assesses whether or not studies could be at risk of test classification bias, that is, when inclusion or exclusion of results classified as intermediate, indeterminate or uninterpretable may systematically influence test accuracy. Only two of the three studies mentioned intermediate, indeterminate or uninterpretable test results, but these were reported differently in each study. In the PELICAN study,[5] 11.1% of women changed PlGF status from low to very low and 3.5% changed from low to normal when a duplicate run of the test was performed. In the PROGNOSIS study,[51] it was stated only that the visit 1 sample was not available or not usable from 34 participants.

QUADAS question 11 assesses whether or not studies could be at risk of attrition bias. In studies of test accuracy, attrition might lead to bias if it differentially impacts on the numbers of true-positive, true-negative, false-positive or false-negative test results. The numbers of exclusions were reported in all three studies. In the PELICAN study,[5] 24 participants were excluded out of 649 enrolled (3.7%) because of a lack of enrolment samples, missing sample barcodes or no outcome data being available (reasons were not stated). The risk of attrition bias in the PELICAN study is unclear because it is not known if the 3.7% of women with missing data would have been equally distributed among the true-positive, false-positive, true-negative and false-negative groups. In the PROGNOSIS study,[51] a relatively large number of women [223 out of 1273 enrolled (17.5%)] were excluded. Reasons given were withdrawal of informed consent (0.9%), failure to meet the inclusion criteria (3.8%), multiple gestation pregnancy (6.1%), lack of an available/usable visit 1 sample (2.7%) and loss to follow-up as they had no PE during the first four visits but no visit after 28 days to confirm lack of PE (4.0%). The risk of attrition bias in the PROGNOSIS study[51] is unclear because it is not known if the 6.7% of women who either lacked a usable visit 1 sample or lacked follow-up after 28 days would have been equally distributed among the true-positive, false-positive, true-negative and false-negative groups. In the study by Álvarez-Fernández *et al.*,[75] 19 out of 281 enrolled women (6.8%) were excluded because of lack of delivery data (4.3%), suspicion of PE without urinalysis (1.4%), or because PE was diagnosed before presentation at triage (1.1%). As with the other studies, the risk of attrition bias is unclear for Álvarez-Fernández *et al.*,[75] as it is not known if the 4.3% of women with missing data would have been equally distributed among the true-positive, false-positive, true-negative and false-negative groups.

Generalisability

The assessment of study generalisability (often referred to as applicability) draws together information on the studies' designs and their participants' characteristics (reported in *Tables 8–11*) and how PE was defined in each study. Some aspects of generalisability are also captured in the QUADAS criteria reported, for example 'patient spectrum bias', which is assessed by QUADAS question 1, and refers to whether or not the study participants would be representative of those likely to be seen in clinical practice.

The studies varied in the amount of information they provided about their participants, and in some cases generalisability is unclear. For example, Álvarez-Fernández *et al.*[75] did not report ethnicity, and none of the published studies reported socioeconomic status.

Only two of the studies included any centres in the UK (PELICAN[5] and PROGNOSIS[51]). A factor that might influence the generalisability of study findings is whether or not management of PE is expectant or based mainly on PE diagnostic criteria. According to the Alere company submission,[6] a more expectant management approach appears to be used in the UK than in the USA (although no source of evidence for this was cited). However, the EAG is not aware of any data that would enable different countries or health systems to be classified according to the degree of expectant management of PE that takes place. As such, it is unclear whether or not the geographical location of a study alone would influence generalisability of its test accuracy outcomes to clinical practice in England and Wales. Most likely to be more important is how PE is defined in each study (see *Definition of pre-eclampsia*).

Population age

The range of median age of women included in the three published studies[5,51,75] was relatively narrow, from 31 to 35 years.

Pregnancy type

Two[5,51] of the three studies focused on women with a singleton pregnancy. The PROGNOSIS study[51] included both single and twin pregnancies, but focused analyses of primary outcomes on the singleton pregnancies. In the PELICAN study,[5] ≥ 90% of the women included had a singleton pregnancy. The results of these studies would probably underestimate the risk of PE in twin pregnancies, as the incidence of PE is higher in twin than in singleton pregnancies.[78]

History of pre-eclampsia

Of the three studies,[5,51,75] only the PELICAN study[5] reported whether or not women had previously experienced PE (the percentages were 20% of women who presented in early gestation and 12% of those who presented in late gestation).

Comorbidities

The studies were inconsistent as to whether or not they included or excluded specific comorbidities. For example, Álvarez-Fernández *et al.*[75] explicitly excluded patients with antiphospholipid syndrome and systemic lupus erythematosus; the PELICAN study[5] explicitly included women with antiphospholipid syndrome or systemic lupus erythematosus, as well as pre-gestational diabetes mellitus or renal disease (Chappell *et al.*[5] stated that there were minimal exclusion criteria in the PELICAN study so as to maximise generalisability), while the PROGNOSIS study[51] did not explicitly mention comorbidities among the inclusion or exclusion criteria.

Ethnicity

One study did not report participants' ethnicity (Álvarez-Fernández *et al.*[75]). The other studies included mostly white or Caucasian participants (54–86%). The PELICAN study[5] was a UK study, but it is unclear whether or not the ethnic mix of participants is fully representative of women attending antenatal clinics in the UK. In the PROGNOSIS study,[51] the proportion of black participants was lower than in the PELICAN study (3–9% compared with 20–24%, respectively), and it is possible that the PROGNOSIS study is less representative than the PELICAN study of a UK population. Although the impact of ethnicity on risk of PE is not fully clear, there appears to be consistent evidence that Asian parents are at lower risk than European parents,[79–81] and that the ethnicity of the father as well as the mother may influence PE risk.[79]

Body mass index

Median BMI was in the range 24.9–31.2 kg/m^2 in the three studies.[5,51,75] The risk of PE increases as BMI increases,[82] although it is unclear whether or not this would influence the accuracy of tests for PE in these studies. Only the PROGNOSIS study[51] reported BMI for women who developed PE and those who did not (24.9–26.2 kg/m^2 for those with PE and 25.9–26.7 kg/m^2 for those without PE).

Socioeconomic status

None of the three studies reported any aspects of the socioeconomic status of their participants (e.g. their educational attainment, employment status or family status) and so it is unclear if any particular socioeconomic groups might have been over-represented or under-represented in any of the studies.

Definition of pre-eclampsia

All three studies reported their definitions of PE, although Álvarez-Fernández *et al.*[75] did not specify the source of their definitions.

The PELICAN study[5] defined hypertensive disorders according to the ACOG practice bulletin, 2002.[14] The 'overall' definition of PE included PE, that was defined similarly to the NICE criteria (cf. ACOG and NICE criteria in *Table 1*), but also superimposed PE and atypical PE, neither of which is explicitly defined by NICE. Thus, the overall definition of PE applied in the PELICAN study is broader than the NICE definition:

- Superimposed PE (ACOG definition[14]): in addition to chronic hypertension plus proteinuria, superimposed PE includes a sudden increase in proteinuria if already present in early gestation, a sudden increase in hypertension, the development of HELLP syndrome, or physical symptoms without proteinuria.
- Atypical PE (ISSHP[15] definition): gestational hypertension without proteinuria, but with other multiorgan involvement or fetal growth restriction.

The PROGNOSIS study[51] also used ACOG (2002[14]) definitions for PE and superimposed PE, thus including broader criteria for superimposed PE than in the NICE definition. The outcome 'rule out or rule in PE' in this study included eclampsia (although no patients experienced this) and HELLP syndrome as well as PE.[51] HELLP syndrome is defined in the same way by NICE and ACOG (see *Table 1*).

Álvarez-Fernández *et al.*[75] used a simple definition of PE that accords with the NICE definition, but goes beyond the NICE definition by also including pre-existing proteinuria with superimposed PE, which consists of a sudden increase in blood pressure and proteinuria.

Population recruitment

As shown above (see *Table 8*), two of the three studies (PELICAN[5] and PROGNOSIS[51]) recruited participants prospectively, while Álvarez-Fernández *et al.*[75] recruited participants retrospectively. None of the studies stated whether or not the recruitment of women was consecutive (i.e. in the order in which they first presented to clinicians). Retrospective studies, and those in which participant recruitment is not consecutive, may be at increased risk of selection bias and could have limited generalisability if preferential selection of certain participants takes place during recruitment. However, it is difficult in this diagnostic assessment to draw any firm conclusions about the adequacy of the recruitment processes and how these might affect studies' generalisability based on the limited information reported by the studies.

Assessment of test accuracy

Relevance of the evidence to the NICE decision problem

The decision problem (see *Chapter 2*) specifies that the angiogenic biomarker tests should be assessed (1) in addition to routine clinical assessment (which includes monitoring of blood pressure and proteinuria) and (2) as an alternative to quantitative proteinuria testing. The four included studies employed the biomarker tests in addition to routine clinical assessment. Therefore, the available evidence only informs the first part of the decision problem; no evidence is available for assessing the accuracy of these biomarker tests as alternatives to proteinuria testing.

Each of the four included studies assessed only one of the tests specified in the NICE scope. Therefore, there are no direct head-to-head comparisons available in relevant populations between any of the four diagnostic tests specified in the NICE scope. Furthermore, evidence from the included studies is limited to

only two of the four biomarker tests specified in the NICE scope, that is the Triage PlGF test and the Elecsys sFlt-1 to PlGF ratio. No studies were identified that compared either the PerkinElmer DELFIA Xpress PlGF test or the BRAHMS Kryptor sFlt-1 to PlGF ratio to usual clinical practice.

The three published studies reporting the Alere and Roche Diagnostics tests differ in their participant characteristics and the test accuracy outcomes they report, for example providing outcomes for different test cut-off points and different gestational timings. For this reason, the EAG considered it inappropriate to conduct meta-analyses to quantitatively combine test accuracy outcomes across studies to improve their precision. Instead, the results of the studies are presented narratively, with tabulation of the test accuracy outcomes [sensitivity and specificity, with confidence intervals (CIs)] together with their source data (true and false positives and true and false negatives), prevalence of PE, and supporting statistics (PPV, NPV, positive and negative likelihood ratios and areas under ROC curves). These outcomes capture the best-available diagnostic test accuracy values that could be employed in our economic model (see *Chapter 5*).

The diagnostic test accuracy outcomes are reported below for each diagnostic test and are then discussed in the context of any key issues of bias or generalisability that were identified at the critical appraisal step above (see *Critical appraisal of the included studies*).

Triage PlGF test (Alere)

Results for this test were available to the NICE Diagnostics Assessment Committee and the EAG from the PELICAN study[5] ($n = 287$ for gestation weeks 20^{+0} to 34^{+6}; $n = 137$ for weeks 35^{+0} to 36^{+6}) and from the PETRA study.[6] As mentioned, at the time of preparing the present report, the PETRA study (Sibai[6]) was confidential and results from this study are therefore not reported here. The impact of the PETRA study on the overall conclusions of the current diagnostic assessment is considered in *Chapter 6*.

The most relevant outcome to the decision problem reported in the PELICAN study is prediction of PE requiring delivery within 14 days of testing (*Tables 13* and *14* and *Figure 3*). Some of these data are used to inform the EAG economic model. Other related test accuracy outcomes (which do not directly inform the economic model but were considered by the NICE Diagnostics Assessment Committee and the EAG as supporting information) were prediction of preterm PE and prediction of delivery independent of the PE diagnosis (*Table 15*). Summary AUC estimates for ROC plots for these outcomes are given in *Table 16*; the AUC data are not used directly in the EAG economic model, but provide an overview of how well the PlGF test performed for different outcomes.

Prognostic accuracy: prediction of pre-eclampsia requiring delivery within 14 days of testing

Three cut-off points were analysed in the PELICAN study (100 pg/ml, 12 pg/ml and the 5th percentile of PlGF concentration adjusted for gestational age). As mentioned in *Table 4*, PlGF concentrations above the 100-pg/ml cut-off point are considered normal and would be unlikely to lead to delivery within 14 days and, as a result, the ≥ 100-pg/ml cut-off point is used to rule out PE, whereas a result of < 100 pg/ml would rule in PE.

The cut-off points of < 100 pg/ml and < 5th percentile of PlGF concentration (test positive) both gave high sensitivity (96%) with good precision (95% CI 89% to 99%) for identifying, among women presenting with suspected PE before week 35 of gestation, those likely to develop PE requiring delivery within 14 days of testing. The cut-off point of < 100 pg/ml (test positive) was reported only in the published paper for the PELICAN study,[5] while the cut-off point of ≥ 100 pg/ml (test negative) was reported only in the product insert[40] (see *Table 13*). Both analyses gave nearly identical sensitivity and specificity (see *Figure 3*).

For women presenting up to week 35 of gestation, the cut-off point of 12 pg/ml yielded lower sensitivity (63%) than the cut-off points of 100 pg/ml or the 5th percentile concentration for gestational age (see *Table 13*). AUC estimates from ROC analyses (see *Table 16*) support the conclusion that the Triage PlGF test has good accuracy for predicting PE requiring delivery within 14 days in women presenting with suspected PE before 35 weeks of gestation.

TABLE 13 Triage PlGF test accuracy for predicting PE requiring delivery within 14 days during weeks 20^{+0} to 34^{+6} of gestation

Study and test cut-off point[a]	TP (*n*)	TN (*n*)	FP (*n*)	FN (*n*)	Total (*n*)	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	Positive LR (95% CI)	Negative LR (95% CI)	Prevalence, % (95% CI)
PELICAN;[5] < 100 pg/ml[b]	73	118	93	3	287	0.96 (0.89 to 0.99)	0.56 (0.49 to 0.63)	0.44 (0.36 to 0.52)	0.98 (0.93 to 0.995)	2.2 (1.9 to 2.6)	0.07 (0.02 to 0.22)	26.48 (21.47 to 31.99)
PELICAN;[5] ≥ 100 pg/ml[c]	72	118	94	3	287	0.960 (0.888 to 0.992)	0.557 (0.487 to 0.625)	0.434 (0.357 to 0.513)	0.975 (0.929 to 0.995)	2.17 (1.85 to 2.54)	0.07 (0.02 to 0.22)	26.13 (21.15 to 31.62)
PELICAN;[5] < 5th percentile	73	115	96	3	287	0.96 (0.89 to 0.99)	0.55 (0.48 to 0.61)	0.43 (0.36 to 0.51)	0.98 (0.93 to 0.995)	2.1 (1.8 to 2.5)	0.07 (0.02 to 0.22)	26.48 (21.47 to 31.99)
PELICAN;[5] < 12 pg/ml	48	190	21	28	287	0.63 (0.51 to 0.74)	0.90 (0.85 to 0.94)	0.70 (0.57 to 0.80)	0.87 (0.82 to 0.91)	6.4 (4.1 to 9.9)	0.41 (0.30 to 0.55)	26.48 (21.47 to 31.99)

FN, false negative; FP, false positive; LR, likelihood ratio; TN, true negative; TP, true positive.
a Data from the PETRA study (100 pg/ml cut-off point) were also available to the EAG and the NICE Diagnostics Assessment Committee, but are confidential and are not reproduced here.
b Main report[5] states that this was an exploratory analysis.
c Reported in the product insert[40] only.

TABLE 14 Triage PlGF test accuracy for predicting PE requiring delivery within 14 days during weeks 35^{+0} to 36^{+6} of gestation

Study and test cut-off point	TP (*n*)	TN (*n*)	FP (*n*)	FN (*n*)	Total (*n*)	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	Positive LR (95% CI)	Negative LR (95% CI)	Prevalence, % (95% CI)
PELICAN;[5] < 5th percentile	47	45	25	20	137	0.70 (0.58 to 0.81)	0.64 (0.52 to 0.75)	0.65 (0.53 to 0.76)	0.69 (0.57 to 0.80)	2.0 (1.4 to 2.8)	0.46 (0.31 to 0.71)	48.91 (40.27 to 57.58)
PELICAN;[5] < 12 pg/ml	15	64	6	52	137	0.22 (0.13 to 0.34)	0.91 (0.82 to 0.97)	0.71 (0.48 to 0.89)	0.55 (0.46 to 0.64)	2.6 (1.1 to 6.3)	0.85 (0.73 to 0.98)	48.91 (40.27 to 57.58)

FN, false negative; FP, false positive; LR, likelihood ratio; TN: true negative; TP, true positive.

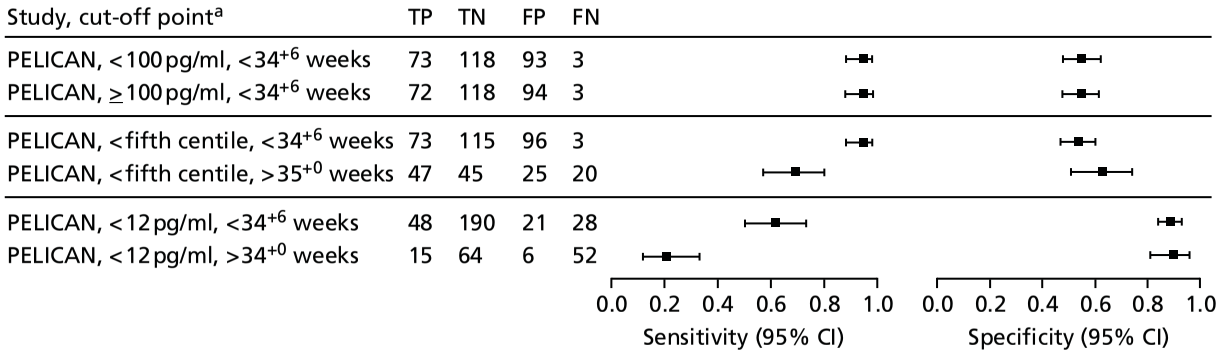

FIGURE 3 Accuracy of the Triage PlGF test at three cut-off points for predicting PE requiring delivery within 14 days. a, Data from the PETRA study (Sibai[6]) (100 pg/ml cut-off point, < 35^{+0} weeks) were also considered by the EAG and the NICE Diagnostics Assessment Committee, but are confidential and are not reproduced here. FN, false negative; FP, false positive; TN, true negative; TP, true positive.

TABLE 15 Other test accuracy outcomes for the Triage PlGF test during weeks 20^{+0} to 34^{+6} of gestation

Study; outcome; cut-off point[a]	TP (*n*)	TN (*n*)	FP (*n*)	FN (*n*)	Total (*n*)	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	Positive LR (95% CI)	Negative LR (95% CI)	Prevalence, % (95% CI)
PELICAN;[5] preterm PE; < 100 pg/ml[b]	108	109	58	12	287	0.900 (0.832 to 0.947)	0.653 (0.575 to 0.725)	0.651 (0.573 to 0.723)	0.901 (0.833 to 0.948)	2.59 (2.09 to 3.22)	0.15 (0.09 to 0.27)	41.81 (36.04 to 47.75)
PELICAN;[5] delivery ≤ 14 days; ≥ 100 pg/ml[b]	78	116	88	5	287	0.940 (0.865 to 0.980)	0.569 (0.498 to 0.638)	0.470 (0.392 to 0.549)	0.959 (0.906 to 0.986)	2.18 (1.84 to 2.57)	0.11 (0.04 to 0.25)	28.92 (23.74 to 34.54)
PELICAN;[5] preterm delivery; < 12 pg/ml[b]	65	135	4	83	287	0.439 (0.358 to 0.523)	0.971 (0.928 to 0.992)	0.942 (0.858 to 0.984)	0.619 (0.551 to 0.684)	15.26 (5.71 to 40.78)	0.58 (0.50 to 0.67)	51.57 (45.62 to 57.48)

FN, false negative; FP, false positive; LR, likelihood ratio; TN, true negative; TP, true positive.

a Data from the PETRA study (delivery ≤ 14 days and delivery ≤ 7 days with a 100-pg/ml cut-off point) were also available to the EAG and the NICE Diagnostics Assessment Committee but are confidential and are not reproduced here.

b Reported in the product insert[40] only.

TABLE 16 Receiver operator characteristic analysis AUC values for the Triage PlGF test during 20^{+0} to 34^{+6} weeks of gestation

Study[a]	Outcome	AUC
PELICAN[5]	PE requiring delivery ≤ 14 days	0.87 (SE 0.03)
	Preterm PE[b]	0.862 (95% CI 0.818 to 0.907)

SE, standard error.
a ROC data from the PETRA study (Sibai[6]) were also available to the EAG and the NICE Diagnostics Assessment Committee but are confidential and are not reproduced here.
b Reported only in the Triage PlGF test product insert.

The 100 pg/ml cut-off point was not tested for women presenting later than 34^{+6} weeks. Both the 5th percentile for gestational age and the 12 pg/ml cut-off points had poor sensitivity (70% and 22%, respectively) for predicting PE requiring delivery within 14 days when women presented after week 34^{+6} (see *Table 14* and *Figure 3*).

As noted (see *Critical appraisal of the included studies*), the EAG did not find any strong reasons to suspect that this outcome might be subject to bias, except for clinical review bias as assessed in QUADAS question 9. The studies diagnosed PE solely according to PlGF test values, whereas in clinical practice the test values would be interpreted alongside hypertension, proteinuria and/or clinical signs or symptoms. However, it is not clear if this difference would lead systematically to an under- or overestimation of test accuracy. The PELICAN[5] study reported that the required sample size was calculated for 'accurate estimation of sensitivity (within 10%) and specificity (within 6%)'. In terms of generalisability, the PELICAN study was conducted in the UK and Ireland, and the PELICAN study stated that reasons for suspecting PE included headache, visual disturbances, epigastric or right upper quadrant pain, hypertension, dipstick proteinuria or suspected fetal growth restriction. This appears consistent with how suspected PE would probably be defined in practice in England and Wales. The definition of actual PE in the PELICAN study included women with superimposed PE and atypical PE.

Other related test accuracy outcomes

Other relevant test accuracy outcomes reported are summarised in *Table 15*. The results from the PELICAN study[5] suggest that the PlGF test with a cut-off of 100 pg/ml has a high sensitivity (90%), with reasonable precision (95% CI 83% to 95%) at diagnosing preterm PE. The results also suggest that the PlGF test with a cut-off point of 100 pg/ml has high sensitivity (94%), with reasonable precision (95% CI 87% to 98%) at predicting delivery within 14 days of testing, independent of the PE diagnosis. In contrast, a cut-off point of 12 pg/ml had poor sensitivity (44%) but good specificity (97%) for predicting preterm delivery (unspecified timescale) independent of the PE diagnosis.

Elecsys sFlt-1 to PlGF ratio test (Roche Diagnostics)

Diagnostic and prognostic accuracy outcomes for the Elecsys sFlt-1 to PlGF ratio test are available in two studies that assessed three test cut-off points: 23, 38, and 85. The studies are not directly comparable because they differed in the cut-off points employed and the time periods during which the rule-in or rule-out of PE applied. The test accuracy outcomes are therefore grouped within each study in the tables presented here. The most relevant outcomes to the NICE decision problem reported in the included studies are:

- rule-out of PE (broadly defined – including HELLP syndrome) within 1 week of testing using a cut-off point of 38 (PROGNOSIS study[51]) (*Table 17* and *Figure 4*)
- rule-in of PE (broadly defined – including HELLP syndrome) within 4 weeks of testing using a cut-off point of 38 (PROGNOSIS study[51]) (see *Table 17* and *Figure 4*).
- rule-out of PE within an unspecified time period (assumed by the EAG to be 3 weeks) using cut-off points of 23 and 85 (Álvarez-Fernández *et al.*[75]) (*Table 18* and *Figure 5*).

TABLE 17 Test accuracy outcomes for the Elecsys sFlt-1 to PlGF ratio test: cut-off point 38

Outcome	TP (*n*)	TN (*n*)	FP (*n*)	FN (*n*)	Total (*n*)	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	Positive LR (95% CI)	Negative LR (95% CI)	Prevalence (95% CI)
PROGNOSIS publication[51] and product insert;[83] presentation 24^{+0}–36^{+6} weeks												
Rule out PE[a] within 1 week: development cohort[51]	NR	NR	NR	NR	500	0.882 (0.725 to 0.967)	0.800 (0.761 to 0.836)	NR	0.989 (0.973 to 0.997)	NR	NR	NR
Rule out PE[a] within 1 week: validation cohort[51]	NR	NR	NR	NR	550	0.800 (0.519 to 0.957)	0.783 (0.746 to 0.817)	NR	0.993 (0.979 to 0.999)	NR	NR	NR
Rule out PE[a] within 1 week: combined cohorts[51,83]	NR	NR	NR	NR	1050	0.857 (0.728 to 0.941)	0.791 (0.765 to 0.816)	0.167 (0.123 to 0.219)	0.991 (0.982 to 0.996)	NR	NR	NR
Rule in PE[a] within 4 weeks: development cohort[51]	NR	NR	NR	NR	500	0.746 (0.625 to 0.845)	0.831 (0.793 to 0.865)	0.407 (0.319 to 0.499)	NR	NR	NR	NR
Rule in PE[a] within 4 weeks: validation cohort[51]	NR	NR	NR	NR	550	0.662 (0.540 to 0.770)	0.831 (0.794 to 0.863)	0.367 (0.284 to 0.457)	NR	NR	NR	NR
Rule in PE[a] within 4 weeks: combined cohorts[51,83]	NR	NR	NR	NR	1050	0.703 (0.619 to 0.778)	0.831 (0.805 to 0.855)	0.386 (0.326 to 0.450)	0.949 (0.931 to 0.963)	NR	NR	NR

FN, false negative; FP, false positive; LR, likelihood ratio; NR, not reported; TN, true negative; TP, true positive.
a Composite outcome of PE/eclampsia/HELLP syndrome (although no cases of eclampsia occurred).

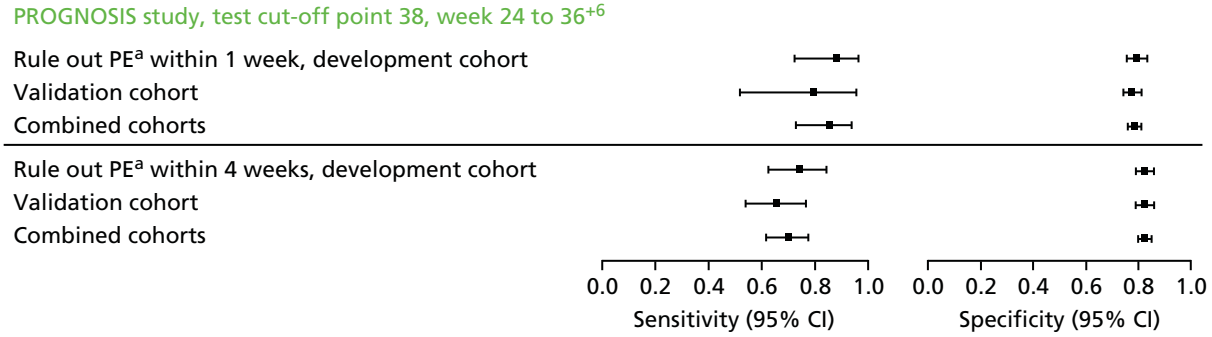

FIGURE 4 Accuracy of the Elecsys sFlt-1 to PlGF ratio test at cut-off point 38 for rule-in and rule-out of PE/eclampsia/HELLP syndrome. Composite diagnostic outcome of PE/eclampsia/HELLP syndrome. FN, false negative; FP, false positive; NR, not reported; TN, true negative; TP, true positive.

Rule out pre-eclampsia within 1 week or rule in within 4 weeks: cut-off point 38

The PROGNOSIS study[51] analysed two cohorts. The development cohort ($n = 500$) aimed to derive a cut-off-based prediction model for the sFlt-1 to PlGF ratio and the validation cohort ($n = 550$) aimed to test the model. The median cut-off point from the model was 38.2 for 1-week rule-out of PE and 37.5 for 4-week rule-in; a common cut-off point of 38 was then validated. The study was reported to have 90% power at an alpha of 0.05 for the validation cohort to show a NPV $\geq$ 96% for ruling out PE within 1 week and a PPV $\geq$ 25% for ruling in PE within 4 weeks.[51] The results of the PROGNOSIS study are presented for the development and validation cohorts and for the two cohorts combined ($n = 1050$) in *Table 17* and *Figure 4*. Sensitivity for ruling out PE within 1 week (cut-off point < 38, i.e. test negative) was relatively high (80–88%) but appears rather uncertain, with wide CIs, especially for the validation cohort, for which the lower bound of the CI is 52%. Specificity was 78–80%, with very narrow CIs, suggesting high precision of the estimates. Despite the uncertainty in the sensitivity estimate, the high specificity and high NPV of 99.3% (see *Table 17*) are supportive of the sFlt-1 to PlGF ratio at a cut-off point of 38 for ruling out PE within 1 week of testing.

The sensitivity of the sFlt-1 to PlGF ratio test for the test ruling in PE within 4 weeks (cut-off point > 38, i.e. test positive) was lower than for the rule-out algorithm (ranging from 66% to 75%), with moderate uncertainty indicated by relatively wide CIs such that the lower 95% confidence bound for the validation cohort is only 54%. Specificity for ruling in PE within 4 weeks was relatively high (83%) with high precision (95% CI 79% to 86%). However, the PPV was only 38.6%, indicating that nearly two-thirds of patients diagnosed would be false positives (see *Table 17* and *Figure 4*).

For both the rule-in and rule-out analyses, the sensitivity of the test was higher in the development cohort than in the validation cohort. Analyses of the combined development and validation cohorts provide the most precise estimates of sensitivity and specificity, and these are the data that inform the EAG economic model for both the rule-in and rule-out of PE using the sFlt-1 to PlGF ratio test (see *Chapter 5*).

Rule out pre-eclampsia within 3 weeks, cut-off points 23 and 85

The study by Álvarez-Fernández *et al.*[75] analysed sFlt-1 to PlGF ratio cut-off points of 23 and 85 for predicting PE (see *Table 18* and *Figure 5*). Their aim was to determine whether or not the original cut-off point of 85, which had been developed from a case–control study of women with normal pregnancies and those with PE (Verlohren *et al.*[84]), could be improved for women with suspected PE presenting up to 34 weeks of gestation ($n = 62$). However, it was not reported whether or not the analysis was powered statistically to achieve a specified level of predictive accuracy.[75] The new cut-off point of 23 had considerably higher sensitivity than the existing cut-off point of 85 (92% vs. 56%), although both had moderate uncertainty, and specificity was lower for the 23 cut-off point (81% vs. 97%). The NPV of

TABLE 18 Test accuracy outcomes for the Elecsys sFlt-1 to PlGF ratio test: cut-off points 23 and 85

Outcome; cut-off point	TP (*n*)	TN (*n*)	FP (*n*)	FN (*n*)	Total (*n*)	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	Positive LR (95% CI)	Negative LR (95% CI)	Prevalence, % (95% CI)
Álvarez-Fernández* et al.:[75] *presentation 20–34 weeks[a]												
PE: rule out within 3 weeks; cut-off point: 23	23	30	7	2	62	0.920 (0.725 to 0.986)[b]	0.811 (0.643 to 0.914)[b]	0.767 (0.573 to 0.894)[b]	0.938 (0.778 to 0.989)[b]	4.86 (2.47 to 9.57)	0.10 (0.03 to 0.38)	40.3 (28.05 to 53.55)
PE: rule out within 3 weeks; cut-off point: 85	NR	NR	NR	NR	62	0.560 (0.353 to 0.750)	0.973 (0.842 to 0.999)	0.933 (0.660 to 0.997)	0.766 (0.616 to 0.872)	20.7 (2.91 to 147)	0.45 (0.29 to 0.70)	40.3 (28.05 to 53.55)

FN, false negative; FP, false positive; LR, likelihood ratio; NR, not reported; TN, true negative; TP, true positive.
a Outcomes are also reported for presentation > 34–41 weeks, but are not included here as they exceed the gestational age range specified in the NICE scope.
b 95% CI values reported by the authors and those calculated by the reviewer differ (by less than ± 1.5%): the authors' data are presented here.

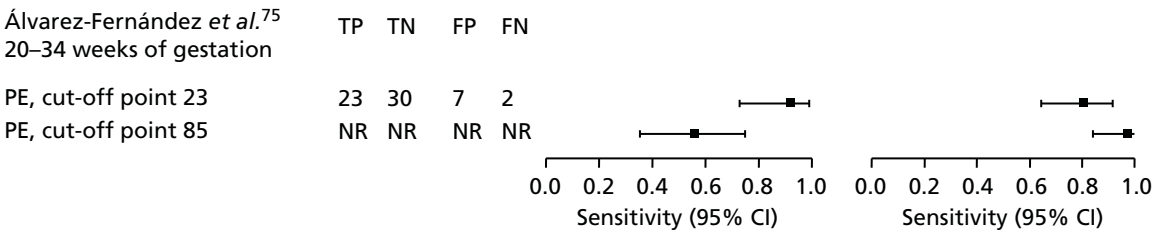

FIGURE 5 Accuracy of the Elecsys sFlt-1 to PlGF ratio test at cut-off points 23 and 85 for predicting PE. FN, false negative; FP, false positive; NR, not reported; TN, true negative; TP, true positive.

93.8% suggests that the 23 cut-off point is appropriate for ruling out PE, although there is moderate uncertainty around this predictive value (95% CI 77.8% to 98.9%) (see *Table 18*).

In summary, the test accuracy data used to inform the EAG economic model (see *Chapter 5*) are taken from the PROGNOSIS study for ruling out PE within 1 week of testing and ruling in PE within 4 weeks of testing (see *Table 17*). The other related outcomes do not directly inform the economic model, but are provided here as supporting information. As noted (see *Critical appraisal of the included studies*), the EAG did not find any strong reasons to suspect that the 1-week rule-out and 4-week rule-in outcomes from the PROGNOSIS study might be subject to bias, except for a high risk of clinical review bias, as assessed in QUADAS question 9. The studies on the Elecsys test diagnosed PE solely according to sFlt-1 to PlGF ratio test values, whereas in clinical practice the test values would be interpreted alongside hypertension, proteinuria and/or clinical signs or symptoms. However, it is not clear if this difference would lead systematically to an under- or over-estimation of test accuracy. In terms of generalisability, the PROGNOSIS study was multinational including one UK centre, whereas the study by Álvarez-Fernández *et al.*[75] was conducted in Spain. The criteria for suspecting PE employed in the PROGNOSIS study (see above, *Table 11*) appear reflective of those that would give rise to suspicion of PE in clinical practice in England and Wales. Regarding the definition of actual PE, PROGNOSIS employed a wider definition than NICE, notably in that it included superimposed PE, HELLP syndrome and eclampsia in the 'overall' PE definition (although no women actually developed eclampsia). Other key features of the PROGNOSIS study that might limit the generalisability of the results are that analyses focused on singleton pregnancies, and that women fell into a relatively narrow age range, being mostly in their late twenties and thirties. However, similar limitations also apply to the study of Álvarez-Fernández *et al.*[75]

The AUC estimates from the ROC analyses (*Table 19*) suggest that, for the PROGNOSIS study,[51] the sFlt-1 to PlGF ratio test tended to have greater accuracy (than clinical information alone for ruling in or ruling out PE (AUC range 0.82 to 0.90 compared with 0.75 to 0.79), although it should be noted that the CIs for the AUC estimates overlap, and adding clinical information to the test result (AUC range 0.85–0.91) did not appear to further improve predictive accuracy in a consistent way.

Comparison of the Triage PlGF test (Alere) and the Elecsys sFlt-1 to PlGF ratio test (Roche Diagnostics)

The Triage PlGF test and Elecsys sFlt-1 to PlGF ratio tests cannot be compared directly in terms of their performance, as the evidence available for each test relates to different test accuracy outcomes. Gencay *et al.*[85] compared diagnostic accuracy of these tests, but this was a case–control study without a suspected PE population, so was excluded from our systematic review of test accuracy. Although AUC estimates from ROC curve analyses can help to compare how well different tests perform, the studies are unbalanced in the extent to which AUC estimates are presented for different outcomes, precluding comparisons (different gestational ages, different primary outcomes; compare *Tables 16* and *19*). For example, in the PETRA study (Triage PlGF test) the outcome was preterm PE necessitating delivery within $\leq$ 14 days or $\leq$ 7 days among women presenting with suspected PE between 20 and 35 weeks of gestation, while in the PROGNOSIS study (Elecsys sFlt-1 to PlGF ratio test) outcomes were rule out PE within 1 week or rule in PE within 4 weeks among women presenting between weeks 24 and 36^{+6}.

TABLE 19 Receiver operator characteristic analysis AUC values for the Elecsys sFlt-1 to PlGF ratio test (and clinical data when reported)

Study	Outcome (predicted by sFlt-1 to PlGF ratio unless stated)	AUC (95% CI)
Zeisler* et al.,[51] *PROGNOSIS study		
Presentation 24+0–36+6 weeks		
Rule out PE[a] within 1 week		
sFlt-1 to PlGF ratio		
Development cohort		0.898 (0.836 to 0.960)
Validation cohort		0.861 (0.798 to 0.924)
Combined cohorts		0.884 (0.829 to 0.924)
sFlt-1 to PlGF ratio plus clinical data		
Development cohort		0.914 (0.852 to 0.975)
Validation cohort		0.859 (0.785 to 0.933)
Clinical data without sFlt-1 to PlGF ratio		
Development cohort		0.793 (0.715 to 0.871)
Validation cohort		0.749 (0.623 to 0.875)
Rule in PE[a] within 4 weeks		
sFlt-1 to PlGF ratio		
Development cohort		0.861 (0.809 to 0.913)
Validation cohort		0.823 (0.773 to 0.873)
Combined cohorts		0.842 (0.803 to 0.874)
sFlt-1 to PlGF ratio plus clinical data		
Development cohort		0.897 (0.852 to 0.942)
Validation cohort		0.849 (0.800 to 0.897)
Clinical data without sFlt-1 to PlGF ratio		
Development cohort		0.794 (0.742 to 0.847)
Validation cohort		0.785 (0.732 to 0.838)
***Álvarez-Fernández* et al.[75]**		
Presentation 20–34 weeks[b]		
PE, all women		0.903 (0.815 to 0.991)
PE, women with single gestation only		0.907 (0.798 to 1.000)

SE, standard error.
a Composite outcome of PE/eclampsia/HELLP syndrome (although no cases of eclampsia occurred).
b Outcomes are also reported for presentation > 34 to 41 weeks, but are not included here as they exceed the gestational age range specified in the NICE scope.

Comparison of the PerkinElmer and Thermo Fisher Scientific tests

In support of the current diagnostic assessment, PerkinElmer provided evidence showing a good correlation between PlGF concentrations measured by the DELFIA Xpress PlGF test and Elecsys PlGF assays,[86] while Thermo Fisher Scientific provided evidence showing good correlations between PlGF concentrations, sFlt-1 concentrations, and the sFlt-1 to PlGF ratios measured by the BRAHMS Kryptor and Elecsys assays.[48,87] However, only one of these documents, a study by Andersen *et al.*,[48] compared diagnostic accuracy of the tests. None of these documents meets the scope for the current diagnostic assessment (and hence they were excluded from our systematic review of test accuracy), as the biomarker samples were not obtained from women with suspected PE. The EAG notes that, although the biomarker test measurements correlated well across tests, slopes of the correlations generally differed from unity:

sFlt-1 to PlGF ratios measured with the BRAHMS Kryptor assay were systematically higher than those measured with the Elecsys assay,[48,87] while PlGF concentrations measured by the DELFIA Xpress PlGF test were systematically lower than those measured with the Elecsys PlGF assay.[86]

Given that no studies on the DELFIA Express PlGF test and the BRAHMS Kryptor sFlt-1 to PlGF ratio met the inclusion criteria for the review of test accuracy, the EAG is unable to determine the diagnostic accuracy of these two tests. A study by Anderson *et al.*[48] has compared the accuracy of the Elecsys and BRAHMS Kryptor sFlt-1 to PlGF ratio tests for predicting PE, in a case–control study that included women with PE ($n = 39$, i.e. a relatively small sample size) and women with healthy pregnancies ($n = 76$). While it may be tempting to use the results from Andersen *et al.*[48] as illustrative of how the Elecsys and BRAHMS Kryptor sFlt-1 to PlGF ratio tests might perform in women with suspected PE, the EAG cautions against this, as both sensitivity and specificity would be overestimated to an unknown extent, which might differ between the tests. In addition, the study was conducted in Denmark, included only singleton pregnancies, and was retrospective (thereby at possible risk of selection bias).

As the study by Anderson *et al.*[48] is the only diagnostic accuracy comparison that has been made between the BRAHMS Kryptor and Elecsys sFlt-1 to PlGF ratio assays, a brief summary of the comparison is provided here with the proviso that this should be interpreted as illustrative only.

Andersen *et al.*[48] made a number of comparisons between the BRAHMS Kryptor and the Elecsys sFlt-1 to PlGF ratio assays, with post hoc subgroup analyses for early-onset and late-onset PE and for non-obese (BMI of < 30 kg/m^2) and obese (BMI of ≥ 30 kg/m^2) women. Sensitivity, specificity, PPVs and NPVs, and areas under the ROC curve for the sFlt-1 to PlGF ratio are concisely tabulated by Andersen *et al.*[48] in their paper, for test cut-off points of > 85, < 33 and > 110, although the EAG has not extracted these data. At all cut-off points, the sensitivity of the BRAHMS Kryptor sFlt-1/PlGF assay was either higher than or similar to that of the Elecsys sFlt-1 to PlGF ratio. At all cut-off points, the specificity of the BRAHMS Kryptor sFlt-1/PlGF assay was lower than that of the Elecsys sFlt-1 to PlGF ratio for the overall gestational period or late gestation, but both tests had similar specificity during early gestation (i.e. diagnosis up to 34 weeks). For both companies' tests, both sensitivity and specificity were highest during early gestation.

Receiver operator characteristic analyses conducted by Andersen *et al.*[48] indicated that the BRAHMS Kryptor PlGF assay performed marginally better than the Elecsys PlGF assay, while the BRAHMS Kryptor sFlt-1 to PlGF ratio had similar accuracy to the Elecsys sFlt-1 to PlGF ratio. Both sFlt-1 to PlGF ratio tests showed best performance (AUC > 0.92) for early-onset PE (i.e. before week 34 of gestation). Both of the sFlt-1 to PlGF ratio tests performed better (AUC > 0.8) in non-obese women than in obese women (AUC < 0.77).

Repeating the caveats above, the EAG emphasises that these findings may not necessarily reflect the relative performance of these tests if applied in a population of women with suspected PE.

Ongoing studies

The EAG identified 14 ongoing studies on the use of the PlGF and/or sFlt-1 to PlGF ratio tests in studies of PE, most of which were summarised in documents provided as part of the companies' submissions. Seven of these studies were considered irrelevant, as they do not include women with suspected PE. The remaining seven ongoing studies, that is, those that are potentially most relevant to the current diagnostic assessment, are listed in *Appendix 5*. It appears unlikely that these ongoing studies would further inform the current diagnostic assessment, as four are not expected to provide results until 2016, and it is unclear if the remaining three would meet the eligibility criteria. At the time of this report going to press, results from these ongoing studies were not available.

Summary of diagnostic test performance

- Four studies met the inclusion criteria for the systematic review of test accuracy; these assessed the Triage PlGF test (two studies) and Elecsys sFlt-1 to PlGF ratio (two studies).
- One of the included studies on the Triage test, PETRA, was confidential when the present report was prepared. The PETRA study is excluded from the present report, but was available to the EAG and the NICE Diagnostics Assessment Committee.
- No test accuracy studies of the DELFIA Xpress PlGF test or BRAHMS Kryptor sFlt-1 to PlGF ratio test were found.
- No test accuracy studies that assessed index tests as an alternative to quantitative proteinuria testing (part 2 of the NICE decision problem) were identified.
- None of the included published studies evaluated more than one test, meaning that head-to-head comparisons of tests are not available.
- The primary studies included in the test accuracy review reported different outcomes for each test (prognostic accuracy for the Triage PlGF test; diagnostic accuracy for the Elecsys sFlt-1 to PlGF ratio test).
- Meta-analysis was not possible because of heterogeneity of outcomes, with different test cut-off points reported in the studies.
- The Triage PlGF test has high prognostic sensitivity (96% in the PELICAN study) for predicting PE requiring delivery within 14 days of testing.
- The Elecsys sFlt-1 to PlGF ratio has good diagnostic sensitivity (85.7%) for rule-out of PE within 1 week of testing and good specificity (83.1%) for rule-in of PE within 4 weeks, but with a high false-positive rate (PPV 38.6%).
- The definitions of PE used in the studies differed from the NICE definition, appearing to include a wider range of women than the NICE definition would permit; notably the main Elecsys test study (PROGNOSIS) included HELLP syndrome in the PE definition.
- Overall, the studies were deemed unlikely to be at high risk of bias. An exception is a high risk of clinical review bias in all three published studies, as the studies diagnosed PE solely according to biomarker test results, whereas in clinical practice the biomarker test results would be interpreted alongside hypertension, proteinuria and/or other signs or symptoms. However, it is unclear whether or not this difference would have led to systematic under- or overestimation of test accuracy outcomes.

Chapter 5 Economic analysis

This section assesses the current state of evidence on the cost-effectiveness of the PlGF test or the sFlt-1 to PlGF ratio test when used either in addition to standard clinical assessment or as an alternative to quantitative proteinuria testing, based on a systematic review of economic analyses and critical appraisal of company submissions. Based on the evidence identified, a de novo economic model was developed to estimate the cost-effectiveness of the biomarker tests. Parameters for the model were identified through a systematic review of economic studies (see *Systematic review of economic studies*), a review of HRQoL studies (see *Review of health-related quality-of-life studies*) and targeted searches for data on relevant costs and resources (see *Results of the review of health-related quality-of-life studies*).

The economic analysis was informed by the Project Advisory Group, which provided comments on draft versions of the protocol, economic model and final report. Four members of the NICE Diagnostics Assessment Committee for this assessment (see *Acknowledgements*) also provided comments on the draft economic model.

Methods for reviews of economic and HRQoL studies

Systematic review of economic studies

The methods detailed in *Chapter 3* were used to systematically review the cost and cost-effectiveness literature. The populations, interventions and comparators included are the same as for the systematic review of test accuracy (as described in *Inclusion and exclusion criteria*), with the exception of study design and outcomes. As the systematic review of test accuracy searches applied no study type filters (see *Chapter 3*), the results of these searches were suitable for identifying cost-effectiveness evidence by applying the relevant economic evaluation inclusion criteria.

Studies were included if they were full economic evaluations, assessing both costs and consequences, or cost studies for the specified index tests. Outcomes included are those consistent with full economic evaluations and cost studies, including intermediate outcomes (budget impact, cost per patient or cost per case of PE correctly managed), or final outcomes (life-years or QALYs gained). Each step of the review was conducted by one health economist and checked by a second health economist, with any disagreements resolved by discussion.

Review of HRQoL studies

The EAG undertook a series of sequential searches designed to identify HRQoL data in gestational hypertension, PE and general pregnancy. The goal of these searches was to identify European Quality of Life-5 Dimensions (EQ-5D) utility values for use in the economic model or to identify utility values that could be mapped to EQ-5D using published algorithms, in line with the NICE reference case.[88] The eligible instruments were identified using Oxford University's Health Economics Research Centre database of mapping studies.[89] The following instruments were eligible for inclusion: EQ-5D, Short Form questionnaire-36 items (SF-36) (using all scales), Short Form questionnaire-12 items (SF-12), Short Form questionnaire-6 items (SF-6D), Nottingham Health Profiles and the Health Assessment Questionnaire. The relevant population is women who are pregnant and/or who have gestational hypertension or PE and their neonates. Non-intervention study designs were preferred, unless the intervention directly addressed hypertensive disorders of pregnancy (including PE). Studies assessing specific symptoms of pregnancy, rather than general pregnancy HRQoL, that were not potentially related to gestational hypertension or PE (such as urinary incontinence or emesis) were excluded. Studies in subpopulations of pregnant women that are not directly related to gestational hypertension or PE were also excluded (such as human immunodeficiency virus, thyroid conditions and cancer). Clinical outcomes relevant to the model were defined as delivery before 35 weeks of gestation, delivery after 35 weeks of gestation, NICU stay, hospital stay, birth by induction, birth by caesarean section, standard vaginal birth, severe complications of birth, mild PE and severe PE.

