HTA Reference No. 17/10/01 Final version, July 17 2017

Biomarker tests to help diagnose preterm labour in women with intact membranes

Name of the External Assessment Group (EAG) and project lead

Peninsula Technology Assessment Group (PenTAG), University of Exeter Medical School

Jo Varley-Campbell

Research Fellow

South Cloisters, St. Luke's Campus, Heavitree Road, Exeter, EX1 2LU, United Kingdom

Tel: 01392 725706

Email address: j.varley-campbell@exeter.ac.uk

Plain English summary

Preterm (premature) birth, or the delivery of a baby before 37 weeks, is the main cause of death or illness in newborns. In England and Wales in 2013, the infant mortality rate for babies born preterm was 21.1 deaths per 1,000 live preterm births, which is significantly lower than for babies born at term or posterm (1.4 infant deaths per 1,000 live births).[1] Preterm labour may be caused by a combination of different things including infection, overstretching of the uterus, cervical weakness and placental problems. Other causes include diabetes, high blood pressure, drug use and smoking.[2, 3] Long-term physical and neuro-developmental issues are not uncommon for babies that survive being born early.

Being able to identify women who are likely to have a preterm birth allows doctors to start interventions that may help improve the survival of the baby and mother. Interventions that may delay or stop labour include tocolytics (anti-contraction medicine), cervical stitch, progesterone therapy (hormones) and the insertion of an Arabin pessary (cone-shaped device that supports the cervix). Whilst interventions that may help the survival of the baby include a transfer to a specialist hospital whilst the baby is still in the womb and corticosteroids and magnesium sulphate (medicine given to the mother to help the baby).

Current tests used in clinical practice to predict whether a woman showing signs of labour is likely to deliver preterm include transvaginal ultrasound measurement of cervical length, fetal fibronectin, Actim Partus and clinical assessment. PartoSure is a new test which also claims to be able to predict preterm birth. This work will look at the evidence for how accurate the biomarker tests (PartoSure, Actim Partus and fetal fibronectin) are for predicting preterm birth and how effective and cost-effective these tests would be if used in the NHS.

1 Background

1.1 Preterm Labour

Preterm, as defined by the World Health Organisation (WHO), are babies born alive before 37^{+0} weeks.[4]

Preterm birth can lead to short-term health problems for a newborn baby, such as difficulties with breathing and feeding. Preterm birth may also result in long-term issues, such as learning disabilities and behavioural problems.[5]

1.1.1 Epidemiology

Of all births in England and Wales in 2013, 7.3% were preterm, 89.3% were term and 3.4% were post-term.[1]

The WHO sub-categorises preterm birth based on gestational age as:[4]

- extremely preterm: <28 weeks
- very preterm: 28 weeks to <32 weeks
- moderate to late preterm: 32 weeks to <37 weeks

Of the 7.3% of babies born preterm in England and Wales in 2013, 4.6% were extremely pre-term, 10.8% were very preterm and 84.6% were moderate to late preterm.[1]

Worldwide, it is estimated that there are 15 million preterm births every year.[4] More specifically in the UK, in 2013 there were 95,960 preterm live births (under 22 weeks to 37 weeks).[1]

In England and Wales, there is a significantly lower infant mortality rate for babies born at full-term (1.4 deaths per 1,000 live births) compared to those born preterm (21.1 deaths per 1,000 live births).[1] In babies born preterm in 2013, 99% resulted in an infant death when the baby was born at less than 22 weeks and under 1000g, while at 22 weeks gestation 90% of preterm births resulted in an infant death. By 24 weeks gestation, infant deaths occurred in 41% of preterm births and by 27 weeks gestation infant death occurred in only 11%.[1] Survival rates improve as gestation increasing with only 1% of babies born at 34 weeks gestation resulting in an infant death.