The sequential approach for identifying HRQoL studies was followed and, unless stated otherwise, all steps were conducted by one health economist and checked by a second health economist, with any disagreements resolved through discussion:

1. References already identified through the systematic review of test accuracy (see *Chapter 4*, *Quantity and quality of research available*) were searched for HRQoL studies. Two health economists independently marked any titles and abstracts that appeared to be relevant. Following this process, only one study, reported by Shmueli *et al.*,[90] was identified (described in detail in *Results of the review of health-related quality-of-life studies*).
2. Information on HRQoL was sought from the company submissions provided for this diagnostic assessment.
3. The NICE Guideline on Hypertensive Disorders in Pregnancy (CG107)[13] was scrutinised to identify any further HRQoL data.
4. Finally (as the preceding searches yielded limited relevant information), additional systematic searches of bibliographic databases were conducted for HRQoL data in pregnant women or women with hypertensive disorders of pregnancy.

The systematic searches were conducted in MEDLINE, MEDLINE In-Process & Other Non-Indexed Citations, EMBASE, and DelphiS. The search strategy (see *Appendix 6*) was designed in MEDLINE and adapted for the other databases. The inclusion and exclusion criteria for eligibility screening of titles and abstracts are given in *Box 1*. The same eligibility criteria were also used for screening full-text records, with an exception that more stringent criteria for HRQoL outcomes were applied at full-text screening. Studies could be

BOX 1 Title and abstract eligibility criteria for systematic searches of studies on HRQoL

Inclusion criteria

Population

Women with PE or gestational hypertension; or a general pregnancy/post-partum population experiencing any events that could be relevant to HRQoL estimation in PE or gestational hypertension (e.g. mode of delivery, hospitalisation).

HRQoL outcomes

SF-36 (all subscales)[a], SF-12, SF-6D, EQ-5D, HUI (1, 2 or 3), HAQ, QALY.

Clinical outcomes

Search was not limited by clinical outcomes.

Exclusion criteria

Population

Conditions not specifically relevant to PE or gestational hypertension (e.g. thyroid disease, human immunodeficiency virus).

Reference type

Protocols, letters, conference abstracts, case reports.

HAQ, Health Assessment Questionnaire; HUI, Health Utilities Index.

a Studies in which it was unclear whether or not all SF-36 subscales were available were included at abstract screening.

excluded at full-text screening if the HRQoL scales were incomplete (e.g. if six scales of the SF-36 instead of eight were reported, or if only the mental component scale of the SF-12 was reported), as the mapping algorithm to obtain EQ-5D scores requires full reporting of the scale to be mapped. In addition, at full-text screening, studies that reported only summary QALYs (i.e. cost-effectiveness results) without utility scores were excluded.

Methods for data extraction and critical appraisal of economic studies

Data extraction and critical appraisal of economic evaluations was undertaken by one health economist and checked by a second using a predesigned standard data extraction form that had been pilot tested on previous projects to facilitate ease of use by reviewers and minimise errors (see *Appendix 7* for an example completed data extraction form). The included economic evaluations were critically appraised using a checklist based upon those proposed by Drummond *et al.*,[91] Philips *et al.*[92] and the NICE reference case.

Results of the review of economic studies

Of the 1972 potentially relevant references originally identified in the searches of test accuracy studies (see *Chapter 4*, *Quantity and quality of research available*), six[33,39,93–96] appeared to provide information about economic studies and were retrieved for further scrutiny (*Figure 6*). Three further documents reporting economic analyses were provided by companies in confidence, as part of the NICE Diagnostics

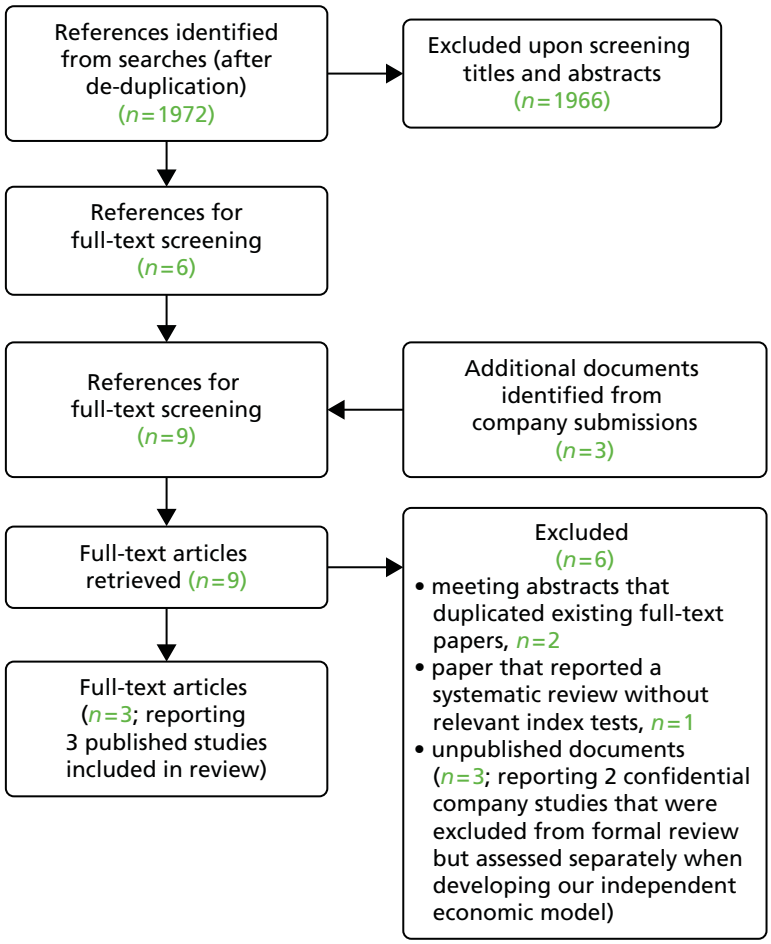

FIGURE 6 Flow chart for the identification of economic studies.

Assessment process,[97–99] and these were also retrieved. After scrutiny, two references were excluded because they were meeting abstracts[93,94] that duplicated existing full-text references, and one was excluded because it was a systematic review that did not include any relevant index tests.[95] This systematic review, by Meads *et al.*,[95] reviewed diagnostic methods for PE, but did not include any of this assessment's index tests. The remaining six references were three published full-text articles[33,39,96] that reported on three unique studies, and three confidential unpublished references that reported on two economic analyses provided by companies submitting evidence under the NICE Diagnostic Assessments process: Alere, Inc.[97,98] and Roche Diagnostics.[99]

Being confidential, the unpublished company documents[97–99] were not included in the systematic review and are not reported here. These documents were, however, taken into consideration by the EAG when planning and developing our economic model (see *Methods for the independent economic analysis*). The three published economic studies included in the systematic review[33,39,96] are described and critically assessed below. Implications of excluding the confidential studies are considered in the *Discussion* (see *Chapter 6*).

The three published studies were all cost analyses focusing on potential savings in health sector resources through improved accuracy of diagnosis of PE. None of these studies formally evaluated maternal or fetal outcomes (other than admission to intensive care or to a special care baby unit, which were included in the cost analysis). In each of the studies, the majority of the saving associated with improved accuracy of PE diagnosis is realised by reduction in false-positive diagnoses. *Table 20* provides an overview of the characteristics of the included cost analyses and a brief summary of their base-case results.

Hadker et al.[39,96]

Critical appraisal of the studies reported by Hadker *et al.* in 2010[39] and 2013[96] has been combined as they used the same model, populated with identical clinical inputs, to address UK[39] and German[96] health-care payer perspectives, respectively. The discussion of the validity and generalisability of the results focuses on the analysis that is more relevant to the UK setting.[39] The UK analysis is described as a budget impact model, although both studies might be better described as cost analyses, which are concerned with determining potential changes in resource use (hence cost) associated with improved diagnostic performance achieved using the sFlt-1 to PlGF ratio (primarily through increased specificity, given the relatively low prevalence of PE in the population being assessed). The population studied was all women receiving obstetric care, assessed for PE beyond 20 weeks of gestation.

Sensitivity and specificity for usual clinical assessment were based on values reported in a systematic review by Meads *et al.*[95] The parameter values used in 2010 by Hadker *et al.*[39] are reported as pooled averages (the paper does not state whether these are means or medians) of the values reported by Meads *et al.* No detail was provided on the method of pooling and there was no indication of which diagnostic methods reported by Meads *et al.* were included in the pooled estimate. Sensitivity for diagnostic tests in Meads *et al.* ranged from 9% to 66% and specificity ranged from 74% to 96% across a range of tests. Sensitivity and specificity for the sFlt-1 to PlGF ratio were taken from a single publication by Verlohren *et al.*,[84] which was excluded from our review of test accuracy (see *Chapter 3*) because the study population had already been diagnosed with PE (i.e. not suspected PE). Neither study reports how this paper was identified or selected for use in the model, and no quality assessment of these or other inputs into the model was provided. The model assumes that all patients who initially test negative on the sFlt-1 to PlGF ratio test receive two repeat tests, and it includes these in resource-use estimates, but does not adjust diagnostic outcomes for repeated testing. Identification or diagnosis of PE in the model appears to be based entirely on the modelled outcome of the sFlt-1 to PlGF ratio test and does not take account of the other maternal symptoms or risk factors (i.e. the presence or absence of clinical signs of PE).

Hadker *et al.*[39] provided little information on resource use assumptions. The study reports that these were based on published literature and expert opinion, but does not explain how sources were identified. The text of the paper states that 'PE management costs include physician office visits, physical exams, regular

TABLE 20 Characteristics of included economic studies

Characteristics	Study[a]		
	Hadker *et al.*[39]	Hadker *et al.*[96]	Schnettler *et al.*[33]
Publication year	2010	2013	2013
Country	UK	Germany	USA
Study type	Cost analysis	Cost analysis	Cost analysis
Population	Women > 20 weeks of gestation receiving obstetric care	Women > 20 weeks of gestation receiving obstetric care	Women < 34 weeks of gestation with suspected PE
Intervention(s)	Intervention: standard care + sFlt-1 to PlGF ratio (Elecsys) test (the diagnostic threshold adopted in the study was a sFlt1 to PlGF ratio of ≥ 85)	Intervention: standard care + sFlt-1 to PlGF ratio (Elecsys) test (diagnostic threshold adopted in the study was a sFlt1 to PlGF ratio of ≥ 85)	Intervention: standard care + sFlt1 to PlGF (Elecsys) ratio test (diagnostic threshold adopted in the study was a sFlt1 to PlGF ratio of ≥ 85)
	Comparator: standard care (diagnose PE based on clinical signs/symptoms/findings)	Comparator: standard care (diagnose PE based on clinical signs/symptoms/findings)	Comparator: standard care (diagnose PE based on clinical signs/symptoms/findings)
Model type	Decision tree	Decision tree	Not clear
Intervention effect	Intervention: sensitivity, 0.82; specificity, 0.95	Intervention: sensitivity, 0.82; specificity, 0.95	Intervention: sensitivity, 0.76; specificity, 0.94
	Comparator: sensitivity, 0.46; specificity, 0.83	Comparator: sensitivity, 0.46; specificity, 0.83	Comparator: sensitivity, 0.94; specificity ,0.36
Base-case results	Overall cost reduction of £945 per patient, from £2726 to £1781	Overall cost reduction of €637 per patient, from €1579 to €942	Overall cost reduction of US$1215 per patient, from US$3022 to US$1807
	Budget impact, per patient, by diagnostic outcome (difference between standard practice and standard practice plus test): TP = £159; TN = £185; FP = –£1204; FN = –£85	Budget impact, per patient, by diagnostic outcome (difference between standard practice and standard practice plus test): TP = €114; TN = €121; FP = –€791; FN = –€82	Budget impact, per patient, by diagnostic outcome (difference between standard practice and standard practice plus test): TP = US$392; TN = –US$166; FP = –US$904; FN = –US$1391

FN, false negative; FP, false positive; TN, true negative; TP, true positive.
a Data from a study by Hunter[97] are confidential and not reproduced here (these data were available to the EAG and NICE Diagnostics Assessment Committee).

blood pressure checks, blood and urine tests, and cardiotocography as well as hospital stays for day assessments, intensive care, inpatient monitoring and delivery or termination of pregnancy', but provides no further information on the assumed frequency of these events or the proportions of women affected. In a table summarising cost assumptions, two published studies are referenced (Meads *et al.*[95] and Murphy and Stirrat[100]). However, no detail was provided on which resource use data were extracted from these sources. As a result it is difficult to judge the validity of the resource use or cost assumptions included in the model (*Table 21*). The cost year for the UK analysis was 2009.

In addition to the base-case analysis reported in *Table 20*, Hadker *et al.*[39] present deterministic sensitivity and scenario analyses conducted on disease prevalence, sensitivity and specificity of usual clinical assessment, the proportion of high-risk patients and the cost of the new test. Separate scenario analyses that involved increasing the prevalence of PE, increasing the sensitivity of usual clinical assessment and an increased cost for the sFlt-1 to PlGF ratio test had minimal impact on the estimated saving associated with the addition of the diagnostic test to current clinical assessment (£969, £955 and £928, respectively, compared with £945 per patient in the base case). Increased specificity of usual clinical assessment resulted in a large decrease in the estimated saving (to £208 per patient), while increasing or decreasing the proportion of high-risk patients increased costs identically for usual clinical assessment and the new

TABLE 21 Parameter inputs for Hadker *et al.*:[39] UK health-care payer perspective

Parameter	Data
Population	Women > 20 weeks of gestation receiving obstetric care
PE prevalence in study population (%)	4.03[a]
Disease severity/maternal outcome (% total population)	
Mild PE	93.60
Severe PE	4.75[a]
Eclampsia	1.65[a]
Death	0.00[a]
Neonatal outcome	Not reported
Costs by disease severity/maternal outcome	***Costs at 2008/9 price base***
Mild PE	
Drug costs	£28.25
Management costs	
True positive	£9576.25
False negative	£4480.38
Severe PE	
Drug costs	£127.30
Management costs	
True positive	£14,545.49
False negative	£11,308.87
Eclampsia	
Drug costs	£163.19
Management costs	
True positive	£21,340.12
False negative	£17,122.77
No PE, but with risk factors	
Management costs	
False positive	£9576.25
Aspirin for women with PE risk factors	£2.74

a Bhattacharya and Campbell[101] reported a retrospective analysis of cases with a diagnosis of hypertension associated with proteinuria in Grampian, Scotland, between 1981 and 2000 (identified from Aberdeen Maternity and Neonatal Databank). In total, 4188 cases were identified out of 103,896 deliveries, yielding overall incidence of 4.03% (varying from 1.16% in 1995 to 8.32% in 1984. The study states that there was a decline in numbers over time).

test, but did not affect the estimated saving. In each of these analyses uncertainty was assessed by varying parameters by a fixed (arbitrary) proportion rather than using any statistically derived measure of variation (such as 95% confidence limits for PE prevalence), using data from Bhattacharya and Campbell,[101] and sensitivity and specificity of usual clinical assessment derived from Meads *et al.*[95]

Overall, this study indicates that improved diagnostic performance, using the Elecsys sFlt-1 to PlGF ratio test for identifying PE in women after 20 weeks of gestation, may be associated with reduced resource use and that the majority of this benefit would be realised by reducing false-positive results. This study was conducted for a general population of pregnant women, not for those with suspected PE, and is therefore not directly relevant to the current decision problem.

Schnettler et al.[33]

Schnettler *et al.*[33] reported a study comparing observed management of women presenting to hospital with suspected PE at less than 34 weeks of gestation and modelled management using the Elecsys sFlt-1 to PlGF ratio diagnostic test. Current clinical assessment was based on a combination of blood pressure, urinary protein excretion, levels of alanine aminotransferase (ALT) and platelet counts with the modelled management based on these same measures combined with the sFlt-1 to PlGF ratio, and treating clinicians were unaware of the PlGF and sFlt1 values. The paper does not indicate specific threshold values for hypertension or proteinuria that would lead to diagnosis of PE, but does provide threshold values for elevated transaminases (ALT or aspartate aminotransferase $\geq$ 80 units/l) and low platelet count ($\leq 100 \times 10^9$/l). The sFlt-1 to PlGF ratio threshold value adopted in the study was $\geq$ 85 (referenced to Rana *et al.*[102]). Resource-use information was collected over a 2-week period following enrolment to the study and characterised according to test outcome (true/false positive and true/false negative). Resource use was valued using institutional charges (converted to US$ using the 2012 charge-to-cost conversion factor) for triage evaluations, maternal and fetal radiological studies, laboratory tests, admissions, consultations, deliveries and miscellaneous items (such as intravenous tubing and fluids). The cost of the diagnostic test was US$101.14 (sourced from Hadker *et al.*[39]).

Sensitivity and specificity of usual assessment for predicting PE were based on diagnoses observed in normal clinical practice for the 149 women enrolled into the study. These are reported in *Table 20* and indicate high sensitivity, but poor specificity, for usual clinical assessment. Diagnostic outcome, using the sFlt-1 to PlGF ratio, is reported (in terms of true/false-positive and true/false-negative status), but it is not clear from the study whether this classification was based entirely on the diagnostic test results or on a combination of the diagnostic test and usual clinical assessment (and if so, what combination algorithm was used). Sensitivity and specificity of the diagnostic test (estimated by the EAG from values reported by Schnettler *et al.*[33]) are reported in *Table 20*, showing poorer sensitivity than usual clinical assessment but substantially improved specificity. The number of true negatives increased from 35 (24% of the total population) with usual clinical assessment to 92 (62% of the total population) using the diagnostic test at the expense of an increase in false-negative results from 3 (2% of the total) to 12 (8% of the total). *Table 22* presents the prevalence of PE in the population studied by Schnettler *et al.*[33] and key outcomes reported by diagnostic test outcome.

As the study that informed the model was a retrospective cohort, patients who had false-positive results with the Elecsys sFlt-1 to PlGF ratio test might not have been treated with the same intensity as patients who had a true-positive result. Likewise, treatment intensity for false-negative results did not closely match that for true-positive results. This is likely to be a result of different lengths of admission to hospital. A woman with a false-positive test result may be hospitalised for a short period of time with suspicion of PE and a woman with a false-negative result may spend some time out of hospital before her symptoms force admission.

The study reported very limited sensitivity analyses. These were undertaken with respect to the sFlt-1 to PlGF ratio cut-off value (which varied between 5 and 200; the results of this analysis are not reported in the study) and costs by test outcomes (i.e. patients' true-positive, false-positive, true-negative or false-negative status, which were varied between 50% and 200%). Schnettler *et al.*[33] do not appear to have considered uncertainty in the estimation of test outcomes, from individual events used in the resource-use estimation (including neonatal outcomes) or in the unit costs applied (other than indirectly through percentage variation applied to costs by test outcomes).

Overall, this study indicates that improved diagnostic performance for identifying PE in women presenting with suspected PE up to 34 weeks of gestation may be associated with reduced resource use, and that the majority of this benefit would be realised by reducing the number of false positives. In this study this potential financial saving was associated with a diagnostic strategy that was inferior in terms of correctly identifying false-negative results (12 compared with three using standard clinical assessment). Schnettler *et al.*[33] reported that women in the false-negative group were predominantly obese (mean BMI 36.8 kg/m^2,

TABLE 22 Parameter inputs for Schnettler *et al.*[33]

Parameter	Data			
Population	Women (< 34 weeks of gestation) presenting at hospital for evaluation of suspected PE			
PE prevalence in study population (%)	34.23[a]			
Disease severity/maternal outcome				
Mild PE	Not reported			
Severe PE				
Eclampsia				
Death				
Neonatal outcome (all births)	***TP* (n = 39)**	***FP* (n = 6)**	***TN* (n = 92)**	***FN* (n = 12)**
Time to delivery (weeks)	0.3	4.2	6.4	0.6
Gestational age at delivery (weeks)	30	35	37	32
Birthweight (g)	1275	2210	3115	1725
Neonatal outcome (births within 2 weeks)	***TP* (n = 36)**	***FP* (n = 0)**	***TN* (n = 0)**	***FN* (n = 12)**
NICU admission (*n*)	36	0	0	11
NICU (days)	18	0	0	17
Perinatal death (%)	0	0	0	1
Resource use by test outcome	***TP* (n = 39)**	***FP* (n = 6)**	***TN* (n = 92)**	***FN* (n = 12)**
Triage visits (*n*)	1	1	1	1
Caesarean section (%)	33	4	55	8
Antepartum admission (%)	25	4	34	9
Duration of antepartum admission (days)	3	3	4	3
Outpatient visits (*n*)	0	0	45	3
Tests of fetal well-being (*n*)	2	2	4	2
Magnetic resonance imaging (%)	1	0	6	1
Magnetic resonance angiogram (%)	2	0	6	0
Chest radiography (*n*)	8	0	2	1
Complete blood count (*n*)	8	2	1	7
Coagulation panel (*n*)	8	1	2	5
Serum chemistries (*n*)	37	7	8	24
Urine studies (*n*)	1	2	2	3
Consultations (*n*)	1	0	0	1
Cost by test outcome (per patient)	***TP* (n = 39)**	***FP* (n = 6)**	***TN* (n = 92)**	***FN* (n = 12)**
Cost	US$4325	US$2998	US$471	US$3273

FN, false negative; FP, false positive; TN: true negative; TP, true positive.
a Of the 176 cases presenting, 18 re-enrolments and nine cases without complete data were excluded. Observed prevalence in the remaining cases was 51 out of 149.

compared with 29.9, 27.8 and 33.9 kg/m^2 for true-positive, false-positive and true-negative cases, respectively). This study is relevant to the decision problem as stated by NICE, as it was conducted in pregnant women being evaluated for suspected PE, but it does not cover the full range of gestation indicated in the scope and has a small sample size. Moreover, the study was conducted in a single hospital in the USA and may lack generalisability to the UK NHS context.

Summary of economic studies

A summary critical appraisal checklist for quality assessment of the included studies is shown in *Table 23*.

Overall, the three published economic evaluations appear applicable to the UK NHS in terms of the patient population and comparator, although Schnettler *et al.*[33] had a different health-care system and setting than is usual in the UK, and Hadker *et al.*[93] was a study in Germany that has unclear relevance to the UK NHS setting. In all three studies, the modelling methodology, model structure and assumptions were appropriate, and data inputs were described, although these were not fully justified in two studies.[39,96] Of the three published studies, none was based on a systematic review, none measured health benefits in QALYs, none used standardised and validated generic instruments, none described and justified the resource costs used, none assessed uncertainty through appropriate sensitivity analyses and none reported whether or not their model had been validated.

TABLE 23 Critical appraisal checklist of economic studies included in the systematic review

	Study		
Item[a]	Hadker *et al.*[39]	Hadker *et al.*[96]	Schnettler *et al.*[33]
1. Is there a clear statement of the decision problem?	Yes	Yes	Yes
2. Is the comparator routinely used in UK NHS?	Yes	Yes	Yes
3. Is the patient group in the study similar to those of interest in UK NHS?	Yes	Yes	Yes
4. Is the health-care system comparable to UK?	Yes	Yes	No
5. Is the setting comparable to the UK?	Yes	Unclear	No
6. Is the perspective of the model clearly stated?	Yes	Yes	Yes
7. Is the study type appropriate?	Yes	Yes	Yes
8. Is the modelling methodology appropriate?	Yes	Yes	Yes
9. Is the model structure described and does it reflect the disease process?	Yes	Yes	Yes
10. Are assumptions about model structure listed and justified?	Yes	Yes	Yes
11. Are the data inputs for the model described and justified?	Yes, described. Not fully justified	Yes, described. Not fully justified	Yes
12. Is the effectiveness of the intervention established based on a systematic review?	No	No	No
13. Are health benefits measured in QALYs?	No	No	No
14. Are health benefits measured using a standardised and validated generic instrument?	No	No	No
15. Are the resource costs described and justified?	No (insufficient detail)	No	No (insufficient detail)
16. Have the costs and outcomes been discounted?	Not applicable	Not applicable	Not applicable
17. Has uncertainty been adequately assessed?	Unclear[b]	Unclear[c]	No[d]
18. Has the model been validated?	Not reported	Not reported	Not reported

a Data from a study by Hunter[97] are confidential and not reproduced here (these data were critically appraised by the EAG and were available to the NICE Diagnostics Assessment Committee).
b Limited deterministic sensitivity/scenario analyses conducted on disease prevalence, sensitivity/specificity of standard practice, proportion of high-risk patients and cost of new test.
c Limited deterministic sensitivity/scenario analyses conducted.
d Very limited sensitivity analysis, with respect to sFlt-1 to PlGF ratio cut-off value (varied between 5 and 200, results not reported in study) and costs (varied between 50% and 200%).

Results of the review of HRQoL studies

HRQoL evidence from searches of test accuracy

Of the 1972 references originally identified in the searches of test accuracy studies (see *Chapter 4*, *Quantity and quality of research available*), only one study, by Shmueli *et al.*,[90] reported relevant information about HRQoL. Shmueli *et al.*[90] constructed a model of a screening programme that allowed some low-impact prevention of PE through management with aspirin, calcium, folate and vitamins in an Israeli health system context. The HRQoL in the model is tied to neonatal mortality, mortality in the first year of an infant's life, and an assumption of higher rates of diabetes mellitus in children born to mothers with PE based on studies cited from the literature.[103–105] These three studies cited by Shmueli *et al.*[90] were all from Israel. One study was based on maternal mortality data from 1964 to 1976,[103] and the other two studies were reported only in conference abstracts,[104,105] only one of which was available to the EAG.[105] The available abstract was for long-term outcomes of maternal gestational diabetes mellitus, not PE, on offspring.[105] None of the data used in Shmueli *et al.*'s[90] model to determine HRQoL differences appear generalisable to the UK, and none appears relevant to the current decision problem.

HRQoL evidence from company submissions

As mentioned, Alere and Roche Diagnostics each produced economic models to simulate the addition of their diagnostic test to current clinical practice. These were critically appraised by the EAG, but were specified as confidential by the companies and are not presented here. Both models were cost models and did not include HRQoL. PerkinElmer and Thermo Fisher Scientific did not provide any economic models or any documents that contained relevant data on HRQoL.

HRQoL evidence in the NICE Guideline on Hypertensive Disorders in Pregnancy

The Guideline Development Group (GDG) for the NICE Guideline on hypertensive disorders in pregnancy[13] created several models of gestational hypertension and PE treatment: a model of prophylactic aspirin use for women suspected of having PE (see appendix H in the guideline), a model comparing expectant monitoring or immediate induction for the management of gestational hypertension (see appendix I), a model comparing expectant monitoring to immediate induction for the management of women with PE (see appendix J) and further models for proteinuria measured via dipstick testing (see appendix K) and via automated urinalysis (see appendix L).

In the prophylactic aspirin model of women with suspected PE, the GDG searched the Harvard Cost-Effectiveness Registry for HRQoL data for normotensive women. Only one study was found, by Sonnenberg *et al.*,[106] reporting costs and health effects of contraception usage. The utility values in the study were derived using time trade-off (TTO) methods (rather than the EQ-5D, which is the preferred method of utility measurement in NICE appraisals) from a convenience sample of unreported size.[106] The EAG believes that these utility values are inadequate. The GDG aspirin model assumed that women who developed PE have the same quality of life as normotensive women and assumed that all children discharged alive would live a normal healthy life up to 80 years and have 27.7 discounted QALYs. No utility decrements were assigned for hospitalisations or adverse outcomes related to birth, although it is unlikely that these events would have a significant effect on cost-effectiveness because of the transitory nature of hospitalisation and shared characteristics with regard to hospitalisations within the comparator arms.

The model for hypertension in pregnancy assumed a lower utility score for women with severe disease based on utility for intensive care stay derived from a study on the treatment of severe hospital infections (not pregnancy or hypertensive disorders of pregnancy).[107] Only patients who developed severe disease had lower HRQoL. Further examination of the referenced study shows that the use of utility data from the study is inconsistent in CG107.[13] CG107 reports that severe complications of PE had a utility of 0.019, derived from an Edwards *et al.*[107] utility value divided by 52 and multiplied by two to represent the amount of time spent in intensive care for severe complications of PE. However, the utility decrement presented in

Edwards *et al.*[107] is 0.402. Performing the operations referenced by the GDG on this value does not produce 0.019. It is unclear how the GDG developed the utility measure used in appendix I for severe complications of PE. The PE model in appendix J of the NICE guideline[13] used the same assumptions.

Models in appendices K and L of NICE guideline CG107[13] added maternal death with an assumption of 24.8 discounted QALYs lost for a maternal death. Clinicians consulted by the EAG indicated that maternal death occurs in a vanishingly small proportion of pregnancies and probably did not warrant inclusion in modelling.

All of the models in the NICE guideline[13] assumed full health for mothers and infants after successful birth and loss of full health in the case of mortality. As acknowledged by the GDG, the assumption that women who develop PE have perfect health for the rest of their lives is likely to be an overestimate. Many women who develop gestational hypertension or PE are overweight or already suffering from hypertension before pregnancy. Likewise, the incidence of both developmental disability and cerebral palsy is higher among children born prematurely.[108–110] As acknowledged by the GDG, the guideline models probably overestimate QALYs for these children. In addition, the models used utility scores derived by two different methods, TTO and EQ-5D. It is well known that utility scores from different instruments or methods of measurement produce different results that may not be comparable.[88] The EAG considers that the utility scores used in NICE CG107[13] are inadequate for the purposes of the current assessment. *Table 24* presents the utility values directly from Sonnenberg *et al.*,[106] and Edwards *et al.*,[107] without any modifications; the values do not perfectly match those reported in the NICE guideline[13] and do not perfectly match after performing operations referenced by the NICE guideline.

HRQoL evidence from additional systematic searches

The systematic searches identified nine potentially relevant studies (*Figure 7*). Seven of these were identified directly from database searches. Among the excluded references there were two studies by Bijlenga *et al.*[111,112] that did not report sufficient data to map SF-36 to EQ-5D, and did not report raw scores for EQ-5D. The lead author of these studies was contacted and provided unpublished data for each of the subscales of SF-36 and summary scores for EQ-5D for each trial arm. These data have been included in the review of HRQoL, bringing the total number of included studies to nine (*Table 25*).

Only three studies reported EQ-5D, while six reported HRQoL outcomes that could be mapped to EQ-5D. As more than one source may be necessary to capture HRQoL for women during pregnancy, delivery and the post-partum period, the use of these values in the EAG economic model is assessed further below when considering relevant model parameters (see *Derivation of utility estimates from health-related quality of life*).

Of the nine studies, only one was in the UK,[119] but five were of European populations.[111,112,115,116,119] Four studies had relatively small patient numbers (fewer than 250 total participants).[116–118,120] Only the Petrou *et al.*[119] and the Bijlenga *et al.*[111,112] studies utilised the EQ-5D questionnaire. Petrou *et al.* used the EQ-5D in a general pregnancy population, with two groups of patients classified by spontaneous or non-spontaneous birth.

TABLE 24 Utility values from studies cited by NICE guideline CG107[13]

Study details [author, year; country (instrument)]	Number in state	Health state described	Utility score
Sonnenberg *et al.*, 2004;[106] USA (TTO)	NR	General pregnancy	0.9625[a]
Edwards *et al.*, 2006;[107] UK (EQ-5D)	NR	People in ICU for severe infection (assumption referencing unpublished data for 'unconscious' state)	–0.402
	NR	Patient in HDU/general ward	0.712

HDU, high-dependency unit; ICU, intensive care unit; NR, not reported.
a Utility score calculated by subtracting reported decrement, 0.0375, from 1.

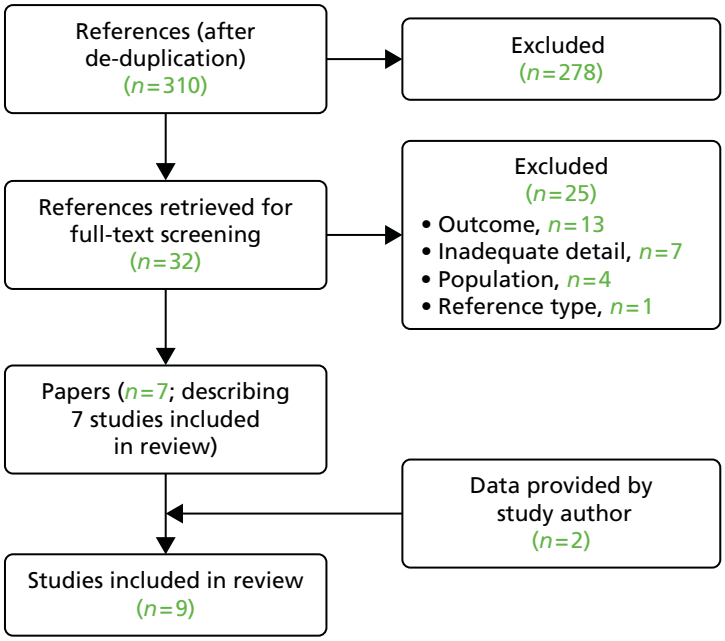

FIGURE 7 Flow chart for the identification of studies on HRQoL.

TABLE 25 Characteristics of candidate HRQoL studies for potential use in the EAG economic model

Author, year	Total number of participants for study (*n*)	Country/territory	Instrument	Health state(s) described
Chang *et al.*, 2013[113]	10,184[a]	Taiwan	SF-36	6 months post partum (general pregnancy population)
Chang *et al.*, 2010[114]	3173[a]	Taiwan	SF-36	6 months post partum (general pregnancy population)
Bijlenga *et al.*, 2011[112]	690	The Netherlands	SF-36 and EQ-5D	36 weeks pregnant; 6 weeks post partum; 6 months post partum (gestational hypertension and/or PE population)
Bijlenga *et al.*, 2011[111]	574	The Netherlands	SF-36 and EQ-5D	36 weeks pregnant; 6 weeks post partum; 6 months post partum (IUGR population)
Hoedjes *et al.*, 2011[115]	128	The Netherlands	SF-36	6 weeks post partum; 12 weeks post partum (mild and severe PE population)
Jansen *et al.*, 2007[116]	141	The Netherlands	SF-36	1 week post partum (vaginal delivery; planned caesarean section; emergency caesarean section populations), 1 and 6 weeks post partum (general pregnancy population)
Torkan *et al.*, 2009[117]	100	Iran	SF-36	6 weeks post partum; 12 weeks post partum (vaginal delivery and caesarean section general pregnancy populations)
Ngai and Ngu, 2013[118]	203	Hong Kong	SF-12	Prenatal (unclear time), 6 weeks post partum, 6 months post partum (general pregnancy population)
Petrou *et al.*, 2009[119]	493	UK	EQ-5D	6 months post partum (spontaneous birth and non-spontaneous birth general pregnancy populations)

a Only the non-intervention arm was included.

Review of economic evidence in the company submissions

Four companies participated in the current diagnostic assessment. Alere and Roche Diagnostics provided economic evidence and economic models, together with information on the sensitivity and specificity of their tests (i.e. the Triage PlGF test and Elecsys sFlt-1 to PlGF ratio test). Although all companies reported the costs of their biomarker tests, PerkinElmer and Thermo Fisher Scientific submitted very limited economic data and did not provide economic models. As noted in *Chapter 4*, PerkinElmer did not provide any information on the sensitivity and specificity of the DELFIA Xpress PlGF test and no relevant test accuracy evidence for this test was identified by the EAG. Thermo Fisher Scientific provided a published study on the sensitivity and specificity of the BRAHMS Kryptor sFlt-1 to PlGF ratio test compared with the Elecsys sFlt-1 to PlGF ratio test,[48] but, as noted in *Chapter 4*, the study was done in a small case–control population (39 with PE and 74 normotensive control subjects), and thus did not meet the inclusion criteria for our systematic review of test accuracy.

The economic models and their supporting documentation submitted by Alere[97,98] and Roche Diagnostics[99] were assessed in detail by the EAG so as to potentially inform the development of the EAG's economic model (see *Methods for the independent economic analysis*). As part of this process, critical summaries of the submissions by Alere and Roche Diagnostics were prepared by the EAG in an earlier version of this report that was made available to the NICE Diagnostics Assessment Committee. These summaries of the companies' submissions contain confidential information and are not included here.

Background to the independent economic analysis

The cost and budget impact models identified in the systematic review of economic evaluations and those submitted by Alere and Roche Diagnostics suffered from a number of limitations: potentially inadequate time horizons; potentially unrepresentative management of gestational hypertension and PE; lack of inclusion of adverse outcomes associated with delivery; lack of long-term costs; and omission of HRQoL-related outcomes. The EAG developed a de novo economic model to address some of these limitations.

The EAG identified potential model structures and model parameters from among the models reviewed and through NICE guidelines, in order to more fully capture the costs and outcomes associated with adding a diagnostic test to current practice in the management of women suspected of having PE. The model will allow the assessment of the cost-effectiveness of the Triage PlGF test and Elecsys sFlt-1 to PlGF ratio test compared with current practice.

The following sections describe the decision problem, the model structure and logic, the parameters used in the model and the initial model results.

Decision problem

The scope developed by NICE states the decision problem as follows:

- What is the cost-effectiveness of diagnostic testing, in addition to standard clinical assessment, for the diagnosis of PE in the second and third trimesters of pregnancy in women presenting with suspected PE between 20 weeks and 36^{+6} weeks of pregnancy?
- What is the cost-effectiveness of PlGF-based diagnostic testing as a replacement for quantitative proteinuria tests in the diagnosis of PE in the second and third trimesters of pregnancy?

The first part of the decision problem is assessed in the base-case analysis, while the second part of the decision problem is assessed in sensitivity analysis only because of lack of any relevant diagnostic accuracy data according to the systematic review of test accuracy.

Strategies and comparators

The reference standard defined in the NICE scope is standard clinical assessment for PE, which is guided by a combination of maternal blood pressure measurement, proteinuria tests, assessment of clinical symptoms

suggestive of the condition and ultrasound fetal growth measurements. The index tests included in the scope for the assessment are diagnostic tests for PE (Triage PlGF test, Elecsys sFlt-1 to PlGF ratio, DELFIA Xpress PlGF test and BRAHMS sFlt-1 Kryptor/PlGF plus Kryptor sFlt-1 to PlGF ratio) in addition to standard clinical assessment. However, as indicated in *Chapter 4*, *Assessment of test accuracy*, evidence of diagnostic test accuracy was only identified for two of the tests: Triage PlGF test and Elecsys sFlt-1 to PlGF ratio. The remaining tests, DELFIA Xpress PlGF test and BRAHMS Kryptor sFlt-1 to PlGF ratio, are not included in the current economic analysis. In addition, no clinical test accuracy evidence to inform the assessment of biomarker tests as alternatives to proteinuria testing was identified (see *Chapter 4*, *Assessment of test accuracy*). Thus, the second part of the decision problem cannot be addressed reliably in the economic analysis.

Methods for the independent economic analysis

Description of the decision analytical model

The EAG developed a decision analytical model to assess the cost-effectiveness of diagnostic tests based on PlGF or sFlt-1 to PlGF ratio test results when used in addition to standard clinical assessment compared with standard clinical assessment alone, in accordance with the scope of the appraisal issued by NICE.

The model was structured to include outcomes identified in the scope issued by NICE for this diagnostic assessment. Suitable data on the sequelae of alternative approaches to the management of suspected PE in pregnancy (in terms of maternal and neonatal morbidity and/or mortality as well as associated resource use) were identified in our systematic review of test accuracy evidence, from company submissions and through targeted searches. The model evaluates costs (GBP using a 2014 price base) from the perspective of the NHS and Personal Social Services. Outcomes in the model are expressed as QALYs. The time horizon of the analysis was < 1 year, corresponding with the length of prebirth monitoring and immediate post-partum monitoring, so no discounting was applied to costs or benefits, in line with current guidance.[69,88,121]

Modelling approach and model structure

The model developed for this assessment was a decision tree, incorporating the management of clinical symptoms of suspected PE, timing and mode of delivery, and maternal and neonatal outcomes. A decision tree was considered more appropriate than a Markov model because events in the model deal with only one pregnancy, as repeat pregnancies are highly uncertain (may not occur, unknown timing and difficult to control for risk factors), removing the need for modelling recurrent events (an advantage of a Markov model) and because there was a paucity of data on long-term outcomes in PE and preterm birth, eliminating the Markov advantage of covering complex future events. With the limitations of the data listed above, a decision tree model structure is sufficient to represent the decision problem.

The decision tree can be considered to have four main structural components:

- Risk stratification (high, intermediate or low risk of PE) of women with suspected PE, determined on the basis of clinical signs, symptoms or findings (*Box 2*) in the case of standard clinical assessment or the same signs, symptoms and clinical findings with the addition of a PlGF-based diagnostic test.
- Pre-eclampsia management (identified as expectant management or immediate delivery based on key symptoms of PE or emergent eclampsia). Expectant management includes monitoring of clinical signs, symptoms and findings (assumed to follow the NICE CG for management of suspected PE[122]), active management of conditions such as hypertension (assumed to follow the NICE CG for management of hypertension in pregnancy[13]) and planned delivery at 37 weeks of gestation. Immediate delivery relates to a requirement to deliver within 24 hours irrespective of gestational age because of clinical findings indicating severe risk to a pregnant woman or fetus, and is the assumed treatment for PE detected after 35 weeks of gestation.
- Maternal outcomes (in terms of admission to intensive care, extended hospital stay and morbidity associated with PE).
- Fetal and neonatal outcomes (in terms of admission to intensive care, extended hospital stay, and morbidity and mortality associated with fetal conditions that may be associated with the underlying cause of maternal PE and/or with early delivery).

BOX 2 Clinical signs, symptoms or findings indicating suspected PE

- New-onset elevated blood pressure.
- Aggravation of pre-existing hypertension.
- New-onset protein in urine.
- Aggravation of pre-existing proteinuria.
- Abnormal uterine perfusion.
- Suspected IUGR.
- Headache.
- Oedema.
- Epigastric pain.
- Visual disturbance.
- Sudden weight gain.
- Low platelets.
- Elevated liver transaminases.

Figure 8 shows a simplified schematic outline of the model, indicating the main structural components. The model distinguishes between suspected PE presenting up to 35 weeks and that presenting from 35 to 37 weeks of gestation. This distinction has been adopted to take account of differing accuracy of biomarker tests according to gestational age. (This relates specifically to sensitivity and specificity of the Triage PlGF test which are reported for pregnancies presenting up to 35 weeks of gestation, for pregnancies between 35 and 37 weeks of gestation and those beyond 37 weeks of gestation. Women beyond 37 weeks of gestation are outside the scope of this assessment; sensitivity and specificity of the Elecsys sFlt-1 to PlGF ratio test are reported for gestation ages between 24 and 37 weeks.)

Risk stratification

The probability of patients with suspected PE being identified as being at high, intermediate or low risk of PE is based on disease prevalence (i.e. the proportion with PE within the suspected PE population), as well as the reported sensitivity and specificity of each diagnostic strategy. These outcomes were modelled based on sensitivity and specificity for standard clinical assessment (using signs/symptoms and clinical findings, i.e. the reference standard) and for assessments including a diagnostic test (signs/symptoms and clinical findings plus a relevant biomarker test, i.e. the index test) reported in studies included in the systematic review of test accuracy (see *Chapter 4*).

In the simple case, when patients are categorised as either positive or negative on the basis of a test with a single cut-off point (assuming a positive test result indicates high risk, a negative test result indicates low risk and, therefore, there is no intermediate risk group) the probability of positive and negative results can be calculated using the formulae below:

$$\text{pPositive} = \text{prevalence} \times \text{sensitivity} + (1 - \text{prevalence}) \times (1 - \text{specificity}) \quad (1)$$

$$\text{pNegative} = \text{prevalence} \times (1 - \text{sensitivity}) + (1 - \text{prevalence}) \times \text{specificity}. \quad (2)$$

These calculations apply for standard clinical assessment where women are identified either as positive or negative for PE based on a combination of signs/symptoms and clinical findings.

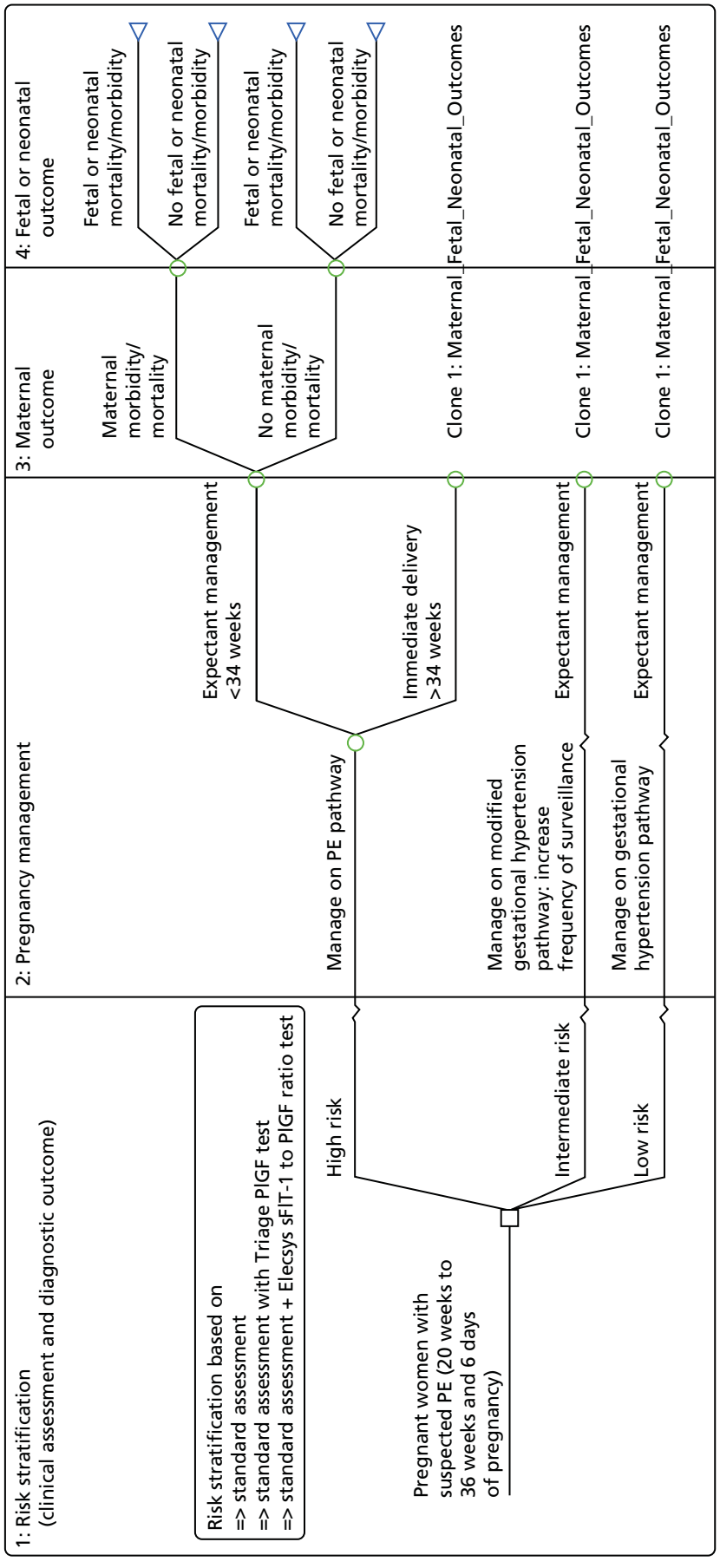

FIGURE 8 Schematic outline of the EAG economic model.

In a situation where a high-risk cut-off point is used to rule in PE and a low-risk cut-off point is used to rule out PE, as is done for the Triage PlGF test, the calculation is modified as shown below:

$$\text{pHighRisk} = \text{prevalence} \times \text{sensitivity}_{\text{RuleIn}} + (1 - \text{prevalence}) \times (1 - \text{specificity}_{\text{RuleIn}}) \quad (3)$$

$$\text{pLowRisk} = \text{prevalence} \times (1 - \text{sensitivity}_{\text{RuleOut}}) + (1 - \text{prevalence}) \times \text{specificity}_{\text{RuleOut}} \quad (4)$$

$$\text{pIntermediate} = 1 - \text{pHighRisk} - \text{pLowRisk}. \quad (5)$$

A similar calculation is applied for the Elecsys sFlt-1 to PlGF ratio test, which has a specific rule-in cut-off point (for PE within 4 weeks) and a specific rule-out cut-off point (for PE within 1 week).

Expectant monitoring/immediate delivery

Clinical guidelines indicate that expectant monitoring is the preferred strategy for pregnancies presenting with suspected PE (particularly those presenting prior to 34 weeks of gestation), with the principal objective of reducing the risk of neonatal respiratory distress syndrome or other complications due to prematurity, while maintaining close monitoring to minimise maternal risk.[123] The ACOG clinical guidance[16] indicates that immediate delivery should be offered for PE between 34 and 36 completed weeks when there is evidence of specific abnormal maternal or fetal findings. Abnormal maternal findings include severe hypertension that is refractory to treatment, persistently abnormal haematological or biochemical findings, severe persistent right upper quadrant or epigastric pain (that is unresponsive to treatment and not accounted for by other diagnosis), pulmonary oedema and cerebral or visual disturbance. The fetal findings include abnormal cardiotocography or evidence of fetal compromise on ultrasound.

The model assumes that women presenting with PE before 35 weeks will be managed using expectant monitoring when there are none of the above signs of increased risk for the mother or neonate. From 35 weeks of gestational age onwards, the model assumes that the pregnancy will be managed by immediate delivery. The time to delivery for both strategies is determined by the PELICAN study.[5] These assumptions are in line with NICE guidance.[13]

Disease status

Maternal and fetal outcomes in the model are assumed to be primarily related to the underlying condition (i.e. the presence or absence of PE). As a result, the outcome components [3: maternal outcome and 4: fetal outcome (see *Figure 8*)] of the model are preceded by an evaluation of true disease status. This is determined on the basis of the proportion of each risk group (high/intermediate/low) identified as having PE as shown:

$$\begin{aligned}
\text{pDisease}_{\text{HighRisk}} &= \frac{\text{prevalence} \times \text{sensitivity}_{\text{RuleIn}}}{\text{pHighRisk}} \\
\text{pDisease}_{\text{LowRisk}} &= \frac{\text{prevalence} \times (1 - \text{sensitivity}_{\text{RuleOut}})}{\text{pLowRisk}} \\
\text{pDisease}_{\text{Intermediate}} &= \frac{\text{prevalence} \times (1 - \text{sensitivity}_{\text{RuleIn}} - (1 - \text{sensitivity}_{\text{RuleOut}}))}{\text{pIntermediate}},
\end{aligned} \quad (6)$$

where $\text{pDisease}_{\text{RiskLevel}}$ is the probability of PE (disease) among women identified as being at the given level of risk during the risk stratification stage of the model.

Maternal outcome

The NICE guideline CG107 (*Box 3*) provides advice on the timing of birth for women with PE.

BOX 3 Advice on the timing of birth for women with PE from NICE CG107

The NICE guideline states the following (see Timing of birth in NICE CG107):[13]

> *Manage pregnancy in women with PE conservatively (that is, do not plan same-day delivery of the baby) until 34 weeks.*
>
> *Consultant obstetric staff should document in the woman's notes the maternal (clinical, biochemical and haematological) and fetal thresholds for elective birth before 34 weeks in women with PE.*
>
> *Consultant obstetric staff should write a plan for antenatal fetal monitoring during birth.*
>
> *Offer birth to women with PE before 34 weeks, after discussion with neonatal and anaesthetic teams and a course of corticosteroids has been given if:*
>
> - *severe hypertension develops refractory to treatment*
> - *maternal or fetal indications develop as specified in the consultant plan.*
>
> *Recommend birth for women who have PE with severe hypertension after 34 weeks when their blood pressure has been controlled and a course of corticosteroids has been completed (if appropriate).*
>
> *Offer birth to women who have PE with mild or moderate hypertension at 34^{+0} to 36^{+6} weeks depending on maternal and fetal condition, risk factors and availability of neonatal intensive care.*
>
> *Recommend birth within 24–48 hours for women who have PE with mild or moderate hypertension after 37^{+0} weeks.*
>
> *National Institute of Health and Clinical Excellence (2010). Adapted from CG107: Hypertension in Pregnancy: Diagnosis and Management.[13] Available from www.nice.org.uk/Guidance/cg107. Reproduced with permission from NICE. The material was accurate at the time of going to press*

Point 6 above recommends that women presenting with PE and mild or moderate hypertension from 34 to 37 weeks be offered immediate delivery. This recommendation was based on the findings of the Dutch Hypertension and Pre-eclampsia Intervention Trial at Term (HYPITAT) I study.[124] The HYPITAT I study was criticised by a member of the NICE CG107 guideline group for including hypertension as one of the outcomes contributing to a composite outcome of maternal morbidity and for not including neonatal outcomes.[125] Awareness of the limitations of the HYPITAT I study may have influenced the formulation of the HYPITAT II study.[124] Clinical experts consulted by the EAG made clear that neonatal outcomes were at least as important as maternal outcomes, and a recently published editorial in the *British Medical Journal* also supports the importance of neonatal outcomes.[123] Consequently, for women presenting between 34 and 37 weeks, we have replaced the values in NICE CG107 (see appendix I)[13] that were from the HYPITAT I study[124] to reflect maternal and neonatal outcomes measured in the HYPITAT II study.[126]

The HYPITAT II study compared expectant monitoring for hypertensive disorders of pregnancy (including PE) in women between 34 and 37 weeks of gestation against induction within 2 days of presentation. In one arm of the trial, women were assigned to induction within 2 days. In the other trial arm, women were expectantly monitored until 37 weeks, at which point these women were induced. Induction at 37 weeks is broadly in line with NICE CG107 on management of women with PE.[13]

Figure 9 shows a simplified schematic for the delivery and maternal outcome component of the model. This stage of the model has been developed following model outlines presented in the NICE *Hypertension*

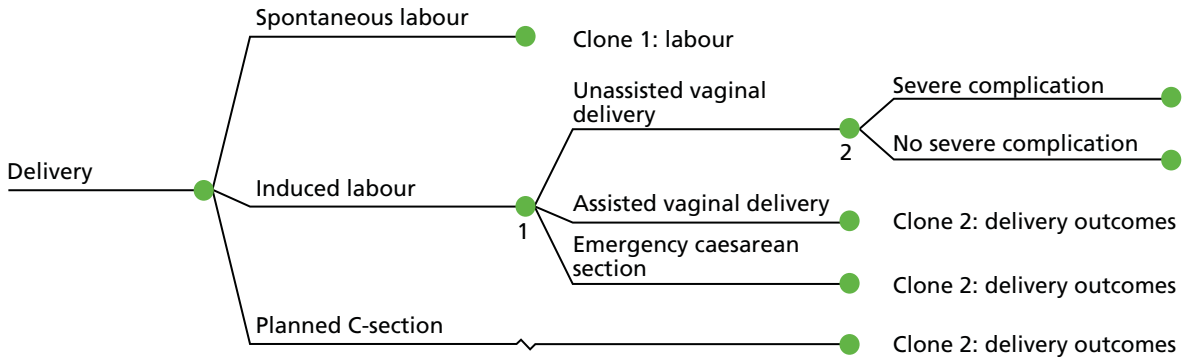

FIGURE 9 Delivery and maternal outcome subtree.

in Pregnancy: The Management of Hypertensive Disorders During Pregnancy CG107[13] (relating to aspirin for prevention of PE, gestational hypertension and PE management; these last two models were primarily used to evaluate cost-effectiveness of immediate delivery compared with expectant management).

The maternal outcome component begins with delivery whether as a result of spontaneous labour (which is expected to be the minority of deliveries in women presenting with suspected PE before 37 weeks of gestation), induced labour or planned caesarean section. This branch corresponds to section 3 of the model in *Figure 8*. This branch has been included in the model on the expectation that delivery costs and outcomes are likely to be highly influenced by mode of delivery, and that the balance of the modes of delivery is likely to differ for women presenting before 35 weeks of gestation and those presenting between 35 and 37 weeks of gestation. We have adopted a slightly different structure to the NICE gestational hypertension and PE management models, by modelling spontaneous and induced deliveries separately. Induction of delivery will be associated with a higher cost than spontaneous delivery because of the need to administer medication to induce labour and a requirement for maternal monitoring during induction. Each of these modes of delivery may be associated with a risk of conversion to assisted/instrumental vaginal delivery or to emergency caesarean section, and the probability of these outcomes may differ according to whether or not labour was initially spontaneous or induced.

Each mode of delivery is associated with a risk of a severe adverse event associated with the progression of severity of PE during the delivery, which results in convulsions. These adverse events confer higher maternal risk, higher risk of admission to intensive or high-dependency care and a requirement for administration of anticonvulsive therapy. The model assumes that women who do not experience convulsions are transferred to the ward following delivery and those who do not experience any further adverse events have a normal length of stay for the given mode of delivery.

Fetal and neonatal outcomes

Figure 10 shows a simplified schematic for the fetal outcome component of the model. This branch corresponds to section 4 of the model in *Figure 8*. As noted for maternal outcomes, this stage of the model has been developed with reference to CG107.[13] The model takes a simplified approach to

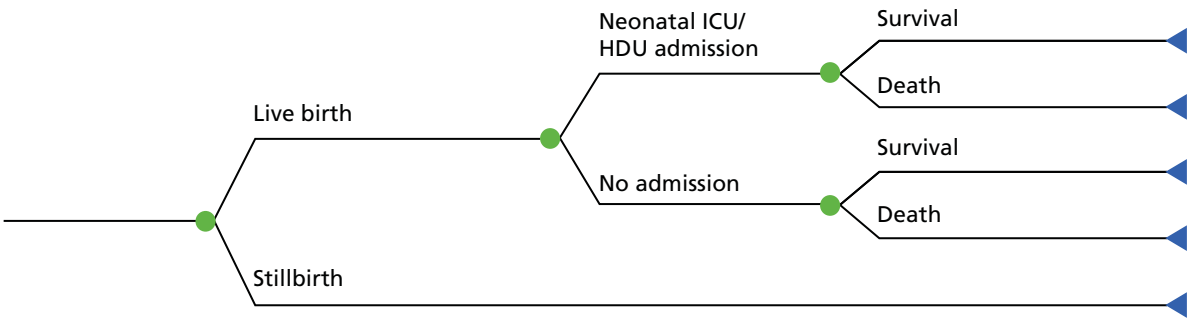

FIGURE 10 Fetal and neonatal outcome subtree. ICU, intensive care unit; HDU, high-dependency unit.

assessing fetal and neonatal outcomes, in which we do not directly model morbidity in terms of clinical manifestations (such as respiratory distress syndrome or sepsis) but instead we model groups of outcomes that may be associated with increased resource use, developmental deficits or differences in HRQoL, and fetal or neonatal mortality. We regard this as a reasonable approach to ensure tractability of the modelling task, but inevitably it involves some simplification of the clinical picture.

The first branch in the neonatal outcome component of the model establishes whether or not the labour results in a live birth or a stillbirth. This is included in the expectation that the probability of stillbirth is likely to be higher in early deliveries, whether or not these occur because of PE. The following branch in the model relates to admission to neonatal high-dependency or intensive care. The probability of admission to neonatal intensive or high-dependency care is expected to be related to gestational age, presence or absence of PE, principal cause of early delivery (maternal condition vs. fetal distress), mode of delivery and the presence or absence of complication(s) during delivery. As a result, this subtree is applied for all risk groups in the model, with or without PE. However, the probability of experiencing neonatal adverse outcomes of delivery may vary according to the factors mentioned above.

The economic model developed for this assessment has been subjected to a number of validation procedures, including assessment of the clinical validity and credibility of the model structure and data used to populate it. The structure of the model has been presented to, and discussed in detail with, a range of clinical experts (see *Acknowledgements*). The model has also been exposed to a range of technical validation exercises. Each component of the model has been tested against published sources (using their data) to establish technical and internal validity of model calculations, as reported in the next section.

Model parameters

The following sections report parameters included in the model. For ease of reference, a list of all model parameters and their sources is provided in *Appendix 8*, and a list of the key model assumptions and their justification is provided in *Appendix 9*.

The model parameters include diagnostic test accuracy, clinical inputs (such as onset of labour, mode of delivery and birth outcomes) and health sector costs (including costs of biomarker tests, antenatal management, delivery and costs of complications). When possible, clinical data were sourced from the PELICAN study,[5] which was conducted in the UK. Targeted searches were conducted to find alternative data sources for model parameters where these were not reported in the PELICAN study.[5] Resource-use assumptions for costing diagnostic and management strategies are presented in full below (based on the companies' suggested approaches to diagnostic testing, NICE CG107[13] and expert opinion). Unit costs were taken from *National Schedule of Reference Costs 2013–14 for NHS Trusts and NHS Foundation Trusts*,[127] *NHS Payment by Results Tariff 2013/14*,[128] and the *British National Formulary*.[120] Targeted searches were conducted to identify unit costs in cases where these sources were inadequate.

Event probabilities

Diagnostic test accuracy estimates used in the model are taken from our systematic review of clinical test accuracy evidence (see *Chapter 4*).

For the Triage PlGF, test diagnostic accuracy is taken from the PELICAN study.[5] The model evaluates both the rule-in and rule-out thresholds (PlGF cut-off point of 12 pg/ml to identify high-risk pregnancies with suspected PE and PlGF cut-off point of 100 pg/ml to identify low-risk pregnancies). Test sensitivity and specificity at the PlGF < 12 pg/ml cut-off point (for pregnancies at less than 35 weeks and weeks 35^{+0} to 36^{+6} of gestation) and at the PlGF < 100 pg/ml cut-off point (for pregnancies at less than 35 weeks) for the PELICAN study[5] are reported in *Table 13*. Sensitivity and specificity at the PlGF < 100 pg/ml cut-off point for pregnancies between 35^{+0} and 36^{+6} weeks of gestation are not reported in a consistent manner in the study publication (i.e. are not reported for PE requiring delivery within 14 days but as PE requiring delivery before 37 weeks). The EAG has based the sensitivity and specificity estimates for the rule-out cut-off point for pregnancies between 35^{+0} and 36^{+6} weeks of gestation on data presented in figure 2 of the study report by Chappell *et al.*,[5] which reports diagnostic outcome by PlGF concentration [categorised as 'abnormal (low)' or 'normal']. These values

correspond to the sensitivity and specificity reported for the < 5th percentile cut-off point (see *Table 14* in *Chapter 4*). *Table 26* reports the sensitivity and specificity used at each PlGF concentration cut-off point for pregnancies at less than 35 weeks of gestation and those between 35^{+0} and 36^{+6} weeks of gestation in the model. Exact CIs for proportions were calculated for test sensitivity, specificity and disease prevalence (see *Appendix 8*, *Summary of model inputs*) and were used in the EAG sensitivity analyses. The EAG assumed that sensitivity and specificity for the Triage PlGF test, evaluating PE requiring delivery within 14 days, were comparable to sensitivity and specificity for the Elecsys sFlt-1 to PlGF ratio test, evaluating any development of PE within 4 weeks. We based this assumption on the similarity of sensitivity and specificity values for the Triage PlGF test using a < 100 pg/ml cut-off point for predicting PE requiring delivery within 14 days in *Table 13* and the < 100 pg/ml cut-off point for diagnosis for any preterm PE in *Table 15*.

For the Elecsys sFlt-1 to PlGF ratio test, diagnostic accuracy is taken from the PROGNOSIS study.[51] The model evaluates both the rule-in and rule-out thresholds, at a cut-off point of 38, for presentation between 24^{+0} and 36^{+6} weeks of gestation. *Table 27* reports the sensitivity and specificity for the rule-in and rule-out thresholds, estimated in the combined development and validation cohorts in the PROGNOSIS study.[51]

Onset of labour, mode of delivery and birth outcomes for women and neonates in women presenting with gestational hypertension or pre-eclampsia up to 34 weeks of gestation

Targeted searches were undertaken to identify studies reporting details of onset of labour, mode of delivery and birth outcomes for women and neonates with PE, gestational hypertension and presenting for assessment up to 34 weeks of gestation, and for studies reporting onset of labour, mode of delivery and birth outcomes in deliveries before 34 weeks of gestation. The EAG identified two studies reporting this information in a manner suitable for inclusion in the model. One, conducted in the UK, is the PELICAN trial reported by Chappell *et al.*,[5] which provided onset of labour, mode of delivery and a range of maternal and fetal outcomes (the last in an

TABLE 26 Diagnostic accuracy of the Triage PlGF test applied in the EAG economic model

Test role	TP	TN	FP	FN	Sensitivity, *n*/*N* (%)	Specificity, *n*/*N* (%)	Prevalence,[a] *n*/*N* (%)
Gestational age at testing: weeks 20^{+0}–34^{+6} (data from* Table 13*)							
Rule in[b]	48	190	21	28	48/76 (63.14)	190/211 (90.05)	76/287 (26.48)
Rule out[c]	72	118	94	3	72/75 (96)	118/212 (55.66)	75/287 (26.13)
Gestational age at testing: weeks 35^{+0}–36^{+6} (data from* Table 14*)							
Rule in[b]	15	64	6	52	15/67 (22.39)	64/70 (91.43)	67/137 (48.91)
Rule out[c]	47	45	25	20	47/67 (70.15)	45/70 (64.29)	67/137 (48.91)

FN, false negative; FP, false positive; TN, true negative; TP, true positive.
a Prevalence of PE requiring delivery within 14 days.
b The rule-in cut-off point relates to a PlGF concentration of ≤ 12 pg/ml.
c The rule-out cut-off point relates to a PlGF concentration of ≥ 100 pg/ml. As noted previously low-risk pregnancies are identified as those with a 'normal' PlGF (≥ 100 pg/ml for gestational age 20^{+0} to 34^{+6} weeks and < 5th percentile for gestational age 35^{+0} to 36^{+6} weeks) with the estimated proportion of low-risk pregnancies, within a population with suspected PE, calculated based on the prevalence of PE, and test sensitivity and specificity ($pLowRisk = prevalence \times (1 - sensitivity_{RuleOut}) + (1 - prevalence) \times specificity_{RuleOut}$).

TABLE 27 Diagnostic accuracy of the Elecsys sFlt-1 to PlGF ratio test applied in the EAG economic model

Test role	Sensitivity	Specificity
Gestational age at testing: weeks 20^{+0}–36^{+6} (data from* Table 17*)		
Rule in	0.703 (95% CI 0.619 to 0.778)	0.831 (95% CI 0.805 to 0.855)
Rule out	0.857 (95% CI 0.728 to 0.941)	0.791 (95% CI 0.765 to 0.816)

online supplement to the published study) for gestational age groups < 35^{+0} weeks, 35^{+0}–36^{+6} weeks and ≥ 37^{+0} weeks. *Table 28* reports data extracted from this study for the < 35^{+0} weeks of gestational age group.

The second study, EPIPAGE,[110] reported outcomes by gestational week, between 30 and 34 weeks, for births occurring in maternity units in nine French regions in 1997. The study enrolled 2467 infants. Data are relatively complete for outcomes evaluated while in the hospital (including stillbirths, in-hospital deaths, and admission to and duration of neonatal intensive care), but provide comparatively little information on mode of delivery. *Table 29* reports data extracted from this study.

TABLE 28 Onset of labour, mode of delivery and fetal/neonatal outcomes for women presenting before 35 weeks of gestation in the PELICAN study

Parameter	Value
Gestational age at assessment (weeks)	31
Fetal/neonatal outcome (probability)	
Stillbirth/fetal death	0.024
In-hospital (neonatal) death	0.007
Admission to neonatal intensive care (for > 48 hours)	0.040
With PE	0.048
Without PE	0.027
Characteristics of labour/delivery (probability)	
Onset of labour	
Spontaneous	0.148
Induction	0.380
Planned caesarean section	0.472
Mode of delivery	
Spontaneous vaginal delivery	0.265
Assisted vaginal delivery	0.114
Caesarean section	0.621
Gestational age at delivery (weeks)	36.7

TABLE 29 Neonatal outcomes reported by the EPIPAGE study

	Gestational age at birth (weeks)					
Neonatal outcome	30	31	32	33	34	Pooled % (SE)
Stillbirth, % (*n*/*N*)	17.4 (88/507)	13.2 (84/635)	11.4 (100/878)	5.1 (11/214)	5.3 (13/243)	11.9 (0.0067)
In-hospital death, % (*n*/*N*)	8.1 (34/419)	4.5 (25/551)	2.7 (21/778)	2.5 (5/203)	0.4 (1/230)	3.9 (0.0044)
Admission to neonatal intensive care, % (*n*/*N*)	84.8 (341/402)	78.9 (416/527)	64.8 (474/731)	45.4 (83/183)	27.2 (58/213)	66.7 (0.0106)
Duration (days) of stay in intensive care, mean (SD)	15 (17)	10 (13)	7 (11)	3 (7)	2 (10)	8.46 (3.43)
Caesarean delivery, % (*n*/*N*)	59.7 (245/410)	61.2 (332/542)	63.7 (488/766)	55.4 (109/197)	47.0 (104/221)	59.8 (0.01084)

SD, standard deviation; SE, standard error.

Onset of labour, mode of delivery and birth outcomes in women presenting with gestational hypertension or pre-eclampsia between 34 and 37 weeks of gestation

We conducted targeted searches to identify studies reporting details of onset of labour, mode of delivery and birth outcomes for women and neonates with PE, gestational hypertension and presenting for assessment between 34 and 37 weeks of gestation. This included appraisal of data sources used in the development and population of the economic models reported in the NICE CG for management of hypertension in pregnancy.[122] The searches identified four studies reporting this information in a manner suitable for inclusion in the model. As noted, the publication of the PELICAN trial by Chappell *et al.*[5] reports this information for the appropriate gestational age groups. *Table 30* reports data extracted from this study for the 35^{+0} to 36^{+6} weeks of gestational age group.

Table 31 reports neonatal outcomes incorporated into the model for births occurring between 35 and 37 weeks of gestation.[129] These data are from a secondary analysis of outcomes between 35 and 37 weeks of gestation from a US multicentre randomised controlled trial (conducted between 1992 and 1995) of

TABLE 30 Onset of labour, mode of delivery and fetal/neonatal outcomes for women presenting after 35 weeks of gestation in the PELICAN study

Parameter	Value
Gestational age at assessment (weeks)	36
Fetal/neonatal outcome (probability)	
Stillbirth/fetal death	0.000
In-hospital (neonatal) death	0.000
Admission to neonatal intensive care (for > 48 hours)	0.066
With PE	0.077
Without PE	0.050
Characteristics of labour/delivery (probability)	
Onset of labour	
Spontaneous	0.184
Induction	0.551
Planned caesarean section	0.265
Mode of delivery	
Spontaneous vaginal delivery	0.412
Assisted vaginal delivery	0.099
Caesarean section	0.489
Gestational age at delivery (weeks)	37.3

TABLE 31 Neonatal outcomes for births occurring between 35 and 37 weeks of gestation in the CPEP trial

	Gestational age at birth (weeks)			
Neonatal outcome	35	36	37	Pooled estimate for PE
Stillbirth	Not reported			Not applicable
In-hospital death				
Admission to intensive care, % (*n*/*N*)	57.1 (16/28)	33.3 (14/42)	25.6 (22/86)	33.3 (0.409)
Duration (days) of stay in intensive care, mean (SD)	5.3 (4)	10.3 (8.6)	5.7 (5)	6.87 (2.02)

SD, standard deviation.

Calcium for Pre-eclampsia Prevention (CPEP) in healthy nulliparous women with singleton pregnancies. The study enrolled 4589 women, with complete follow-up available for 4293. Of these, 3229 remained normotensive, while 1064 developed gestational hypertension or PE. Of these, 379 normotensive women and 156 women with gestational hypertension or PE delivered between 35 and 37 weeks of gestation. Outcomes by week of gestation are reported in *Table 31*.

The data presented in the PELICAN study[5] do not distinguish characteristics of labour and delivery for women with suspected PE who require immediate delivery and those who are offered expectant monitoring. Two trials (HYPITAT I[124] and HYPITAT II[126]) were identified that reported these characteristics separately, for women with gestational hypertension or PE who were randomised to immediate delivery (within 48 hours) and those who were randomised to expectant monitoring. The HYPITAT I study has been criticised on the basis of its composite outcome of maternal morbidity and for not including neonatal outcomes. The HYPITAT II study compared expectant monitoring for hypertensive disorders of pregnancy (including PE) in women between 34 and 37 weeks of gestation, and reported a wider range of maternal and neonatal outcomes. In one arm of the trial, women were assigned to induction within 2 days. In the other trial arm, women were expectantly monitored until 37 weeks, at which point these women were induced. Induction at 37 weeks is broadly in line with the NICE CG107[13] on management of women with PE.

The primary neonatal outcome for the HYPITAT II study[126] was the probability of neonatal respiratory distress syndrome. After appraising data from the HYPITAT II study,[126] and consulting clinical experts to determine whether or not neonatal morbidity had any probable effects on quality of life for mothers or neonates, the EAG decided only to model the rate of NICU admission to capture the effects of neonatal morbidity. This decision was supported by three experts. Experts indicated that long-term differences in morbidity between neonates born before 35 weeks and from 35 to 37 weeks of gestation were not likely to be significant. Thus, the EAG, with support of expert opinion, has assumed that NICU costs adequately capture the effects of neonatal morbidity.

The primary composite outcome for women in the HYPITAT II trial was a combination of thromboembolic processes, pulmonary oedema, HELLP syndrome, eclampsia and placental abruption. There were very few events, with HELLP syndrome being the most common.[126] Experts indicated that there were unlikely to be any significant long-term differences in maternal morbidity based on data in the HYPITAT II trial.[126]

Cost of biomarker tests and antenatal management

Alere submitted two different costs for the Triage PlGF test; the lower cost was reported in the economic model reported by Duckworth *et al.*[98] and the higher cost was provided as a separate value in response to NICE questions during scoping. The lower cost per test was based solely on the cost of an Alere Triage PlGF cassette of 25 tests (£1000) and does not include any costs for buying equipment, additional consumables or equipment maintenance. Given that implementation will vary according to whether or not hospital and community locations will own and maintain a Triage MeterPro PlGF testing machine, the EAG has used in our model the higher cost per test provided in the Duckworth *et al.*[98] cost model (confidential data). In the base-case analysis, the cost of the test assumes that all tests will be conducted in a central laboratory. Near-patient testing, for which it is assumed that the test is delivered in a midwife-led day-case unit, is explored as a scenario analysis in *Scenario analyses*.

Roche Diagnostics submitted two costs for the Elecsys sFlt-1 to PlGF ratio test (confidential data). The EAG has used the higher of these costs per test in our economic model. The testing strategy in Hadker *et al.*[39] assumes that testing is repeated twice for those who initially test negative. Similarly, the product insert recommends re-testing[83] and the economic model submitted by Roche Diagnostics assumes one potential re-test. The base-case analysis assumes no re-testing, as there are no published sensitivity or specificity values available for multiple testing.

The economic model is partially based on the NICE *Hypertension in Pregnancy: The Management of Hypertensive Disorders During Pregnancy* guideline (CG107).[13] CG107 defines management of gestational hypertension and its associated resource use by whether patients have mild, moderate or severe hypertension. *Table 32* shows the prescribed resource use for each hypertension category.

TABLE 32 Resource use associated with management of hypertension in pregnancy

Degree of hypertension	Location of monitoring/treatment	Type of care	Frequency	Activities/assessments
Mild[a]	Community	Monitoring	Weekly	Blood pressure
			Weekly	Proteinuria test (dipstick)
			Weekly	Standard blood tests
Moderate[b]	Community	Monitoring	Twice weekly	Blood pressure
			Twice weekly	Proteinuria test (dipstick)
			Weekly	Standard blood tests
			Once (continue if proteinuria at subsequent visits)	Kidney function + electrolytes + full blood count + transaminases + bilirubin
		Treatment	Once	Oral labetalol
Severe[c]	Hospital[d]	Admission	Per day (3 days)	Hospitalisation cost
		Monitoring	> 4 times per day	Blood pressure
			Daily	Proteinuria test
			Weekly	Standard blood tests + kidney function + electrolytes + full blood count + liver function tests + bilirubin
		Treatment	Once	Oral labetalol

a 140/90–149/99 mmHg.
b 150/100–159/109 mmHg.
c 160/110 mmHg or higher.
d Until blood pressure ≤ 159/109 mmHg, then manage as moderate.

In order to identify unit costs for resource use for the clinical pathway developed in the guideline, *National Schedule of Reference Costs 2013–14 for NHS Trusts and NHS Foundation Trusts* [127] was searched along with the most recent online version of the *British National Formulary* (accessed June 2015).[120] The EAG assumed that the costs of dipstick proteinuria testing would be subsumed within the cost of a routine outpatient antenatal appointment as the cost for the test is negligible. Similarly, the costs of blood pressure monitoring and proteinuria monitoring were subsumed within the cost of a hospital stay for patients with severe hypertension. The length of stay for severe hypertension was based on expert opinion. The cost of this stay was derived from *NHS Payment by Results Tariff 2013/14*.[128] Based on data from the PELICAN trial, reported by Chappell *et al.*,[5] women had varying lengths and intensities of monitoring based on their hypertension status. Women with mild to moderate hypertension were monitored for 8 weeks, while women with severe hypertension were hospitalised for 3 days and then monitored for moderate hypertension for 7 weeks. The EAG assumed that women managed under the gestational hypertension pathway receive two oral labetalol prescriptions.

It was unclear which cost classifications in the NHS reference costs were most appropriate for the costs of blood tests. In order to assess which cost classifications would be appropriate to use for costing blood tests, an internet search was conducted to identify these values. A study by Akhtar and Chung[130] was identified that provided cost estimates for three varieties of blood test panels: full bloods, liver function tests, and urea and electrolytes. The cost values from these tests were applied to the guideline pathway resource use descriptions. *Table 33* presents the unit costs identified for the NICE CG107 pathway;[13] location of monitoring/treatment and type of care, and the frequency of each, correspond to the values given in *Table 32*.

TABLE 33 Unit costs for hypertension in the NICE pregnancy management pathway

Degree of hypertension	Activities/assessments	Unit cost	Source	Notes
Mild[a]	Blood pressure	£49	*National Schedule of Reference Costs 2013–14 for NHS Trusts and NHS Foundation Trusts*[127]	NZ16Z (outpatient procedure)
	Proteinuria test (dipstick)			
	Standard blood tests	£2.65	Akhtar and Chung[130]	Full bloods
Moderate[b]	Blood pressure	£49	*National Schedule of Reference Costs 2013–14 for NHS Trusts and NHS Foundation Trusts*[127]	NZ16Z (outpatient procedure)
	Proteinuria test (dipstick)			
	Standard blood tests	£2.65	Akhtar and Chung[130]	Full bloods
	Kidney function + electrolytes + full blood count + transaminases + bilirubin	£4.90	Akhtar and Chung[130]	Liver function test, and urea and electrolytes
	Oral labetalol	£6.99	BNF (56 edition)[120]	
Severe[c,d]	Hospitalisation cost	£740	*NHS Payment by Results Tariff 2013/14*[128]	NZ07C (non-elective inpatient short stay)
	Blood pressure			
	Proteinuria test			
	Standard blood tests + kidney function + electrolytes + full blood count + transaminases + bilirubin	£7.55	Akhtar and Chung[130]	Full bloods, liver function test, and urea and electrolytes
	Oral labetalol	£6.99	BNF (56 edition)[120]	

BNF, *British National Formulary*.
a 140/90–149/99 mmHg.
b 150/100–159/109 mmHg.
c 160/110 mmHg or higher.
d Until blood pressure ≤ 159/109 mmHg, then manage as moderate.

Resource use for management of women diagnosed with PE is higher than for those diagnosed only with gestational hypertension. In accordance with the NICE CG107,[13] women with PE who have any level of hypertension are to be admitted to hospital. Women with moderate or severe hypertension should be treated with oral labetalol to keep diastolic blood pressure between 80 and 100 mmHg and systolic blood pressure < 150 mmHg. In the model, we have assumed that women receive oral labetalol until delivery, according to expert clinicians' advice. At a dose of 100–200 mg twice daily, 56 100-mg pills should be sufficient for this time period. Women diagnosed with PE do not receive further proteinuria testing.

To determine the length of stay for women diagnosed with PE, the time to delivery for women with early-onset PE (< 35 weeks) and the time to delivery for PE with an onset between 35 and 37 weeks from the PELICAN trial were used.[5] For women presenting with PE before 35 weeks, it was assumed that the length of hospitalisation was 9 days, while women presenting with PE between 35 and 37 weeks would be hospitalised for 4 days.[5] *Table 34* presents the unit costs and activity frequencies for management of women with a diagnosis of PE. Payment by results tariffs were used for cost of hospital stay because the data set more closely corresponds to the expected length of stay of the model population, as defined by the PELICAN study and expert opinion. Experts informed the EAG that women examined for suspected severe PE would be expected to be hospitalised for 3 days. The payment by results tariff assesses stays of 0–4 days, with the ability to extend costs with additional days. This feature allows the same cost data to be used for shorter and longer stays, maintaining consistent costs between those who stay 4 and 9 days in the model. If NHS reference costs are used there are only two applicable cost values: £491 for a short stay (NES NZ16Z), and £1036 for a long stay (stated as an average length of 1 day in NHS reference costs, NES NZ16Z), which is undesirably inflexible.

TABLE 34 Unit costs for management of women diagnosed with PE

Degree of hypertension	Activities/assessments	Unit cost	Source	Notes
Mild[a]	Hospitalisation cost	£2248 (9 days) £740 (4 days)	*NHS Payment by Results Tariff 2013/14*,[128] HRG code NZ07C	Assume 9 days until delivery for PE based on the PELICAN trial before 35 weeks, and 4 days after[5]
	Standard blood tests + kidney function + electrolytes + full blood count + liver function tests + bilirubin	£7.55	Akhtar and Chung[130]	Full bloods, liver function test, and urea and electrolytes
Moderate[b]	Hospitalisation cost	£2248 (9 days) £740 (4 days)	*NHS Payment by Results Tariff 2013/14*,[128] HRG code NZ07C	Assume 9 days until delivery for PE before 35 weeks, and 4 days after[5]
	Standard blood tests + kidney function + electrolytes + full blood count + liver function tests + bilirubin	£7.55	Akhtar and Chung[130]	Full bloods, liver function test, and urea and electrolytes
	Oral labetalol	£6.99	BNF (56 edition)[120]	
Severe[c,d]	Hospitalisation cost	£2248 (9 days) £740 (4 days)	*NHS Payment by Results Tariff 2013/14*,[128] HRG code NZ07C	Assume 9 days until delivery for PE before 35 weeks, and 4 days after[5]
	Standard blood tests + kidney function + electrolytes + full blood count + liver function tests + bilirubin	£7.55	Akhtar and Chung[130]	Full bloods, liver function test, and urea and electrolytes
	Oral labetalol	£6.99	BNF (56 edition)[120]	

BNF, *British National Formulary*; HRG, Healthcare Resource Group.
a 140/90 to 149/99 mmHg.
b 150/100 to 159/109 mmHg.
c 160/110 mmHg or higher.
d Until blood pressure ≤ 159/109 mmHg, then manage as moderate.

The NICE CG107[13] modelled aspirin use in PE based on evidence that there was some benefit for women at risk of PE. The EAG consulted clinical experts to assess whether or not it was necessary to model aspirin treatment. The consensus was that during the time period that the EAG is modelling (20–37 weeks of gestation) women at high risk of PE should have already begun receiving aspirin and would not receive a new prescription for aspirin during this time. Therefore, we did not model aspirin therapy.

The EAG also sought further clarification on the use of ultrasound in monitoring and diagnosing PE. Clinical experts indicated that ultrasound usage is determined primarily by the status of the fetus, something that was unknown in the studies of the diagnostic tests. In addition, the experts gave a wide range of values for how often ultrasound should be conducted, varying from once every 2 weeks to four times per week. The EAG believes that frequency of ultrasound is likely to be highly variable, and unlikely to be a major cost driver in the model; therefore, frequency of ultrasound was not modelled.

Cost of birth and maternal and neonatal outcomes by timing of birth

The EAG model assumes that the unit costs associated with birth are not dependent on whether the mother has hypertension or PE. The frequency of different birth types (i.e. spontaneous, assisted or caesarean) is associated with whether a birth is managed through immediate delivery or expectant monitoring. However, the unit cost of a delivery type is fixed. To calculate the different costs for different

types of deliveries, the EAG created weighted averages for each type of delivery, including births with no complications and births with many complications based on NHS Reference Costs.[127] Upon expert advice, the EAG assumed that there was no difference in rates of complications between births to mothers with gestational hypertension or PE and women with normal pregnancies. In addition to the costs for delivery type, costs of maternal and neonatal intensive care, high-dependency unit stays and ward stays were calculated in a similar manner, based on NHS reference costs.[127] The *British National Formulary* was consulted for drug costs[120] (*Table 35*).

TABLE 35 Unit costs associated with birth and in-hospital monitoring independent of birth timing

Outcome	Unit cost	Source	Notes or reference costs codes
Spontaneous birth (normal)	£1506	*National Schedule of Reference Costs 2013–14 for NHS Trusts and NHS Foundation Trusts*[127]	NZ30A–NZ30C
Spontaneous birth (assisted delivery)	£1988	*National Schedule of Reference Costs 2013–14 for NHS Trusts and NHS Foundation Trusts*[127]	NZ40A–NZ40C
Induced birth (normal)	£2133	*National Schedule of Reference Costs 2013–14 for NHS Trusts and NHS Foundation Trusts*[127]	NZ31A–NZ32C
Induced birth (assisted delivery)	£3033	*National Schedule of Reference Costs 2013–14 for NHS Trusts and NHS Foundation Trusts*[127]	NZ42A–NZ44C
Planned caesarean section	£3182	*National Schedule of Reference Costs 2013–14 for NHS Trusts and NHS Foundation Trusts*[127]	NZ50A–NZ50C
Emergency caesarean section	£4013	*National Schedule of Reference Costs 2013–14 for NHS Trusts and NHS Foundation Trusts*[127]	NZ51A–NZ50C
First 5 days of hospitalisation for antenatal or postnatal monitoring	£740	*NHS Payment by Results Tariff 2013/14*,[128] HRG code NZ07C	NZ07C
Each additional day of antenatal or postnatal monitoring	£377	*NHS Payment by Results Tariff 2013/14*,[128] HRG code NZ07C	NZ07C
Maternal critical care, intensive care unit	£1449	*National Schedule of Reference Costs 2013–14 for NHS Trusts and NHS Foundation Trusts*[127]	XC04Z
Neonatal critical care, high-dependency unit	£839	*National Schedule of Reference Costs 2013–14 for NHS Trusts and NHS Foundation Trusts*[127]	XA02Z
Neonatal critical care, intensive care unit	£1118	*National Schedule of Reference Costs 2013–14 for NHS Trusts and NHS Foundation Trusts*[127]	XA01Z
Magnesium sulphate (intravenous)	£16.98 (4-g injection) £7.30 (1-g infusion)[a]	BNF 2015[120]	One dose of 4-g injection and then 1 g/hour infusion for at least 24 hours
Labetalol (oral) 100 mg twice daily with food, increased to maximum 200 mg with titration over 14-day interval	£6.99	BNF 2015[120]	100 mg, 56 tablets

BNF, *British National Formulary*; HRG, Healthcare Resource Group.
a Costed as five 10-ml (5-g) ampoules at £1.46 each.

Derivation of utility estimates from HRQoL

As reported in *Review of health-related quality-of-life studies*, the EAG conducted a series of systematic searches to identify HRQoL data in women with gestational hypertension, PE or outcomes of pregnancy that are related to gestational hypertension and PE. The searches sought to identify studies reporting EQ-5D data or data that could be mapped to EQ-5D. Instruments that could be mapped to EQ-5D were identified using the Oxford University Health Economics Research Centre Database of Mapping Studies.[89] The purpose of the review was to identify utility scores to represent patient quality of life in the health states used in the model, and then apply these utility scores to time spent in these health states to produce QALYs. The values and health states found through the review are presented in *Table 36*.

The searches identified seven studies, and follow-up with authors resulted in two additional studies being included (see *Table 36*). Only three instruments were identified among studies included for potential use in the model: the SF-36, SF-12 and EQ-5D. The EAG used an algorithm from Ara and Brazier[131] to map the eight domains of the SF-36 to obtain EQ-5D-3L scores, and an algorithm from Sullivan and Ghushchyan[132] to map SF-12 scores to obtain EQ-5D-3L scores (no mapping algorithms to the EQ-5D-5L were available at the time of this assessment). Seven studies reported SF-36 data.[111–117] One study reported SF-12 data.[118] Three studies reported EQ-5D values: Petrou *et al.*[119] and two studies by Bijlenga *et al.*[111,112] Both studies by Bijlenga *et al.*[111,112] also included SF-36 data. All data that could be were mapped to EQ-5D-3L and are shown in *Table 37*. The characteristics of these HRQoL studies are reported in *Table 25*.

In general, the studies showed that women largely recovered to pre-pregnancy HRQoL by 6 months post partum, regardless of the method of delivery and whether or not hypertensive disorder of pregnancy was present. This led the EAG to decide to model HRQoL outcomes only over the time period between birth and 6 months post partum. Mapping both Chang *et al.*[113,114] studies gave utility scores that were above 0.92 at 6 months for general pregnancy (see *Table 36*). Mapping Hoedjes *et al.*[115] gave utility scores that were over 0.93 at 12 weeks post partum for women with mild PE. This appears to support the theory that there is little or no utility difference between women with an average pregnancy and a pregnancy complicated by mild PE at 6 months post partum. The Chang *et al.* studies[113,114] were conducted in Taiwan, and the Hoedjes *et al.* study[115] was conducted in the Netherlands, but mapping both from SF-36 to EQ-5D supports that the utility score for women at 6 months post partum is approximately 0.92. A close value, 0.89, was obtained by mapping SF-12 values to EQ-5D utility scores at 6 months post partum from Ngai *et al.*,[118] a Hong Kong study (see *Table 37*). Likewise, the Dutch study, by Jansen *et al.*,[116] produced a mean utility score of 0.89 when mapped from SF-36 for women at 6 weeks post partum. The mapped values appear to support that women have an almost full recovery at 6 weeks post partum, and a full recovery to pre-pregnancy HRQoL by 6 months.

With the exception of the Iranian study by Torkan *et al.*,[117] which used the SF-36 (see *Table 36*), the mapped values from SF-36 and SF-12 appeared to be consistent across studies. The Torkan *et al.* study was small, and the environment in which post-partum women live in Iran is unlikely to be very comparable to that in the UK.

All the utility scores presented in *Tables 36* and *37* are from complete cases. All patients for whom any data were missing were excluded from the data analysis. It is unclear what effect this could have on HRQoL, but there is potential that some of the estimates may be biased if some of the data that have been excluded from the studies, owing to not having follow-up for all times or for having any missing values, are not missing at random. The number of missing data was not reported in all studies. This is a potential limitation for all of the HRQoL data used in the economic model.

The EAG evaluated the utility scores in *Table 37* and selected values for use in the model as shown in *Table 38*. While there were directly measured EQ-5D values available for baseline health in each of the three time periods contained in the model, most of the decrements (the drivers of difference between the diagnostic tests) were derived from SF-36 mapped to EQ-5D using Ara and Brazier's algorithm.[131] Values produced through mapping were remarkably consistent across all health states; this was not observed in

TABLE 36 EQ-5D utility values mapped from SF-36

Study details (authors, year; country/territory)	Number in health state	Health state	SF-36 domain								EQ-5D mapping[a]
			Physical functioning	Role physical	Bodily pain	General health	Vitality	Social functioning	Role emotional	Mental health	
Chang *et al.*, 2013;[113] Taiwan	10,184	6 months post partum	97.54	91.98	86.76	77.21	62.48	88.93	83.76	70.83	0.9204
Chang *et al.*, 2010;[114] Taiwan	3173	6 months post partum	97.37	91.43	87.68	77.79	63.10	88.86	81.97	71.52	0.9237
Bijlenga *et al.*, 2011;[112] the Netherlands	690	36 weeks pregnant (PE or GH)	47.56	21.40	66.89	74.14	52.64	68.50	83.53	76.95	0.6932
	225	6 weeks post partum (induced)	85.97	52.88	75.21	76.87	57.51	76.11	81.48	80.77	0.8676
	305	6 weeks post partum (expectant monitoring)	84.6	50.01	75.67	74.19	57.36	75.87	78.82	79.56	0.8593
	192	6 months post partum (induction)	86.08	53.12	75.17	77.08	57.52	76.13	81.70	80.87	0.8683
	275	6 months post partum (expectant monitoring)	84.57	49.94	75.68	74.12	57.36	75.87	78.75	79.53	0.8590
Bijlenga *et al.*, 2011;[111] the Netherlands	547	36 weeks pregnant (IUGR)	57.95	32.81	68.99	74.33	54.68	70.07	77.55	73.25	0.7245
	165	6 weeks post partum (induced)	85.72	52.36	75.29	76.38	57.48	76.07	81.00	80.55	0.8661
	141	6 weeks post partum (expectant monitoring)	85.66	52.22	75.32	76.25	57.48	76.05	80.87	80.49	0.8657
	237	6 months post partum (induction)	84.35	49.48	75.75	73.70	57.34	75.83	78.33	79.34	0.8577
	192	6 months post partum (expectant monitoring)	84.34	49.45	75.76	73.67	57.34	75.83	78.30	79.33	0.8576

Study details (authors, year; country/territory)	Number in health state	Health state	SF-36 domain								EQ-5D mapping[a]
			Physical functioning	Role physical	Bodily pain	General health	Vitality	Social functioning	Role emotional	Mental health	
Hoedjes *et al.*, 2011;[115] the Netherlands	33	6 weeks post partum (mild PE)	86.40	56.30	77.10	76.80	57.60	78.00	79.80	81.90	0.8783
	95	6 weeks post partum (severe PE)	76.90	41.90	63.70	67.80	48.40	58.80	63.60	69.50	0.7524
	33	12 weeks post partum (mild PE)	90.50	76.20	86.90	76.40	67.40	89.40	87.60	84.90	0.9325
	95	12 weeks post partum (severe PE)	86.90	68.60	83.30	67.10	60.20	78.00	75.00	75.90	0.8717
Jansen *et al.*, 2007;[116] the Netherlands	71	1 week post partum (vaginal delivery)	59.40	26.90	44.70	76.50	55.80	64.60	70.40	81.90	0.6766
	36	1 week post partum (elective caesarean section)	45.00	18.00	38.80	75.90	53.80	53.00	74.70	78.70	0.5895
	34	1 week post partum (emergency caesarean section)	32.80	16.40	35.60	72.20	42.40	47.00	54.00	73.80	0.5167
	141	1 week post partum	48.80	22.00	40.90	75.20	51.70	57.20	67.00	79.00	0.6132
	NR	6 weeks post partum	85.40	73.80	78.20	78.00	68.20	86.20	82.70	86.30	0.8885
Torkan *et al.*, 2009;[117] Iran	50	6 weeks post partum (vaginal delivery)	79.50	42.50	62.80	73.10	62.90	68.20	50.60	75.10	0.7746
	50	12 weeks post partum (vaginal delivery)	88.40	59.50	71.90	75.70	61.10	70.50	62.00	74.70	0.8362
	50	6 weeks post partum (caesarean section)	77.20	33.00	62.10	77.80	54.40	63.20	38.00	66.70	0.7441
	50	12 weeks post partum (caesarean section)	81.50	50.00	70.70	73.80	64.30	71.50	60.60	71.50	0.7991

GH, gestational hypertension; NR, not reported.
a The mapping equation for SF-36 to EQ-5D is derived from Ara and Brazier.[131]

TABLE 37 EQ-5D utility values derived directly and mapped from SF-12

Study details (author, year; country/territory)	Number in health state	Health state	Component summary score		EQ-5D mapping[a]	EQ-5D direct
			Physical	Mental		
Ngai and Ngu 2013;[118] Hong Kong	203	General pregnancy	43.1	48.7	0.8230	–
	203	6 weeks post partum	47.7	47.00	0.8518	–
	203	6 months post partum	51.4	48.1	0.8914	–
Bijlenga *et al.*, 2011;[112] the Netherlands[b]	616	36 weeks pregnant (PE or GH)	–	–	–	0.5947
	222	6 weeks post partum (induced)	–	–	–	0.7430
	192	6 weeks post partum (expectant monitoring)	–	–	–	0.7366
	300	6 months post partum (induction)	–	–	–	0.7435
	272	6 months post partum (expectant monitoring)	–	–	–	0.7365
Bijlenga *et al.*, 2011;[111] the Netherlands[c]	457	36 weeks pregnant (IUGR)	–	–	–	0.6326
	160	6 weeks post partum (induced)	–	–	–	0.7421
	138	6 weeks post partum (expectant monitoring)	–	–	–	0.7417
	234	6 months post partum (induction)	–	–	–	0.7354
	192	6 months post partum (expectant monitoring)	–	–	–	0.7353
Petrou *et al.*, 2009;[119] UK	493	6 months post partum (spontaneous birth)	–	–	–	0.8670
	493	6 months post partum (non-spontaneous birth)	–	–	–	0.8470

GH, gestational hypertension.
a The mapping equation for SF-12 to EQ-5D is derived from Sullivan and Ghuschchyan.[132]
b This study provided both SF-36 data that have been mapped to EQ-5D in *Table 36* and directly measured EQ-5D data.
c This study provided only directly measured EQ-5D data.

the directly measured EQ-5D values. The EAG decided to use mapped utilities for most utility values to maintain a consistent scale across the health states. Petrou *et al.*'s[119] directly measured EQ-5D value was used for a utility decrement because otherwise the long-term effects of different modes of delivery could not be adequately captured. In addition, Petrou *et al.*'s[119] directly measured EQ-5D utilities are more similar to Bijlenga *et al.*'s[111,112] mapped utility scores than Bijlenga *et al.*'s[111,112] directly measured utilities, which should minimise scale effects. The mapped utilities and directly measured EQ-5D utilities behave similarly, but do not appear to have the same scale, as can be seen in the marked differences between the utility scores obtained from both methods in the Bijlenga *et al.* studies.[111,112] The utility scores reported directly from EQ-5D in the Bijlenga *et al.*[111,112] studies were consistently approximately 0.1 lower than utility scores mapped from SF-36 in the same study (compare *Tables 36* and *37*).