1.1.2 Risk factors

There are various risk factors that may increase the chance of preterm labour. The most prevalent risk factor is previous preterm birth.[2, 6] Other risk factors include: previous late miscarriage, previous excisional cervical treatment, uterine abnormalities, black ethnicity, maternal smoking, vaginal infections, low body mass index (BMI), multiple gestations and low socioeconomic status.[2] Maternal demographics associated with preterm birth include: high levels of stress, low socioeconomic status, extremes of maternal age, and single marital status.[2]

1.1.3 Management

For women presenting with symptoms of preterm labour, baseline clinical investigations such as digital cervical examination, speculum examination, biochemical tests or ultrasound may help to distinguish women who are in preterm labour from those who are not.[7, 8]

Symptoms of suspected preterm labour include: painful contractions or cramps, abdominal and low back pain, and increased vaginal discharge.[7] Symptoms do not always result in progression to established labour and birth; they may occur but then settle allowing the pregnancy to continue towards term. It is understood that over 90% of women presenting with symptoms of preterm labour do not go on to deliver in the next two weeks and of these 50% will continue with pregnancy until full term.[2, 9] It is important to determine whether preterm labour is the cause of the symptoms and assess the risk of preterm delivery, to allow appropriate management to begin as soon as possible.[7]

Women presenting with symptoms of preterm labour and intact membranes will be diagnosed as being in threatened preterm labour if one or more of the following assessments are positive:

- Fetal fibronectin ≥50 ng/ml
    - Current NICE guidelines recommend the fetal fibronectin test used qualitatively with a threshold of 50 ng/ml. However it is understood in clinical practice the quantitative fetal fibronectin (Rapid fFN 10Q Cassette Kit) test is commonly used (different thresholds other than 50 ng/ml may be used with the quantitative test).
- Transvaginal ultrasound cervical length <15mm
    - Cervical length is understood to be infrequently used for women presenting with symptoms of threatened preterm labour as it requires specialist equipment and skilled personnel, not routinely available.

See Figure 1, in Appendix 1 for the NICE guidance care pathway.

Typical management depending on the test results include[5]:

- If test results are negative and the symptoms of preterm labour have settled, the woman would be discharged home with routine follow up in community and advised to return if symptoms reappear.
- If test results are negative but symptoms of preterm labour continue, the woman could be admitted for observation or discharged home. If symptoms were managed successfully, the woman would be discharged home.
- If the test results are positive, the woman would be admitted, treated as appropriate and monitored.

Typical treatments for preterm labour include tocolytics to delay preterm delivery and strategies to improve the outcome of the preterm baby, typically corticosteroids. There is, however, variation in practice and little agreement about whether attempting to delay delivery improves outcomes.[7]

1.2 Diagnostic tests for preterm labour

1.2.1 PartoSure

PartoSure (Parsagen Diagnostics Inc.) is a CE marked point-of-care diagnostic test that detects placental alpha microglobulin-1 (PAMG-1) in the vaginal secretions of pregnant women. PAMG-1 (placental alpha microglobulin-1) is a protein released from decidual cells into the amniotic cavity throughout pregnancy which is present in cervicovaginal discharge when labour and delivery are imminent.[7] It is found in very high concentrations in amniotic fluid and in very low concentrations in normal vaginal discharge. Studies have demonstrated a strong correlation between the presence of PAMG-1 in cervico-vaginal discharge and imminent delivery.[10]

PartoSure is designed to be used in conjunction with clinical assessment to assess the risk of preterm delivery in ≤7 or ≤14 days from the time of sample collection. A positive PartoSure test result indicates that delivery is likely to occur within 7 days, whereas a negative test result suggests that delivery is not likely within the next 7 to 14 days. Clinicians are able to use the results of the test to determine whether a woman can be sent safely home or whether they should be admitted to hospital for treatments to either delay birth and/or improve neonatal outcomes. The notified indication is in pregnant women with symptoms of preterm labour who have intact membranes, minimal cervical dilation (≤3 cm) and who are between 20^{+0} weeks and 36^{+6} weeks gestation.[7]