The health states in the model are determined by the length of time that a woman would expect to remain in the state. The time periods spent in each health state are aligned as closely as possible with when HRQoL measurements were taken in their source studies. Baseline utility and initial decrements associated with mode of birth were assumed to last for 3 weeks, and from 3 to 12 weeks utility scores were assumed to be equivalent to utility scores measured at 6 weeks in Bijlenga *et al.*[112] As the HRQoL studies in *Tables 36* and *37* show that women had mostly recovered from adverse effects of birth at

TABLE 38 Utility scores and decrements used in the EAG economic model

Health state	Utility scores	Utility decrement	QALYs[a]	Source
Birth to 3 weeks post partum (baseline, vaginal delivery)	0.6766	–	0.0389	Jansen *et al.*,[116] the Netherlands (SF-36)
Birth to 3 weeks post partum (caesarean section)[b]	0.5895	0.0871	0.0050	Jansen *et al.*,[116] the Netherlands (SF-36)
Birth to 3 weeks post partum (emergency caesarean section)[b]	0.5167	0.1599	0.0092	Jansen *et al.*,[116] the Netherlands (SF-36)
3 weeks post partum to 12 weeks post partum	0.8676	–	0.1496	Bijlenga *et al.*,[112] the Netherlands (SF-36)
12 weeks post partum to 6 months post partum	0.8683	–	0.2171	Bijlenga *et al.*,[112] the Netherlands (SF-36)
Decrement for 3 weeks to 6 months post partum (non-spontaneous delivery)[c]	–	0.0200	0.0084	Petrou *et al.*,[119] UK (EQ-5D)

a All QALYs were calculated assuming a 365.25-day year, with 30.44 days per month.
b Decrement is applied for 3 weeks.
c Decrement is applied for 9 weeks and 3 months.

6 weeks, we believe that this distribution of health states is justified. Utility scores for beyond 12 weeks post partum were derived from 6-month utilities in Bijlenga *et al.*[112]

Multiplying utility scores and decrements by time in the state produces QALYs. For example, referring to *Table 38*, a woman who had a non-induced vaginal birth would have an expected QALY gain of 0.4056 (0.0389 + 0.1496 + 0.2171) and a woman who had an emergency caesarean section would have an expected QALY gain of 0.3880 (0.0389 + 0.1496 + 0.2171 – 0.1599 – 0.0084). The model assumes that the differences between the diagnostic tests relate to the differences in frequencies of different modes of birth in the model. All utilities are assumed constant over the time period during which they occur.

Results of the independent economic analysis

Model validation

The initial technical validation of the model was conducted by replication of published analyses reviewed in *Results of the review of economic studies* and of analyses developed for (and reported in) the NICE CG on *Hypertension in Pregnancy: The Management of Hypertensive Disorders during Pregnancy*.[122]

For technical validation against the study by Hadker *et al.*[39] the simplified model structure illustrated in *Figure 8* was populated with the reported data and evaluated for the cost analysis reported in the paper. This validation yielded identical results to the published paper. The technical validation against the model developed for the NICE CG on *Hypertension in Pregnancy: The Management of Hypertensive Disorders during Pregnancy* did not yield identical results to those reported in the appendices to the guideline,[122] although the results were similar. Identification of reasons for the differences in results was made difficult by a lack of clarity in the reporting of the model structures used in the analyses and some inconsistencies in the reporting of data used to populate the model.

A second stage of technical validation of the model used data from key clinical studies included in our systematic review of test accuracy (see *Chapter 4*). Based on the company-recommended cut-off points, these studies provided data on the diagnostic accuracy for ruling in and ruling out PE using the Triage test (PELICAN study[5]) and the Elecsys sFlt-1 to PlGF ratio test (PROGNOSIS study;[133] Álvarez-Fernández *et al.*[75]). The simplified model structure was populated with sensitivity and specificity data for women presenting before 34 weeks and between 34 and 37 weeks of gestation. The model predictions for the number of

women with PE (among the suspected PE population) classified into the high-, intermediate- and low-risk categories were compared with those reported in 2 × 2 contingency tables extracted from the studies. The model predictions were identical to those presented in the clinical studies, indicating technical validity of the calculations in the model.

Base-case cost-effectiveness results

This section reports the cost-effectiveness results for women presenting for assessment of suspected PE, prior to 35 weeks and between 35 and 37 weeks of gestation, using the PlGF-based tests in addition to standard clinical assessment as compared against standard clinical assessment alone. The results for costs and QALYs are presented for each diagnostic strategy with incremental costs and QALYs calculated compared with the next best alternative (based on dominance and extended dominance).

Cost-effectiveness results for suspected pre-eclampsia presenting before 35 weeks of gestation

The cost-effectiveness results for women presenting for assessment of suspected PE prior to 35 weeks of gestation, using each diagnostic strategy, are presented in *Table 39*. In the base case, total costs vary between £6048 for the Triage test and £8945 for standard clinical assessment. Both strategies including biomarker tests are cost-saving compared with standard clinical assessment, with the cost reductions per patient varying between £2896 for the Triage PlGF test and £2488 for the Elecsys sFlt-1 to PlGF ratio test. Total QALYs for each diagnostic strategy are similar, with no more than 0.00076 QALYs separating the most clinically effective diagnostic strategy and the least clinically effective diagnostic strategy for women suspected of PE before 35 weeks of gestation. Given that the utility data have a high degree of uncertainty as a result of being derived primarily from mapping from SF-36, the differences in HRQoL for this cohort of patients may not be clinically significant. The base case indicates that including a biomarker test in the assessment of suspected PE, prior to 35 weeks of gestation, is cost-saving and may yield slightly better clinical outcomes.

Cost-effectiveness results for suspected pre-eclampsia presenting between 35 and 37 weeks of gestation

In the base-case analysis the cost differences are much smaller for women with suspected PE presenting between 35 and 37 weeks (*Table 40*), than for those presenting before 35 weeks, and there is no

TABLE 39 Base-case cost-effectiveness results for women presenting before 35 weeks

	Costs		QALYs	
Strategy	Total	Increment	Total	Increment
Triage PlGF test	£6048		0.39445	
Elecsys sFlt-1 to PlGF ratio test	£6456	£408	0.39434	–0.00011
Standard assessment	£8945	£2896	0.39368	–0.00076

TABLE 40 Base-case cost-effectiveness results for women presenting between 35 and 37 weeks

	Costs		QALYs	
Strategy	Total	Increment	Total	Increment
Triage PlGF test	£3393		0.3954	
Elecsys sFlt-1 to PlGF ratio test	£3584	£191	0.3954	0
Standard assessment	£3758	£365	0.3954	0

difference between any of the strategies in HRQoL. This is because HRQoL is dependent on the type of delivery in the model, and there are no differences between the strategies after 35 weeks. The Triage PlGF test is still the least costly diagnostic assessment, but the difference between the Triage PlGF test and the Elecsys sFlt-1 to PlGF ratio test is now only £191, and standard assessment is only £365 more expensive than the Triage PlGF test.

Sensitivity analyses

Deterministic sensitivity analyses for suspected pre-eclampsia presenting before 35 weeks of gestation

This section shows the results of deterministic sensitivity analyses applied to the base-case analysis for women presenting for assessment of suspected PE prior to 35 weeks of gestation. Variables included in the sensitivity analyses were those that the EAG considered most likely to influence diagnostic outcome (test sensitivity and specificity, and disease prevalence), key cost variables (including costs of the diagnostic tests) and model parameters that are associated with the greatest degree of uncertainty. When lower and upper limits are specified, these values are derived from *Model parameters*.

Factors influencing diagnostic outcome

Sensitivity analysis on the prevalence of PE *Table 41* indicates that the size of the cost saving from including biomarker tests in the assessment of suspected PE reduces as the prevalence value of PE increases between 21.5% and 32% in women with suspected PE presenting before 35 weeks of gestation.

Sensitivity analysis on diagnostic test sensitivity and specificity *Table 42* shows the effect on costs and QALYs for each strategy, compared with their base-case values when varying test sensitivity and specificity between the limits of their 95% CIs. As would be expected, increased sensitivity is associated with increased cost (more cases of PE identified and hospitalised prior to delivery) while increased specificity is associated with lower cost (fewer cases without PE inappropriately identified as high risk and hospitalised prior to delivery). Greater variation in cost is generally associated with variation in specificity.

Table 42 does not show sensitivity analyses for sensitivity and specificity of the rule-out criteria for the strategies with biomarker tests, as patients identified as being at intermediate risk in these strategies are treated following the gestational hypertension pathway and, therefore, variation in these test characteristics is not associated with any variation in costs or outcomes in the base case.

Scenario analyses with alternative management pathways: patients identified as being at intermediate risk follow the PE management pathway In the base-case analysis, patients with intermediate test results are assumed to be managed in accordance with the gestational hypertension

TABLE 41 Sensitivity analysis on the prevalence of PE in women with suspected PE presenting before 35 weeks of gestation

Strategy	Cost	Change from base case	QALY	Change from base case	Cost	Change from base case	QALY	Change from base case
	Lower prevalence value = 0.215				***Upper prevalence value = 0.32***			
Triage PlGF test	£5768	–£281	0.39449	0.00004	£6357	£309	0.39440	–0.00005
Elecsys sFlt-1 to PlGF ratio test	£6192	–£265	0.39438	0.00004	£6747	£291	0.39429	–0.00005
Standard assessment	£8818	–£127	0.39371	0.00003	£9084	£139	0.39366	–0.00002

TABLE 42 Sensitivity analysis on the diagnostic test accuracy for women presenting before 35 weeks of gestation

Test accuracy	Cost	Change from base case	QALY	Change from base case	Cost	Change from base case	QALY	Change from base case
	Lower sensitivity value = 0.838				***Upper sensitivity value = 0.988***			
Sensitivity of standard assessment	£8885	–£60	0.39373	0.00005	£8972	£27	0.39366	–0.00002
	Lower specificity value = 0.263				***Upper specificity value = 0.46***			
Specificity of standard assessment	£9424	£479	0.39357	–0.00011	£8420	–£525	0.39380	0.00012
	Lower sensitivity value = 0.513				***Upper sensitivity value = 0.739***			
Sensitivity of Triage PlGF test (rule in)	£5979	–£69	0.39450	0.00005	£6111	£62	0.39440	–0.00005
	Lower specificity value = 0.852				***Upper specificity value = 0.937***			
Specificity of Triage PlGF test (rule in)	£6293	£245	0.39439	–0.00006	£5860	–£188	0.39449	0.00004
	Lower sensitivity value = 0.619				***Upper sensitivity value = 0.778***			
Sensitivity of Elecsys sFlt-1 to PlGF ratio test (rule in)	£6407	–£49	0.39437	–0.00008	£6500	£44	0.39430	–0.00015
	Lower specificity value = 0.805				***Upper specificity value = 0. 855***			
Specificity of Elecsys sFlt-1 to PlGF ratio test (rule in)	£6589	£132	0.39430	–0.00015	£6334	–£122	0.39436	–0.00009

pathway, so that women with moderate hypertension receive enhanced monitoring in the community along with treatment for their hypertension, while those with severe hypertension are hospitalised for 3 days for assessment and stabilisation of their condition and then discharged to enhanced monitoring in the community.

Two alternative treatment strategies were considered for patients identified by the biomarker tests as being at intermediate risk, that is, those who fall between the rule-in and rule-out criteria. In the first alternative management strategy, all patients with intermediate test results are assigned to the PE management pathway and are hospitalised until delivery (after an average of 9 days of hospitalisation). Cost-effectiveness results for this first alternative management strategy are reported in *Table 43*. In this scenario, diagnostic strategies including biomarker tests remain cost-saving compared with standard clinical assessment. However, the

TABLE 43 Sensitivity analysis on intermediate test results for women presenting before 35 weeks of gestation with all women managed by hospitalisation

Strategy	Cost	Change from base case	QALY	Change from base case
Triage PlGF test	£7987	£1939	0.39391	–0.00054
Elecsys sFlt-1 to PlGF ratio test	£6750	£294	0.39422	–0.00012
Standard assessment	£8945	£0	0.39368	0.00000

strategy using the Elecsys sFlt-1 to PlGF ratio test is associated with a £1237 lower cost than the strategy using the Triage PlGF test, whereas the Triage PlGF test was £408 less expensive in the base case.

In the second alternative management strategy, only patients with PE with intermediate test results are assigned to the PE management pathway. This assumes that women with PE who have an intermediate test will have perfect identification and management: this is a best-case scenario, whereas the first alternative management strategy may be considered a highly conservative or worst-case scenario. Cost-effectiveness results for the second alternative management strategy are reported in *Table 44*. The costs of these diagnostic strategies, including the costs of biomarker tests, are closer to their base-case values, but are still higher. In this sensitivity analysis, the costs of the strategy using the Triage PlGF test are lower than those for the strategy using the Elecsys sFlt-1 to PlGF ratio test (although the difference is slightly smaller than in the base case).

Factors influencing costs

Sensitivity analysis on costs of biomarker tests A sensitivity analysis was conducted in which the costs of the biomarker tests for diagnosing PE (confidential data) were doubled and trebled. Changing the cost of the test does not increase the cost of any other factors. As a result, the change from the base case for each test directly corresponds to the increase in price of the test. In both cases, the clinical assessment strategies including the biomarker tests remained cost-saving, with relatively little reduction in the size of the modelled cost-saving.

Sensitivity analysis on probability of admission and length of stay in neonatal intensive care There is a substantial degree of uncertainty over the model inputs related to admission to neonatal care, both the probability of admission and length of stay, particularly for early deliveries (i.e. those resulting from true-positive, false-positive or true-negative test results). Rather than vary the probability of admission using a CI derived from the data presented by the EPIPAGE study[110] (given the uncertainty involved in extrapolating from the EPIPAGE study[110] as it is a nearly 20-year-old study in a French population), we used the proportion of cases admitted to a NICU at the earliest birth week (30 weeks) and the proportion admitted at the latest birth week (34 weeks) in the EPIPAGE study as the lower and upper limits, respectively. Similarly, for a sensitivity analysis on duration of stay, the EAG has used the length of stay for cases admitted to a NICU at 30 weeks as the lower limit and the value at 34 weeks for the upper limit. *Table 45* reports cost-effectiveness results for these sensitivity analyses. The difference between the diagnostic strategies that include biomarker tests and standard clinical assessment roughly doubles when varying the probability of admission from the lower to the upper limit. As NICU stay has no effect on maternal HRQoL, HRQoL data for all assessment strategies remain unchanged.

The analysis on duration of stay in a NICU following early delivery because of PE and/or positive diagnostic test result shows that length of NICU stay has a greater effect on costs than the probability of NICU admission.

These sensitivity analyses were repeated for the probability of admission and length of admission to NICU for cases with a negative test result without PE (i.e. true negatives). *Table 46* reports the results of these two analyses. Probability of NICU admission and length of NICU admission before 35 weeks of gestation had a much smaller effect on costs in women with a negative test than in women with a positive test.

TABLE 44 Sensitivity analysis on managing only PE cases (among intermediate results) using the PE management pathway before 35 weeks of gestation

Strategy	Cost	Change from base case	QALY	Change from base case
Triage PlGF test	£6240	£192	0.39431	–0.00014
Elecsys sFlt-1 to PlGF ratio test	£6546	£90	0.39427	–0.00007
Standard assessment	£8945	£0	0.39368	0.00000

TABLE 45 Neonatal intensive care unit admission and length of stay before 35 weeks of gestation following a positive test result

	Probability of NICU admission				Days in NICU			
	0.272		0.848		2 days		15 days	
Strategy	Cost	Change from base case	Cost	Change from base case	Cost	Change from base case	Cost	Change from base case
Triage PlGF test	£4967	–£1081	£6544	£496	£4654	–£1394	£7460	£1412
Elecsys sFlt-1 to PlGF ratio test	£5213	–£1243	£7026	£570	£4853	–£1603	£8079	£1623
Standard assessment	£6588	–£2356	£10,025	£1080	£5906	–£3038	£12,021	£3076

TABLE 46 Neonatal intensive care unit admission and length of stay before 35 weeks of gestation following a true-negative test result

	Probability of NICU admission				Days in NICU			
	0.049		0.106		2 days		6 days	
Strategy	Cost	Change from base case	Cost	Change from base case	Cost	Change from base case	Cost	Change from base case
Triage PlGF test	£6001	–£47	£6109	£61	£6002	–£47	£6189	£140
Elecsys sFlt-1 to PlGF ratio test	£6412	–£44	£6512	£56	£6413	–£43	£6586	£130
Standard assessment	£8926	–£19	£8969	£24	£8926	–£19	£9000	£56

Uncertainty in the distribution of hypertension severity There is a large degree of uncertainty in the distribution of hypertension within the population of women included in the model. The distribution in the model is based on values reported by Duckworth *et al.*,[98] which are based on data from Anumba *et al.*[134] The proportion of women with and without PE in two categories of hypertension related to the clinical management pathway (moderate and severe) were varied between their 95% confidence limits, with the remaining population assumed to have mild or no hypertension. For the lower estimate, both moderate and severe hypertension were set to their lower 95% confidence limit. Similarly for the upper estimate: both categories were set to their upper 95% confidence limit. *Table 47* reports the cost-effectiveness results for this sensitivity analysis. There were no changes in QALYs from the base case and changes in costs caused by changes in the distribution of hypertension severity were all less than £35. This suggests that the distribution of hypertension severity has a limited impact on the potential difference in costs between the diagnostic assessment strategies.

Deterministic sensitivity analyses for suspected pre-eclampsia presenting between 35 and 37 weeks of gestation

This section describes the results of the same sensitivity analyses as reported above, but for women presenting with PE between 35 and 37 weeks of gestational age. All results are presented as changes in costs only. As the differences in HRQoL in the model are driven by the type of delivery, and there are no differences between the tests on the type of delivery that occurs for women presenting with PE from 35 to 37 weeks of gestation, there are no differences in HRQoL in the model. When lower and upper limits are specified, these values are derived from *Model parameters*.

TABLE 47 Sensitivity analysis on the distribution of hypertension for women presenting before 35 weeks of gestation

	Lower limits for proportions of moderate and severe hypertension				Upper limits for proportions of moderate and severe hypertension			
Strategy	Cost	Change from base case	QALY	Change from base case	Cost	Change from base case	QALY	Change from base case
Triage PlGF test	£6015	–£33	0.3945	0.0000	£6082	£33	0.3945	0.0000
Elecsys sFlt-1 to PlGF ratio test	£6426	–£30	0.3943	0.0000	£6487	£30	0.3943	0.0000
Standard assessment	£8932	–£13	0.3937	0.0000	£8957	£13	0.3937	0.0000

Factors influencing diagnostic outcome

Table 48 reports the outcomes of sensitivity analyses assessing the impact of prevalence of disease and influence of sensitivity and specificity for the Triage PlGF test, Elecsys sFlt-1 to PlGF ratio test and standard clinical assessment. As in the before-35-weeks model, the rankings of the interventions do not change from the base case, but the magnitude of the change in costs is much smaller. None of the analyses altered costs by more than £50.

Sensitivity analysis on prevalence of PE *Table 48* indicates that the difference in total cost for any of the diagnostic strategies is relatively small when varying the prevalence of PE between its lower and upper limit. The greatest variation is seen for the Elecsys sFlt-1 to PlGF ratio test. Unlike the sensitivity analysis on prevalence of PE for women presenting up to 35 weeks of gestation, there is no consistent pattern in the size of the difference between costs for the biomarker tests and standard clinical assessment. For the Triage PlGF test, the difference in cost compared with standard clinical assessment increases slightly with higher prevalence, while the reverse is found for the Elecsys sFlt-1 to PlGF ratio test.

Sensitivity analysis on diagnostic test sensitivity and specificity *Table 49* reports sensitivity analyses on diagnostic test accuracy for women presenting between 35 and 37 weeks of gestation. As expected, improved sensitivity is associated with increased costs and increased specificity with lower costs. However, the cost differences, compared with the base case, are small. Unlike the sensitivity analysis conducted for women presenting up to 35 weeks of gestation, there is no discernible pattern of greater variability associated with variation in specificity.

Sensitivity analyses with alternative management pathways: patients identified as being at intermediate risk follow the PE management pathway There is uncertainty regarding the approach to managing women who fall between the rule-in and rule-out criteria with the biomarker tests. As indicated previously, the base-case analysis assumes that women with intermediate test results are managed according to the gestational hypertension pathway. *Table 50* presents the same alternative management strategies for women presenting between 35 and 37 weeks of gestation, as previously considered for women presenting

TABLE 48 Sensitivity analyses on the prevalence of PE for women presenting between 35 and 37 weeks of gestation

Strategy	Cost	Change from base case	Cost	Change from base case
	Lower limit prevalence = 0.403		***Upper limit prevalence = 0.576***	
Triage PlGF test	£3364	–£28	£3421	£29
Elecsys sFlt-1 to PlGF ratio test	£3534	–£49	£3633	£50
Standard assessment	£3724	–£34	£3792	£34

TABLE 49 Sensitivity analysis on the diagnostic test accuracy for women presenting between 35 and 37 weeks of gestation

Test accuracy	Cost	Change from base case	Cost	Change from base case
	Lower limit sensitivity = 0.838		***Upper limit sensitivity = 0.988***	
Sensitivity of standard assessment	£3726	–£32	£3772	£14
	Lower limit specificity = 0.263		***Upper limit specificity = 0.46***	
Specificity of standard assessment	£3791	£33	£3722	–£36
	Lower limit sensitivity = 0.131		***Upper limit sensitivity = 0.342***	
Sensitivity of Triage PlGF test (rule in)	£3364	–£28	£3429	£36
	Upper limit specificity = 0.823		***Upper limit specificity = 0.968***	
Specificity of Triage PlGF test (rule in)	£3425	£32	£3374	–£19
	Lower limit sensitivity = 0.619		***Upper limit sensitivity = 0.778***	
Sensitivity of Elecsys sFlt-1 to PlGF ratio test (rule in)	£3558	–£26	£3607	£23
	Lower limit specificity = 0.805		***Upper limit specificity = 0.855***	
Specificity of Elecsys sFlt-1 to PlGF ratio test (rule in)	£3593	£9	£3575	–£8

TABLE 50 Alternative management strategies for women with intermediate test results between 35 and 37 weeks of gestation

Strategy	Cost	Change from base case
Manage all intermediate results using the PE management pathway		
Triage PlGF test	£3634	£241
Elecsys sFlt-1 to PlGF ratio test	£3645	£61
Standard assessment	£3758	£0
Manage only PE cases (among intermediate results) using the PE management pathway		
Triage PlGF test	£3539	£146
Elecsys sFlt-1 to PlGF ratio test	£3631	£47
Standard assessment	£3758	£0

up to 35 weeks of gestation. Changing the management strategies had the most significant effect on the Triage PlGF test, with costs increasing by £241 for the first alternative management strategy and by £146 for the second. Managing all women with suspected PE and a test in the intermediate range using the PE pathway reduces the difference between the Triage PlGF test and the Elecsys sFlt-1 to PlGF ratio test to £11. The second sensitivity analysis also narrows the gap between the Triage PlGF test and the Elecsys sFlt-1 to PlGF ratio test by £99. As noted for women presenting up to 35 weeks of gestation, these scenarios may be considered as worst case (manage all intermediate cases using the PE pathway) and best case (only manage PE cases within the intermediate group using the PE pathway).

Factors influencing costs

Sensitivity analysis on costs of biomarker tests A sensitivity analysis was conducted in which the costs of the biomarker tests for diagnosing PE (confidential data) were doubled and trebled. As noted above, changing the cost of the test does not increase the cost of any other factors. As a result, the change from the base case for each test directly corresponds to the increase in price of the test. While the strategies including the biomarker tests remained cost-saving in this sensitivity analysis, the size of the cost saving was substantially reduced when test costs were increased.

Sensitivity analysis on probability of admission and length of stay in neonatal intensive care

Table 51 shows that changing the probability of NICU admission or the length of NICU stay for neonates of women suspected of having PE between 35 and 37 weeks of gestational age has the least effect on overall costs for the Triage PlGF test, followed by the Elecsys sFlt-1 to PlGF ratio test and then standard clinical assessment. None of the assessment strategies had a decrease in overall costs of more than £60, and none showed an increase of more than £200. Applying the upper limit value for NICU stay had the greatest effect on costs.

Varying the probability of NICU stay for women without PE who test negative (i.e. true negatives) had the opposite effect, with changes having the greatest cost impact on the Triage PlGF test, followed by the Elecsys sFlt-1 to PlGF ratio test and then standard clinical assessment. None of the cost differences associated with variation in the probability of admission to NICU in this group was > £40. The sensitivity analysis results for changing the length of stay in women suspected of having PE who test negative produced identical results to the analysis in women testing positive.

Uncertainty in the distribution of hypertension severity

As noted previously, there is a large degree of uncertainty in the distribution of hypertension within the population of women included in the model. The sensitivity analysis on the distribution of hypertension in women with and without PE, conducted for the population of women presenting with suspected PE up to 35 weeks of gestation, was repeated for weeks 35–37 using the upper and lower 95% confidence limits reported by Duckworth *et al.*[98]

The results of the sensitivity analysis are reported in *Table 52* and, similar to the analysis conducted in women suspected of having PE before 35 weeks of gestation, indicate that the distribution of severity of hypertension had little impact on costs for each diagnostic strategy.

Scenario analyses

This section reports two scenario analyses that were conducted. The first examined the impact of processing and analysing the PlGF-based test results in a near-patient setting instead of in a central laboratory. This is based on an assumption that the Triage test could be employed in a midwifery day unit. The second

TABLE 51 Sensitivity analyses on the cost-related probability of NICU admission and length of stay for neonates born to women with suspected PE between 35 and 37 weeks of gestation

	Probability of NICU admission				Length of stay in NICU			
	0.049		0.106		2 days		6 days	
Strategy	Cost	Change from base case	Cost	Change from base case	Cost	Change from base case	Cost	Change from base case
Sensitivity analyses of probability of admission to NICU and NICU length of stay following early delivery because of PE and/or positive diagnostic test result								
Triage PlGF test	£3354	–£39	£3443	£50	£3337	–£55	£3559	£166
Elecsys sFlt-1 to PlGF ratio test	£3541	–£42	£3638	£54	£3527	–£57	£3755	£171
Standard assessment	£3698	–£60	£3835	£77	£3692	–£66	£3955	£197
Sensitivity analyses on probability of admission to NICU and NICU length of stay following a negative test result, without PE								
Triage PlGF test	£3370	–£23	£3428	£36	£3337	–£55	£3559	£166
Elecsys sFlt-1 to PlGF ratio test	£3562	–£21	£3616	£32	£3527	–£57	£3755	£171
Standard assessment	£3749	–£9	£3772	£14	£3692	–£66	£3955	£197

TABLE 52 Sensitivity analysis on the distribution of hypertension (with and without PE) in women with suspected PE presenting between 35 and 37 weeks of gestation

	Lower limits for moderate and severe		Upper limits for moderate and severe	
Distribution of hypertension severity	Cost	Change from base case	Cost	Change from base case
Triage PlGF test	£3370	–£22	£3415	£22
Elecsys sFlt-1 to PlGF ratio test	£3570	–£14	£3598	£14
Standard assessment	£3752	–£5	£3763	£5

examined the potential impact of using a PlGF-based biomarker test as a replacement for quantitative proteinuria testing as part of standard clinical assessment.

Scenario analysis on the cost impact of adopting near-patient testing with the Triage PlGF test (Alere)

In all the relevant studies (PETRA, PELICAN and PROGNOSIS) the PlGF-based biomarker tests were carried out in a central laboratory. There is, therefore, no information available on whether or not these tests could be cost-effective if employed in a near-patient setting, that is, in a midwifery day unit.

Table 53 reports a scenario analysis indicating the possible cost impact of adopting near-patient testing with the Triage PlGF test. The analysis assumes that the testing occurs in a midwifery day unit, with sufficient throughput to ignore the overhead costs of maintaining a centrifuge (required to derive the anticoagulated plasma samples), as well as the acquisition and maintenance costs for the test system.[40] As a result, the Triage PlGF test is costed in this scenario at the same unit cost as adopted for the base case: the average cost of an individual test (confidential data). It is further assumed, because of the adoption of near-patient testing in the midwifery day unit, that no women are required to be admitted overnight while awaiting the return of test results.

Owing to the greater complexity of the Elecsys sFlt-1 to PlGF ratio tests and expertise required in running them, this test can be conducted only in a central laboratory (i.e. it is not a near-patient test). The costs of the Elecsys sFlt-1 to PlGF ratio test and standard clinical assessment therefore remain at the base-case

TABLE 53 Scenario analysis: cost impact of near-patient testing

		Presenting			
		Before 35 weeks of gestation		Between 35 and 37 weeks of gestation	
Strategy	Proportion of women admitted overnight awaiting test results	Total	Difference from base case	Total	Difference from base case
Triage PlGF Test	0.0	£6048	£0	£3393	£0
Elecsys sFlt-1 to PlGF ratio	0.1	£6477	£21	£3604	£21
Standard assessment	0.1	£8965	£21	£3778	£21
	0.2	£8986	£41	£3799	£41
	0.3	£9007	£62	£3820	£62
	0.4	£9028	£83	£3841	£83
	0.5	£9048	£104	£3861	£104

values (Elecsys data confidential; £0 for standard clinical assessment). However, for each of these strategies, we assume that a proportion of women may be admitted awaiting test results. Moreover, we have assumed that more women may require overnight admission with standard clinical assessment. This is based on clinical advice that the quantitative proteinuria test is the test that is most commonly associated with such unscheduled overnight stays.

The EAG has not identified any reliable sources of information on the proportion of women requiring overnight stay while awaiting test results. As a result we have selected a single value for the Elecsys sFlt-1 to PlGF ratio test (10% of women being assessed for suspected PE requiring overnight stay while awaiting test results) and a range for standard clinical assessment (10–50%). The excess cost of an overnight stay was estimated as £207. This is the difference between the cost reported for a day-case antenatal routine observation (£284) and a non-elective short-stay admission for antenatal routine observation (£491) in *National Schedule of Reference Costs 2013–14 for NHS Trusts and NHS Foundation Trusts*.[127]

As would be expected, given the assumptions underlying this scenario analysis, the adoption of near-patient testing (with the intention of avoiding unnecessary overnight stays associated with delayed turnaround of test results) has no impact on the cost of the strategy including the Triage PlGF test. In contrast, if 10% of women tested under the strategy including the Elecsys test are admitted overnight because of delayed turnaround of test results, the cost of this strategy increases by £21 for each 10% increase in the proportion of women admitted. As indicated in *Table 53*, the costs of standard clinical assessment increase in increments of approximately £21 as the proportion of women admitted overnight awaiting test results increases.

Overall, this scenario analysis suggests that the cost saving, from the NHS perspective, of avoiding overnight admissions while waiting for test results in this population may be modest, and that a strategy of near-patient testing purely on these grounds would need to be justified by a careful assessment of the projected savings weighed against the likely cost of acquiring and maintaining additional equipment required to provide the Triage PlGF test. There may be other costs falling on patients and their families that need to be considered, but these are outside the scope of this assessment. It should also be noted that this scenario analysis has not attempted to capture any HRQoL benefit that may be associated with the avoidance of these overnight admissions, just as the base case has not assigned disutility to hospital admissions. Owing to the short length of stay, any effects on HRQoL are likely to be extremely limited.

Scenario analysis on cost impact of replacing quantitative proteinuria assessment with biomarker tests

As indicated in *Strategies and comparators* (and in *Chapter 4*, *Assessment of test accuracy*), no clinical test evidence to inform the assessment of biomarker tests as alternatives to proteinuria testing was identified in our systematic review of clinical evidence of diagnostic test accuracy. As a result, we cannot provide a reliable economic analysis of this scenario. In the absence of a reliable base case for replacing quantitative proteinuria with biomarker tests, we conducted a simple cost-based scenario analysis, similar to that used for near-patient testing.

During the scoping stage of this assessment, quantitative proteinuria testing was identified by clinical experts as a possible factor that was leading to delays in diagnostic assessment of women with suspected PE, with a proportion being unnecessarily admitted for overnight stays awaiting results of the quantitative proteinuria test. *Table 54* reports a scenario analysis of the cost impact of overnight stays because of delayed turnaround of diagnostic tests, if PlGF-based biomarker tests replace quantitative proteinuria testing. This scenario analysis assumes no delay in turnaround for biomarker tests (i.e. no associated overnight stays for patients awaiting test results). To allow for the possibility that a test strategy including biomarkers in place of quantitative proteinuria may have poorer diagnostic performance, we have repeated the analysis for the biomarker tests using the lower limits for the 95% CIs for sensitivity and specificity. For standard assessment, we assume a range of 10–50% for the proportion of women required to stay overnight awaiting diagnostic test results.

TABLE 54 Scenario analysis: cost impact of replacing the quantitative proteinuria test with a biomarker test for assessment of suspected PE

Strategy	Proportion of women admitted overnight awaiting test results	Presenting			
		Before 35 weeks of gestation		Between 35 and 37 weeks of gestation	
		Total	Difference from base case	Total	Difference from base case
Triage PlGF test	0.0	£6048[a]	£0	£3393	£0
		£6223[b]	£175	£3396	£3
Elecsys sFlt-1 to PlGF ratio test	0.0	£6456[a]	£0	£3584	£0
		£6540[b]	£83	£3567	–£17
Standard	0.1	£8965	£21	£3778	£21
	0.2	£8986	£41	£3799	£41
	0.3	£9007	£62	£3820	£62
	0.4	£9028	£83	£3841	£83
	0.5	£9048	£104	£3861	£104

a Sensitivity and specificity of strategy including biomarker test at values used in the base case.
b Sensitivity and specificity of strategy including biomarker test set at lower limit of 95% CI (to consider robustness of cost estimates to diagnostic accuracy of test strategy).

Comparing the base-case costs for strategies with biomarker tests instead of proteinuria testing against standard clinical assessment shows that cost savings with the proteinuria replacement strategy increase as the proportion of women requiring an overnight stay under standard clinical assessment increases. The assumption of poorer diagnostic performance for strategies including biomarker tests results in increased costs for those strategies, although the strategies remain cost-saving compared with standard clinical assessment in both groups of women (those presenting up to 35 weeks of gestation, and also those presenting between 35 and 37 weeks of gestation).

As indicated in the previous analysis, the estimated cost savings by avoiding overnight stays that are solely due to the delayed turnaround of diagnostic tests are limited and are unlikely, on their own, to provide justification for change in clinical practice. The potential benefits from improved diagnostic test specificity are likely to be substantially greater than the cost savings accruing from avoided overnight stays.

Chapter 6 Discussion

At the time of preparing the final version of this report, some of the data considered by the NICE Diagnostics Assessment Committee and the EAG were confidential. These have been excluded from the present report, as indicated clearly. The majority of the confidential data were from the PETRA study,[6] which, at the time of this assessment, remained unpublished, and also from unpublished academic documents provided by Duckworth *et al.*,[98] Hunter *et al.*[97] and Roche Diagnostics.[99] The PETRA data[6] do not directly inform our economic analysis because the PELICAN study[5] was considered to be a more relevant study in terms of its population characteristics and outcomes. The Duckworth *et al.*,[98] Hunter *et al.*[97] and Roche Diagnostics[99] studies primarily informed EAG discussions about developing the structure of our economic model. The only parameters for the model that we obtained from the confidential studies were the cost for the Triage PlGF test and distributions for hypertension severity from Duckworth *et al.*,[98] neither of which had a meaningful impact on results when varied in sensitivity analyses.

Statement of principal findings

Clinical effectiveness (test accuracy)

Four studies met the inclusion criteria for the systematic review of test accuracy. These assessed the Triage PlGF test (PETRA study[6] and PELICAN study[5]) and the Elecsys sFlt-1 to PlGF ratio (PROGNOSIS study[133] and Álvarez-Fernández *et al.* study[75]). As noted above, the PETRA study on the Triage PlGF test was confidential at the time of preparing the current report and has been excluded, although it was available to the EAG and NICE Diagnostics Assessment Committee. Critical appraisal of the three published studies on these tests suggested that the studies were probably at low risk of bias. An exception is a high risk of clinical review bias in all three studies, as the studies diagnosed PE solely according to biomarker test results, whereas in clinical practice the biomarker test results would be interpreted alongside hypertension, proteinuria and/or other signs or symptoms. However, it is unclear whether or not this difference would have led to systematic under- or overestimation of test accuracy outcomes.

The decision problem for this diagnostic assessment was not fully met by the available evidence. No test accuracy studies were found for the PerkinElmer DELFIA Xpress PlGF test or BRAHMS Kryptor sFlt-1 to PlGF ratio test. In addition, no test accuracy studies were identified that assessed index tests as an alternative to quantitative proteinuria testing (part 2 of the NICE decision problem). All the identified evidence from primary studies relates to use of the biomarker tests where the test assays were performed in a laboratory; none of the studies reported assays that could be done in an antenatal clinic (near-patient) setting.

Based on the available published evidence, the Triage PlGF test has high prognostic sensitivity for predicting PE requiring delivery within 14 days of testing, while the Elecsys sFlt-1 to PlGF ratio has high diagnostic sensitivity for rule-out of PE within 1 week of testing and good specificity for rule-in of PE within 4 weeks, although it has a high false-positive rate. However, the primary studies included in the test accuracy review reported different outcomes for each test (prognostic accuracy for the Triage PlGF test; diagnostic accuracy for the Elecsys sFlt-1 to PlGF ratio test), which makes direct comparisons difficult.

For the Triage PlGF test cut-off points of < 100 pg/ml and < 5th percentile of PlGF concentration (test positive), the PELICAN study gave high sensitivity (96%) with good precision (i.e. narrow 95% CIs) for identifying women likely to develop PE requiring delivery within 14 days when presenting with suspected PE at up to 35 weeks of gestation. Diagnostic accuracy outcomes for the Elecsys sFlt-1 to PlGF ratio are for three test cut-off points: 23, 38 and 85; however, the majority of data are from the PROGNOSIS study, which employed the 38 cut-off point. The PROGNOSIS study outcomes suggest that the Elecsys sFlt-1 to PlGF ratio can rule out PE within 1 week of testing in approximately 99% of patients (based on the NPVs for two study cohorts) and has reasonable specificity for ruling in PE within 4 weeks of testing (specificity 83% for two study cohorts, although with a likelihood of false positives: PPVs were approximately 40% for both study cohorts).

The three published studies used definitions of PE that were different from the NICE definition, appearing to include a wider range of women than the NICE definition would permit; notably, the main Elecsys test study (PROGNOSIS) included HELLP syndrome in the definition of PE.

None of the included studies evaluated more than one test, meaning that head-to-head comparisons of tests are not available. Meta-analysis was not possible because of heterogeneity of outcomes, with different test cut-off points reported in the studies.

Cost-effectiveness

The systematic review of economic studies identified a small number of cost models. No models assessed health benefits to mothers or neonates. Targeted systematic searches for HRQoL studies identified a similarly sparse evidence base for HRQoL in gestational hypertension and PE. Most studies were in general pregnancy and post-partum populations and very few used the EQ-5D or any other preference-based utility instrument. Utility values used in the cost-effectiveness model were primarily derived from SF-36 data mapped to EQ-5D. The mapping equation appeared to overestimate utility when compared with directly measured EQ-5D scores. However, it is unclear whether this reflects limitations of the EQ-5D or of the mapping process (see *Uncertainties*).

The cost-effectiveness model found that both the Triage PlGF test and the Elecsys sFlt-1 to PlGF ratio test were cost-saving compared with standard clinical assessment. The differences in QALYs were very small, requiring four decimal places to show a difference in the base-case analyses for prior to 35 weeks of gestational age and finding no difference between diagnostic assessments for women with suspected PE presenting between 35 and 37 weeks of gestation. The cost differences between the Triage PlGF test and the Elecsys sFlt-1 to PlGF ratio test were slightly in favour of the Triage PlGF test, for women presenting both before and after 35 weeks of gestation.

Sensitivity analyses required up to five decimal points (four decimal places minimum) to show a difference in QALYs for women presenting with suspected PE before 35 weeks of gestation. The most influential parameters in the model were associated with the probability and cost of stay for neonates in the NICU. Most other parameters had very small effects on the model results.

Owing to a lack of available evidence, the accuracy and cost-effectiveness of the DELFIA Xpress PlGF test and of the BRAHMS Kryptor sFlt-1 to PlGF ratio could not be assessed.

Strengths and limitations of the assessment

Strengths of the assessment

The systematic reviews and economic analysis presented in this report have been carried out independent of competing interests, and were based on methods specified a priori in a peer-reviewed protocol, consistent with the NICE scope and decision problem for this diagnostic assessment. All studies included in the systematic review of test accuracy were critically appraised using a standard approach to identify possible threats to validity and generalisability. A multidisciplinary advisory group commented on the research protocol and on a draft of the final report. Additional clinical experts (see *Acknowledgements*) also commented on a draft version of the economic model.

The de novo economic model developed by the EAG is based on recognised guidelines. The model structure and data inputs are presented in the current report with explanatory rationale. The economic model is based on data identified from systematic searches for test accuracy, economic studies and HRQoL evidence, and other best available information. The model structure has been subjected to comment and external validation by experts. Additional validation checks were undertaken for model input parameters using published models. The model has been subjected to deterministic and scenario sensitivity analyses to test the robustness of this model compared with alternative data inputs.

Limitations of the assessment

The current assessment is dependent on a relatively limited evidence base, meaning that only four primary studies of test accuracy met the inclusion criteria, and these report on only two of the biomarker tests specified in the decision problem. No evidence to address the second part of the decision problem (exploring the effectiveness of the biomarker tests as a replacement for quantitative proteinuria testing) was identified. Meta-analysis of the primary evidence was not feasible, meaning that the majority of test accuracy information included in this report comes primarily from three studies, namely PETRA and PELICAN for the Triage PlGF test and PROGNOSIS for the Elecsys sFlt-1 to PlGF ratio test.

Data from the PETRA study are confidential and, therefore, have not been presented in detail in this report. However, this does not influence our conclusions, as the test accuracy data for the Triage PlGF test in the economic model were obtained from the PELICAN study, which we considered to be more relevant to a UK population.

The EAG cost-effectiveness model and subsequent analyses based on the model outputs have several limitations. The EAG was forced to make a number of assumptions, with clear justifications, due to the lack of data. The model was designed to perform a probabilistic sensitivity analysis, as indicated in the protocol. However, the absence of evidence directly comparing tests and standard care, or an evidence structure that could support robust, bivariate meta-analysis of diagnostic accuracy data precluded the possibility of conducting an appropriate probabilistic sensitivity analysis. While variation in the individual components of sensitivity and specificity for standard care or the diagnostic tests could be introduced into the model, no data exist to inform on the correlation between these parameters for each test and standard clinical assessment or between the tests and standard assessment. The results of such a probabilistic sensitivity analysis would be of little benefit in informing decision-making, and indeed may be detrimental by mischaracterising uncertainty. The model drew data from a wide range of data sets, which may not all be relevant or generalisable to the UK. Owing to lack of adequate diagnostic effectiveness data, only the Triage PlGF test and the Elecsys sFlt-1 to PlGF ratio test were included in the base-case analysis.

There was a lack of data on utility scores in PE. By necessity, this required expanding the searches to include pregnant and post-partum women who did not have PE. The model requires some assumptions with regard to whether or not these data are applicable, but it is better to include estimates and acknowledge some uncertainty than to exclude measurements and in essence assume no uncertainty in utility differences. In this model, the differences in utility scores are very small, and it is unknown whether or not utility scores measured from women with PE would cause different conclusions. The short time horizon of the model, and rarity of maternal and neonatal mortality in the studies used to populate the models make utility scores unlikely to be a driving factor in the model. No data were available on long-term outcomes of birth for our model population; consequently, this model has not overcome this specific limitation of other models in this treatment area.

Searches were limited to studies published in the English language. This was a pragmatic decision made when developing the review protocol because the current diagnostic assessment is specifically focused on clinical practice in England and Wales. Non-English-language studies would be unlikely to be generalisable to the UK clinical setting because the management of women suspected of having PE varies by country.

Uncertainties

None of the identified primary studies of test accuracy specifically included defined populations or subgroups relevant to the high PE risk subgroups specified in the NICE scope (chronic hypertension, pre-existing or gestational diabetes mellitus, renal conditions and autoimmune conditions). It is therefore unclear whether or not the reported accuracy of the biomarker tests would be generalisable to these high-risk subgroups. The available evidence on test accuracy is also primarily from studies that excluded multiple (e.g. twin) pregnancies, which (although not specified in the NICE scope) are at increased risk of PE compared with

singleton pregnancies. There appears to be anecdotal evidence[48] that accuracy of the biomarker tests may be lower in women with high BMI, although it is unclear whether or not this evidence, which is from a case–control study of women already diagnosed, would apply to a population with suspected PE.

As noted in *Strengths and limitations of the assessment*, there are a number of limitations to the primary studies of test accuracy that may have a bearing on the generalisability of the results. In particular, the studies tended to use wider definitions of PE than that specified by NICE. The trials identified by the systematic review of diagnostic studies also had a distinct lack of data for women presenting with suspected PE between 20 and 30 weeks of gestation, and a lack of data on long-term neonatal and maternal outcomes for births complicated by PE. Women suspected of PE between 20 and 30 weeks of gestation have the highest risk of adverse maternal and neonatal morbidity and mortality, but this population is currently underexplored. Many of the children born in this time period could have significant lifelong morbidity, with potential corresponding HRQoL effects on their parents. In an attempt to identify the size of this population the EAG conducted targeted searches that identified the EPICURE studies.[108,109] These studies indicated that the proportion of women affected by PE before 30 weeks is very small.

A lack of information about some of the intermediate outcomes (time to test result, test failure rate, time to diagnosis, time to onset of PE, proportion of women returned to less intensive follow-up, length of inpatient stay and time to delivery) outlined in the NICE scope means that the impact of the tests on these outcomes could not be assessed in the EAG economic model. Although data about clinical outcomes other than those related to diagnosing PE (maternal and fetal morbidity and mortality, and emergency admission) were reported in some studies, heterogeneity between studies meant that no useful assessment of the effects of tests on these outcomes could be made.

Another factor that might potentially affect generalisability is that the majority of patients included in the studies were of white ethnicity, which is unlikely to reflect the multiethnic case mix encountered in antenatal clinics in England and Wales or the wider UK. However, it is unclear whether or not variation in ethnicity would affect PE diagnosis or management.

Our searches identified a lack of data on utility values in women with gestational hypertension and PE, and the availability of HRQoL data for women with PE is poor. Only nine studies were identified that were of potential use in modelling, and only two of these studies were conducted specifically in a population with gestational hypertension or PE. All of these studies had data sets based on complete cases, and reporting on missing data was inconsistent. There is potential for bias in the HRQoL data due to utility values when the data may not be missing at random. The unavailability of data that would allow consistent utility decrements to be applied forced the EAG to rely primarily on EQ-5D data mapped from SF-36, which Bijlenga *et al.*[112] seemed to indicate may overestimate HRQoL in women with gestational hypertension or PE. It is also plausible that EQ-5D may underestimate utility in pregnant and post-partum women if it does not fully capture positive aspects of being a new mother.

None of the reports for the identified studies defined the proportions of patients who would be managed as having mild, moderate or severe gestational hypertension or the proportions who would be managed as having mild, moderate or severe PE patients. In the model, we have used proportions of hypertension that are mild, moderate or severe cited by Duckworth *et al.*[98] as part of the confidential Alere evidence submission. These were credited to a published paper by Anumba *et al.*,[134] however, we were unable to find these data in the Anumba *et al.* paper, so cannot verify their source. Sensitivity analyses that adjusted for different proportions of mild, moderate and severe hypertension had little effect on the results.

There is a paucity of data for long-term maternal and neonatal outcomes in women with gestational hypertension, and in the general population of pregnant women with preterm births.

Chapter 7 Conclusions

Implications for service provision

The PlGF and sFlt-1 to PlGF ratio tests are currently used in few UK hospitals. However, the results of the current review suggest that there would be clinical benefit and cost savings of using the Triage PlGF test or the Elecsys sFlt-1 to PlGF ratio test in addition to standard clinical assessment for women suspected of having PE presenting between 20 and 37 weeks of gestation. Sensitivity analyses indicate that replacing quantitative proteinuria testing with the PlGF or sFlt-1 to PlGF ratio biomarker tests, or conducting the biomarker tests in a near-patient (e.g. antenatal clinic) setting (as opposed to a central laboratory) would have negligible impact on cost-effectiveness. It is likely that the most appropriate location and type of testing will vary by local needs, local acquisition and maintenance costs for the test equipment. Investment in equipment and training will be required for any of the biomarker tests to be employed in NHS practice (including those that were not included in the base-case analysis, that is, the DELFIA Xpress PlGF test and the BRAHMS Kryptor sFlt-1 to PlGF ratio test).

Suggested research priorities

The following research priorities (highest priority first) have been identified during the current diagnostic assessment:

- Observational research studies would be helpful to clarify long-term fetal, neonatal and maternal outcomes for women diagnosed with PE.
- Observational research studies would be helpful to clarify maternal and neonatal utilities, specifically utilities associated with PE, suspected PE, and neonatal adverse events (e.g. developmental delays, cerebral palsy and other disabilities).

Given the high risk of clinical review bias identified in the included studies of test accuracy, pragmatic research studies would be helpful to clarify how the PlGF test and sFlt-1 to PlGF ratio test influence key decisions in a clinical setting, and whether or not clinicians using these tests would require specific guidance to ensure that they interpret test results appropriately. Such studies could be of a prospective test accuracy design, but observational studies may also be appropriate.

A prospective head-to-head comparison of the Triage PlGF test and Elecsys immunoassay sFlt-1 to PlGF ratio test, and other relevant PlGF and sFlt-1 to PlGF ratio tests, such as the DELFIA Express PlGF test and BRAHMS Kryptor sFlt-1 to PlGF ratio test, would be helpful to clarify more precisely which test(s) could be most cost-saving for the NHS. However, this would require that the tests being compared employ the same diagnostic or prognostic endpoints and cover the same periods of gestation, which is not currently the case for the Triage PlGF test and Elecsys immunoassay sFlt-1 to PlGF ratio test.

In order to address the current knowledge gaps, research studies to address any of the above research priorities should ensure that the study population is women with suspected PE (i.e. not case–control studies comparing pre-eclamptic and healthy women); be designed in such a way that the risks of bias associated with test accuracy studies are minimised; pragmatically reflect UK clinical practice; include women between 20 and 30 weeks of gestation (in addition to other gestational age groups); and employ definitions of PE that are consistent with those employed in UK clinical practice.

Acknowledgements

We thank Karen Welch, Information Scientist, Southampton Health Technology Assessments Centre, who developed the search strategy, ran bibliographic searches and downloaded the search results into a bibliographic database. We are also grateful to Dr Emma Loveman, Effective Evidence LLC, for co-ordinating the early scoping stage of the project; Dr Keith Cooper for contributing to selection of quality-of-life studies; and Dr Jill Colquitt, Effective Evidence LLC, for acting as an internal editor for the draft report.

We also thank the following members of the project's advisory group who offered clinical advice and comments on draft versions of the protocol, economic model and project report: Ginny Brunton, Midwife, Institute of Education, University College London; Professor Lelia Duley, Director, Nottingham Clinical Trials Unit, Queen's Medical Centre, University of Nottingham; Dr Sam Oddie, Consultant Neonatologist, Neonatal Unit, Bradford Teaching Hospitals NHS Foundation Trust; Steve Robson, Professor of Fetal Medicine, Institute of Cellular Medicine, Newcastle University; and Dr David Williams, Consultant Obstetric Physician, Institute for Women's Health, University College London Hospital.

In addition, the following members of the NICE Diagnostics Assessment Committee kindly provided comments on a draft version of the economic model: Ann Marie Barnard, Chief Executive Officer, Action on Pre-eclampsia; Dr Jenny Myers, Consultant Obstetrician, Maternal and Fetal Health Research Centre, St Mary's Hospital, Manchester and the University of Manchester; Nigel Simpson, Senior Lecturer in Obstetrics and Gynaecology, Faculty of Medicine and Health, University of Leeds; and Suzanne Thomas, Midwifery lead for the Manchester Placenta Clinic and St Mary's Rainbow Clinic.

Contributions of authors

Geoff K Frampton (Senior Research Fellow, Evidence Synthesis) contributed to developing the research protocol, assessed test accuracy studies for inclusion, extracted data from and critically appraised included studies, synthesised evidence, drafted and edited the final report and project managed the review.

Jeremy Jones (Principal Research Fellow, Health Economics) led the economic analysis, assessed economic and HRQoL studies for inclusion, synthesised health economic evidence, critically appraised included studies, developed the health economic model, drafted the report and acted as the project guarantor.

Micah Rose (Research Fellow, Health Economics) assessed test accuracy, economic and HRQoL studies for inclusion, extracted data from and critically appraised included studies, developed the economic analysis, synthesised health economic evidence and drafted the report.

Liz Payne (Research Fellow, Evidence Synthesis) contributed to developing the research protocol, assessed test accuracy studies for inclusion, extracted data from and critically appraised included studies and drafted the report.

Data sharing statement

All data relevant to this technology assessment report are provided in the accompanying appendices, or may be obtained on request from the corresponding author. Note that the current report does not include confidential data that were considered during the NICE diagnostics assessment. Confidential data cannot be shared, but their implications for the conclusions of the diagnostic assessment are clearly stated in the current report.

References

1. Duhig KE, Shennan AH. Recent advances in the diagnosis and management of pre-eclampsia. *F1000Prime Rep* 2015;**7**:24.
2. Powe CE, Levine RJ, Karumanchi SA. Preeclampsia, a disease of the maternal endothelium: the role of antiangiogenic factors and implications for later cardiovascular disease. *Circulation* 2011;**123**:2856–69. http://dx.doi.org/10.1161/CIRCULATIONAHA.109.853127
3. Chappell LC, Enye S, Seed P, Briley AL, Poston L, Shennan AH. Adverse perinatal outcomes and risk factors for preeclampsia in women with chronic hypertension: a prospective study. *Hypertension* 2008;**51**:1002–9. http://dx.doi.org/10.1161/HYPERTENSIONAHA.107.107565
4. Forest JC, Charland M, Massé J, Bujold E, Rousseau F, Lafond J, Giguère Y. Candidate biochemical markers for screening of pre-eclampsia in early pregnancy. *Clin Chem Lab Med* 2012;**50**:973–84. http://dx.doi.org/10.1515/cclm.2011.820
5. Chappell LC, Duckworth S, Seed PT, Griffin M, Myers J, Mackillop L, *et al.* Diagnostic accuracy of placental growth factor in women with suspected preeclampsia: a prospective multicenter study. *Circulation* 2013;**128**:2121–31. http://dx.doi.org/10.1161/CIRCULATIONAHA.113.003215
6. Sibai B. *Pre-Eclampsia Triage by Rapid Assay of Novel Biomarkers of Placental Function and Maternal Adaptation (PETRA)*. Confidential company submission. San Diego, CA: Alere, Inc.; 2014.
7. NHS Choices. *Pre-eclampsia*. 2013. URL: www.nhs.uk/conditions/Pre-eclampsia/Pages/Introduction.aspx (accessed 23 May 2016).
8. Brunelli VB, Prefumo F. Quality of first trimester risk prediction models for pre-eclampsia: a systematic review. *BJOG* 2015;**122**:904–14. http://dx.doi.org/10.1111/1471-0528.13334
9. Health and Social Care Information Centre. *Hospital Episode Statistics, Admitted Patient Care, England 2012–13*. Leeds: Health and Social Care Information Centre; 2013.
10. National Institute for Health and Care Excellence. *Hypertension in Pregnancy*. Quality Standard 35. London: National Institute for Health and Care Excellence; 2013.
11. Knight M, Kenyon S, Brocklehurst P, Neilson J, Shakespeare J, Kurinczuk J, editors. EMBRRACE-UK. *Saving Lives, Improving Mothers' Care: Lessons Learned to Inform Future Maternity Care from the UK and Ireland Confidential Enquiries into Maternal Deaths and Morbidity 2009–12*. Oxford: National Perinatal Epidemiology Unit, University of Oxford; 2014.
12. APEC. *Questions and Answers*. 2015. URL: http://action-on-pre-eclampsia.org.uk/public-area/questions-answers/ (accessed 8 July 2015).
13. National Institute for Health and Care Excellence. *Hypertension in Pregnancy: The Management of Hypertensive Disorders During Pregnancy*. Clinical Guideline 107. London: National Institute for Health and Care Excellence; 2010.
14. American Congress of Obstetricians and Gynecologists. ACOG practice bulletin: clinical management guidelines for obstetrician-gynecologists. Number 33: diagnosis and management of preeclampsia and eclampsia. *Obstet Gynecol* 2002;**99**:159–67. http://dx.doi.org/10.1097/00006250-200201000-00028
15. Brown MA, Lindheimer MD, de Swiet M, Van Assche A, Moutquin JM. The classification and diagnosis of the hypertensive disorders of pregnancy: statement from the international Society for the Study of Hypertension in Pregnancy (ISSHP). *Hypertens Pregnancy* 2001;**20**:ix–xiv. http://dx.doi.org/10.3109/10641950109152635

16. American Congress of Obstetricians and Gynecologists. Hypertension in pregnancy. Report of the American Congress of Obstetricians and Gynecologists' Task Force on Hypertension in Pregnancy. *Obstet Gynecol* 2013;**122**:1122–31.

17. Aloizos S, Seretis C, Liakos N, Aravosita P, Mystakelli C, Kanna E, Gourgiotis S. HELLP syndrome: Understanding and management of a pregnancy-specific disease. *J Obstet Gynaecol* 2013;**33**:331–7. http://dx.doi.org/10.3109/01443615.2013.775231

18. Buchbinder A, Sibai BM, Caritis S, Macpherson C, Hauth J, Lindheimer MD, *et al.* Adverse perinatal outcomes are significantly higher in severe gestational hypertension than in mild preeclampsia. *Am J Obstet Gynecol* 2002;**186**:66–71. http://dx.doi.org/10.1067/mob.2002.120080

19. Morikawa M, Yamada T, Yamada T, Cho K, Yamada H, Sakuragi N, Minakami H. Pregnancy outcome of women who developed proteinuria in the absence of hypertension after mid-gestation. *J Perinat Med* 2008;**36**:419–24. http://dx.doi.org/10.1515/JPM.2008.062

20. Akaishi R, Yamada T, Morikawa M, Nishida R, Minakami H. Clinical features of isolated gestational proteinuria progressing to pre-eclampsia: retrospective observational study. *BMJ Open* 2014;**4**:e004870. http://dx.doi.org/10.1136/bmjopen-2014-004870

21. Mustafa R, Ahmed S, Gupta A, Venuto RC. A comprehensive review of hypertension in pregnancy. *J Pregnancy* 2012;**2012**:105918. http://dx.doi.org/10.1155/2012/105918

22. Bellamy L, Casas J, Hingorani A, Williams D. Pre-eclampsia and risk of cardiovascular disease and cancer in later life: systematic review and meta-analysis. *BMJ* 2007;**335**:974–7. http://dx.doi.org/10.1136/bmj.39335.385301.BE

23. Williams D, Craft N. Pre-eclampsia. *BMJ* 2012;**345**:e4437. http://dx.doi.org/10.1136/bmj.e4437

24. Fisher D, Baird J, Payne L, Lucas P, Kleijnen J, Roberts H, Law C. Are infant size and growth related to burden of disease in adulthood? A systematic review of literature. *Int J Epidemiol* 2006;**35**:1196–210. http://dx.doi.org/10.1093/ije/dyl130

25. O'Reilly M, Sozo F, Harding R. Impact of preterm birth and bronchopulmonary dysplasia on the developing lung: Long-term consequences for respiratory health. *Clin Exp Pharmacol Physiol* 2013;**40**:765–73. http://dx.doi.org/10.1111/1440-1681.12068

26. Stern C, Trapp EM, Mautner E, Deutsch M, Lang U, Cervar-Zivkovic M. The impact of severe preeclampsia on maternal quality of life. *Qual Life Res* 2014;**23**:1019–26. http://dx.doi.org/10.1007/s11136-013-0525-3

27. Postma IR, Groen H, Easterling TR, Tsigas EZ, Wilson ML, Porcel J, Zeeman GG. The brain study: cognition, quality of life and social functioning following preeclampsia; an observational study. *Pregnancy Hypertens* 2013;**3**:227–34. http://dx.doi.org/10.1016/j.preghy.2013.06.003

28. Engelhard IM, van Rij M, Boullart I, Ekhart TH, Spaanderman ME, van den Hout MA, Peeters LL. Posttraumatic stress disorder after pre-eclampsia: an exploratory study. *Gen Hosp Psychiatry* 2002;**24**:260–4. http://dx.doi.org/10.1016/S0163-8343(02)00189-5

29. East C, Conway K, Pollock W, Frawley N, Brennecke S. Women's experiences of preeclampsia: Australian action on preeclampsia survey of women and their confidants. *J Pregnancy* 2011;**2011**:375653. http://dx.doi.org/10.1155/2011/375653

30. Delahaije DH, Dirksen CD, Peeters LL, Smits LJ. Anxiety and depression following preeclampsia or hemolysis, elevated liver enzymes, and low platelets syndrome. A systematic review. *Acta Obstet Gynecol Scand* 2013;**92**:746–61. http://dx.doi.org/10.1111/aogs.12175

31. Hoedjes M, Berks D, Vogel I, Franx A, Bangma M, Darlington AS, *et al.* Postpartum depression after mild and severe preeclampsia. *J Womens Health* 2011;**20**:1535–42. http://dx.doi.org/10.1089/jwh.2010.2584

32. National Institute for Health and Care Excellence. *Antenatal Care*. London: National Institute for Health and Care Excellence, 2008.

33. Schnettler WT, Dukhovny D, Wenger J, Salahuddin S, Ralston SJ, Rana S. Cost and resource implications with serum angiogenic factor estimation in the triage of pre-eclampsia. *BJOG* 2013;**120**:1224–32. http://dx.doi.org/10.1111/1471-0528.12259

34. Bramham K, Briley AL, Seed P, Poston L, Shennan AH, Chappell LC. Adverse maternal and perinatal outcomes in women with previous preeclampsia: a prospective study. *Am J Obstet Gynecol* 2011;**204**:512.e1–9. http://dx.doi.org/10.1016/j.ajog.2011.02.014

35. National Institute for Health and Care Excellence. *NICE Pathway: Hypertension in Pregnancy*. London: National Institute for Health and Care Excellence; 2016. URL: https://pathways.nice.org.uk/pathways/hypertension-in-pregnancy (accessed 25 May 2016).

36. National Institute for Health and Care Excellence. *Pre-eclampsia: The Triage PlGF test, Elecsys Immunoassay sFlt-1/PlGF Ratio, DELFIA Xpress PlGF 1-2-3 Test and BRAHMS sFlt-1 Kryptor/PlGF plus Kryptor PE Ratio to Aid the Assessment of Suspected Pre-eclampsia: Final Scope*. Diagnostics Assessment Programme. London: National Institute for Health and Care Excellence; 2015.

37. Wilkinson H, Trustees and Medical Advisers. Saving mothers' lives. Reviewing maternal deaths to make motherhood safer: 2006–8. *BJOG* 2011;**118**:1402–3. http://dx.doi.org/10.1111/j.1471-0528.2011.03097.x

38. Shennan AH, Redman C, Cooper C, Milne F. Are most maternal deaths from pre-eclampsia avoidable? *Lancet* 2012;**379**:1686–7. http://dx.doi.org/10.1016/S0140-6736(11)60785-X

39. Hadker N, Garg S, Costanzo C, Miller JD, Foster T, van der Helm W, Creeden J. Financial impact of a novel pre-eclampsia diagnostic test versus standard practice: a decision-analytic modeling analysis from a UK healthcare payer perspective. *J Med Econ* 2010;**13**:728–37. http://dx.doi.org/10.3111/13696998.2010.535285

40. Alere, Inc. *The Triage® PLGF Test*. Product insert. San Diego, CA: Alere, Inc.; 2015.

41. Molvarec A, Gullai N, Stenczer B, Fügedi G, Nagy B, Rigó J. Comparison of placental growth factor and fetal flow Doppler ultrasonography to identify fetal adverse outcomes in women with hypertensive disorders of pregnancy: an observational study. *BMC Pregnancy Childbirth* 2013;**13**:161. http://dx.doi.org/10.1186/1471-2393-13-161

42. Benton SJ, Hu Y, Xie F, Kupfer K, Lee SW, Magee LA, von Dadelszen P. Angiogenic factors as diagnostic tests for preeclampsia: a performance comparison between two commercial immunoassays. *Am J Obstet Gynecol* 2011;**205**:469.e1–469.e8. http://dx.doi.org/10.1016/j.ajog.2011.06.058

43. PerkinElmer, Inc. *DELFIA Xpress PlGF 1–2–3 Test to Aid the Assessment of Suspected Pre-eclampsia: Company's Response to Request for Information from NICE*. Company's response to request for information from the National Institute for Health and Care Excellence Diagnostics Assessment Programme. Turku: PerkinElmer, Inc.; 2015.

44. PerkinElmer, Inc. *Performance Evaluation Study for 6007–0030 DELFIA Xpress PlGF Kit at Wallac Oy*. Confidential company report. Turku: PerkinElmer, Inc.; 2014.

45. Roche Diagnostics. *Submission of Further Information on the Roche ELECSYS® Immunoassay sFlt-1/PlGF Ratio Test for Scoping of the Diagnostics Assessment Programme, 'Pre-eclampsia: The Triage PlGF test, Elecsys Immunoassay sFlt-1/PlGF Ratio, DELFIA Xpress PlGF 1–2–3 Test and BRAHMS sFlt-1 Kryptor / PlGF plus Kryptor PE Ratio to Aid the Assessment of Suspected Pre-eclampsia'*. Company's response to a request for information from the National Institute for Health and Care Excellence Diagnostics Assessment Programme. Mannheim: Roche Diagnostics GmbH; 2015.

46. Roche Diagnostics. *Elecsys PlGF test*. Product insert. Mannheim: Roche Diagnostics GmbH; 2014.

47. Chaiworapongsa T, Romero R, Korzeniewski SJ, Kusanovic JP, Soto E, Lam J, *et al.* Maternal plasma concentrations of angiogenic/antiangiogenic factors in the third trimester of pregnancy to identify the patient at risk for stillbirth at or near term and severe late preeclampsia. *Am J Obstet Gynecol* 2013;**208**:287.e1–287.e15. http://dx.doi.org/10.1016/j.ajog.2013.01.016

48. Andersen LB, Frederiksen-Møller B, Work Havelund K, Dechend R, Jørgensen JS, Jensen BL, *et al.* Diagnosis of preeclampsia with soluble Fms–like tyrosine kinase 1/placental growth factor ratio: an inter–assay comparison. *J Am Soc Hypertens* 2015;**9**:86–96. http://dx.doi.org/10.1016/j.jash.2014.11.008

49. Verlohren S, Herraiz I, Lapaire O, Schlembach D, Zeisler H, Calda P, *et al.* New gestational phase-specific cutoff values for the use of the soluble fms-like tyrosine kinase-1/placental growth factor ratio as a diagnostic test for preeclampsia. *Hypertension* 2014;**63**:346–52. http://dx.doi.org/10.1161/HYPERTENSIONAHA.113.01787

50. Roche Diagnostics. *Elecsys sFlt-1 test: Extended Intended Use*. Product insert. Mannheim: Roche Diagnostics GmbH; 2015.

51. Zeisler H, Llurba E, Chantraine F, Vatish M, Staff AC, Sennström M, *et al.* Predictive Value of the sFlt-1:PlGF Ratio in Women with Suspected Preeclampsia. *N Engl J Med* 2016;**374**:13–22. http://dx.doi.org/10.1056/NEJMoa1414838

52. Stepan H, Herraiz I, Schlembach D, Verlohren S, Brennecke S, Chantraine F, *et al.* Implementation of the sFlt-1/PlGF ratio for prediction and diagnosis of pre-eclampsia in singleton pregnancy: implications for clinical practice. *Ultrasound Obstet Gynecol* 2015;**45**:241–6. http://dx.doi.org/10.1002/uog.14799

53. Thermo Fisher Scientific. *ThermoFisher Scientific Response to Request for Information from NICE Diagnostics Assessment Programme on 'Pre-eclampsia: The Triage PlGF Test (and Alternative Technologies Identified During Scoping) to Aid the Assessment of Suspected Pre-eclampsia (Provisional Title)'*. Company's response to a request for information from the National Institute for Health and Care Excellence Diagnostics Assessment Programme. Hennigsdorf: Thermo Fisher Scientific GmbH; 2015.

54. Thermo Fisher Scientific. *BRAHMS sFlt-1 Kryptor*. Product insert. Hennigsdorf: Thermo Fisher Scientific GmbH; 2015.

55. Thermo Fisher Scientific. *BRAHMS PlGF plus Kryptor*. Product insert. Hennigsdorf: Thermo Fisher Scientific GmbH; 2015.

56. Verlohren S, Herraiz I, Lapaire O, Schlembach D, Moertl M, Zeisler H, *et al.* The sFlt-1/PlGF ratio in different types of hypertensive pregnancy disorders and its prognostic potential in preeclamptic patients. *Am J Obstet Gynecol* 2012;**206**:58.e1–58.e8. http://dx.doi.org/10.1016/j.ajog.2011.07.037

57. Vellanki K. Pregnancy in chronic kidney disease. *Adv Chronic Kidney Dis* 2013;**20**:223–8. http://dx.doi.org/10.1053/j.ackd.2013.02.001

58. Lu J, Zhao YY, Qiao J, Zhang HJ, Ge L, Wei Y. A follow-up study of women with a history of severe preeclampsia: relationship between metabolic syndrome and preeclampsia. *Chin Med J* 2011;**124**:775–9.

59. Cheng YW, McLaughlin GB, Esakoff TF, Block-Kurbisch I, Caughey AB. Glucose challenge test: Screening threshold for gestational diabetes mellitus and associated outcomes. *J Matern Fetal Neonatal Med* 2007;**20**:903–8. http://dx.doi.org/10.1080/14767050701739384

60. Alere, Inc. *The Alere Triage® PlGF Test to Aid the Diagnosis of Pre-term Pre-eclampsia and Prediction of Adverse Pregnancy Outcome:* Company's response to a request for information from the National Institute for Health and Care Excellence Diagnostics Assessment Programme. San Diego, CA: Alere, Inc.; 2015.

61. Roche Diagnostics. *Information on the Roche ELECSYS® Immunoassay sFlt-1/PlGF Ratio Test for Scoping of the Diagnostics Assessment Programme, 'Pre-eclampsia: The Triage PlGF test and Alternative Tests Identified at Scoping'*. Company's response to a request for information from the National Institute for Health and Care Excellence Diagnostics Assessment Programme. Mannheim: Roche Diagnostics GmbH; 2015.

62. Macaskill P, Gatsonis C, Deeks JJ, Harbord RM, Takwoingi Y. Analysing and Presenting Results. In Deeks JJ, Bossuyt PM, Gatsonis C, editors. *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy*. Version 1.0. London: Cochrane; 2010.

63. Centre for Reviews and Dissemination. *Systematic Reviews: CRD's Guidance for Undertaking Reviews in Health Care*. 3rd edn. York: York Publishing Services Ltd, Centre for Reviews and Dissemination; 2009.

64. Bossuyt P, Davenport C, Deeks J, Hyde C, Leeflang M, Scholten R. Interpreting Results and Drawing Conclusions. In Deeks JJ, Bossuyt PM, Gatsonis C, editors. *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy*. Version 0.9. London: Cochrane; 2013.

65. Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: the PRISMA statement. *BMJ* 2009;**339**:b2535. http://dx.doi.org/10.1136/bmj.b2535

66. Southampton Health Technology Assessments Centre. *The Triage PlGF Test, Elecsys Immunoassay sFlt-1 / PlGF Ratio, DELFIA Xpress PlGF 1–2-3 Test and BRAHMS sFlt-1 Kryptor/PlGF plus Kryptor PE Ratio to Aid the Assessment of Suspected Pre-eclampsia: Systematic Review and Economic Evaluation. Protocol for Evidence Assessment and Analysis Report Commissioned by the NIHR HTA programme on Behalf of the National Institute for Health and Care Excellence*. Southhampton: Southampton Health Technology Assessments Centre; 2015. URL: www.nice.org.uk/guidance/gid-dt26/documents/the-triage-plgf-test-elecsys-immunoassay-sflt1plgf-ratio-delfia-xpress-plgf-123-test-and-brahms-sflt1-kryptor-plgf-plus-kryptor-pe-ratio-to-aid-the-assessment-of-suspected-preeclampsia-final-protocol2 (accessed 25 May 2016).

67. Reitsma JB, Rutjes AWS, Whiting P, Vlassov VV, Leeflang MMG, Deeks JJ. Assessing Methodological Quality. In Deeks JJ, Bossuyt PM, Gatsonis C, editors. *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0.0*. London: Cochrane; 2009.

68. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003;**3**:25. http://dx.doi.org/10.1186/1471-2288-3-25

69. National Institute for Health and Care Excellence. *Diagnostics Assessment Programme Manual*. London: National Institute for Health and Care Excellence 2011.

70. Chappell L, Duckworth S, Griffin M, Seed P, Redman C, Shennan A. OP007. PLGF in combination with other commonly utilised tests and other biomarkers for predicting need for delivery for pre-eclampsia within 14 days in women presenting prior to 35 weeks' gestation. *Pregnancy Hypertens* 2013;**3**:64–5. http://dx.doi.org/10.1016/j.preghy.2013.04.023

71. Chappell L, Duckworth S, Griffin M, Seed P, Redman C, Shennan A. 636: Plasma placental growth factor (PlGF) in the diagnosis of women with pre-eclampsia requiring delivery within 14 days: the PELICAN study. *Am J Obstet Gynecol* 2013;**208**(Suppl. 1):S269–70. http://dx.doi.org/10.1016/j.ajog.2012.10.802

72. Chappell LC, Duckworth S, Griffin M, Tarft H, Seed P, Redman CW, Shennan AH, PELICAN study consortium. OS100. Plasma placental growth factor (PLGF) measurement in women presenting with suspected pre-eclampsia: the pelican study. *Pregnancy Hypertens* 2012;**2**:233–4. http://dx.doi.org/10.1016/j.preghy.2012.04.101

73. Zeisler H, Llurba E, Chantraine F, Vatish M, Staff A, Sennstrom M, *et al. sFlt-1/PlGF for Prediction of Short-Term Outcome in Women with Suspected Preeclampsia*. Unpublished manuscript (confidential document) submitted to the National Institute for Health and Care Excellence Diagnostics Assessment Programme. Mannheim: Roche Diagnostics GmbH; 2015.

74. Zeisler H, Llurba E, and Chantraine F. *Prediction of Short-Term Outcome in Pregnant Women with Suspected Preeclampsia: The PROGNOSIS Study*. The 20th World Congress on Controversies in Obstetrics, Gynecology & Infertility, Paris, France, 4–7 December 2014.

75. Álvarez-Fernández I, Prieto B, Rodríguez V, Ruano Y, Escudero AI, Álvarez FV. New biomarkers in diagnosis of early onset preeclampsia and imminent delivery prognosis. *Clin Chem Lab Med* 2014;**52**:1159–68. http://dx.doi.org/10.1515/cclm-2013-0901

76. Álvarez-Fernández I, Prieto B, Ruano Y, Escudero A, Álvarez FV. sFlt-1/PlGF ratio as biomarker in the diagnosis of preeclampsia in singleton and multiple pregnancies. *Biochim Clin* 2013;**37**:S400.

77. Hund M, Allegranza D, Schoedl M, Dilba P, Verhagen-Kamerbeek W, Stepan H. Multicenter prospective clinical study to evaluate the prediction of short-term outcome in pregnant women with suspected preeclampsia (PROGNOSIS): study protocol. *BMC Pregnancy Childbirth* 2014;**14**:324. http://dx.doi.org/10.1186/1471-2393-14-324

78. Bdolah Y, Lam C, Rajakumar A, Shivalingappa V, Mutter W, Sachs BP, *et al.* Twin pregnancy and the risk of preeclampsia: bigger placenta or relative ischemia? *Am J Obstet Gynecol* 2008;**198**:428.e1–6. http://dx.doi.org/10.1016/j.ajog.2007.10.783

79. Caughey AB, Stotland NE, Washington AE, Escobar GJ. Maternal ethnicity, paternal ethnicity, and parental ethnic discordance: predictors of preeclampsia. *Obstet Gynecol* 2005;**106**:156–61. http://dx.doi.org/10.1097/01.AOG.0000164478.91731.06

80. Xiao J, Shen F, Xue Q, Chen G, Zeng K, Stone P, *et al.* Is ethnicity a risk factor for developing preeclampsia? An analysis of the prevalence of preeclampsia in China. *J Hum Hypertens* 2014;**28**:694–8. http://dx.doi.org/10.1038/jhh.2013.148

81. Anderson NH, Sadler LC, Stewart AW, Fyfe EM, McCowan LM. Ethnicity, body mass index and risk of pre-eclampsia in a multiethnic New Zealand population. *Aust N Z J Obstet Gynaecol* 2012;**52**:552–8. http://dx.doi.org/10.1111/j.1479-828X.2012.01475.x

82. Sohlberg S, Stephansson O, Cnattingius S, Wikström AK. Maternal body mass index, height, and risks of preeclampsia. *Am J Hypertens* 2012;**25**:120–5. http://dx.doi.org/10.1038/ajh.2011.175

83. Roche Diagnostics. *Elecsys sFlt-1 test*. Product insert. Mannheim: Roche Diagnostics GmbH; 2015.

84. Verlohren S, Galindo A, Schlembach D, Zeisler H, Herraiz I, Moertl MG, *et al.* An automated method for the determination of the sFlt-1/PlGF ratio in the assessment of preeclampsia. *Am J Obstet Gynecol* 2010;**202**:161.e1–161.e11. http://dx.doi.org/10.1016/j.ajog.2009.09.016

85. Gencay M, Denk B, Dinkel C, Pfeffer M, Hund M. *Comparison of Elecsys® sFlt-1/PlGF assays of Roche and Alere Triage® PLGF Assay for Accurate Pre-eclampsia Diagnosis During Pregnancy*. The 23rd European Congress of Obstetrics and Gynaecology, Glasgow, UK, 1–10 May 2014.

86. PerkinElmer, Inc. *Performance Evaluation Study for 6007–0030 DELFIA Xpress PlGF Kit at Kings*. Confidential company report. Turku: PerkinElmer, Inc.; 2015.

87. Thermo Fisher Scientific. *Correlation Data: Kryptor – Elecsys (PlGF and sFlt-1)*. Confidential company report submitted to the National Institute for Health and Care Excellence Diagnostics Assessment Programme. Hennigsdorf: Thermo Fisher Scientific GmbH; 2015.

88. National Institute for Health and Care Excellence. *Guide to the Methods of Technology Appraisal 2013*. London: National Institute for Health and Care Excellence; 2013.

89. Dakin H. Review of studies mapping from quality of life or clinical measures to EQ-5D: an online database. *Health Qual Life Outcomes* 2013;**11**:151. http://dx.doi.org/10.1186/1477-7525-11-151

90. Shmueli A, Meiri H, Gonen R. Economic assessment of screening for pre-eclampsia. *Prenat Diagn* 2012;**32**:29–38. http://dx.doi.org/10.1002/pd.2871

91. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. *Methods for the Economic Evaluation of Health Care Programmes*. 3rd edn. Oxford: Oxford University Press; 2005.

92. Philips Z, Ginnelly L, Sculpher M, Claxton K, Golder S, Riemsma R, *et al.* Review of guidelines for good practice in decision-analytic modelling in health technology assessment. *Health Technol Assess* 2004;**8**(36). http://dx.doi.org/10.3310/hta8360

93. Hadker N, Garg S, Costanzo C, Miller JD, van der Helm W, Foster T, *et al.* Financial impact of a novel preeclampsia diagnostic test vs. standard care: a decision-analytic modeling analysis from a German health care payer perspective. *Value Health* 2009;**12**:A321. http://dx.doi.org/10.1016/S1098-3015(10)74576-9

94. Hunter R, Duckworth S, Seed P, Shennan A, Chappell L. PP051. Budget impact analysis of maternal plasma PlGF concentrations in women with suspected pre-eclampsia: The potential for improved health service usage. *Pregnancy Hypertens* 2013;**3**:85. http://dx.doi.org/10.1016/j.preghy.2013.04.078

95. Meads CA, Cnossen JS, Meher S, Juarez-Garcia A, ter Riet G, Duley L, *et al.* Methods of prediction and prevention of pre-eclampsia: systematic reviews of accuracy and effectiveness literature with economic modelling. *Health Technol Assess* 2008;**12**(6). http://dx.doi.org/10.3310/hta12060

96. Hadker N, Garg S, Costanzo C, van der Helm W, Creeden J. Are there financial savings associated with supplementing current diagnostic practice for preeclampsia with a novel test? Learnings from a modeling analysis from a German payer perspective. *Hypertens Pregnancy* 2013;**32**:105–19. http://dx.doi.org/10.3109/10641955.2011.638958

97. Hunter R. *Budget Impact Analysis of the Triage® PlGF Test for Women < 35 Weeks Gestation with Suspected but not Diagnosed Pre-eclampsia*. Unpublished report (confidential document) prepared by University College London for Alere. London: University College London; 2013.

98. Duckworth S, Chappell LC, Seed PT, Mackillop L, Shennan AH, Hunter R. *Placental Growth Factor (PlGF) in Women with Suspected Pre-eclampsia Prior to 35 Weeks' Gestation: A Budget Impact Analysis*. Unpublished manuscript (confidential document) submitted to the National Institute for Health and Care Excellence Diagnostics Assessment Programme. San Diego, CA: Alere, Inc.; 2015.

99. Roche Diagnostics. *sFlt-1 PlGF Ratio Cost Analysis – User Guide v(2 1)*. Confidential company submission to the National Institute for Health and Care Excellence Diagnostics Assessment Programme. Mannheim: Roche Diagnostics GmbH; 2015.

100. Murphy DJ, Stirrat GM. Mortality and morbidity associated with early-onset preeclampsia. *Hypertens Pregnancy* 2000;**19**:221–31. http://dx.doi.org/10.1081/PRG-100100138

101. Bhattacharya S, Campbell DM. The incidence of severe complications of preeclampsia. *Hypertens Pregnancy* 2000;**24**:181–90. http://dx.doi.org/10.1081/PRG-200059873

102. Rana S, Powe CE, Salahuddin S, Verlohren S, Perschel FH, Levine RJ, *et al.* Angiogenic factors and the risk of adverse outcomes in women with suspected preeclampsia. *Circulation* 2012;**125**:911–19. http://dx.doi.org/10.1161/CIRCULATIONAHA.111.054361

103. Calderon-Margalit R, Friedlander Y, Yanetz R, Deutsch L, Manor O, Harlap S, Paltiel O. Late stillbirths and long-term mortality of mothers. *Obstet Gynecol* 2007;**109**:1301–8. http://dx.doi.org/10.1097/01.AOG.0000264548.10767.ea

104. Calderon-Margalit R, Friedlander Y, Manor O, Meiner V, Sagi Y, Tsadok MA, *et al. Cardiovascular Health of Mothers and Offspring following Pregnancies Complicated with Preeclampsia*. The 15th Annual Meeting of the Israel Medical Association Society for Research, Prevention and Treatment of Atherosclerosis Conference, Eilat, Israel, 14–17 October 2010.

105. Tsadok MA, Friedlander Y, Paltiel O, Ezra Y, Manor O, Siscovick D, *et al.* 72: Long term consequences of maternal GDM on their offspring from the Jerusalem perinatal study birth cohort. *Am J Obstet Gynecol* 2008;**199**:S33. http://dx.doi.org/10.1016/j.ajog.2008.09.097

106. Sonnenberg FA, Burkman RT, Hagerty CG, Speroff L, Speroff T. Costs and net health effects of contraceptive methods. *Contraception* 2004;**69**:447–59. http://dx.doi.org/10.1016/j.contraception.2004.03.008

107. Edwards SJ, Campbell HE, Plumb JM. Cost-utility analysis comparing meropenem with imipenem plus cilastatin in the treatment of severe infections in intensive care. *Eur J Health Econ* 2006;**7**:72–8. http://dx.doi.org/10.1007/s10198-005-0333-y

108. Costeloe K, Hennessy E, Gibson AT, Marlow N, Wilkinson AR. The EPICure study: outcomes to discharge from hospital for infants born at the threshold of viability. *Pediatrics* 2000;**106**:659–71. http://dx.doi.org/10.1542/peds.106.4.659

109. Moore T, Hennessy EM, Myles J, Johnson SJ, Draper ES, Costeloe KL, Marlow N. Neurological and developmental outcome in extremely preterm children born in England in 1995 and 2006: the EPICure studies. *BMJ* 2012;**345**:e7961. http://dx.doi.org/10.1136/bmj.e7961

110. Marret S, Ancel PY, Marpeau L, Marchand L, Pierrat V, Larroque B, *et al.* Neonatal and 5-year outcomes after birth at 30–4 weeks of gestation. *Obstet Gynecol* 2007;**110**:72–80. http://dx.doi.org/10.1097/01.AOG.0000267498.95402.bd

111. Bijlenga D, Boers KE, Birnie E, Mol BWJ, Vijgen SC, Van der Post JA, *et al.* Maternal health-related quality of life after induction of labor or expectant monitoring in pregnancy complicated by intrauterine growth retardation beyond 36 weeks. *Qual Life Res* 2011;**20**:1427–36. http://dx.doi.org/10.1007/s11136-011-9891-x

112. Bijlenga D, Koopmans CM, Birnie E, Mol BW, van der Post JA, Bloemenkamp KW, *et al.* Health-related quality of life after induction of labor versus expectant monitoring in gestational hypertension or preeclampsia at term. *Hypertens Pregnancy* 2011;**30**:260–74. http://dx.doi.org/10.3109/10641955.2010.486458

113. Chang PJ, Lin CC, Chen YC, Chuang CH, Tseng YC, Hsieh WS, *et al.* Use of herbal dietary supplement si-wu-tang and health-related quality of life in postpartum women: a population-based correlational study. *Evid Based Complement Alternat Med* 2013;**2013**:790474. http://dx.doi.org/10.1155/2013/790474

114. Chang PJ, Tseng YC, Chuang CH, Chen YC, Hsieh WS, Hurng BS, *et al.* Use of Sheng-Hua-Tang and health-related quality of life in postpartum women: A population-based cohort study in Taiwan. *Int J Nurs Stud* 2010;**47**:13–19. http://dx.doi.org/10.1016/j.ijnurstu.2009.06.002

115. Hoedjes M, Berks D, Vogel I, Franx A, Duvekot JJ, Steegers EA, Raat H. Poor health-related quality of life after severe preeclampsia. *Birth* 2011;**38**:246–55. http://dx.doi.org/10.1111/j.1523-536X.2011.00477.x

116. Jansen AJ, Essink-Bot ML, Duvekot JJ, van Rhenen DJ. Psychometric evaluation of health-related quality of life measures in women after different types of delivery. *J Psychosom Res* 2007;**63**:275–81. http://dx.doi.org/10.1016/j.jpsychores.2007.06.003

117. Torkan B, Parsay S, Lamyian M, Kazemnejad A, Montazeri A. Postnatal quality of life in women after normal vaginal delivery and caesarean section. *BMC Pregnancy Childbirth* 2009;**9**:4. http://dx.doi.org/10.1186/1471-2393-9-4

118. Ngai FW, Ngu SF. Quality of life during the transition to parenthood in Hong Kong: a longitudinal study. *J Psychosom Obstet Gynaecol* 2013;**34**:157–62. http://dx.doi.org/10.3109/0167482X.2013.852534

119. Petrou S, Morrell J, Spiby H. Assessing the empirical validity of alternative multi-attribute utility measures in the maternity context. *Health Qual Life Outcomes* 2009;**7**:40. http://dx.doi.org/10.1186/1477-7525-7-40

120. Joint Formulary Committee. *British National Formulary* (online). London: BMJ Group and Pharmaceutical Press. URL: www.medicinescomplete.com (accessed 25 May 2016).

121. HM Treasury. *The Green Book: Appraisal and Evaluation in Central Government*. 2015. URL: www.gov.uk/government/publications/the-green-book-appraisal-and-evaluation-in-central-governent (accessed 25 May 2016).

122. National Collaborating Centre for Women's and Children's Health. *Hypertension in Pregnancy: The Management of Hypertensive Disorders during Pregnancy* [revised reprint 2011]. London: National Collaborating Centre for Women's and Children's Health; 2010.

123. Chappell LC, Milne F, Shennan A. Is early induction or expectant management more beneficial in women with late preterm pre-eclampsia? *BMJ* 2015;**350**:h191. http://dx.doi.org/10.1136/bmj.h191

124. Koopmans CM, Bijlenga D, Groen H, Vijgen SM, Aarnoudse JG, Bekedam DJ, *et al.* Induction of labour versus expectant monitoring for gestational hypertension or mild pre-eclampsia after 36 weeks' gestation (HYPITAT): a multicentre, open-label randomised controlled trial. *Lancet* 2009;**374**:979–88. http://dx.doi.org/10.1016/S0140-6736(09)60736-4

125. Gomez-Arriaga PI, Herraiz I, Lopez-Jimenez EA, Escribano D, Denk B, Galindo A. Uterine artery Doppler and sFlt-1/PlGF ratio: prognostic value in early-onset pre-eclampsia. *Ultrasound Obstet Gynecol* 2014;**43**:525–32. http://dx.doi.org/10.1002/uog.13224

126. Broekhuijsen K, van Baaren GJ, van Pampus MG, Ganzevoort W, Sikkema JM, Woiski MD, *et al.* Immediate delivery versus expectant monitoring for hypertensive disorders of pregnancy between 34 and 37 weeks of gestation (HYPITAT-II): an open-label, randomised controlled trial. *Lancet* 2015;**385**:2492–501. http://dx.doi.org/10.1016/S0140-6736(14)61998-X

127. Department of Health. *National Schedule of Reference Costs 2013–14 for NHS Trusts and NHS Foundation Trusts*. London: Department of Health; 2014.

128. Department of Health. *NHS Payment by Results Tariff 2013/14*. London: Department of Health; 2014. URL: www.gov.uk/government/publications/payment-by-results-pbr-operational-guidance-and-tariffs (accessed 25 May 2016).

129. Habli M, Levine RJ, Qian C, Sibai B. Neonatal outcomes in pregnancies with preeclampsia or gestational hypertension and in normotensive pregnancies that delivered at 35, 36 or 37 weeks of gestation. *Am J Obstet Gynecol* 2007;**197**:406.e1–406.e7. http://dx.doi.org/10.1016/j.ajog.2007.06.059

130. Akhtar W, Chung Y. Saving the NHS one blood test at a time. *BMJ Qual Improv Rep* 2014;**2**:u204012.w1749.

131. Ara R, Brazier J. Deriving an algorithm to convert the eight mean SF-36 dimension scores into a mean EQ-5D preference-based score from published studies (where patient level data are not available). *Value Health* 2008;**11**:1131–43. http://dx.doi.org/10.1111/j.1524-4733.2008.00352.x

132. Sullivan PW, Ghushchyan V. Mapping the EQ-5D index from the SF-12: US general population preferences in a nationally representative sample. *Med Decis Making* 2006;**26**:401–9. http://dx.doi.org/10.1177/0272989X06290496

133. Hirashima C, Ohkuchi A, Takahashi K, Suzuki H, Matsubara S, Suzuki M. A novel three-step approach for predicting the imminent onset of preeclampsia within 4 weeks after blood sampling at 19–31 weeks of gestation. *Hypertens Res* 2014;**37**:519–25. http://dx.doi.org/10.1038/hr.2014.31

134. Anumba DO, Lincoln K, Robson SC. Predictive value of clinical and laboratory indices at first assessment in women referred with suspected gestational hypertension. *Hypertens Pregnancy* 2010;**29**:163–79. http://dx.doi.org/10.3109/10641950902968635

135. Atzeni I, Paoletti AM, Deiana SF, Meloni A, Guerriero S, Parodo G, *et al.* Placental growth factor (PLGF): correlations with placental function. *Int J Gynaecol Obstet* 2012;**119**:S782–3. http://dx.doi.org/10.1016/S0020-7292(12)61963-8

136. Benton S, Knudsen U, Hu Y, Vittinghus E, Allen J, Konborg C, *et al.* Angiogenic factor imbalance is not specific to the maternal syndrome of pre-eclampsia. *Reprod Sci* 2012;**19**:317A–18A.

137. Benton S, Hu Y, Fang X, Knudsen AB, Kronborg CS, Vittinghus E, *et al.* Placental growth factor as a diagnostic for early onset pre-eclampsia and normotensive intrauterine growth restriction of placental origin. *Pregnancy Hypertens* 2010;**1**(Suppl. 1):S24. http://dx.doi.org/10.1016/S2210-7789(10)60103-1

138. Benton S, Hu YX, Xie F, Kupfer K, Lee SW, Magee L, *et al.* Placental growth factor quantified on rapid assay as a diagnostic test for early-onset pre-eclampsia. *Placenta* 2011;**32**:A84.

139. Benton SJ, Hu Y, Xie F, Kupfer K, Lee SW, Magee LA, von Dadelszen P. Can placental growth factor in maternal circulation identify fetuses with placental intrauterine growth restriction? *Am J Obstet Gynecol* 2012;**206**:163.e1–7. http://dx.doi.org/10.1016/j.ajog.2011.09.019

140. Bersinger NA, Odegard RA. Serum levels of macrophage colony stimulating, vascular endothelial, and placenta growth factor in relation to later clinical onset of pre-eclampsia and a small-for-gestational age birth. *Am J Reprod Immunol* 2005;**54**:77–83. http://dx.doi.org/10.1111/j.1600-0897.2005.00290.x

141. Boucoiran I, Rey E, Delvin E, Audibert F. Repeated measures of PlGF, PP13 and ADAM12 at first and second trimester for preeclampsia screening. *Am J Obstet Gynecol* 2012;**206**:S337. http://dx.doi.org/10.1016/j.ajog.2011.10.781

142. Boucoiran I, Thissier-Levy S, Wu Y, Wei SQ, Zhong-Cheng L, Delvin E, *et al.* Risk for preeclampsia and intrauterine growth restriction: effective value of PlGF, Sflt-1 and Inhibin A in singleton and multiple pregnancies. *Am J Obstet Gynecol* 2012;**206**(Suppl. 1):S336–7. http://dx.doi.org/10.1016/j.ajog.2011.10.780

143. Boucoiran I, Thissier-Levy S, Wu Y, Wei SQ, Luo ZC, Delvin E, *et al.* Risks for preeclampsia and small for gestational age: predictive values of placental growth factor, soluble fms-like tyrosine kinase-1, and inhibin A in singleton and multiple-gestation pregnancies. *Am J Perinatol* 2013;**30**:607–12.

144. Bramham K, Via P, Laivuori H, Poston L, Mistry HD, Predo CLC. Prediction and mechanisms of superimposed pre-eclampsia in women with chronic hypertension. *Reprod Sci* 2014;**21**(Suppl. 1):187A.

145. Bramham K, Seed P, Nelson-Piercy C, Lightstone L, Ashford L, Butler J, *et al.* Diagnostic accuracy of placental growth factor in women with chronic kidney disease or hypertension and suspected preeclampsia: A prospective cohort study. *Pregnancy Hypertens* 2015;**5**:21. http://dx.doi.org/10.1016/j.preghy.2014.10.046

146. Calabrese S, Cardellicchio M, Mazzocco M, Taricco E, Martinelli A, Cetin I. Placental growth factor (PLGF) maternal circulating levels in normal pregnancies and in pregnancies at risk of developing placental insufficiency complications. *Reprod Sci* 2012;**19**:211A–12A.

147. Cappellini F, Algeri P, Ornaghi S, Bernasconi D, Signorini S, Brambilla P, *et al.* Diagnostic accuracy of sFlt-1, PlGF, sFlt-1/PlGF ratio and PTX3 in preeclampsia. *Biochim Clin* 2012;**36**:583.

148. Caruhel P, Guerin N, Demirdjian G. Development of a new highly sensitive automated placental growth factor assay, using BRAHMS KRYPTOR compact PLUS. *Prenatal Diagnosis* 2013;**33**(Suppl. 1):54–5.

149. Cetin I, Mazzocco M, Giardini V, Calabrese S, Algeri P, Martinelli A, *et al.* Placental growth factor for the prediction of fetal outcomes in pregnancies at risk and in IUGR. *J Matern Fetal Neonatal Med* 2014;**27**:135.

150. Chaiworapongsa T, Romero R, Than NG, Hernandez-Andrade E, Savasan ZA, Tarca AL, *et al.* A prospective cohort study demonstrates that maternal plasma concentrations of angiogenic/anti-angiogenic factors have prognostic value in women presenting with suspected preeclampsia to the obstetrical triage area. *Reprod Sci* 2012;**19**(Suppl. 1):212A.

151. Chaiworapongsa T, Romero R, Savasan ZA, Kusanovic JP, Ogge G, Soto E, *et al.* Maternal plasma concentrations of angiogenic/anti-angiogenic factors are of prognostic value in patients presenting to the obstetrical triage area with the suspicion of preeclampsia. *J Matern Fetal Neonatal Med* 2011;**24**:1187–207. http://dx.doi.org/10.3109/14767058.2011.589932

152. Chaiworapongsa T, Romero R, Savasan ZA, Ogge G, Soto E, Dong Z, *et al.* Maternal plasma concentrations of angiogenic/anti-angiogenic factors are of prognostic value in patients suspected to have preeclampsia presenting to the obstetrical triage. *Reprod Sci* 2011;**18**:T-004.

153. Chaiworapongsa T, Romero R, Korzeniewski SJ, Cortez JM, Pappas A, Tarca AL, *et al.* Plasma concentrations of angiogenic/anti-angiogenic factors have prognostic value in women presenting with suspected preeclampsia to the obstetrical triage area: a prospective study. *J Matern Fetal Neonatal Med* 2014;**27**:132–44. http://dx.doi.org/10.3109/14767058.2013.806905

154. Cho Y, Park HJ, Kim SH, Kim JY, Lee KJ, Cha DH. Screening models for early detection of late-onset preeclampsia with various markers in low-risk pregnancy population. *Placenta* 2013;**34**:A42. http://dx.doi.org/10.1016/j.placenta.2013.06.126

155. Crispi F, Llurba E, Domínguez C, Martín-Gallán P, Cabero L, Gratacós E. Predictive value of angiogenic factors and uterine artery Doppler for early- versus late-onset pre-eclampsia and intrauterine growth restriction. *Ultrasound Obstet Gynecol* 2008;**31**:303–9. http://dx.doi.org/10.1002/uog.5184

156. Daubert C, Morfeld C, Schild R. PP052. Diagnosis of pre-eclampsia and prediction of time to delivery in women with suspected pre-eclampsia presenting before 36 weeks' gestation. *Pregnancy Hypertens* 2013;**3**:85–6. http://dx.doi.org/10.1016/j.preghy.2013.04.079

157. De Oliveira L, Peraçoli JC, Peraçoli MT, Korkes H, Zampieri G, Moron AF, Sass N. sFlt-1/PlGF ratio as a prognostic marker of adverse outcomes in women with early-onset preeclampsia. *Pregnancy Hypertens* 2013;**3**:191–5. http://dx.doi.org/10.1016/j.preghy.2013.02.003

158. De Vivo A, Baviera G, Giordano D, Todarello G, Corrado F, D'anna R. Endoglin, PlGF and sFlt-1 as markers for predicting pre-eclampsia. *Acta Obstet Gynecol Scand* 2008;**87**:837–42. http://dx.doi.org/10.1080/00016340802253759

159. Diab AE, El-Behery MM, Ebrahiem MA, Shehata AE. Angiogenic factors for the prediction of pre-eclampsia in women with abnormal midtrimester uterine artery Doppler velocimetry. *Int J Gynaecol Obstet* 2008;**102**:146–51. http://dx.doi.org/10.1016/j.ijgo.2008.02.016

160. Doherty A, Carvalho JC, Drewlo S, El-Khuffash A, Downey K, Dodds M, Kingdom J. Altered hemodynamics and hyperuricemia accompany an elevated sFlt-1/PlGF ratio before the onset of early severe preeclampsia. *J Obstet Gynaecol Can* 2014;**36**:692–700. http://dx.doi.org/10.1016/S1701-2163(15)30511-9

161. Dröge L, Herraìz I, Zeisler H, Schlembach D, Stepan H, Küssel L, *et al.* Maternal serum sFlt-1/PlGF ratio in twin pregnancies with and without pre-eclampsia in comparison with singleton pregnancies. *Ultrasound Obstet Gynecol* 2015;**45**:286–93. http://dx.doi.org/10.1002/uog.14760

162. Engels T, Pape J, Schoofs K, Henrich W, Verlohren S. Automated measurement of sFlt1, PlGF and sFlt1/PlGF ratio in differential diagnosis of hypertensive pregnancy disorders. *Hypertens Pregnancy* 2013;**32**:459–73. http://dx.doi.org/10.3109/10641955.2013.827205

163. Engels T, Pape J, Schoofs K, Denk B, Beinder E, Verlohren S. P9. The elecsys assay for PlGF, sFlt1 and their ratio (sFlt1/PlGF) as an aid in differential diagnosis of pregnancy-related hypertensive disorders. *Pregnancy Hypertens* 2011;**1**:276–7. http://dx.doi.org/10.1016/j.preghy.2011.08.070

164. Erez O, Romero R, Espinoza J, Fu W, Todem D, Kusanovic JP, *et al.* The change in concentrations of angiogenic and anti-angiogenic factors in maternal plasma between the first and second trimesters in risk assessment for the subsequent development of preeclampsia and small-for-gestational age. *J Matern Fetal Neonatal Med* 2008;**21**:279–87. http://dx.doi.org/10.1080/14767050802034545

165. Espinoza J, Romero R, Nien JK, Gomez R, Kusanovic JP, Gonçalves LF, *et al.* Identification of patients at risk for early onset and/or severe preeclampsia with the use of uterine artery Doppler velocimetry and placental growth factor. *Am J Obstet Gynecol* 2007;**196**:326.e1–13. http://dx.doi.org/10.1016/j.ajog.2006.11.002

166. Forest JC, Thériault S, Massé J, Bujold E, Giguère Y. Soluble Fms-like tyrosine kinase-1 to placental growth factor ratio in mid-pregnancy as a predictor of preterm preeclampsia in asymptomatic pregnant women. *Clin Chem Lab Med* 2014;**52**:1169–78. http://dx.doi.org/10.1515/cclm-2013-0955

167. Garcia-Tizon Larroca S, Tayyar A, Poon LC, Wright D, Nicolaides KH. Competing risks model in screening for preeclampsia by biophysical and biochemical markers at 30–3 weeks' gestation. *Fetal Diagn Ther* 2014;**36**:9–17. http://dx.doi.org/10.1159/000362518

168. Gaziano E, Wothe D, Nunnelly P, Jones K, Martin B, Sunderji S. Prototype sVEGFR1 and PlGF automated immunoassays aid in the diagnosis of preterm preeclampsia. *Clin Chem* 2008;**54**:A36.

169. Ghosh SK, Raheja S, Tuli A, Raghunandan C, Agarwal S. Can maternal serum placental growth factor estimation in early second trimester predict the occurrence of early onset preeclampsia and/or early onset intrauterine growth restriction? A prospective cohort study. *J Obstet Gynaecol Res* 2013;**39**:881–90. http://dx.doi.org/10.1111/jog.12006

170. Ghosh SK, Raheja S, Tuli A, Raghunandan C, Agarwal S. Is serum placental growth factor more effective as a biomarker in predicting early onset preeclampsia in early second trimester than in first trimester of pregnancy? *Arch Gynecol Obstet* 2013;**287**:865–73. http://dx.doi.org/10.1007/s00404-012-2662-2

171. Ghosh SK, Raheja S, Tuli A, Raghunandan C, Agarwal S. Serum PLGF as a potential biomarker for predicting the onset of preeclampsia. *Arch Gynecol Obstet* 2012;**285**:417–22. http://dx.doi.org/10.1007/s00404-011-1960-4

172. Gomez-Arriaga PI, Herraiz I, Lopez-Jimenez EA, Gomez-Montes E, Denk B, Galindo A. Uterine artery Doppler and sFlt-1/PlGF ratio: usefulness in diagnosis of pre-eclampsia. *Ultrasound Obstet Gynecol* 2013;**41**:530–7. http://dx.doi.org/10.1002/uog.12400

173. Gullai N, Stenczer B, Molvarec A, Veresh Z, Nagy B, Rigo J. OS085. Decreased maternal circulating PLGF is a significant predictor of length of pregnancy in women with hypertensive disorders of pregnancy. *Pregnancy Hypertens* 2012;**2**:224. http://dx.doi.org/10.1016/j.preghy.2012.04.086

174. Gullai N, Stenczer B, Molvarec A, Fügedi G, Veresh Z, Nagy B, Rigó J. Evaluation of a rapid and simple placental growth factor test in hypertensive disorders of pregnancy. *Hypertens Res* 2013;**36**:457–62. http://dx.doi.org/10.1038/hr.2012.206

175. Hanita O, Alia NN, Zaleha AM, Nor Azlin MI. Serum soluble FMS-like tyrosine kinase 1 and placental growth factor concentration as predictors of preeclampsia in high risk pregnant women. *Malays J Pathol* 2014;**36**:19–26.

176. Hanses T, Korpimaki T. Performance of a next generation PLGF 1–2–3 assay. *J Matern Fetal Neonatal Med* 2014;**27**:347–8.

177. Hassan MF, Rund NM, Salama AH. An Elevated Maternal Plasma Soluble fms-Like Tyrosine Kinase-1 to Placental Growth Factor Ratio at Midtrimester Is a Useful Predictor for Preeclampsia. *Obstet Gynecol Int* 2013;**2013**:202346. http://dx.doi.org/10.1155/2013/202346

178. Hund M, Verhagen-Kamerbeek W, Reim M, Messinger D, van der Does R, Stepan H. Influence of the sFlt-1/PlGF ratio on clinical decision-making in women with suspected preeclampsia – the PreOS study protocol. *Hypertens Pregnancy* 2015;**34**:102–15. http://dx.doi.org/10.3109/10641955.2014.982331

179. Hund M, Verhagen-Kamerbeek WDJ, Reim M. *PreOS (Preeclampsia Open Study): A Multicenter, Prospective, Open, Non-Interventional Study Evaluating the Influence of the Angiogenic Biomarkers sFlt-1 and PlGF on Decision-Making of Physicians in Pregnant Women with Suspicion of Preeclampsia*. 18th World Congress on Controversies in Obstetrics, Gynecology and Infertility, Vienna, Austria, 24–27 October 2013.

180. Hund M, Allegranza D, Schoedl M, Verhagen-Kamerbeek WDJ. *Prognosis Study: Prediction of Short-Term Outcome in Pregnant Women with Suspected Preeclampsia Study Using the Angiogenic Biomarkers sFlt-1/PlGF*. 18th World Congress on Controversies in Obstetrics, Gynecology and Infertility, Vienna, Austria, 24–27 October 2013.

181. Husse S, Schaarschmidt W, Jank A, Denk B, Stepan H. P28. Predictive value of the sFlt-1/PlGF ratio in pregnancies at risk for hypertensive disorders. *Pregnancy Hypertens* 2011;**1**:285–6. http://dx.doi.org/10.1016/j.preghy.2011.08.089

182. Hyde C, Thornton S. Does screening for pre-eclampsia make sense? *BJOG* 2013;**120**:1168–70. http://dx.doi.org/10.1111/1471-0528.12309

183. Kaufmann I, Rusterholz C, Hösli I, Hahn S, Lapaire O. Can detection of late-onset PE at triage by sflt-1 or PlGF be improved by the use of additional biomarkers? *Prenat Diagn* 2012;**32**:1288–94. http://dx.doi.org/10.1002/pd.3995

184. Khalil MI, Sagr ER, Bahyan R, Elrifaei RM, Alzahrani MH, Tamimi W. How accurate are placental growth factor, urate, lactate dehydrogenase and proteinuria in diagnosing preeclampsia and its severity? *Pregnancy Hypertens* 2014;**4**:156–63. http://dx.doi.org/10.1016/j.preghy.2014.02.002

185. Kim SY, Ryu HM, Yang JH, Kim MY, Han JY, Kim JO, *et al.* Increased sFlt-1 to PlGF ratio in women who subsequently develop preeclampsia. *J Korean Med Sci* 2007;**22**:873–7. http://dx.doi.org/10.3346/jkms.2007.22.5.873

186. Kjos K, Ghashghaei R, Klisser K, Woelkers D. Prediction of adverse outcomes with mid-pregnancy serum placental growth factor levels and uterine artery Dopplers in high risk pregnancies. *Pregnancy Hypertens* 2015;**5**:80. http://dx.doi.org/10.1016/j.preghy.2014.10.161

187. Klein E, Schlembach D, and Ramoni A. *PreOS (Preeclampsia open Study), a Multicentre, Prospective, Non-interventional Study Evaluating the Influence of the SFlt-1/PlGF Ratio on Physician Decision-Making in Pregnant Women with Suspicion of Preeclampsia*. The 20th World Congress on Controversies in Obstetrics, Gynecology & Infertility, Paris, France, 4–7 December 2014.

188. Klein E, Schlembach D, Ramoni A, Langer E, Bahlmann F, Grill S, *et al. Influence of the sFlt-1/PlGF Ratio on Clinical Decision Making in Pregnant Women with Suspicion of Preeclampsia: Results from the Preeclampsia Open Study (PreOS)*. Unpublished manuscript (confidential document) submitted to the National Institute for Health and Care Excellence Diagnostics Assessment Programme. Mannheim: Roche Diagnostics GmbH; 2015.

189. Knudsen UB, Kronborg CS, von Dadelszen P, Kupfer K, Lee SW, Vittinghus E, *et al.* A single rapid point-of-care placental growth factor determination as an aid in the diagnosis of preeclampsia. *Pregnancy Hypertens* 2012;**2**:8–15. http://dx.doi.org/10.1016/j.preghy.2011.08.117

190. Kusanovic JP, Romero R, Chaiworapongsa T, Erez O, Mittal P, Vaisbuch E, *et al.* A prospective cohort study of the value of maternal plasma concentrations of angiogenic and anti-angiogenic factors in early pregnancy and midtrimester in the identification of patients destined to develop preeclampsia. *J Matern Fetal Neonatal Med* 2009;**22**:1021–38. http://dx.doi.org/10.3109/14767050902994754

191. Lai J, Garcia-Tizon Larroca S, Peeva G, Poon LC, Wright D, Nicolaides KH. Competing risks model in screening for preeclampsia by serum placental growth factor and soluble fms-like tyrosine kinase-1 at 30–3 weeks' gestation. [Erratum published in *Fetal Diagn Ther* 2014;**36**:48]. *Fetal Diagn Ther* 2014;**35**:240–8. http://dx.doi.org/10.1159/000359968

192. Leaños-Miranda A, Campos-Galicia I, Isordia-Salas I, Rivera-Leaños R, Romero-Arauz JF, Ayala-Méndez JA, Ulloa-Aguirre A. Changes in circulating concentrations of soluble fms-like tyrosine kinase-1 and placental growth factor measured by automated electrochemiluminescence immunoassays methods are predictors of preeclampsia. *J Hypertens* 2012;**30**:2173–81. http://dx.doi.org/10.1097/HJH.0b013e328357c0c9

193. Lehnen H, Schaefer S, Puchooa A, Reineke T, Menke-Mollers I. Critical assessment of s-Flt1 (soluble Fms-like tyrosine kinase 1)/PlGF (placentals growth factor) ratio for preeclampsia at delivery. *Pregnancy Hypertens* 2010;**1**:S71. http://dx.doi.org/10.1016/S2210-7789(10)60276-0

194. Lehnen H, Mosblech N, Reineke T, Puchooa A, Menke-Möllers I, Zechner U, Gembruch U. Prenatal clinical assessment of sflt-1 (soluble fms-like tyrosine kinase-1)/plgf (placental growth factor) ratio as a diagnostic tool for preeclampsia, pregnancy-induced hypertension, and proteinuria. *Geburtshilfe Frauenheilkd* 2013;**73**:440–5. http://dx.doi.org/10.1055/s-0032-1328601

195. Lim JH, Kim SY, Park SY, Yang JH, Kim MY, Ryu HM. Effective prediction of preeclampsia by a combined ratio of angiogenesis-related factors. *Obstet Gynecol* 2008;**111**:1403–9. http://dx.doi.org/10.1097/AOG.0b013e3181719b7a

196. Madazli R, Kuseyrioglu B, Uzun H, Uludag S, Ocak V. Prediction of preeclampsia with maternal mid-trimester placental growth factor, activin A, fibronectin and uterine artery Doppler velocimetry. *Int J Gynaecol Obstet* 2005;**89**:251–7. http://dx.doi.org/10.1016/j.ijgo.2005.02.008

197. Martínez-Ruiz A, Sarabia-Meseguer MD, Pérez-Fornieles J, Vílchez JA, Tovar-Zapata I, Noguera-Velasco JA. Placental growth factor, soluble fms-like tyrosine kinase 1 and progesterone as diagnostic biomarkers for ectopic pregnancy and missed abortion. *Clin Biochem* 2014;**47**:844–7. http://dx.doi.org/10.1016/j.clinbiochem.2014.03.013

198. Mathur P, Maru L, Dave A. PLGF assay as a novel biomarker for prediction of early onset preeclampsia. *BJOG* 2014;**12**:190.

199. Mazzocco MI, Calabrese S, Cardellicchio M, Martinelli A, Cetin I. Could PLGF be a useful marker to predict fetal outcomes in pregnancies at risk? *Reprod Sci* 2014;**21**(Suppl. 1):304A–5A.

200. McElrath TF, Lim KH, Pare E, Rich-Edwards J, Pucci D, Troisi R, Parry S. Longitudinal evaluation of predictive value for preeclampsia of circulating angiogenic factors through pregnancy. *Am J Obstet Gynecol* 2012;**207**:407.e1–7. http://dx.doi.org/10.1016/j.ajog.2012.08.010

201. Meiri H, Huppertz B, Cetin I. Development of early non-invasive markers and means for the diagnosis and progression monitoring of preeclampsia and tailoring putative therapies (project pregenesys 037244). *Placenta* 2011;**32**:S1–3. http://dx.doi.org/10.1016/j.placenta.2010.12.003

202. Mijal RS, Holzman CB, Rana S, Karumanchi SA, Wang J, Sikorskii A. Mid-pregnancy levels of angiogenic markers as indicators of pathways to preterm delivery. *J Matern Fetal Neonatal Med* 2012;**25**:1135–41. http://dx.doi.org/10.3109/14767058.2011.625458

203. Molvarec A, Szarka A, Walentin S, Szucs E, Nagy B, Rigó J. Circulating angiogenic factors determined by electrochemiluminescence immunoassay in relation to the clinical features and laboratory parameters in women with pre-eclampsia. *Hypertens Res* 2010;**33**:892–8. http://dx.doi.org/10.1038/hr.2010.92

204. Moore A, Young H, Keller J, Ojo L, Yan J, Simas TM, *et al.* Angiogenic biomarkers for the prediction of pregnancy complications in women with suspected preeclampsia. *Am J Obstet Gynecol* 2012;**206**(Suppl. 1):S326–27. http://dx.doi.org/10.1016/j.ajog.2011.10.754

205. Moore AG, Young H, Keller JM, Ojo LR, Yan J, Simas TA, Maynard SE. Angiogenic biomarkers for prediction of maternal and neonatal complications in suspected preeclampsia. *J Matern Fetal Neonatal Med* 2012;**25**:2651–7. http://dx.doi.org/10.3109/14767058.2012.713055

206. Moore Simas TA, Crawford SL, Bathgate S, Yan J, Robidoux L, Moore M, Maynard SE. Angiogenic biomarkers for prediction of early preeclampsia onset in high-risk women. *J Matern Fetal Neonatal Med* 2014;**27**:1038–48. http://dx.doi.org/10.3109/14767058.2013.847415

207. Moore Simas TA, Crawford SL, Solitro MJ, Frost SC, Meyer BA, Maynard SE. Angiogenic factors for the prediction of preeclampsia in high-risk women. *Am J Obstet Gynecol* 2007;**197**:244.e1–8. http://dx.doi.org/10.1016/j.ajog.2007.06.030

208. Myers JE, Kenny LC, McCowan LME, Chan EHY, Dekker GA, Poston L, *et al.* Comparison of two placental growth factor immunoassays as an aid in the diagnosis of preterm preeclampsia. *Reprod Sci* 2012;**19**(Suppl. 1):217A.

209. Myers JE, Kenny LC, McCowan LM, Chan EH, Dekker GA, Poston L, *et al.* O8. Comparison of two placental growth factor immunoassays as an aid in the diagnosis of preterm preeclampsia. *Pregnancy Hypertens* 2011;**1**:260–1. http://dx.doi.org/10.1016/j.preghy.2011.08.040

210. National Horizon Scanning Centre. *DELFIA® Xpress PlGF kit antenatal screening test to predict risk of pre-eclampsia*. Birmingham: National Horizon Scanning Centre; 2012.

211. North RA, Kenny LC, McCowan LME, Chan EHY, Dekker GA, Poston L, *et al.* Prediction of the preterm preeclampsia in a prospective cohort of nulliparous women. *Reprod Sci* 2012;**19**(Suppl. 1):218A.

212. Ohkuchi A, Hirashima C, Matsubara S, Suzuki H, Takahashi K, Arai F, *et al.* Alterations in placental growth factor levels before and after the onset of preeclampsia are more pronounced in women with early onset severe preeclampsia. *Hypertens Res* 2007;**30**:151–9. http://dx.doi.org/10.1291/hypres.30.151

213. Ohkuchi A, Hirashima C, Suzuki H, Takahashi K, Yoshida M, Matsubara S, Suzuki M. Evaluation of a new and automated electrochemiluminescence immunoassay for plasma sFlt-1 and PlGF levels in women with preeclampsia. *Hypertens Res* 2010;**33**:422–7. http://dx.doi.org/10.1038/hr.2010.15

214. Ohkuchi A, Hirashima C, Takahashi K, Suzuki H, Matsubara S, Suzuki M. Onset threshold of the plasma levels of soluble fms-like tyrosine kinase 1/placental growth factor ratio for predicting the imminent onset of preeclampsia within 4 weeks after blood sampling at 19–31 weeks of gestation. *Hypertens Res* 2013;**36**:1073–80. http://dx.doi.org/10.1038/hr.2013.95

215. Ohkuchi A, Hirashima C, Matsubara S, Suzuki H, Takahashi K, Usui R, Suzuki M. Serum sFlt1:PlGF ratio, PlGF, and soluble endoglin levels in gestational proteinuria. *Hypertens Pregnancy* 2009;**28**:95–108. http://dx.doi.org/10.1080/10641950802419895

216. Ohkuchi A, Hirashima C, Matsubara S, Takahashi K, Matsuda Y, Suzuki M. Threshold of soluble fms-like tyrosine kinase 1/placental growth factor ratio for the imminent onset of preeclampsia. *Hypertension* 2011;**58**:859–66. http://dx.doi.org/10.1161/HYPERTENSIONAHA.111.174417

217. Park HJ, Kim SH, Jung YW, Shim SS, Kim JY, Cho YK, *et al.* Screening models using multiple markers for early detection of late-onset preeclampsia in low-risk pregnancy. *BMC Pregnancy Childbirth* 2014;**14**:35. http://dx.doi.org/10.1186/1471-2393-14-35

218. PerkinElmer, Inc. *DELFIA Xpress PlGF 1–2-3*. Product insert. Waltham, MA: PerkinElmer, Inc.; 2015.

219. Powers RW, Roberts JM, Plymire DA, Pucci D, Datwyler SA, Laird DM, *et al.* Low placental growth factor across pregnancy identifies a subset of women with preterm preeclampsia: type 1 versus type 2 preeclampsia? *Hypertension* 2012;**60**:239–46. http://dx.doi.org/10.1161/HYPERTENSIONAHA.112.191213

220. Powers RW, Jeyabalan A, Clifton RG, Van DP, Hauth JC, Klebanoff MA, *et al.* Soluble fms-Like tyrosine kinase 1 (sFlt1), endoglin and placental growth factor (PlGF) in preeclampsia among high risk pregnancies. *PLOS ONE* 2010;**5**:e13263. http://dx.doi.org/10.1371/journal.pone.0013263

221. Prefumo F. Re: Uterine artery Doppler and sFlt-1/PlGF ratio: prognostic value in early-onset pre-eclampsia. P. I. Gómez-Arriaga, I. Herraiz, E. A. López-Jiménez, D. Escribano, B. Denk and A. Galindo. Ultrasound Obstet Gynecol 2014; **43**: 525–32. *Ultrasound Obstet Gynecol* 2014;**43**:488–9. http://dx.doi.org/10.1002/uog.13362

222. Rana S, Hacker M, Merport A, Salahuddin S, Verlohren S, Perschel F, Karumanchi SA. PP059. Angiogenic factors and risk of preeclampsia related adverse outcomes in twin pregnancies. *Pregnancy Hypertens* 2012;**2**:273–4. http://dx.doi.org/10.1016/j.preghy.2012.04.170

223. Rana S, Hacker MR, Modest AM, Salahuddin S, Lim KH, Verlohren S, *et al.* Circulating angiogenic factors and risk of adverse maternal and perinatal outcomes in twin pregnancies with suspected preeclampsia. *Hypertension* 2012;**60**:451–8. http://dx.doi.org/10.1161/HYPERTENSIONAHA.112.195065

224. Rana S, Schnettler WT, Powe C, Wenger J, Salahuddin S, Cerdeira AS, *et al.* Clinical characterization and outcomes of preeclampsia with normal angiogenic profile. *Hypertens Pregnancy* 2013;**32**:189–201. http://dx.doi.org/10.3109/10641955.2013.784788

225. Rana S, Cerdeira AS, Wenger J, Salahuddin S, Lim KH, Ralston SJ, *et al.* Plasma concentrations of soluble endoglin versus standard evaluation in patients with suspected preeclampsia. *PLOS ONE* 2012;**7**:e48259. http://dx.doi.org/10.1371/journal.pone.0048259

226. Rana S, Salahuddin S, Hegde S, Lim K-H, Karumanchi A, Powe C. Prospective evaluation of angiogenic proteins as predictors of adverse maternal and fetal outcomes. *Am J Obstet Gynecol* 2011;**204**:S310. http://dx.doi.org/10.1016/j.ajog.2010.10.814

227. Rasanen J, Quinn MJ, Laurie A, Bean E, Roberts CT, Nagalla SR, Gravett MG. Maternal serum glycosylated fibronectin as a point-of-care biomarker for assessment of preeclampsia. *Am J Obstet Gynecol* 2015;**212**:82.e1–9. http://dx.doi.org/10.1016/j.ajog.2014.07.052

228. Redman C, Sargent I, Meacher H, Simms C, Breth KU, Kronborg C, *et al.* The triage PLGF test: Evaluation of circulating levels of placental growth factor in plasma as an aid to the diagnosis of preterm pre-eclampsia. *J Matern Fetal Neonatal Med* 2010;**23**:630–1.

229. Redman C, Lodge T, Meacher H, Marks J, Simms C, Sargent I. Triage (R) PLGF test: point-of-care assay of plasma placental growth factor to diagnose preeclampsia. *Adv Perinat Med* 2010;181–5.

230. Redman C, Lodge T, Meacher H, Marks J, Simms C, Sargent I. Triage® PlGF test: point-of-care assay of plasma placental growth factor to diagnose pre-eclampsia. *Pregnancy Hypertens* 2010;**1**:S23. http://dx.doi.org/10.1016/S2210-7789(10)60099-2

231. Ris-Stalpers C, Lahsinoui HH, Boussata S, van der Post JA. OS065. Placental growth factor as a diagnostic and prognostic test for placental complications of pregnancy. *Pregnancy Hypertens* 2012;**2**:212. http://dx.doi.org/10.1016/j.preghy.2012.04.066

232. Ris C, van der Post J. Placental growth factor as a diagnostic and prognostic test for placental complications of pregnancy. *Int J Gynaecol Obstet* 2012;**119**:S463. http://dx.doi.org/10.1016/S0020-7292(12)61004-2

233. Rizos DA, Eleftheriades M, Haliassos A, Akalestos A, Rizou M, Vlazakis V, *et al.* Measurement of soluble fms-like tyrosine kinase 1 (sFlt-1) and placental growth factor (PlGF) with Roche Elecsys, throughout normal pregnancy and in pregnancies complicated by hypertension and/or IUGR. *Clin Chem* 2009;**55**:A241.

234. Rohra DK, Zeb A, Qureishi RN, Azam SI, Khan NB, Zuberi HS, *et al.* Prediction of pre-eclampsia during early pregnancy in primiparas with soluble fms-like tyrosine kinase-1 and placental growth factor. *Natl Med J India* 2012;**25**:68–73.

235. Romero R, Nien JK, Espinoza J, Todem D, Fu W, Chung H, *et al.* A longitudinal study of angiogenic (placental growth factor) and anti-angiogenic (soluble endoglin and soluble vascular endothelial growth factor receptor-1) factors in normal pregnancy and patients destined to develop preeclampsia and deliver a small for gestational age neonate. *J Matern Fetal Neonatal Med* 2008;**21**:9–23. http://dx.doi.org/10.1080/14767050701830480

236. Rutherford D, Lee G, Adno A, Levidiotis V, Hennessy A, Makris A. The role of neutrophil gelatinase-associated lipocalin, soluble fms-like tyrosine kinase-1, placental growth factor and endoglin in predicting preeclampsia. *Pregnancy Hypertens* 2010;**1**(Suppl. 1):23. http://dx.doi.org/10.1016/S2210-7789(10)60101-8

237. Rutherford D, Lee G, Spicer T, Cleland B, Adno V, Suranyi M, *et al.* The role of neutrophil gelatinase associated lipocalin, soluble FMS-like tyrosine kinase-1, placental growth factor and endoglin in predicting pre-eclampsia. *Nephrology* 2010;**15**(Suppl. 4.):46.

238. Saleh L, Verdonk K, Danser AH, Steegers EA, Russcher H, van den Meiracker AH, Visser W. 8C.01: Sflt-1 and PlGF measurements and their ratio for the diagnosis and prognosis of preeclampsia in a high-risk cohort. *J Hypertens* 2015;**33**(Suppl. 1):e110. http://dx.doi.org/10.1097/01.hjh.0000467646.70117.6e

239. Schaarschmidt W, Rana S, Stepan H. PP052. The course of sFlt-1 and PLGF reflects different progression pattern in early- versus late-onset preeclampsia and HELLP syndrome. *Pregnancy Hypertens* 2012;**2**:269. http://dx.doi.org/10.1016/j.preghy.2012.04.163

240. Schaarschmidt W, Herraiz I, Lapaire O, Schlembach D, Zeisler H, Calda P, *et al.* The routine measurement of the sFlt1/PlGF ratio allows differential diagnosis of hypertensive pregnancy disorders and a prediction of pregnancy outcome. *Arch Gynecol Obstet* 2010;**282**(Suppl. 1):82–3.

241. Schiettecatte J, Russcher H, Anckaert E, Mees M, Leeser B, Tirelli AS, *et al.* Multicenter evaluation of the first automated diagnostic preeclampsia tests: Elecsys Sflt-1 and Pigf. *Acta Clin Belg* 2011;**66**:162.

242. Schiettecatte J, Russcher H, Anckaert E, Mees M, Leeser B, Tirelli AS, *et al.* Multicenter evaluation of the first automated Elecsys sFlt-1 and PlGF assays in normal pregnancies and preeclampsia. *Clin Biochem* 2010;**43**:768–70. http://dx.doi.org/10.1016/j.clinbiochem.2010.02.010

243. Schoofs K, Engels T, Pape J, Henrich W, Stepan H, Verlohren S. PP051. Prediction of preeclampsia with the sFlt-1/PLGF ratio: impact of the slope of repeated measurements. *Pregnancy Hypertens* 2012;**2**:269. http://dx.doi.org/10.1016/j.preghy.2012.04.162

244. Schoofs K, Grittner U, Engels T, Pape J, Denk B, Henrich W, Verlohren S. The importance of repeated measurements of the sFlt-1/PlGF ratio for the prediction of preeclampsia and intrauterine growth restriction. *J Perinat Med* 2014;**42**:61–8. http://dx.doi.org/10.1515/jpm-2013-0074

245. Shaker OG, Shehata H. Early prediction of preeclampsia in high-risk women. *J Womens Health* 2011;**20**:539–44. http://dx.doi.org/10.1089/jwh.2010.2378

246. Shawkat E, Johnstone E, Nice D, Sayce A, Hayden K, Myers J. Clinical evaluation of placental growth factor (PlGF) for the management of suspected placental pathology in high-risk pregnancies. *Pregnancy Hypertens* 2015;**5**:57. http://dx.doi.org/10.1016/j.preghy.2014.10.111

247. Sibiude J, Dionne MD, Guibourdenche J, Le Ray C, Anselem O, Serreau R, *et al.* PP116. Circulating PlGF can help predict preeclampsia and adverse outcome in patients with suspected preeclampsia or fetal growth restriction. *Pregnancy Hypertens* 2012;**2**:302–3. http://dx.doi.org/10.1016/j.preghy.2012.04.227

248. Sibiude J, Guibourdenche J, Dionne MD, Le Ray C, Anselem O, Serreau R, *et al.* Placental growth factor for the prediction of adverse outcomes in patients with suspected preeclampsia or intrauterine growth restriction. *PLOS ONE* 2012;**7**:e50208. http://dx.doi.org/10.1371/journal.pone.0050208

249. Southcombe JH, Benton SJ, Hu Y, von Dadelszen P, Child T, Snider JV, *et al.* Measurement of sST2 is comparable to PlGF in the diagnosis of early-onset pre-eclampsia. *Pregnancy Hypertens* 2013;**3**:115–17. http://dx.doi.org/10.1016/j.preghy.2013.02.002

250. Staff AC, Harsem NK, Braekke K, Hyer M, Hoover RN, Troisi R. Maternal, gestational and neonatal characteristics and maternal angiogenic factors in normotensive pregnancies. *Eur J Obstet Gynecol Reprod Biol* 2009;**143**:29–33. http://dx.doi.org/10.1016/j.ejogrb.2008.11.003

251. Stenczer B, Molvarec A, Veresh Z, Gullai N, Nagy GR, Walentin S, *et al.* Circulating levels of the anti-angiogenic thrombospondin 2 are elevated in pre-eclampsia. *Acta Obstet Gynecol Scand* 2011;**90**:1291–5. http://dx.doi.org/10.1111/j.1600-0412.2011.01220.x

252. Stenczer B, Molvarec A, Gullai N, Fügedi G, Nagy B, Rigó J. P22. Evaluation of a new, simple and rapid placental growth factor test for the evaluation of hypertensive disorders in pregnancy. *Pregnancy Hypertens* 2011;**1**:283. http://dx.doi.org/10.1016/j.preghy.2011.08.083

253. Stenczer B, Gullai N, Szabo G, Fugedi G, Hanuszka B, Molvarec A, *et al.* Evaluation of a new, simple and rapid placental growth factor test for the evaluation of hypertensive disorders in pregnancy. Abstracts for the International Federation of Placenta Associations (IFPA) meeting 2011. Geilo, Norway, 14–17 September 2011. *Placenta* 2011;**32**:A85.

254. Stenczer B, Molvarec A, Gullai N, Veresh Z, Nagy B, Rigo J. PP056. Placental growth factor is a better predictor of preterm birth than uterine or umbilical artery doppler in hypertensive disorders of pregnancy. *Pregnancy Hypertens* 2012;**2**:272. http://dx.doi.org/10.1016/j.preghy.2012.04.167

255. Stepan H, Schaarschmidt W, Jank A, Kratzsch J. Angiogenic factors as routine aid in diagnosis in cases with suspected preeclampsia: first clinical experience. *Arch Gynecol Obstet* 2010;**282**(Suppl. 1):78.

256. Stepan H, Unversucht A, Wessel N, Faber R. Predictive value of maternal angiogenic factors in second trimester pregnancies with abnormal uterine perfusion. *Hypertension* 2007;**49**:818–24. http://dx.doi.org/10.1161/01.HYP.0000258404.21552.a3

257. Strunz-McKendry T, Allegranza D, Hund M, Posnett J. *UK Cost-Effectiveness Analysis of Incorporating the sFlt-1/PlGF Ratio Test into Standard Clinical Practice for Diagnosis of Preeclampsia and/or HELLP Syndrome*. The 20th World Congress on Controversies in Obstetrics, Gynecology and Infertility, Paris, France, 4–7 December 2014.

258. Stubert J, Ullmann S, Bolz M, Bohlmann M, Külz T, Gerber B, Reimer T. PP117. Measurement of sFlt-1 and PlGF for prediction of pregnancy associated diseases. *Pregnancy Hypertens* 2012;**2**:303. http://dx.doi.org/10.1016/j.preghy.2012.04.228

259. Stubert J, Ullmann S, Bolz M, Külz T, Dieterich M, Richter DU, Reimer T. Prediction of preeclampsia and induced delivery at < 34 weeks gestation by sFLT-1 and PlGF in patients with abnormal midtrimester uterine Doppler velocimetry: a prospective cohort analysis. *BMC Pregnancy Childbirth* 2014;**14**:292. http://dx.doi.org/10.1186/1471-2393-14-292

260. Su YN, Lee CN, Cheng WF, Shau WY, Chow SN, Hsieh FJ. Decreased maternal serum placenta growth factor in early second trimester and preeclampsia. *Obstet Gynecol* 2001;**97**:898–904. http://dx.doi.org/10.1097/00006250-200106000-00007

261. Sunderji S, Gaziano E, Wothe D, Rogers LC, Sibai B, Karumanchi SA, Hodges-Savola C. Automated assays for sVEGF R1 and PlGF as an aid in the diagnosis of preterm preeclampsia: a prospective clinical study. *Am J Obstet Gynecol* 2010;**202**:40.e1–7. http://dx.doi.org/10.1016/j.ajog.2009.07.025

262. Sunderji S, Sibai B, Wothe D, Hodges-Savola C, Rogers L, Golden W, *et al.* Differential diagnosis of preterm preeclampsia and hypertension using prototype automated assays for Svegf R1 and Plgf: a prospective clinical study. *Am J Obstet Gynecol* 2008;**199**:S81. http://dx.doi.org/10.1016/j.ajog.2008.09.277

263. Sundrani D, Khot V, Pisal H, Mehendale S, Wagh G, Joshi A, Joshi S. Gestation dependant changes in angiogenic factors and their associations with fetal growth measures in normotensive pregnancy. *PLOS ONE* 2013;**8**:e54153. http://dx.doi.org/10.1371/journal.pone.0054153

264. Teixeira PG, Reis ZS, Andrade SP, Rezende CA, Lage EM, Velloso EP, *et al.* Presymptomatic prediction of preeclampsia with angiogenic factors, in high risk pregnant women. *Hypertens Pregnancy* 2013;**32**:312–20. http://dx.doi.org/10.3109/10641955.2013.807818

265. Thermo Fisher Scientific. *Company Submission Literature Review*. Company submission to the National Institute for Health and Care Excellence Diagnostics Assessment Programme. Hennigsdorf: Thermo Fisher Scientific GmbH; 2015.

266. Thermo Fisher Scientific. *Prenatal Screening: Medical Trials. Global Medical Affairs*. Confidential company submission to the National Institute for Health and Care Excellence Diagnostics Assessment Programme. Hennigsdorf: Thermo Fisher Scientific GmbH; 2015.

267. Thermo Fisher Scientific. *BRAHMS Kryptor Compact PLUS User Manual Version 4*. Hennigsdorf: Thermo Fisher Scientific GmbH; 2015.

268. Thornton J. Angiogenic factors combined with clinical risk factors to predict preterm pre-eclampsia in nulliparous women: a predictive test accuracy study. *BJOG* 2014;**121**:507. http://dx.doi.org/10.1111/1471-0528.12564

269. Tidwell SC, Ho HN, Chiu WH, Torry RJ, Torry DS. Low maternal serum levels of placenta growth factor as an antecedent of clinical preeclampsia. *Am J Obstet Gynecol* 2001;**184**:1267–72. http://dx.doi.org/10.1067/mob.2001.113129

270. van Helden J, Weiskirchen R. Analytical evaluation of the novel soluble fms-like tyrosine kinase 1 and placental growth factor assays for the diagnosis of preeclampsia. *Clin Biochem* 2015;**48**:1113–19. http://dx.doi.org/10.1016/j.clinbiochem.2015.06.020

271. Vatten LJ, Eskild A, Nilsen TI, Jeansson S, Jenum PA, Staff AC. Changes in circulating level of angiogenic factors from the first to second trimester as predictors of preeclampsia. *Am J Obstet Gynecol* 2007;**196**:239.e1–6. http://dx.doi.org/10.1016/j.ajog.2006.10.909

272. Verdonk K, Visser W, Russcher H, Danser AH, Steegers EA, van den Meiracker AH. Differential diagnosis of preeclampsia: remember the soluble fms-like tyrosine kinase 1/placental growth factor ratio. *Hypertension* 2012;**60**:884–90. http://dx.doi.org/10.1161/HYPERTENSIONAHA.112.201459

273. Verlohren S, Herraiz I, Lapaire O, Schlembach D, Moertl M, Zeisler H, *et al.* L13. The routine measurement of the sFlt1/PlGF ratio allows differential diagnosis of hypertensive pregnancy disorders and has prognostic potential in preeclamptic patients. *Pregnancy Hypertens* 2011;**1**:245–6. http://dx.doi.org/10.1016/j.preghy.2011.08.014

274. Verlohren S, Herraiz I, Lapaire O, Schlembach D, Zeisler H, Calda P, *et al.* Characterization of the sFlt1/PlGF ratio in different types of hypertensive pregnancy disorders in the second and third trimester of pregnancy and correlation with pregnancy duration. *Pregnancy Hypertens* 2010;**1**(Suppl. 1):S23. http://dx.doi.org/10.1016/S2210-7789(10)60100-6

275. Verlohren S, Herraiz I, Lapaire O, Schlembach D, Zeisler H, Calda P, *et al.* Gestational phase specific cut-off values for the use of the SFlt-1/PlGF-ratio as a diagnostic test for preeclampsia. *Reproductive Sciences* 2014;**21**:403A.

276. Villa PM, Hämäläinen E, Mäki A, Räikkönen K, Pesonen AK, Taipale P, *et al.* Vasoactive agents for the prediction of early- and late-onset preeclampsia in a high-risk cohort. *BMC Pregnancy Childbirth* 2013;**13**:110. http://dx.doi.org/10.1186/1471-2393-13-110

277. Wald NJ, Bestwick JP, George LM, Wu T, Morris JK. Screening for pre-eclampsia using serum placental growth factor and endoglin with Down's syndrome Quadruple test markers. *J Med Screen* 2012;**19**:60–7. http://dx.doi.org/10.1258/jms.2012.011114

278. Wang A, Holston AM, Yu KF, Zhang J, Toporsian M, Karumanchi SA, Levine RJ. Circulating anti-angiogenic factors during hypertensive pregnancy and increased risk of respiratory distress syndrome in preterm neonates. *J Matern Fetal Neonatal Med* 2012;**25**:1447–52. http://dx.doi.org/10.3109/14767058.2011.640368

279. Widmer M, Cuesta CB, Khan K, Gulmezoglu AM, Karumanchi SA, Lindheimer MD. [93-OR]: Accuracy of angiogenic biomarkers for predicting preeclampsia: An observational study. *Pregnancy Hypertens* 2015;**5**:50. http://dx.doi.org/10.1016/j.preghy.2014.10.097

280. Woods A, Dekker G. [92-OR]: Placental growth factor measurement as a marker of subsequent disease and harm in placental insufficiency conditions. *Pregnancy Hypertens* 2015;**5**:50. http://dx.doi.org/10.1016/j.preghy.2014.10.096

281. Yu Y, Jenkins AJ, Nankervis AJ, Hanssen KF, Scholz H, Henriksen T, *et al.* Anti-angiogenic factors and pre-eclampsia in type 1 diabetic women. *Diabetologia* 2009;**52**:160–8. http://dx.doi.org/10.1007/s00125-008-1182-x

282. Sharp A, Agarwal U, Alfirevic Z. *PlGF Bedside Testing in Standard Clinical Care: 6 Months Experience from a Large Teaching Hospital*. The 3rd International Conference on Fetal Growth, Oxford, UK, 1–3 October 2014.

283. Department of Health. *NHS Payment by Results: Guidance and Tariff for 2008–9*. London: Department of Health; 2009.

Appendix 1 Search strategy for review of test accuracy studies

Database, host, years searched, date searched	Literature search strategy (number of references)	Results
MEDLINE(R) (via Ovid) without revisions 1996 to March week 1 2015; search limited year 2000 onwards; searched 10 March 2015	1. Pre-Eclampsia/ (11,761) 2. (preeclamp* or pre eclamp*).tw. (14,776) 3. (tox?emi* adj5 pregnan*).tw. (188) 4. gestosis.tw. (266) 5. (pregnan* adj3 hypertensi*).tw. (4530) 6. (gestation* adj3 hypertensi*).tw. (1591) 7. ((maternal or maternity) adj3 hypertens*).tw. (641) 8. Hypertension, Pregnancy-Induced/ (1679) 9. or/1-8 (19,665) 10. (PlGF and (triage or test* or assay* or immunoassay* or diagnos* or detect* or surveillance or screen* or measur* or analys* or analyz* or determin* or sensitivity or specificity or accuracy or accurate or assessment* or predict* or positive or negative or electrochemiluminescen*)).tw. (776) 11. ("Placenta* growth factor" and (triage or test* or assay* or immunoassay* or diagnos* or detect* or surveillance or screen* or measur* or analys* or analyz* or determin* or sensitivity or specificity or accuracy or accurate or assessment* or predict* or positive or negative or electrochemiluminescen*)).tw. (1011) 12. Vascular Endothelial Growth Factor Receptor-1/bl [Blood] (542) 13. ("VEGFR1" or "VEGFR 1").tw. (1738) 14. Early Diagnosis/ or Diagnosis/ (16,177) 15. Diagnostic Tests, Routine/ or Diagnostic Equipment/ or "Diagnostic Techniques, Obstetrical and Gynecological"/ or Diagnostic Services/ (6892) 16. Maternal Serum Screening Tests/ (94) 17. Serologic Tests/ (7089) 18. Pregnancy Proteins/an, du [Analysis, Diagnostic Use] (185) 19. Membrane Proteins/bl, du [Blood, Diagnostic Use] (1119) 20. Biological Markers/bl, du [Blood, Diagnostic Use] (67,716) 21. "fms-like tyrosine kinase*".tw. (1365) 22. (("FLT 1" or "sFLT 1" or "FLT1" or "sFLT1") and (triage or test* or assay* or immunoassay* or diagnos* or detect* or screen* or measur* or analys* or analyz* or determin* or sensitivity or specificity or accuracy or accurate or "prognostic assessment*" or predict* or positive or negative or electrochemiluminescen*)).tw. (1779) 23. ("soluble fms-like tyrosine kinase" and (triage or test* or assay* or immunoassay* or diagnos* or detect* or measur* or analys* or analyz* or determin* or sensitivity or specificity or accuracy or accurate or assessment* or predict* or positive or negative or electrochemiluminescen*)).tw. (392) 24. elecsys.af. (520) 25. roche.af. (12,349) 26. alere.af. (86) 27. delfia.af. (243) 28. brahms.af. (202) 29. kryptor.af. (73) 30. thermo.af. (4450) 31. or/10-30 (119,298) 32. 9 and 31 (1574) 33. limit 32 to (english language and yr="2000 -Current") (1427) 34. limit 33 to humans (1366)	1366

Database, host, years searched, date searched	Literature search strategy (number of references)	Results
MEDLINE In-Process & Other Non-Indexed Citations (via Ovid); 2000 to March week 1 2015; searched 10 March 2015	As per MEDLINE	86
EMBASE (via Ovid) 1996 to 2015 week 10; limited 2000–current; searched 10 March 2015	1. pre-eclampsia/ or "eclampsia and pre-eclampsia"/ (28,376) 2. (preeclamp* or "pre eclamp*").tw. (24,989) 3. (tox?emi* adj5 pregnan*).tw. (277) 4. gestosis.tw. (365) 5. (pregnan* adj3 hypertensi*).tw. (7597) 6. (gestation* adj3 hypertensi*).tw. (2903) 7. ((maternal or maternity) adj3 hypertens*).tw. (1076) 8. maternal hypertension/ (9492) 9. pregnancy toxemia/ (403) 10. or/1-9 (38,140) 11. (PlGF and (triage or alere)).af. (61) 12. (Triage and MeterPro).af. (4) 13. (Elecsys and ("sFlt-1" or "sFlt1" or VEGFR1 or "VEGFR-1" or PlGF or "sFlt-1/PlGF" or "sFlt1/PlGF" or "soluble FMS-like tyrosine kinase-1")).af. (36) 14. (PlGF and (Delfia or Perkin Elmer)).af. (22) 15. ((BRAHMS or Kryptor or Thermo) and (PlGF or "sFlt-1" or "sFlt1" or "sFlt-1/PlGF" or VEGFR1 or "VEGFR-1" or "soluble FMS-like tyrosine kinase-1")).af. (8) 16. or/11-15 (124) 17. 10 and 16 (98) 18. pre-eclampsia/di, pc [Diagnosis, Prevention] (3255) 19. "eclampsia and pre-eclampsia"/di, pc [Diagnosis, Prevention] (21) 20. 18 or 19 (3269) 21. (test* or triage or assay* or immunoassay* or electrochemiluminescen* or detect* or surveillance or screen* or measur* or analys* or analyz* or determin* or sensitivity or specificity or accuracy or accurate or predict*).tw. (8,583,423) 22. ("sFlt-1" or "sFlt1" or VEGFR1 or "VEGFR-1" or PlGF or "sFlt-1/PlGF" or "soluble FMS-like tyrosine kinase-1").tw. (5723) 23. placental growth factor/ (2734) 24. protein tyrosine kinase/ (40,969) 25. vasculotropin receptor 1/ (6123) 26. or/22-25 (50071) 27. 20 and 21 and 26 (201) 28. (pre?eclamp* and diagnos* and test*).ti. (15) 29. (pre?eclamp* and diagnos* and assay*).ti. (1) 30. (pre?eclamp* and diagnos* and immunoassay*).ti. (3) 31. (pre?eclamp* and diagnos* and electrochemiluminescen*).ti, ab. (8) 32. or/28-31 (26) 33. 17 or 27 or 32 (304) 34. limit 33 to (english language and yr="2000 -Current") (280)	280
SCI-EXPANDED – 1970–present; 1970 to March 2015; CPCI-S – 1990–present; 1990 to March 2015	#1 (TS=(preeclamp* or "pre eclamp*" or "pre-eclamp*")) (19,516) #2 (TS=(tox?emia NEAR pregnan*)) (48) #3 (TS=(gestosis)) (108) #4 (TS=(pregnan* NEAR hypertensi*)) (7700) #5 (TS=(gestation NEAR hypertensi*)) (774) #6 (TS=("maternal hypertensi*")) (341) #7 (TS=(maternity NEAR hypertensi*)) (27) #8 #7 OR #6 OR #5 OR #4 OR #3 OR #2 OR #1 (22,992)	284

Database, host, years searched, date searched	Literature search strategy (number of references)	Results
	#9 (TS=(PlGF and (triage or alere))) (16)	
	#10 (TS=((“placental growth factor”) and (triage or alere))) (17)	
	#11 (TS=(Elecsys and (“sFlt-1” or “sFlt1” or VEGFR1 or “VEGFR-1” or PlGF or “sFlt-1/PlGF” or"SFlt1/PLGF” or “soluble FMS-like tyrosine kinase-1”))) (16)	
	#12 (TS=(PlGF and (Delfia or Perkin Elmer))) (4)	
	#13 (TS=((BRAHMS or Kryptor or Thermo) and (PlGF or “sFlt-1” or “sFlt1” or “sFlt-1/PlGF” or “SFlt1/PlGF” or VEGFR1 or “VEGFR-1” or “soluble FMS-like tyrosine kinase-1”))) (1)	
	#14 #13 OR #12 OR #11 OR #10 OR #9 (38)	
	#15 #14 AND #8 (34)	
	#16 (TS=(Diagnos* NEAR test*)) (87,478)	
	#17 (TS=(diagnos* NEAR (test* or assay* or immunoassay* or ectrochemiluminescen*))) (105,489)	
	#18 #17 AND #8 (260)	
	#19 #18 OR #15 AND LANGUAGE: (English) (284)	
The Cochrane Library; inception to 11 March 2015; searched 11 March 2015	#1 MeSH descriptor: [Pre-Eclampsia] this term only (614)	*N* = 85: CDSR, *n* = 5; DARE, *n* = 15; CENTRAL, *n* = 59; HTA, *n* = 3; NHS EED, *n* = 3; and CENTRAL, *n* = 39
	#2 (preeclamp* or pre-eclamp*) (1623)	
	#3 pre near eclamp* (1163)	
	#4 tox?emia near pregnan* (26)	
	#5 gestosis (29)	
	#6 pregnan* near hypertensi* (1230)	
	#7 gestation near hypertensi* (82)	
	#8 matern* near hypertensi* (290)	
	#9 MeSH descriptor: [Hypertension, Pregnancy-Induced] explode all trees (720)	
	#10 #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 (2417)	
	#11 (PlGF and (triage or alere)) (2)	
	#12 (“placental growth factor” and (triage or alere)) (3)	
	#13 (Elecsys and (“sFlt-1” or “sFlt1” or VEGFR1 or “VEGFR-1” or PlGF or “sFlt-1/PlGF” or “SFlt1/PLGF” or “soluble FMS-like tyrosine kinase-1”)) (0)	
	#14 (PlGF and (Delfia or Perkin Elmer)) (1)	
	#15 ((BRAHMS or Kryptor or Thermo) and (PlGF or “sFlt-1” or “sFlt1” or “sFlt-1/PlGF” or “SFlt1/PlGF” or VEGFR1 or “VEGFR-1” or “soluble FMS-like tyrosine kinase-1”)) (0)	
	#16 #11 or #12 or #13 or #14 or #15 (4)	
	#17 #10 and #16 (3)	
	#18 MeSH descriptor: [Pre-Eclampsia] explode all trees and with qualifier(s): [Diagnosis - DI] (68)	
	#19 (test* or triage or assay* or immunoassay* or electrochemiluminescen* or detect* or surveillance or screen* or measur* or analys* or analyz* or determin* or sensitivity or specificity or accuracy or accurate or predict*) (511,393)	

Database, host, years searched, date searched	Literature search strategy (number of references)	Results
	#20 ("sFlt-1" or "sFlt1" or VEGFR1 or "VEGFR-1" or PlGF or "sFlt-1/PlGF" or "soluble FMS-like tyrosine kinase-1") (100) #21 #19 and #20 (93) #22 MeSH descriptor: [Diagnostic Techniques, Obstetrical and Gynecological] explode all trees (2906) #23 #21 and #22 (2) #24 #10 and #21 (22) #25 ((preeclamp* or pre-eclamp* or "pre eclamp") and (diagnos* and test*)):ti (4) #26 ((preeclamp* or pre-eclamp* or "pre eclamp") and (diagnos* and assay*)):ti,ab (1) #27 ((preeclamp* or pre-eclamp* or "pre eclamp") and (diagnos* and immunoassay*)) (5) #28 ((preeclamp* or pre-eclamp* or "pre eclamp") and (diagnos* and electrochemiluminescen*)) (1) #29 (PlGF or "placental growth factor") (27) #30 #10 and #19 and #29 (32) #31 #17 or #18 or #23 or #24 or #25 or #26 or #27 or #28 or #30 Publication Year from 2000 to 2015 (83)	
CRD, DARE, HTA, NHS EED; inception to 11 March 2015; searched 11 March 2015 Also looked at new Canadian and International HTA • www.cadth.ca/en/resources/hta-database • www.crd.york.ac.uk/PanHTA/AboutPage.asp	1. MeSH DESCRIPTOR pre-eclampsia EXPLODE ALL TREES (99) 2. ((preeclamp* or "pre-eclamp*" or "pre eclamp*")) (211) 3. ((pregnan* and (toxaemia or toxemia))) (3) 4. (gestation) AND (hypertensi*) (42) 5. ((maternal or maternity)) AND (hypertensi*) (84) 6. MeSH DESCRIPTOR Hypertension, Pregnancy-Induced EXPLODE ALL TREES (112) 7. (gestosis) (0) 8. (pregnan*) AND (hypertensi*) (197) 9. #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 (345) 10. (PlGF) AND ((triage or test* or assay* or immunoassay* or diagnos* or detect* or surveillance or screen* or measur* or analys* or analyz* or determin* or sensitivity or specificity or accuracy or accurate or assessment* or predict* or positive or negative or electrochemiluminescen*)) (1) 11. (Placenta* growth factor) AND ((triage or test* or assay* or immunoassay* or diagnos* or detect* or surveillance or screen* or measur* or analys* or analyz* or determin* or sensitivity or specificity or accuracy or accurate or assessment* or predict* or positive or negative or electrochemiluminescen*)) (7) 12. MeSH DESCRIPTOR Vascular Endothelial Growth Factor Receptor-1 EXPLODE ALL TREES (5) 13. (vegfr) (6) 14. MeSH DESCRIPTOR EARLY DIAGNOSIS (136) 15. MeSH DESCRIPTOR Diagnosis (76) 16. MeSH DESCRIPTOR Diagnostic Tests, Routine EXPLODE ALL TREES (189) 17. MeSH DESCRIPTOR Diagnostic Equipment EXPLODE ALL TREES (113) 18. MeSH DESCRIPTOR Diagnostic Techniques, Obstetrical and Gynecological EXPLODE ALL TREES (807) 19. MeSH DESCRIPTOR Diagnostic Services EXPLODE ALL TREES (2535) 20. MeSH DESCRIPTOR Maternal Serum Screening Tests EXPLODE ALL TREES (5) 21. MeSH DESCRIPTOR Serologic Tests EXPLODE ALL TREES (146) 22. MeSH DESCRIPTOR Biological Markers (601)	$n = 45$: 44 from main CRD HTA database; and one extra hit from International and Canadian HTA database

Database, host, years searched, date searched	Literature search strategy (number of references)	Results
	23. (fms-like tyrosine kinase*) (3) 24. (((“FLT 1” or “sFLT 1” or “FLT1” or “sFLT1”) and (triage or test* or assay* or immunoassay* or diagnos* or detect* or screen* or measur* or analys* or analyz* or determin* or sensitivity or specificity or accuracy or accurate or “prognostic assessment*” or predict* or positive or negative or electrochemiluminescen*))) (1) 25. ((“soluble fms-like tyrosine kinase” and (triage or test* or assay* or immunoassay* or diagnos* or detect* or measur* or analys* or analyz* or determin* or sensitivity or specificity or accuracy or accurate or assessment* or predict* or positive or negative or electrochemiluminescen*))) (2) 26. (ELECSYS OR ROCHE OR ALERE OR DELFIA OR BRAHMS OR KRYPTOR OR THERMO) (223) 27. #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21 OR #22 OR #23 OR #24 OR #25 OR #26 (4405) 28. #9 AND #27 (44)	

CDSR, Cochrane Database of Systematic Reviews; CENTRAL, Cochrane Central Register of Controlled Trials; CINAHL, Cumulative Index to Nursing and Allied Health Literature; HTA, Health Technology Assessment; MeSH, medical subject heading; NHS EED, NHS Economic Evaluation Database; NIHR, National Institute for Health Research.

Appendix 2 Study selection worksheet for review of test accuracy

Full-text inclusion criteria	First author and RefID:		
	Reviewer 1:	***Reviewer 2:***	
Research type: primary diagnostic research published in English	Yes ↓ Next question	Unclear ↓ Next question	No → EXCLUDE
Population: women presenting with suspected PE between 20 weeks and 36^{+6} weeks	Yes ↓ Next question	Unclear ↓ Next question	No → EXCLUDE
Index test (intervention): • Triage PlGF test • Elecsys immunoassay sFlt-1 to PlGF ratio • DELFIA Xpress PlGF test • BRAHMS sFlt-1 to PlGF ratio	Yes ↓ Next question	Unclear ↓ Next question	No → EXCLUDE
Reference standard (comparator): • Usual clinical assessment[a]	Yes ↓ Next question	Unclear ↓ Next question	No → EXCLUDE
Outcomes: • Diagnostic accuracy[b] • Prognostic accuracy • Other outcomes to be recorded as keyword only (see list below)	Yes ↓ Next question	Unclear ↓ Next question	No → EXCLUDE
FINAL DECISION to be noted in Reference Manger (version 11, Thomson ResearchSoft, San Francisco, CA, USA) User Def field 4	INCLUDE	UNCLEAR	EXCLUDE

a Any definition of usual clinical assessment acceptable unless obviously irrelevant to NHS practice.
b Sensitivity, specificity, predictive values (+ or –), and/or likelihood ratios (+ or –) reported or calculable.

Additional questions/actions – applies for ALL articles whether included or excluded	***No***	***Yes → add the following keyword(s) to Reference Manager keywords field***
Costs or cost-effectiveness reported?	No action	PECOST
Resources reported?	No action	PERES
HRQoL or anxiety reported?	No action	PEQOL
Systematic review or meta-analysis?	No action	PESR
Other relevant (non-diagnostic) outcomes reported?	No action	PEOUTCOME

Study selection worksheet for review of test performance (continued)

Relevant outcomes to note (keyword as PEOUTCOME) only if research type, population, index test and reference standard criteria are also met:

- Time to test result.
- Test failure rate.
- Time to diagnosis.
- Proportion diagnosed with PE.
- Time to onset of PE and/or eclampsia.
- Proportion returned to less intensive follow-up.
- Length of inpatient hospital stay.
- Time to delivery.
- Maternal morbidity and mortality (see below).
- Fetal and neonatal morbidity and mortality (see below).
- Emergency admissions.
- HRQoL including anxiety.

Relevant morbidity outcomes

Maternal	Fetal/neonatal
Biochemical abnormalities	Breathing difficulties
Disseminated intravascular coagulation/thrombosis	Chronic lung disease
Eclampsia	Gestational age at delivery
Emergency caesarean for compromised baby	Growth at delivery
Haematological abnormalities	Intracranial haemorrhage
HELLP syndrome	Late-onset infection
Liver failure	Necrotising enterocolitis
Renal failure	Neonatal length of stay
Severe hypertension	Neonatal resuscitation
Stroke	Preschool developmental delays
	Weight at delivery (very low $\leq$ 1500 g)

Appendix 3 List of references excluded at full-text screening, with reasons, for review of test accuracy

Reference	Not primary diagnostic research	Population outside of scope (not suspected PE)	Test outside of scope	Ineligible comparator and/or no reference standard	No relevant diagnostic accuracy outcomes	2015 meeting abstract; relevance unclear
Andersen *et al.*[48]		✓				
Atzeni *et al.*[135]		✓			✓	
Benton *et al.*[136]		✓				
Benton *et al.*[137]		✓				
Benton *et al.*[138]		✓				
Benton *et al.*[42]		✓				
Benton *et al.*[139]		✓		✓	✓	
Bersinger and Odegard[140]		✓		✓	✓	
Boucoiran *et al.*[141]		✓				
Boucoiran *et al.*[142]		✓				
Boucoiran *et al.*[143]		✓				
Bramham *et al.*[144]		✓			✓	
[a]Bramham *et al.*[145]		✓				
Calabrese *et al.*[146]		✓				
Cappellini *et al.*[147]		✓				
Caruhel *et al.*[148]		✓	✓		✓	
Cetin *et al.*[149]		✓			✓	
Chaiworapongsa *et al.*[150]			✓			
Chaiworapongsa *et al.*[151]			✓			
Chaiworapongsa *et al.*[152]			✓			
Chaiworapongsa *et al.*[47]			✓			
Chaiworapongsa *et al.*[153]			✓			
Chaiworapongsa *et al.*[47]		✓	✓			
Cho *et al.*[154]		✓				
Crispi *et al.*[155]		✓	✓			
[b]Daubert *et al.*[156]						
De Oliviera *et al.*[157]		✓				
De Vivo *et al.*[158]		✓	✓			
Diab *et al.*[159]		✓	✓			
Doherty *et al.*[160]			✓			

Reference	Not primary diagnostic research	Population outside of scope (not suspected PE)	Test outside of scope	Ineligible comparator and/or no reference standard	No relevant diagnostic accuracy outcomes	2015 meeting abstract; relevance unclear
Dröge *et al.*[161]		✓				
Duckworth *et al.*[98]	✓					
Engels *et al.*[162]		✓				
Engels *et al.*[163]		✓				
Erez *et al.*[164]		✓	✓		✓	
Espinoza *et al.*[165]			✓			
Forest *et al.*[166]		✓	✓			
Garcia-Tizon *et al.*[167]		✓		✓		
Gaziano *et al.*[168]		✓	✓			
Gencay *et al.*[85]		✓				
Ghosh *et al.*[169]		✓	✓			
Ghosh *et al.*[170]		✓	✓			
Ghosh *et al.*[171]		✓	✓		✓	
Gomez-Arriaga *et al.*[125]		✓				
Gomez-Arriaga *et al.*[172]		✓				
Gullai *et al.*[173]		✓			✓	
Gullai *et al.*[174]		✓				
Hanita *et al.*[175]		✓				
Hanses *et al.*[176]		✓			✓	
Hassan *et al.*[177]		✓	✓			
Hirashima *et al.*[133]		✓				
Hund *et al.*[178]					✓	
Hund *et al.*[77]					✓	
Hund *et al.*[179]					✓	
Hund *et al.*[180]					✓	
Hunter[97]	✓	✓			✓	
Husse *et al.*[181]		✓			✓	
Hyde and Thornton[182]	✓	✓	✓		✓	
Kaufmann *et al.*[183]		✓	✓			
Khalil *et al.*[184]		✓				
Kim *et al.*[185]		✓	✓			
[a]Kjos *et al.*[186]						✓
Klein *et al.*[187]					✓	
Klein *et al.*[188]					✓	
Knudsen *et al.*[189]		✓				
Kusanovic *et al.*[190]		✓	✓			

Reference	Not primary diagnostic research	Population outside of scope (not suspected PE)	Test outside of scope	Ineligible comparator and/or no reference standard	No relevant diagnostic accuracy outcomes	2015 meeting abstract; relevance unclear
Lai *et al.*[191]		✓				
Leaños-Miranda *et al.*[192]		✓				
Lehnen *et al.*[193]		✓			✓	
Lehnen *et al.*[194]		✓				
Lim *et al.*[195]		✓	✓			
Madzali *et al.*[196]		✓	✓			
Martínez-Ruiz *et al.*[197]		✓			✓	
Mathur *et al.*[198]		✓			✓	
Mazzocco *et al.*[199]		✓			✓	
McElrath *et al.*[200]		✓	✓			
Meiri *et al.*[201]	✓	✓	✓		✓	
Mijal *et al.*[202]		✓	✓		✓	
Molvarec *et al.*[203]		✓				
Molvarec *et al.*[41]		✓			✓	
Moore *et al.*[204]			✓		✓	
Moore *et al.*[205]						
Moore Simas *et al.*[206]		✓	✓			
Moore Simas *et al.*[207]		✓	✓			
Myers *et al.*[208,209]		✓				
National Horizon Scanning Centre[210]	✓					
North *et al.*[211]		✓			✓	
Ohkuchi *et al.*[212]		✓	✓		✓	
Ohkuchi *et al.*[213]		✓				
Ohkuchi *et al.*[214]		✓				
Ohkuchi *et al.*[215]			✓		✓	
Ohkuchi *et al.*[216]		✓	✓			
Park *et al.*[217]		✓				
PerkinElmer, Inc. (company submission study report)[86]		✓				
PerkinElmer, Inc. (company submission study report)[44]		✓				
PerkinElmer, Inc. (product insert)[218]					✓	
Powers *et al.*[219]		✓	✓		✓	
Powers *et al.*[220]		✓	✓		✓	
Prefumo[221]	✓					

Reference	Not primary diagnostic research	Population outside of scope (not suspected PE)	Test outside of scope	Ineligible comparator and/or no reference standard	No relevant diagnostic accuracy outcomes	2015 meeting abstract; relevance unclear
Rana *et al.*[222]					✓	
Rana *et al.*[102]					✓	
Rana *et al.*[223]					✓	
Rana *et al.*[224]					✓	
Rana *et al.*[225]			✓		✓	
Rana *et al.*[226]					✓	
[a]Rasanen *et al.*[227]		✓	✓			
Redman *et al.*[228]		✓				
Redman *et al.*[229]		✓			✓	
Redman *et al.*[230]		✓			✓	
Ris-Stalpers *et al.*[231]		✓				
Ris and van der Post[232]		✓				
Rizos *et al.*[233]		✓			✓	
Roche Diagnostics (product insert)[83]		✓				
Roche Diagnostics (model user manual)[99]					✓	
Rohra *et al.*[234]		✓	✓			
Romero *et al.*[235]		✓	✓		✓	
Rutherford *et al.*[236]			✓			
Rutherford *et al.*[237]			✓			
[a]Saleh *et al.*[238]						✓
Schaarschmidt *et al.*[239]		✓			✓	
Schaarschmidt *et al.*[240]		✓			✓	
Schiettecatte *et al.*[241]		✓			✓	
Schiettecatte *et al.*[242]		✓			✓	
Schnettler *et al.*[33]					✓	
Schoofs *et al.*[243]		✓			✓	
Schoofs *et al.*[244]		✓			✓	
Shaker *et al.*[245]		✓	✓			
[a]Shawkat *et al.*[246]						✓
Sibiude *et al.*[247]					✓	
Sibiude *et al.*[248]					✓	
Southcombe *et al.*[249]		✓				
Staff *et al.*[250]		✓	✓		✓	
Stenczer *et al.*[251]		✓			✓	
Stenczer *et al.*[252]		✓				

Reference	Not primary diagnostic research	Population outside of scope (not suspected PE)	Test outside of scope	Ineligible comparator and/or no reference standard	No relevant diagnostic accuracy outcomes	2015 meeting abstract; relevance unclear
Stenczer *et al.*[253]		✓				
Stenczer *et al.*[254]		✓			✓	
[c]Stepan *et al.*[255]						
Stepan *et al.*[52]	✓					
Stepan *et al.*[256]		✓	✓			
Strunz-McKendry *et al.*[257]		✓			✓	
Stubert *et al.*[258]					✓	
Stubert *et al.*[259]		✓				
Su *et al.*[260]		✓	✓			
Sunderji *et al.*[261]			✓			
Sunderji *et al.*[262]		✓	✓			
Sundrani *et al.*[263]		✓	✓		✓	
Teixeira *et al.*[264]		✓	✓			
Thermo Fisher Scientific (company submission, literature review)[265]	✓	✓			✓	
Thermo Fisher Scientific (company submission, main document)[266]		✓			✓	
Thermo Fisher Scientific (company submission, supporting document)[87]		✓			✓	
Thermo Fisher Scientific (test user manual)[267]	✓	✓			✓	
Thermo Fisher Scientific (product insert)[54]	✓	✓			✓	
Thermo Fisher Scientific (product insert)[55]	✓	✓			✓	
Thornton[268]	✓					
Tidwell *et al.*[269]		✓	✓			
[a]van Helden and Weiskirchen[270]		✓		✓		
Vatten *et al.*[271]		✓	✓		✓	
Verdonk *et al.*[272]	✓					
Verlohren *et al.*[273]		✓			✓	
Verlohren *et al.*[84]		✓				
Verlohren *et al.*[274]		✓			✓	
Verlohren *et al.*[275]		✓			✓	
Verlohren *et al.*[49]		✓				
Verlohren *et al.*[56]		✓				

Reference	Not primary diagnostic research	Population outside of scope (not suspected PE)	Test outside of scope	Ineligible comparator and/or no reference standard	No relevant diagnostic accuracy outcomes	2015 meeting abstract; relevance unclear
Villa *et al.*[276]		✓				
Wald *et al.*[277]		✓	✓			
Wang *et al.*[278]		✓	✓		✓	
[a]Widmer *et al.*[279]						✓
[a]Woods and Dekker[280]						✓
Yu *et al.*[281]		✓	✓		✓	

a Identified in search update (July 2015).
b Daubert *et al.*[156] meeting abstract met the inclusion criteria, but reported insufficient information to inform the diagnostic assessment and has therefore been excluded (authors were contacted for further information, but no response was received).
c Stepan *et al.*[255] meeting abstract: authors were contacted to clarify whether or not this abstract met the eligibility criteria; however, no response was received by the EAG and since the abstract presents minimal information it has been excluded.

Appendix 4 Example data extraction form for review of test accuracy

Example data extraction form for the PELICAN study (Chappell *et al.*[5])

Data extraction forms for all included studies are available from the report authors on request.

Reference and design	Diagnostic tests	Participants	Outcome measures
PELICAN study **Primary reference:** Chappell[5] **Publication year:** 2013 **Related documents:** Abstracts: Chappell (2013)[71] Chappell (2013)[70] Chappell (2012)[72] **Product insert:** Alere[40] **Country:** UK and Ireland **Presenting condition:** symptoms or signs of suspected PE between 20^{+0} and 40^{+6} weeks of gestation, including headache, visual disturbances, epigastric or right upper quadrant pain, hypertension, dipstick proteinuria or suspected FGR. **Condition being diagnosed:** *Superimposed PE* (ACOG practice bulletin definition): new-onset proteinuria in women with HT before 20 weeks, a sudden increase in proteinuria if already present in early gestation, a sudden increase in HT, or the development of HELLP syndrome. *Atypical PE* (ISSHP definition): gestational HT without proteinuria but with other multiorgan involvement or FGR (< 10th customised birthweight percentile). *Severe PE:* ACOG practice bulletin definition.	**Index test:** Alere's Triage test to determine PlGF concentrations. **Test methodology:** blood drawn into EDTA; plasma stored at –80 °C within 1 hour. All test meters were programmed to produce a masked result, indicating satisfactory test completion only, without revealing the value. Laboratory staff were unaware of clinical outcomes. **Test timing:** at presentation: $20–34^{+6}$ weeks (primary analysis group); also $35–36^{+6}$ weeks; $37–40^{+6}$ weeks. **Location in care pathway:** additional to standard clinical assessment which included proteinuria testing. **Reference standard:** final adjudicated diagnosis of pregnancy outcome determined by two independent senior physicians (three if necessary to resolve disagreement) based on documented end points required to fulfil the diagnostic criteria. All adjudicators were masked to PlGF values so that the test result could not influence delivery decisions. Pregnancy outcome details for mother and infant were obtained from case notes and electronic database review. Stated (in Alere's product insert only) that diagnosis was made after the complete course of pregnancy, including a post-partum period of 7 days, wherein each woman was placed into one of 18 categories.	**Number of participants:** Recruited, $n = 649$ Analysed, $n = 625$: $20–34^{+6}$ weeks: 287; $35–36^{+6}$ weeks: 137; $37–40^{+6}$ weeks: 201 **Sample attrition/dropout:** total, $n = 24$: No enrolment sample: 13 Missing sample barcode: 4 No outcome data available: 7 **Selection of participants:** from seven consultant-led maternity units as per the inclusion criteria. Not stated whether or not recruitment was strictly consecutive. **Inclusion criteria for study entry:** enrolled (once only) if the health-care provider deemed that the woman required evaluation for suspected PE, was aged ≥ 16 years and had a singleton or twin pregnancy. **Exclusion criteria for study entry:** any woman already meeting diagnostic criteria for confirmed PE. Stated (in discussion) that there were minimal exclusion criteria so as to maximise generalisability. **Risk factors for PE:** not explicitly stated; mixed-risk population (see participant characteristics below). **Method for determining gestational age:** not reported.	**Primary outcome:** diagnostic accuracy of low plasma PlGF (< 5th percentile for gestational age) to predict need for delivery within 14 days of testing in women presenting before 35 weeks of gestation. **Secondary outcomes:** as for primary, except: 1. Presentation during $35–36^{+6}$ weeks. 2. Presentation ≥ 37 weeks (data not extracted). 3. PlGF threshold < 12 pg/ml (limit of detection). **Other relevant outcomes:** diagnostic accuracy of systolic and diastolic BP, dipstick proteinuria, urate and ALT, and combinations of these. Accuracy of PlGF < 5th percentile for diagnosing SGA at any time after test (not extracted). Time to delivery. **Diagnostic cut-off point(s):** Normal: PlGF ≥ 5th percentile for gestational age. Positive, low: < 5th percentile. Positive, very low: < 12 pg/ml. **Study dates:** January 2011 to February 2012.

Reference and design	Diagnostic tests	Participants	Outcome measures
Study design: prospective single cohort recruited from multiple centres. **Number of centres:** 7 **Funding:** Alere; Tommy's Charity (but stated no authors were paid to write the article). **Competing interests:** 5/15 authors received honoraria from Alere for consultancy ($n = 3$) or speaking at an Alere-sponsored meeting ($n = 2$). Of these, one was also a paid consultant to Roche and PerkinElmer and one had a minority shareholding in Metabolomic Diagnostics (Cork, Ireland).			
Participant characteristics (at booking and enrolment unless stated otherwise)			
		20–34^{+6} weeks (n = *287)*	*35–36^{+6} weeks* (n = *137)*
Age (years), median (IQR)		31.9 (27.0–35.9)	32.4 (27.5–35.4)
Gestational age (weeks), median (IQR)		31.0 (27.9–33.4) [31.1 (28.0–33.4)][a]	36.0 (35.4–36.4)
BMI (kg/m^2), median (IQR)		28.6 (24.2–33.6)	28.6 (24.4–32.7)
Nulliparous, *n* (%)		123 (43) [164 (57.1)][a]	60 (44)
Singleton pregnancy, *n* (%)		275 (96)	123 (90)
Ethnicity, *n* (%)			
White		187 (65)	88 (64)
Black		70 (24)	27 (20)
Asian		19 (7)	13 (9)
Other		11 (4)	9 (7)
Never smoked		204 (73)	101 (76)
Gestational age ranges (weeks), *n* (%) (reported in Alere's product insert only)			
20^{+0}–23^{+6}		27 (9.4)	Not reported
24^{+0}–28^{+6}		69 (24.0)	
29^{+0}–31^{+6}		70 (24.4)	
32^{+0}–34^{+6}		121 (42.2)	

Participant characteristics (at booking and enrolment unless stated otherwise)		
Reasons (non-exclusive) for suspected PE, *n* (%)		
New-onset HT	155 (54)	92 (67)
Worsening of underlying HT	56 (20)	21 (15)
New-onset dipstick proteinuria	161 (56)	85 (62)
Persistent epigastric/right upper quadrant pain	18 (6)	8 (6)
Headaches	84 (29)	44 (32)
Suspected FGR	25 (9)	4 (3)
Previous medical history, *n* (%)		
Previous PE	55 (20)	17 (12)
Previous PE requiring delivery < 34/40 weeks[b]	30 (11)	6 (4.4)
Chronic HT	45 (17)	10 (7.9)
SLE/antiphospholipid syndrome	12 (4.5)	0
Pregestational diabetes mellitus	6 (2.2)	4 (3.2)
Renal disease	19 (7.1)	4 (3.2)
BP (mmHg), median (IQR)		
Highest systolic	144 (131–159) [140.0 (129.0–151.0)][a]	144 (132–153)
Highest diastolic	92 (82–100) [90.0 (80.0–96.0)][a]	94 (86–100)
BP in first trimester (mmHg), median (IQR)		
Highest systolic	120 (110–130)	118 (110–127)
Highest diastolic	74 (66–81)	70 (65–80)
Dipstick proteinuria, *n* (%)		
Not tested	38 (13)	19 (14)
Negative	103 (36)	34 (25)
Positive (1+ or greater)	146 (51)	84 (61)
Perinatal outcomes		
	20–34^{+6} weeks (n = *287 unless stated)*	*35–36^{+6} weeks* (n = *137 unless stated)*
PE diagnosis, *n* (%)[c]	176 (61) [178 (62.0)]	81 (59)
Mild PE	25 (9)	24 (18)
Severe PE	76 (26)	31 (23)
Superimposed PE	40 (14[d]) [32 (11.1)]	10 (6)
Atypical PE	32 (11[d]) [37 (12.9)]	15 (12)
Eclampsia	1 (0)	1 (1)
HELLP syndrome	2 (1) [5 (1.7)][a]	0 (0)
No PE diagnosis, *n* (%)[c]	111 (39) [109 (38.0)]	56 (41)
Gestational HT [mild gestational HT]	27 (9) [24 (8.4)][a]	14 (10)
Chronic HT only	28 (10) [29 (10.1)]	9 (7)
Isolated proteinuria only	10 (3)	6 (4)
Isolated SGA (< 10th customised birthweight percentile)	8 (3)	3 (2)

Perinatal outcomes		
Transient HT	14 (5)	17 (12)
Normal	22 (8)	5 (4)
Other	2 (1)	2 (1)
Preterm PE diagnosis, *n* (%) (reported in Alere's product insert only)	120 (41.8)	Not reported
Time to delivery, adjusted hazard ratio compared with PlGF > 5th percentile for gestational age (95% CI)[e]		
< 5th percentile for gestational age	2.31 (1.68 to 3.18)	Not reported
< 12 pg/ml	10.61 (7.09 to 15.89)	Not reported
Adverse maternal outcome, *n* (%)	122 (43)	44 (32)
PE group	91[f]	33
No-PE group	28[f]	11
Maternal mortality	0 (0)	0 (0)
Total number of babies	299	151
PE group	186	91
No-PE group	113	60
Babies with preterm delivery, < 37/40 weeks[b], *n/N* (%)	158/299 (53) [148/287 (51.6)][a]	55/151 (36)
Gestation at delivery, weeks, median, IQR (reported in Alere's product insert only)	36.9 (33.6–38.7)	Not reported
Birthweight percentile, median, IQR (reported in Alere's product insert only)	11.4 (0.8–33.4)	Not reported
SGA < 10th percentile, *n* (%) (reported in Alere's product insert only)	138 (48.1)	Not reported
Babies with any adverse perinatal outcome, *n/N* (%)	69/299 (23)	13/151 (8.6)
PE group	56/186 (30.1)	10/91 (11.0)
No-PE group	13/113 (11.5)	3/60 (5.0)
Fetal mortality, *n/N* (%)	7/299 (2.3)	0/151 (0)
PE group	5/186 (2.7)	0/91 (0)
No-PE group	2/113 (1.8)	0/60 (0)
Neonatal mortality, *n/N* (%)	2/299 (0.7)	0/151 (0)
PE group	2/186 (1.1)	0/91 (0)
No-PE group	0/113 (0)	0/60 (0)
Neonatal unit admission > 48 hours, *n/N* (%)	12/299 (4.0)	10/151 (6.6)
PE group	9/186 (4.8)	7/91 (7.7)
No-PE group	3/113 (2.7)	3/60 (5.0)
Respiratory distress syndrome, *n/N* (%)	46/299 (15.4)	3/151 (2.0)
PE group	38/186 (20.4)	2/91 (2.2)
No-PE group	8/113 (7.1)	1/60 (1.7)
Bronchopulmonary dysplasia, *n/N* (%)	6/299 (2.0)	0/151 (0)
PE group	5/186 (2.7)	0/91 (0)
No-PE group	1/113 (0.9)	0/60 (0)

Perinatal outcomes		
Necrotising enterocolitis, *n/N* (%)	4/299 (1.3)	1/151 (0.7)
PE group	4/186 (2.2)	1/91 (1.1)
No-PE group	0/113 (0)	0/60 (0)

BP, blood pressure; EDTA, ethylenediaminetetraacetic acid; FGR, fetal growth restriction; HT, hypertension; IQR, interquartile range; SGA, small for gestational age; SLE, systemic lupus erythematosus.
a Additional data from Alere's product insert. When numbers differ from those reported in Chappell *et al.*,[5] data from the product insert are presented in square brackets.
b As reported in the study publication; timing not explained.
c In the Alere product insert,[40] severe gestational hypertension is classified as PE, while mild gestational hypertension is not. In the Chappell paper[5] gestational hypertension is classified as non-PE.
d Percentage values 11 and 14 transposed in publication – corrected by reviewer.
e Taken from meeting abstract.[71] Based on Cox proportional hazards analysis controlling for gestational age at enrolment and final diagnosis. Time to delivery in weeks is shown in a graph in the main publication[5] (not extracted).
f Numbers in PE and no-PE groups in Chappell *et al.*[5] supplementary table S1 do not sum to the reported total.

Diagnostic accuracy results

(A) Presentation 20–34^{+6} weeks: prediction of preterm PE (= PE before week 37^{+0})			
PlGF cut-off < 100 pg/ml (data reported in Alere's product insert only)	*Population with PE*	*Population without PE*	*Total*
PlGF test positive	(a) 108	(c) 58	166
PlGF test negative	(b) 12	(d) 109	121
Total	120	167	287
Test accuracy statistics	***Parameter value***	***95% CI***	
Sensitivity a/(a + b)	0.900	0.832 to 0.947	
Specificity d/(c + d)	0.653	0.575 to 0.725	
PPV a/(a + c)	0.651	0.573 to 0.723	
NPV d/(b + d)	0.901	0.833 to 0.948	
Positive likelihood ratio[a] sensitivity/(100 – specificity)	2.59	2.09 to 3.22	
Negative likelihood ratio[a] (100 – sensitivity)/specificity	0.15	0.09 to 0.27	
Disease prevalence[a] a + b/(a + b + c + d)	41.81%	36.04% to 47.75%	
(B) Presentation 20–34^{+6} weeks: prediction of preterm PE (= PE before week 37^{+0}) requiring delivery within 14 days of testing			
PlGF cut-off ≥ 100 pg/ml (data reported in Alere's product insert only)	*Population with PE*	*Population without PE*	*Total*
PlGF test positive	(a) 72	(c) 94	166
PlGF test negative	(b) 3	(d) 118	122
Total	75	212	287
Test accuracy statistics	***Parameter value***	***95% CI***	
Sensitivity a/(a + b)	0.960	0.888 to 0.992	
Specificity d/(c + d)	0.557	0.487 to 0.625	
PPV a/(a + c)	0.434	0.357 to 0.513	
NPV d/(b + d)	0.975	0.929 to 0.995	
Positive likelihood ratio[a] sensitivity/(100 – specificity)	2.17	1.85 to 2.54	
Negative likelihood ratio[a] (100 – sensitivity)/specificity	0.07	0.02 to 0.22	
Disease prevalence[a] a + b/(a + b + c + d)	26.13%	21.15% to 31.62%	

(C) Presentation 20 to 34+6 weeks: Prediction of PE (timing unspecified) requiring delivery within 14 days of testing			
(C1) PlGF cut-off < 5th percentile for gestational age	*Population with PE*	*Population without PE*	*Total*
PlGF test positive	(a) 73	(c) 96	169
PlGF test negative	(b) 3	(d) 115	118
Total	76	211	287
Test accuracy statistics	***Parameter value***	***95% CI***	
Sensitivity a/(a + b)	0.96	0.89 to 0.99	
Specificity d/(c + d)	0.55	0.48 to 0.61	
PPV a/(a + c)	0.43	0.36 to 0.51	
NPV d/(b + d)	0.98	0.93 to 0.995	
Positive likelihood ratio[a] sensitivity/(100 – specificity)	2.1	1.8 to 2.5	
Negative likelihood ratio[a] (100 – sensitivity)/specificity	0.07	0.02 to 0.22	
Disease prevalence[a] a + b/(a + b + c + d)	26.48%	21.47% to 31.99%	
(C2) PlGF cut-off < 12 pg/ml	*Population with PE*	*Population without PE*	*Total*
PlGF test positive	(a) 48	(c) 21	69
PlGF test negative	(b) 28	(d) 190	218
Total	76	211	287
Test accuracy statistics	***Parameter value***	***95% CI***	
Sensitivity a/(a + b)	0.63	0.51 to 0.74	
Specificity d/(c + d)	0.90	0.85 to 0.94	
PPV a/(a + c)	0.70	0.57 to 0.80	
NPV d/(b + d)	0.87	0.82 to 0.91	
Positive likelihood ratio[a] sensitivity/(100 – specificity)	6.4	4.1 to 9.9	
Negative likelihood ratio[a] (100 – sensitivity)/specificity	0.41	0.30 to 0.55	
Disease prevalence[a] a + b/(a + b + c + d)	26.48%	21.47% to 31.99%	
(C3) PlGF cut-off < 100 pg/ml ('exploratory analysis')	*Population with PE*	*Population without PE*	*Total*
PlGF test positive	(a) 73	(c) 93	166
PlGF test negative	(b) 3	(d) 118	121
Total	76	211	287
Test accuracy statistics	***Parameter value***	***95% CI***	
Sensitivity a/(a + b)	0.96	0.89 to 0.99	
Specificity d/(c + d)	0.56	0.49 to 0.63	
PPV a/(a + c)	0.44	0.36 to 0.52	
NPV d/(b + d)	0.98	0.93 to 0.995	
Positive likelihood ratio[a] sensitivity/(100 – specificity)	2.2	1.9 to 2.6	
Negative likelihood ratio[a] (100 – sensitivity)/specificity	0.07	0.02 to 0.22	
Disease prevalence[a] a + b/(a + b + c + d)	26.48%	21.47% to 31.99%	

(D) Presentation 20 to 34+6 weeks: prediction of delivery within 14 days			
PlGF cut-off ≥ 100 pg/ml (data reported in Alere's product insert only)	*Population with PE*	*Population without PE*	*Total*
PlGF test positive	(a) 78	(c) 88	166
PlGF test negative	(b) 5	(d) 116	121
Total	83	204	287
Test accuracy statistics	***Parameter value***	***95% CI***	
Sensitivity a/(a + b)	0.940	0.865 to 0.980	
Specificity d/(c + d)	0.569	0.498 to 0.638	
PPV a/(a + c)	0.470	0.392 to 0.549	
NPV d/(b + d)	0.959	0.906 to 0.986	
Positive likelihood ratio[a] sensitivity/(100 – specificity)	2.18	1.84 to 2.57	
Negative likelihood ratio[a] (100 – sensitivity)/specificity	0.11	0.04 to 0.25	
Disease prevalence[a] a + b/(a + b + c + d)	28.92%	23.74% to 34.54%	
(E) Presentation 20 to 34+6 weeks: prediction of preterm delivery			
PlGF cut-off < 12 pg/ml (data reported in Alere's product insert only)	*Population with PE*	*Population without PE*	*Total*
PlGF test positive	(a) 65	(c) 4	69
PlGF test negative	(b) 83	(d) 135	218
Total	148	139	287
Test accuracy statistics	***Parameter value***	***95% CI***	
Sensitivity a/(a + b)	0.439	0.358 to 0.523	
Specificity d/(c + d)	0.971	0.928 to 0.992	
PPV a/(a + c)	0.942	0.858 to 0.984	
NPV d/(b + d)	0.619	0.551 to 0.684	
Positive likelihood ratio[a] sensitivity/(100 – specificity)	15.26	5.71 to 40.78	
Negative likelihood ratio[a] (100 – sensitivity)/specificity	0.58	0.50 to 0.67	
Disease prevalence[a] a + b/(a + b + c + d)	51.57%	45.62% to 57.48%	

(F) Presentation 35^{+0} to 36^{+6} weeks: prediction of PE (timing unspecified) requiring delivery within 14 days of testing			
(F1) PlGF cut-off < 5th percentile for gestational age	*Population with PE*	*Population without PE*	*Total*
PlGF test positive	(a) 47	(c) 25	72
PlGF test negative	(b) 20	(d) 45	65
Total	67	70	137
Test accuracy statistics	***Parameter value***	***95% CI***	
Sensitivity a/(a + b)	0.70	0.58 to 0.81	
Specificity d/(c + d)	0.64	0.52 to 0.75	
PPV a/(a + c)	0.65	0.53 to 0.76	
NPV d/(b + d)	0.69	0.57 to 0.80	
Positive likelihood ratio[a] sensitivity/(100 – specificity)	2.0	1.4 to 2.8	
Negative likelihood ratio[a] (100 – sensitivity)/specificity	0.46	0.31 to 0.71	
Disease prevalence[a] a + b/(a + b + c + d)	48.91%	40.27% to 57.58%	
(F2) PlGF cut-off < 12 pg/ml	*Population with PE*	*Population without PE*	*Total*
PlGF test positive	(a) 15	(c) 6	21
PlGF test negative	(b) 52	(d) 64	116
Total	67	70	137
Test accuracy statistics	***Parameter value***	***95% CI***	
Sensitivity a/(a + b)	0.22	0.13 to 0.34	
Specificity d/(c + d)	0.91	0.82 to 0.97	
PPV a/(a + c)	0.71	0.48 to 0.89	
NPV d/(b + d)	0.55	0.46 to 0.64	
Positive likelihood ratio[a] sensitivity/(100 – specificity)	2.6	1.1 to 6.3	
Negative likelihood ratio[a] (100 – sensitivity)/specificity	0.85	0.73 to 0.98	
Disease prevalence[a] a + b/(a + b + c + d)	48.91%	40.27% to 57.58%	

a Calculated by reviewer.

Receiver operator characteristic curve summary (presentation 20–34+6 weeks)

Prediction of preterm PE (data from Alere's product insert only)	AUC	95% CI
PlGF	0.862	0.818 to 0.907
Prediction of PE requiring delivery within 14 days of testing	***AUC***	***Standard error***
PlGF	0.87	0.03
SBP	0.67	0.05
DBP	0.66	0.05
Urate	0.68	0.06
ALT	0.61	0.05
Dipstick proteinuria	0.76	0.04
PlGF + SBP + DBP	0.87	0.03
PlGF + SBP + DBP + Urate + ALT	0.87	0.03
SBP + DBP + Urate + ALT	0.70	0.05

DBP, diastolic blood pressure; SBP, systolic blood pressure.

Interpretability of test

Numbers excluded from analysis due to indeterminate test results	None of the reported exclusions was specifically attributed to indeterminate results
Test reproducibility	595 of the 625 samples (95%) were measured in duplicate, at the study site and a central laboratory
	Of these, 170 (29%) were out of the range of detection (12–3000 pg/ml) in ≥ 1 evaluation
	85.4% of women received the same classification in both evaluations
	11.1% moved between low and very low (in either direction) (i.e. between < 5th percentile and < 12 pg/ml classes)
	3.5% moved between low (< 5th percentile) and normal (> 5th percentile)
	Sensitivity and specificity in predicting delivery within 14 days of testing were changed by < 1% when 29 twin pregnancies were excluded
Adverse events associated with testing	Stated there were none

Comments

All diagnostic accuracy outcomes are taken from the main publication[5] or Alere's product insert.[40] Chappell *et al.*[71] (meeting abstract) also reported the accuracy of PlGF < 5th percentile of gestational age for diagnosing PE within 14 days of testing, but their results differ slightly from those reported in the main publication and have not been extracted (sample sizes reported in the abstract[71] are slightly lower, but the diagnostic outcomes are very similar to those of the main publication). Chappell *et al.*[70,72] (meeting abstracts) reported interim results,[72] which have been superseded by those in the main publication and ROC statistics[70] that duplicate those in the main publication.

Critical appraisal criteria (based on Reitsma et al.[67] *adaptation of the QUADAS Tool[68])*

	Item	Description	Judgement
1	Was the spectrum of patients representative of the patients who will receive the test in practice?	Suspected PE in UK maternity units	Yes
2	Is the reference standard likely to classify the target condition correctly?	Clinical assessment of BP and proteinuria (± other factors) is the current gold standard for diagnosing PE	Yes
3	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Not explicitly stated but routine monitoring with ≤ 2-week interval and PE diagnosis unlikely to change once determined	Yes
4	Did the whole sample or a random selection of the sample, receive verification using the intended reference standard?	Whole sample	Yes
5	Did patients receive the same reference standard irrespective of the index test result?	Clinicians applying reference standard assessments were masked to PlGF test results	Yes
6	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	PlGF test was not part of the reference standard; although PlGF was tested in addition to standard clinical assessment (i.e. in addition to the reference standard), PlGF concentrations alone are the outcome (except in some ROC analyses)	Yes
7	Were the reference standard results interpreted without knowledge of the results of the index test?	Clinicians diagnosing PE were masked to PlGF test results	Yes
8	Were the index test results interpreted without knowledge of the results of the reference standard?	Laboratory staff interpreting PlGF tests were unaware of clinical outcomes	Yes
9	Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	The test result alone informed PE diagnosis; in clinical practice the test result would be interpreted alongside hypertension, proteinuria and/or other clinical signs or symptoms	No
10	Were uninterpretable/intermediate test results reported?	Uninterpretable/intermediate test results are unlikely, as the test has specific cut-off points, but reproducibility of test results was explored and indicated that some patients changed PlGF classification status depending upon whether the study centre or central laboratory ran the test	Yes
11	Were withdrawals from the study explained?	Reason for missing outcomes not reported but these formed only 1% of the population	Yes

BP, blood pressure.

Appendix 5 Potentially relevant ongoing studies of test accuracy

Test	Study	Objective	Completion date[a]
Triage PlGF test	MAPPLE audit (Germany, Austria, Australia and the UK; expected $n = 1000$)	Audit of whether or not PlGF testing improves management of hypertension in pregnancy. Unclear whether or not it includes a suspected PE population	Unclear; interim analysis available on 49 women (abstract only)[282] (excluded by EAG as no diagnostic outcomes reported)
	PARROT UK study (seven centres in the UK, $n = 1000$)	Assess PlGF testing in suspected PE to reduce maternal morbidity. Outcomes include time to delivery and fetal and maternal adverse outcomes	Mid-2016
	PARROT Ireland RCT (seven centres in Ireland, $n = 800$)	Similar to PARROT UK. Primary outcome: time to diagnosis; secondary outcomes: time to intervention, cost-effectiveness	Late 2016
Elecsys sFlt-1 to PlGF ratio test	INSPIRE RCT (one centre in the UK, $n = 366$)	Determine whether sFlt-1 to PlGF ratio can decrease hospital admissions for women with suspected PE. Outcomes include bed-days/episodes and health-care visits	2016
Elecsys PlGF test	Unnamed cohort study (King's College London, $n = 10{,}000$)	Assess PlGF (and sFlt-1) for predicting PE in women attending for routine 36–37 week scan. Unclear whether or not it includes a suspected PE population	Not reported (source: Roche Diagnostics company submission[45])
BRAHMS Kryptor sFlt-1 to PlGF ratio test	ROPE cohort study (USA, $n = 616$)	Assess prognostic value of sFlt-1 to PlGF ratio in women with suspected PE. Outcomes include maternal, fetal and neonatal adverse events. Unclear whether or not diagnostic outcomes included	Draft manuscript expected 2015
BRAHMS Kryptor PlGF test	Essen cohort study (Germany, $n = 1200$)	Assess PlGF (and sFlt-1) for predicting PE in gestation weeks 34–36. Outcomes include adverse events. Unclear whether or not it includes a suspected PE population	2016

RCT, randomised controlled trial.
a Results of the listed trials were not available at the time of going to press.

Appendix 6 MEDLINE search strategies for HRQoL studies

Initial strategy

Database	Literature search strategy (number of references)	Results
Database: Ovid MEDLINE(R) 1946 to 2015 and MEDLINE In Process & Other Non-Indexed Citations; cross-searched; searched 4 August 2015	1. Pre-Eclampsia/ (24,813) 2. (preeclamp* or pre eclamp*).tw. (22,892) 3. (tox?emi* adj5 pregnan*).tw. (3290) 4. gestosis.tw. (1195) 5. (pregnan* adj3 hypertensi*).tw. (9047) 6. (gestation* adj3 hypertensi*).tw. (2338) 7. ((maternal or maternity) adj3 hypertens*).tw. (1089) 8. Hypertension, Pregnancy-Induced/ (1865) 9. or/1-8 (40,320) 10. value of life/ (5497) 11. quality adjusted life year/ (7902) 12. quality adjusted life.ti,ab. (7592) 13. (qaly$ or qald$ or qale$ or qtime$).ti,ab. (6231) 14. disability adjusted life.ti,ab. (1676) 15. daly$.ti,ab. (1596) 16. health status indicators/ (20,895) 17. (sf36 or sf 36 or short form 36 or shortform 36 or sf thirtysix or sf thirty six or shortform thirstysix or shortform thirty six or short form thirty six or short form thirtysix or short form thirty six).ti,ab. (18,169) 18. (sf6 or sf 6 or short form 6 or shortform 6 or sf six or sfsix or shortform six or short form six).ti,ab. (1492) 19. (sf12 or sf 12 or short form 12 or shortform 12 or sf twelve of sftwelve or shortform twelve or short form twelve).ti,ab. (3369) 20. (sf16 or sf 16 or short form 16 or shortform 16 or sf sixteen or sfsixteen or shortform sixteen or short form sixteen).ti,ab. (24) 21. (sf20 or sf 20 or short form 20 or shortform 20 or sf twenty of sftwenty or shortform twenty of short form twenty).ti,ab. (344) 22. (euroqol or euro qol or eq5d or eq 5d).ti,ab. (5139) 23. (hql or hqol or h qol or hrqol or hr qol).ti,ab. (9648) 24. (hye or hyes).ti,ab. (55) 25. health$ year$ equivalent$.ti,ab. (39) 26. health utilit$.ab. (1232) 27. (hui or hui1 or hui2 or hui3).ti,ab. (1022) 28. disutil$.ti,ab. (273) 29. rosser.ti,ab. (74) 30. quality of well being.ti,ab. (359) 31. quality of wellbeing.ti,ab. (8) 32. qwb.ti,ab. (186) 33. willingness to pay.ti,ab. (2872) 34. standard gamble$.ti,ab. (734) 35. time trade off.ti,ab. (861) 36. time tradeoff.ti,ab. (227) 37. tto.ti,ab. (697) 38. (index adj2 well being).mp. (611) 39. (quality adj2 well being).mp. (1074) 40. (health adj3 utilit$ ind$).mp. (739) 41. ((multiattribute$ or multi attribute$) adj3 (health ind$ or theor$ or health state$ or utilit$ or analys$)).mp. (274) 42. quality adjusted life year$.mp. (11,387) 43. (15D or 15 dimension$).mp. (1392) 44. (12D or 12 dimension$).mp. (499) 45. rating scale$.mp. (95,898) 46. linear scal$.mp. (942)	47

Database	Literature search strategy (number of references)	Results
	47. linear analog$.mp. (931) 48. visual analog$.mp. (37,956) 49. (categor$ adj2 scal$).mp. (1414) 50. or/10-49 (204,540) 51. (letter or editorial or comment).pt. (1,496,950) 52. 50 not 51 (199,678) 53. 9 and 52 (57) 54. limit 53 to english language (54) 55. limit 54 to yr="2000 -Current" (47)	
EMBASE; 1974 to 2015; searched 4 August 2015	1. preeclampsia/ or "eclampsia and preeclampsia"/ (40,859) 2. (preeclamp* or "pre eclamp*").tw. (31,639) 3. (tox?emi* adj5 pregnan*).tw. (3400) 4. gestosis.tw. (1433) 5. (pregnan* adj3 hypertensi*).tw. (11,993) 6. (gestation* adj3 hypertensi*).tw. (3461) 7. ((maternal or maternity) adj3 hypertens*).tw. (1517) 8. maternal hypertension/ (11,439) 9. pregnancy toxemia/ (3291) 10. or/1-9 (57,825) 11. quality adjusted life year/ (14,357) 12. quality adjusted life.ti,ab. (10,443) 13. (qaly$ or qald$ or qale$ or qtime$).ti,ab. (10,622) 14. disability adjusted life.ti,ab. (1912) 15. daly*.ti,ab. (2013) 16. (sf36 or sf 36 or short form 36 or shortform 36 or sf thirtysix or sf thirty six or shortform thirtysix or shortform thirty six or short form thirty six or short form thirtysix or short form thirty six).ti,ab. (26,757) 17. (sf6 or sf 6 or short form 6 or shortform 6 or sf six or sfsix or shortform six or short form six).ti,ab. (1626) 18. (sf12 or sf 12 or short form 12 or shortform 12 or sf twelve or sftwelve or shortform twelve or short form twelve).ti,ab. (5196) 19. (sf16 or sf 16 or short form 16 or shortform 16 or sf sixteen or sfsixteen or shortform sixteen or short form sixteen).ti,ab. (37) 20. (sf20 or sf 20 or short form 20 or shortform 20 or sf twenty or sftwenty or shortform twenty or short form twenty).ti,ab. (358) 21. (euroqol or "euro qol" or "eq5d" or "eq 5d").ti,ab. (8886) 22. (hql or hqol or "h qol" or hrqol or "hr qol").ti,ab. (14,451) 23. health* year* equivalent*.ti,ab. (39) 24. health utilit*.ti,ab. (1886) 25. (hui or hui1 or hui2 or hui3).ti,ab. (1458) 26. (health adj3 util* adj ind*).mp. (1005) 27. quality adjusted life year*.mp. (16,554) 28. "health related quality of living".ti,ab. (4) 29. "health related quality of life".ti,ab. (34,896) 30. (categor* adj scale*).mp. (921) 31. "quality of life".ti. (64,109) 32. or/11-31 (122,834) 33. 10 and 32 (60) 34. limit 33 to (english language and yr="2000 -Current") (60)	60
The Cochrane Library; searched 4 August 2015	1. Search Name: QOL PREECLAMPSIA 2. Last Saved: 04/08/2015 09:51:47.833 3. ID Search 4. #1 MeSH descriptor: [Pre-Eclampsia] this term only 5. #2 (preeclamp* or pre-eclamp*) 6. #3 pre near eclamp* 7. #4 tox?emia near pregnan* 8. #5 gestosis 9. #6 pregnan* near hypertensi* 10. #7 gestation near hypertensi*	42 results, zero downloaded as nothing relevant (37 CDSR, one DARE, one NHS EED and one Cochrane group)

Database	Literature search strategy (number of references)	Results
	11. #8 matern* near hypertensi* 12. #9 MeSH descriptor: [Hypertension, Pregnancy-Induced] explode all trees 13. #10 #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 Publication Year from 2010 to 2015 14. #11 (sf36 or "sf 36" or "short form 36" or "shortform 36" or "sf thirtysix" or "sf thirty six" or "shortform thirtysix" or "shortform thirty six" or "short form thirty six" or "short form thirtysix" or "short form thirty six") 15. #12 sf near 36 16. #13 ("hui" or "hui1" or "hui-1" or "hui2" "hui-2" or "hui3" or "hui-3") 17. #14 health near utilit* 18. #15 (hql or hqol or "h qol" or "h-qol" or hrqol or "hr qol" or "hr-qol") 19. #16 "health related quality of life" 20. #17 #11 or #12 or #13 or #14 or #15 or #16 21. #18 #10 and #17 Publication Year from 2010 to 2015	
Web of Science QOL Pre-eclampsia; Indexes=SCI-EXPANDED, CPCI-S Timespan=2000–15	1. #16 #15 AND #8 (10) 2. #15 #14 OR #13 OR #12 OR #11 OR #10 OR #9 (40,466) 3. #14 (TS=(hql or hqol or "h qol" or hrqol or "hr qol")) (8457) 4. #13 (TS=("health related quality of life")) (24,618) 5. #12 (TS=("health utility" or "health utilities")) AND LANGUAGE: (English) (1285) 6. #11 (TS=("hui" or "hui1" or "hui2" or "hui3")) AND LANGUAGE: (English) (1185) 7. #10 (TS=(sf36 or "sf 36" or "short form 36" or "shortform 36" or "sf thirtysix" or "sf thirty six" or "shortform thirtysix" or "shortform thirty six" or "short form thirty six" or "short form thirtysix" or "short form thirty six")) (16,491) 8. #9 (TS=(euroqol or "euro qol" or "eq5d" or "eq 5d")) (5030) 9. #8 #7 OR #6 OR #5 OR #4 OR #3 OR #2 OR #1 (24,007) 10. #7 (TS=(maternity NEAR hypertensi*)) (27) 11. #6 (Ts=("maternal hypertensi*")) (362) 12. #5 (TS=(gestation NEAR hypertensi*)) (806) 13. #4 (TS=(pregnan* NEAR hypertensi*)) (108) 14. #3 (TS=(gestosis)) (8030) 15. #2 (TS=(tox?emia NEAR pregnan*)) (52) 16. #1 (TS=(preeclamp* or "pre eclamp*" or "pre-eclamp*")) AND LANGUAGE: (English) (20,388)	10
CRD – pre-eclampsia search just all NHS EED records taken	1. MeSH DESCRIPTOR pre-eclampsia EXPLODE ALL TREES (103) 2. ((preeclamp* or "pre-eclamp*" or "pre eclamp*")) (214) 3. ((pregnan* and (toxaemia or toxemia))) (3) 4. (gestation) AND (hypertensi*) (42) 5. ((maternal or maternity)) AND (hypertensi*) (84) 6. MeSH DESCRIPTOR Hypertension, Pregnancy-Induced EXPLODE ALL TREES (116) 7. (gestosis) (0) 8. (pregnan*) AND (hypertensi*) (199) 9. #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 (349) 10. * FROM 2010 TO 2015 (41,846) 11. #9 AND #10 (157) NHS EED 11 results.	11

CDSR, Cochrane Database of Systematic Reviews; CENTRAL, Cochrane Central Register of Controlled Trials; CINAHL, Cumulative Index to Nursing and Allied Health Literature; HTA, Health Technology Assessment; MeSH, medical subject heading. NHS EED, NHS Economic Evaluation Database; NIHR, National Institute for Health Research.

Revised strategy with wider pregnancy terms and specific HRQoL terms

Database	Search strategy	Results
Ovid MEDLINE(R) 1946 to July week 5 2015 + Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations August 10, 2015; years searched 2000–15; searched 11 August 2015	1. *Pregnancy/ (50,039) 2. Postpartum Period/ (18,440) 3. Peripartum Period/ (480) 4. (maternity or maternal or pregnan* or postpartum or "post-partum" or peripartum or "peri-partum" or puerperium).ti. (257,835) 5. or/1-4 (280,726) 6. (sf36 or sf 36 or short form 36 or shortform 36 or sf thirtysix or sf thirty six or shortform thirtysix or shortform thirty six or short form thirty six or short form thirtysix or short form thirty six).ti,ab. (18,169) 7. (sf6 or sf 6 or short form 6 or shortform 6 or sf six or sfsix or shortform six or short form six or SF6D or "SF 6D" or "SF 6 D").ti,ab. (1959) 8. (sf12 or sf 12 or short form 12 or shortform 12 or sf twelve of sftwelve or shortform twelve or short form twelve).ti,ab. (3369) 9. (hui or hui1 or hui2 or hui3 or "hui-1" or "hui-2" or "hui-3").ti,ab. (1022) 10. "health utility index*".tw. (139) 11. "health utilities index*".tw. (584) 12. (eq5d or "eq5-D").tw. (248) 13. ("euroqol 5D" or "euro qol 5D").tw. (690) 14. "nottingham health profile*".tw. (1013) 15. QALY.tw. (5024) 16. Quality-adjusted life years/ (7902) 17. or/6-16 (35,293) 18. 5 and 17 (194) 19. limit 18 to (english language and yr="2000 -Current") (184) 20. (comment or editorial of letter).pt. (663,343) 21. 19 not 20 (184) 22. remove duplicates from 21 (178)	178
Database: EMBASE 1996 to 2015 week 32; years searched 2000–15; searched 11 August 2015	1. Database: EMBASE <1996 to 2015 Week 32> 2. Search strategy: 3. *pregnancy/ (57,290) 4. *puerperium/ (4936) 5. (maternity or maternal or pregnan* or postpartum or "post-partum" or peripartum or "peri-partum" or puerperium).ti. (171,973) 6. or/1-3 (182,196) 7. (sf36 or sf 36 or short form 36 or shortform 36 or sf thirtysix or sf thirty six or shortform thirtysix or shortform thirty six or short form thirty six or short form thirtysix or short form thirty six).tw. (26,672) 8. (sf6 or sf 6 or short form 6 or shortform 6 or sf six or sfsix or shortform six or short form six).tw. (1112) 9. (sf12 or sf 12 or short form 12 or shortform 12 or sf twelve of sftwelve or shortform twelve or short form twelve).tw. (5196) 10. (eq5d or "eq 5D").mp. (7280) 11. ("euroqol 5D" or "euro qol 5D").mp. (1171) 12. ("health utility index" or "health utilities index").mp. (907) 13. Nottingham Health Profile/ (243) 14. "nottingham health profile".tw. (1071) 15. (hui or hui1 or hui2 or "hui3" or "hui-1" or "hui-2" or "hui-3").mp. (1751) 16. quality adjusted life year/ (14,337) 17. qaly.tw. (8474) 18. or/5-15 (57,598) 19. 4 and 16 (346) 20. limit 17 to (english language and yr="2000 -Current") (332) 21. limit 18 to (conference abstract or conference paper or conference proceeding or "conference review" or editorial or letter or note) (148) 22. 18 not 19 (184)	184
DelphiS – University of Southampton: cross-university resource	Series of simple searches with pregnancy terms combined with HRQoL parameters	Eight unique records identified

Appendix 7 Example data extraction form for full economic analyses

Example data extraction form for the study by Hadker *et al.*[39]

Data extraction forms for all included cost-effectiveness studies are available from the report authors on request.

1	Study	Hadker *et al.*, 2010[39]
2	Research question	To assess the potential clinical and economic benefits of early detection of PE (with a new serum test) vs. standard practice
3	Country/setting	UK, primarily secondary care including specialist prenatal management of PE and hospital admission for delivery. The study population is a general obstetric population including a proportion with pre-existing risk factors for PE, treated with aspirin
4	Funding source	Study sponsored by Roche Diagnostics Ltd
5	Analysis type	CEA, CUA, CBA
6	Study type	Decision analytical model: using a decision tree
7	Perspective	Health-care payer perspective
8	Time horizon	Model time horizon is stated as being from week 12 to term (week 40), but does not appear to include any delivery or post-delivery costs for TN cases (with or without PE risk), although their delivery would be expected to fall within this model time horizon
9	Model assumptions	Implicit assumptions not discussed in the model publication: • the incidence of PE • the distribution of severity of outcome • the sensitivity and specificity of both test strategies are the same in women with PE risk factors and women without PE risk factors Women who have an initial negative biomarker test result are tested a further two times. The diagnostic outcomes of the strategy including the biomarker test were not adjusted for including these repeated tests (i.e. the model used sensitivity and specificity for a single test). This assumption only affects cost
10	Discounting (rate)	No. Not required due to short model time horizon
11	Costing year, currency	Assume costing year is 2008/09 financial year. The costing year is not explicitly stated, but sources for treatment and drug costs are 2008/09 NHS Reference Costs[283] and 2009 *British National Formulary*[120] Currency is UK pounds sterling

1	Study	Hadker *et al.*, 2010[39]
12	Population	Baseline population characteristics are based on information from observational studies, expert opinion and assumption

Assumption	*Source*
Cohort size = 1000	Assumption
Proportion of cohort with risk factors (to be treated with aspirin) = 15%	UK expert opinion
Incidence of PE in general obstetric population = 4.03%	Bhattacharya S, Campbell DM. The incidence of severe complications of preeclampsia. *Hypertens Pregnancy* 2000;**24**:181–190[a]
Distribution of severity • Mild PE = 93.60% • Severe PE = 4.75% • Eclampsia = 1.65% Maternal death = 0.00%	Bhattacharya S, Campbell DM. The incidence of severe complications of preeclampsia. *Hypertens Pregnancy* 2000;**24**:181–190

a A retrospective analysis of cases with a diagnosis of hypertension associated with proteinuria in the Grampian region of Scotland, between 1981 and 2000 (identified from Aberdeen Maternity and Neonatal Databank). A total of 4188 cases were identified out of 103,896 deliveries, yielding overall incidence of 4.03% (varying from 1.16% in 1995 to 8.32 in 1984 – paper states that there was decline in numbers over time).

The model publication did not include any specific reference to neonatal outcomes. There is a mention of intensive care in reference to costing assumptions. However, it is not clear if this refers to maternal or neonatal intensive care

13	Intervention(s), comparator(s)	Intervention = standard clinical assessment + biomarker test Comparator = standard clinical assessment alone
14	Intervention effect	Effectiveness in the model relates to the diagnostic test accuracy (based on sensitivity and specificity) of the two alternative strategies: standard clinical assessment + biomarker test vs. standard clinical assessment alone. Sensitivity and specificity for each strategy, with data source are shown in the table below

Strategy	*Test sensitivity and specificity*	*Source*
Intervention (using reported sensitivity and specificity of test alone, not combination with standard clinical assessment)	Sensitivity = 0.82; specificity = 0.95	Verlohren S, Galindo A, Schlembach D, Zeisler H, Herraiz I, Moertl MG, *et al.* An automated method for the determination of the sFlt-1 to PlGF ratio in the assessment of preeclampsia. *Am J Obstet Gynecol* 2010;**202**:161-e1–11[a]
Comparator (average of sensitivities and specificities of individual test, not combinations. Pooling/ averaging method not reported)	Sensitivity = 0.46; specificity = 0.83	Meads CA, Cnossen JS, Meher S, Juarez-Garcia A, ter Riet G, Duley L *et al.* Methods of prediction and prevention of PE: systematic reviews of accuracy and effectiveness literature with economic modelling. *Health Technol Assess* 2008;**12**(6)

a Study was conducted in a population already diagnosed with PE (i.e. not suspected PE) using a sFlt-1 to PlGF ratio cut-off value of 85.

1	Study	Hadker *et al.*, 2010[39]
15	Health state utilities	Not applicable
16	Intervention cost	Very little detail is provided on costing assumptions. Summary values for initial assessment (12–20 weeks), cost of aspirin therapy, drug cost by disease severity and 'PE management costs' (20–40 weeks) by diagnostic outcome and disease severity (i.e. true/false positive/negative with mild PE/severe PE/eclampsia) – see below for detail Cost of standard clinical assessment = £0 Cost of biomarker test = £31.13 (£93.39 in women with negative initial test, given assumption of two re-tests)

Costs by disease severity/diagnostic outcome	***Costs at 2008/9 price base***
Mild PE	
Drug costs	£28.25
Management costs	
TP	£9576.25
FN	£4480.38
Severe PE	
Drug costs	£127.30
Management costs	
TP	£14,545.49
FN	£11,308.87
Eclampsia	
Drug costs	£21,340.12
Management costs	
TP	£163.19
FN	£17,122.77
No PE, but with risk factors	
Management costs	
FP	£9576.25
TN	£0.00
Aspirin for women with PE risk factors	£2.74

17	Indirect costs	Not included
18	Results	

Per 1000 patients	***Intervention***	***Comparator***	***Incremental***
Costs	£1,780,916	£2,726,224	–£945,309

By diagnostic outcome

Per 1000 patients	***Intervention***	***Comparator***	***Incremental***
Costs			
TP	£361,389	£202,154	£159,236
TN	£872,177	£687,617	£184,560
FP	£503,789	£1,707,802	–£1,204,104
FN	£43,561	£128,652	–£85

1	Study	Hadker *et al.*, 2010[39]

19 Sensitivity analysis

Deterministic univariate sensitivity analyses

Parameter	*Scenario/value*	*Intervention/patient*	*Comparator/patient*	*Difference*
PE incidence	Reduced by 20% from 4.03% to 3.22%	£1771	£2680	£969
Sensitivity of current tests	Improved by 10% from 0.46 to 0.51	£2736	£1781	£955
Specificity of current tests	Improved by 10% from 0.83 to 0.91	£1989	£1781	£208
Proportion of high-risk patients	Reduced by 33% from 0.15 to 0.10	£2644	£1699	£945
Cost of biomarker test	Increased cost of test by 20% from £31.13 to £37.36	£2726	£1798	£928

20 Authors' conclusions

The biomarker test may provide savings for NHS – reduced costs of treatment (following testing) exceed additional costs of the test by some margin

Overall cost reduction of £945 per patient, from £2726 to £1781

Budget impact, per patient, by diagnostic outcome (difference between standard practice and standard practice plus test):

TP = £159

TN = £185

FP = –£1204

FN = –£85

They suggest savings are driven by reduced false-positive and false-negative results compared with current standard of care

Acknowledge limitations in terms of data for standard care (based on pooled average of individual tests [method for pooling/averaging not described explicitly] rather than on combinations of tests) and general data limitations requiring substantial reliance on expert opinion

Source of diagnostic accuracy data for the biomarker test (clinical trial) are acknowledged to be limited and may not reflect standard practice

Further limitations are identified due to lack of inclusion of neonatal outcomes and exclusion of longer-term impacts of PE /early delivery for women and children. They suggest further study is required on longer term direct and indirect costs of PE

21 Reviewer's comments

Reporting of the study suffers from a lack of detail (presumably partly because of the journal word limit)

Insufficient detail is provided on:

- Pooling on sensitivity/specificity for standard assessment
- Resource-use assumptions that underlie costings (only report total costs for assessment, drug costs by severity and 'PE management costs' by diagnostic outcome and disease severity – little or no detail on resource assumptions, thus limiting critical assessment and reproducibility)

Model distinguishes population with and without risk factors for PE – but assumes incidence (and sensitivity/specificity) is the same in both populations. This assumption not discussed

1 Study	Hadker *et al.*, 2010[39]
	There is little or no critical assessment of data sources for parameter inputs or discussion of alternative sources. Overall, the paper does not seem to fully assess uncertainty in the choice of input data or reflect the uncertainty likely to be associated with data inputs
	The authors argue that savings accrue from reductions in both FP and FN test results compared with standard assessment. However, results by diagnostic outcome seem to suggest that the majority of the saving accrues from avoiding FP results; this explains the model's sensitivity in their analyses to the assumed specificity of standard assessment

CBA, cost–benefit analysis; CEA, cost-effectiveness analysis; CUA, cost–utility analysis; FN, false negative; FP, false positive; TN, true negative; TP, true positive.

Appendix 8 List of model inputs

This appendix provides a quick reference to all of the model inputs stated throughout the economic section of this assessment. Any blank values in tables for sensitivity analysis indicate that the relevant parameters were not varied in deterministic sensitivity analyses.

Summary of model inputs

Table 55 provides a summary of the sensitivities and specificities for the Triage PlGF test, the Elecsys sFlt-1 to PlGF ratio test and the standard diagnostic assessment used in the model. Sensitivities and specificities for the Triage PlGF test and Elecsys sFlt-1 to PlGF ratio test were derived from the systematic review of test accuracy conducted by the EAG (see *Chapter 4*). Sensitivity and specificity for standard assessment was not reported in any of the studies included in the systematic review of test accuracy. The parameter values used for sensitivity and specificity for standard assessment in the model were derived from Schnettler *et al.*,[33] a study that was identified and excluded from the systematic review of test accuracy, but included in the systematic review of economic studies. The values from Schnettler *et al.*[33] were chosen because they best represent what experts told us about identification of PE in clinical practice: most cases are identified, but there is some overtreatment. This indicates that the sensitivity of standard assessment is good, while the specificity is poor.

TABLE 55 Parameters used in the EAG model: diagnostic accuracy

Parameter	Value	Range for sensitivity analysis	Source
Standard clinical assessment			
Sensitivity	0.941	0.838–0.988	Schnettler *et al.*[33]
Specificity	0.357	0.263–0.460	
Triage PlGF test			
Rule in: 20^{+0} to 34^{+6} weeks of gestation			PELICAN study[5] (see *Tables 13*, *14* and *26*)
Sensitivity	0.632	0.513–0.739	
Specificity	0.900	0.852–0.937	
Rule out: 20^{+0} to 34^{+6} weeks of gestation			
Sensitivity	0.960	0.888–0.992	
Specificity	0.557	0.487–0.625	
Rule in: 35^{+0} to 36^{+6} weeks of gestation			
Sensitivity	0.224	0.131–0.342	
Specificity	0.914	0.823–0.968	
Rule out: 35^{+0} to 36^{+6} weeks of gestation			
Sensitivity	0.701	0.577–0.807	
Specificity	0.643	0.519–0.754	
Elecsys sFlt-1 to PlGF ratio test			
Rule in: 20^{+0} to 36^{+6} weeks of gestation			PROGNOSIS study[51] and product insert[83] (see *Tables 17* and *27*)
Sensitivity	0.703	0.619–0.778	
Specificity	0.831	0.805–0.855	
Rule out: 20^{+0} to 36^{+6} weeks of gestation			
Sensitivity	0.857	0.728–0.941	
Specificity	0.791	0.765–0.816	

Table 56 reports the prevalence of PE among women suspected of having PE and the distribution of mild, moderate and severe hypertension in women with suspected PE. The prevalence data were derived from the UK PELICAN study.[5] The hypertension data were reported in Duckworth *et al.*[98] as being derived from Anumba *et al.*[134] The EAG was unable to identify the data from the published study by Anumba *et al.*,[134] so we have used the data reported by Duckworth *et al.*[98] in the absence of any better data. As these data are confidential, they are not displayed in the table. The EAG was consulted, and was unable to identify any other estimates of hypertension severity in PE.

Table 57 reports median times to delivery and probabilities of each type of delivery up to 35 weeks' gestation, whereas *Table 58* reports these values for 35–37 weeks of gestation. Time to delivery determines the cost of hospitalisation in the model, while onset of labour and mode of delivery determine the cost of birth and influence maternal HRQoL. Some data inputs required assumptions (see *Model parameters* for further details on the assumptions applied to data in *Table 56*).

TABLE 56 Parameters used in the EAG model: maternal risk factors

Parameter	Value	Range for sensitivity analysis	Source
Prevalence of PE			
Suspected PE presenting up to 35 weeks of gestation	0.265	0.215–0.320	PELICAN study[5] (see *Tables 13*, *14* and *26*)
Suspected PE presenting between 35 and 37 weeks of gestation	0.489	0.403–0.576	
Distribution of women by degree of hypertension, with/out PE			
With PE			Duckworth *et al.*[98] (credited to, but not reported by, Anumba *et al.*[134])
Severe	Confidential data removed		
Moderate	Confidential data removed		
Mild or no hypertension	Not reported		
Without PE			
Severe	Confidential data removed		
Moderate	Confidential data removed		
Mild or no hypertension	Not reported		

TABLE 57 Parameters used in the EAG model: delivery characteristics by gestation timing and diagnostic test outcome (up to 35 weeks of gestation)

Parameter	Value	Range for sensitivity analysis	Source
Pregnancies presenting up to 35 weeks of gestation, testing positive/rule-in for PE (delivering within 7–14 days)			
Median time to delivery (days)	9	3–16	PELICAN study[5]
Onset of labour			
Spontaneous	0.000		Proportion of caesarean deliveries from the EPIPAGE study[110] (see *Table 29*). Absent further data, the EAG assumed all non-caesarean deliveries require induction
Induction	0.402		
Planned caesarean section	0.598		
Mode of delivery			
Non-assisted vaginal delivery	0.52		PELICAN study[5] – in absence of further data, assume the same distribution as reported for all deliveries presenting before 35 weeks
Assisted vaginal delivery	0.225		
Emergency caesarean section	0.254		

TABLE 57 Parameters used in the EAG model: delivery characteristics by gestation timing and diagnostic test outcome (up to 35 weeks of gestation) (*continued*)

Parameter	Value	Range for sensitivity analysis	Source
Suspected PE presenting up to 35 weeks of gestation, testing negative/rule-out for PE			
Median time to delivery (days)			
PE	14	7–42	PELICAN study[5]
No PE	62	14–63	
Onset of labour			
Spontaneous	0.148		PELICAN study[5] (see *Table 28*)
Induction	0.380		
Planned caesarean section	0.472		
Mode of delivery[a]			
Non-assisted vaginal delivery	0.522		PELICAN study[5]
Assisted vaginal delivery	0.225		
Emergency caesarean section	0.254		
Pregnancies presenting up to 35 weeks of gestation with PE (true positive and false negative)			
Probability of severe complication of delivery due to PE	0.064	0.041–0.093	HYPITAT[124]

a The distribution for mode of delivery was estimated from data reported for the PELICAN study.[5] In total, 134 pre-labour caesarean deliveries were reported from a total of 169 caesarean deliveries in women presenting with suspected PE up to 35 weeks of gestation. We inferred that the difference between these two values would represent spontaneous or induced labours that required emergency caesarean delivery. The distribution across modes of delivery (excluding planned caesarean sections) was calculated using the reported figures for spontaneous and assisted vaginal deliveries (72 and 31, respectively) and the inferred number of emergency caesarean deliveries (35).

TABLE 58 Parameters used in the EAG model: delivery characteristics by gestation timing and diagnostic test outcome (35–37 weeks of gestation)

Parameter	Value	Range for sensitivity analysis	Source
Suspected PE presenting between 35 and 37 weeks of gestation, testing positive/rule-in for PE (delivering within 7 days)			
Average time to delivery (days)	4	2–9	PELICAN study[5]
Onset of labour			
Spontaneous	0.184		PELICAN study[5]
Induction	0.551		
Planned caesarean section	0.265		
Mode of delivery			
Non-assisted vaginal delivery	0.568		PELICAN study[5]
Assisted vaginal delivery	0.137		
Emergency caesarean section	0.295		
Suspected PE presenting between 35 and 37 weeks of gestation, testing negative/rule-out for PE			
Average time to delivery (days)			
PE	4	4–16	PELICAN study[5]
No PE	16	6–23	

continued

TABLE 58 Parameters used in the EAG model: delivery characteristics by gestation timing and diagnostic test outcome (35–37 weeks of gestation) (*continued*)

Parameter	Value	Range for sensitivity analysis	Source
Onset of labour			
Spontaneous	0.184		PELICAN study[5] (see *Table 30*)
Induction	0.551		
Planned caesarean section	0.265		
Mode of delivery[a]			
Non-assisted vaginal delivery	0.568		PELICAN study[5]
Assisted vaginal delivery	0.137		
Emergency caesarean section	0.295		
Pregnancies presenting up to 35 weeks of gestation with PE (true positive and false negative)			
Probability of severe complication of delivery due to PE	0.064	0.041–0.093	HYPITAT[124]

a The distribution for mode of delivery was estimated from data reported for the PELICAN study.[5] In total, 36 prelabour caesarean deliveries were reported from a total of 164 caesarean deliveries in women presenting with suspected PE between 35 and 37 weeks of gestation. We inferred that the difference between these two values would represent spontaneous or induced labours that required emergency caesarean delivery. The distribution across modes of delivery (excluding planned caesarean sections) was calculated using the reported figures for spontaneous and assisted vaginal deliveries (54 and 13, respectively) and the inferred number of emergency caesarean deliveries (28).

Table 59 reports the probabilities of fetal death, neonatal intensive care and the length of stay in NICU if a neonate is admitted. These values were derived from the PELICAN study[5] and from the EPIPAGE study.[110]

Unit costs and for the diagnostic assessment tests, antenatal monitoring, hospitalisations and births are all reported in *Table 60*. Most costs were identified through reviewing CG107[13] for modelled values and then searching in current NHS reference costs,[127] Payment by Results Tariffs[128] and the *British National Formulary*[120] for appropriate updated values. Full description of the values chosen is available in *Cost of biomarker tests and antenatal management*.

TABLE 59 Parameters used in the EAG model: fetal and neonatal outcomes

Parameter	Value	Range for sensitivity analysis	Source
Pregnancies presenting up to 35 weeks of gestation, testing positive/rule-in for PE (delivering within 7–14 days)			
Antepartum/intrapartum fetal death	0.023	0.009–0.048	PELICAN study[5] (see *Table 28*)
In-hospital neonatal death	0.007	0.001–0.025	
Admission to neonatal intensive care	0.667	0.272–0.848	EPIPAGE[110] (see *Table 29*)
Duration of stay in neonatal intensive care (days)	8.46	6.60–10.31	
Pregnancies presenting up to 35 weeks of gestation, testing negative/rule-out for PE			
Antepartum/intrapartum fetal death	0.023	0.009–0.048	PELICAN study[5] (see *Table 28*)
In-hospital neonatal death	0.007	0.001–0.025	
Admission to neonatal intensive care	0.074	0.049–0.106	HYPITAT II[126]
Duration of stay in neonatal intensive care (days)	3.0	2.0–6.0	HYPITAT[124]

TABLE 59 Parameters used in the EAG model: fetal and neonatal outcomes (*continued*)

Parameter	Value	Range for sensitivity analysis	Source
Pregnancies presenting between 35 and 37 weeks of gestation, testing positive/rule-in for PE (delivering within 7 days)			
Antepartum/intrapartum fetal death	0.000		PELICAN study[5] (see *Table 30*)
In-hospital neonatal death	0.000		
Admission to neonatal intensive care	0.074	0.049–0.106	HYPITAT II[126]
Duration of stay in neonatal intensive care (days)	3	2–6	HYPITAT[124]
Pregnancies presenting between 35 and 37 weeks of gestation, testing negative/rule-out for PE			
Antepartum/intrapartum fetal death	0.000		PELICAN study[5] (see *Table 30*)
In-hospital neonatal death	0.000		
Admission to neonatal intensive care	0.037	0.020–0.063	HYPITAT II[126]
Duration of stay in neonatal intensive care (days)	3	2–6	HYPITAT[124]

TABLE 60 Parameters used in the EAG model: unit costs

Parameter	Value	Range for sensitivity analysis	Source
Standard diagnostic assessment cost	£0		Common to all diagnostic pathways (costs cancel out)
Triage PlGF test cost	£[a]		Company data
Elecsys sFlt-1 to PlGF ratio test cost	£[a]		Company data
Cost of pregnancy surveillance and management of gestational hypertension in suspected PE presenting up to 35 weeks of gestation – applies to all women testing negative/rule-out for PE			
Median time to delivery (days)			
PE	14	7–42	PELICAN study[5]
No PE	62	14–63	
Mild or no hypertension			
With PE (false negative)	£103.30		See *Tables 32* and *33*
Without PE (true negative)	£413.20		
Moderate hypertension			
With PE (false negative)	£225.08		See *Tables 32* and *33*
Without PE (true negative)	£900.32		
Severe hypertension			
With PE (false negative)	£867.08		See *Tables 32* and *33*
Without PE (true negative)	£1500.38		

continued

TABLE 60 Parameters used in the EAG model: unit costs (*continued*)

Parameter	Value	Range for sensitivity analysis	Source
Cost of PE management in suspected PE presenting up to 35 weeks of gestation – applies to all women testing positive/rule-in for PE (delivering within 7–14 days)			
Median time to delivery (days)	9	3–16	PELICAN study[5]
Mild or no hypertension			
With or without PE	£2315.95		See *Table 34*
Moderate hypertension			
With or without PE	£2322.94		See *Table 34*
Severe hypertension			
With or without PE	£2322.94		See *Table 34*
Cost of pregnancy surveillance and management of gestational hypertension in suspected PE presenting between 35 and 37 weeks of gestation – applies to all women testing negative/rule-out for PE			
Median time to delivery (days)			
PE	9	4–16	PELICAN study[5]
No PE	16	6–23	
Mild or no hypertension			
With PE (false negative)	£51.65		See *Tables 32* and *33*
Without PE (true negative)	£103.30		
Moderate hypertension			
With PE (false negative)	£112.54		See *Tables 32* and *33*
Without PE (true negative)	£218.09		
Severe hypertension			
With PE (false negative)	£754.54		See *Tables 32* and *33*
Without PE (true negative)	£860.09		
Cost of pregnancy surveillance and management of gestational hypertension in suspected PE presenting between 35 and 37 weeks of gestation – applies to all women testing positive/rule-in for PE (delivering within 7 days)			
Median time to delivery (days)	4	2–9	PELICAN study[5]
Mild or no hypertension			
With or without PE	£770.20		See *Table 34*
Moderate hypertension			
With or without PE	£777.19		See *Table 34*
Severe hypertension			
With or without PE	£777.19		See *Table 34*
Costs of delivery			
Spontaneous onset of labour			*NHS Payment by Results Tariff 2013/14*[127] (see *Table 35* for relevant HRG codes)
Normal delivery	£1506		
Assisted delivery	£1988		
Induced onset of labour			
Normal delivery	£2133		
Assisted delivery	£3033		

TABLE 60 Parameters used in the EAG model: unit costs (*continued*)

Parameter	Value	Range for sensitivity analysis	Source
Caesarean delivery			
Planned	£3182		
Emergency	£4013		
Costs of critical care			
Maternal critical care	£1449		*NHS Payment by Results Tariff 2013/14*[127] (see *Table 35* for relevant HRG codes)
NICU	£978.50		

a Data on test costs are confidential and are not presented here.

The chosen inputs for QALYs from *Table 38* are reported in *Table 61*. Full details of the process of identifying HRQoL data are reported in *Chapter 5*, *Review of HRQoL studies*. The full set of mapped EQ-5D utilities are reported in *Tables 36* and *37*. Neonatal HRQoL was not modelled because of lack of data. The full rationale for HRQoL parameter choices is provided in *Chapter 5*, *Derivation of utility estimates from health-related HRQoL*.

TABLE 61 Parameters used in the EAG model: QALYs

Parameter	Value	Source
Baseline QALYs from (vaginal) delivery to 6 months post partum		See *Table 38*
Birth to 3 weeks post partum	0.0389	Jansen *et al.*[116]
3 weeks to 12 weeks post partum	0.1496	Bijlenga *et al.*[112]
12 weeks to 6 months post partum	0.2171	Bijlenga *et al.*[112]
Decrement for caesarean delivery (birth to 3 weeks post partum)		See *Table 38*
Non-emergency caesarean section	0.0050	Jansen *et al.*[116]
Emergency caesarean section	0.0092	Jansen *et al.*[116]
Decrement for non-spontaneous delivery		See *Table 38*
3 weeks to 6 months post partum	0.0084	Petrou *et al.*[119]

Appendix 9 List of model assumptions

The table presents a list of the key model assumptions, with a justification for each of these assumptions.

Section in report	Assumption	Justification
Chapter 5, Description of the decision analytical model	The model assumes that UK guidelines for management of suspected PE, gestational hypertension and PE are followed	CG107[13] is the most appropriate guideline for NHS management of suspected PE
	The model assumes that women presenting with PE before 35 weeks of gestation will be managed using expectant monitoring and women presenting from 35 weeks of gestation to 37 weeks of gestation will be managed using the immediate delivery strategy	Assumption in line with time to delivery in the PELICAN study[5] and in line with CG107[13] management guidance
Chapter 5, Event probabilities	The Triage PlGF test has sensitivity and specificity values for detecting PE requiring delivery within 14 days, while the Elecsys sFlt-1 to PlGF ratio test has sensitivity and specificity for detecting PE within 4 weeks (irrespective of delivery time). Without an assumption of comparability, no comparison of these interventions could be made to each other	The assumption is justified by the similarity of sensitivity and specificity values for the Triage PlGF test using a < 100 pg/ml cut-off value for PE requiring delivery within 14 days in *Table 13* and the < 100 pg/ml cut-off value for diagnosis for any preterm PE in *Table 15*
Chapter 5, Cost of biomarker tests and antenatal management	The EAG has assumed that costs of NICU stay capture the effects of neonatal morbidity for deliveries occurring between 35 and 37 weeks of gestation (the study population)	Low rates of adverse outcomes, and expert opinion supported this decision
	The EAG has assumed a specific cost[a] for the Triage PlGF test, and a specific cost[a] for the Elecsys sFlt-1 to PlGF ratio test	These values are each the higher of the two values presented by the test manufacturers. Choosing the higher values is a conservative assumption
	The base-case analysis assumes that tests are conducted in a central laboratory	This is in line with the data submitted by both Alere and Roche Diagnostics
	The cost of proteinuria dipstick testing and blood pressure measurement is subsumed in the cost of a standard antenatal appointment	NHS costs are done by attendances and procedure codes; given that these procedures are included in standard antenatal management, they should be part of the attendance cost
	The cost of blood pressure monitoring and quantitative proteinuria testing are subsumed in the cost of hospitalisation	As above
	The length of stay for women with severe gestational hypertension is assumed to be 3 days	NICE data on short-stay patients for routine antenatal monitoring indicated a length of stay of 1 day. Experts indicated that this was shorter than seen in practice and the length of stay was modified accordingly
	Women managed on the gestational hypertension pathway were assumed to receive two oral labetalol prescriptions	The standard number of tablets in a package (56) is insufficient to last 8 weeks. Experts indicated that women would continue taking medication after blood pressure had stabilised

Section in report	Assumption	Justification
Chapter 5, Cost of biomarker tests and antenatal management	Women managed on the PE pathway receive one prescription for oral labetalol	The shorter length of monitoring will not use all 56 tablets
	The EAG model assumes that the unit costs associated with birth are not dependent on whether or not the mother has hypertension or PE	Experts indicated that births to women with PE were no more complicated than average births
Chapter 5, Derivation of utility estimates from health-related quality of life	Utility scores for birth were assumed to last for 3 weeks	This reflects that the first period after any birth is the most difficult, but that they begin improving on the way to 6 weeks post partum
	Utility scores and decrements for other periods had assumed lengths of exposure	This assumption was intended to reflect a gradual improvement in quality of life over time, from delivery, in the absence of all data having perfectly matched dates
	All utilities are assumed constant over the period in which they occur	This is a standard modelling assumption
	Differences in utility scores for diagnostic tests are determined by differences in modes of delivery	Assumption based on best-available utility evidence. No data to make assumptions on long-term outcomes for mother or neonate

a Data on test costs are confidential and are not presented here.

EME
HS&DR
HTA
PGfAR
PHR

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

Published by the NIHR Journals Library