In the PartoSure test kit, there is a test strip, a vaginal swab and a plastic vial containing a solvent solution. A health care professional will insert the sterile flocked swab 5–7 cm into the vagina and will remove after 30 seconds, collecting a sample of vaginal discharge. The specimen is then rotated for 30 seconds in the solvent solution in the vial. A test strip is finally added to the vial and left for 5 minutes or until 2 lines are visible (indicating a positive test result and a high risk of delivery within 7 days). If only 1 line is visible, this indicates a negative test result and a low risk of delivery within 7 to 14 days, whilst no lines indicate a failed/invalid test.[5]

PartoSure is designed to be used, in conjunction with clinical assessment, by healthcare professionals. The test is not affected by the presence of vaginal infections, urine, semen and trace amounts of blood, but may not be reliable in the presence of significant discharge of blood, meconium, anti-fungal creams, suppositories, lubricants, moisturisers, talcum powder or baby oil.[7] The test has a limit of detection of 1 ng/ml and a measuring range of 1 to 40,000 ng/ml.[5]

1.2.2 Actim Partus

Phosphorylated insulin-like growth factor binding protein-1 (IGFBP-1) is a protein made by the cells lining the uterus. Phosphorylated IGFBP-1 leaks in small amounts into the cervix when delivery is imminent. Actim Partus is designed to detect phosphorylated IGFBP-1 in the cervical secretions using monoclonal antibodies. It is a CE marked qualitative immunochromatographic test.[7]

In the Actim Partus test kit, there is a sterile polyester swab for specimen collection, a tube of specimen extraction buffer and a dipstick in a sealed foil pouch (no other instruments or consumables are needed). A health care professional collects a sample from the cervical os

using the swab during a sterile speculum examination. The swab is swirled vigorously in the extraction solution for 10-15 seconds. The dipstick is finally added to the extracted sample until the liquid front enters the results area and left for 5 minutes or until 2 lines are visible (indicating a positive test result and a high risk of delivery within 7 days). If only 1 line is visible, this indicates a negative test result and a low risk of delivery within 7 to 14 days, whilst no lines indicate a failed/invalid test.[5]

The Actim Partus test is indicated for use in pregnant women with signs and symptoms of preterm labour and intact amniotic membranes. The test has a limit of detection of 10 ng/ml and a measuring range of 10 to 8000 ng/ml.[5]

1.2.3 Fetal Fibronectin

The fetal fibronectin test is available as a quantitative enzyme-linked immunosorbent assay (ELISA) that is performed in a laboratory, and as a qualitative membrane immunosorbent assay that can be performed near the patient. Examples include:

- Rapid fFN 10Q Cassette Kit; a quantitative laboratory based assay (Hologic)
- Fetal Fibronectin Enzyme Immunoassay; a qualitative laboratory based assay (Hologic)
- Rapid fFN for the TLiIQ System; a qualitative immuno-chromatographic assay with a limit of detection of 50 ng/ml (Hologic)
- QuikCheck Fetal Fibronectin Test; a qualitative test with a limit of detection of 50 ng/ml (Hologic).

The report will assess the clinical effectiveness, cost effectiveness and diagnostic test accuracy of the Rapid fFN 10Q Cassette Kit used at threshold other than 50ng/ml. All other fetal fibronectin tests (used qualitatively with a threshold of 50ng/ml) are comparators.

Rapid fFN 10Q Cassette Kit test is a diagnostic immunochromatographic dipstick test that detects fibronectin, a protein released by fetal cells into the cervix or vagina due to disruption of the fetal membrane during birth. Fetal fibronectin is an adhesive glycoprotein that holds the membranes of the uterus to the fetal membranes. After the 35th week of pregnancy, it begins to break down naturally, and is detectable in vaginal secretions. Elevated levels of fetal fibronectin detected in cervicovaginal secretions (50 ng/ml or more) between 24^{+0} and 34^{+6} weeks of pregnancy is an indicator of preterm birth risk.[7]

Fetal fibronectin testing, to determine likelihood of birth within 48 hours for women who are 30^{+0} weeks pregnant or more, is recommended in NICE's guideline on preterm labour and birth, if transvaginal ultrasound measurement of cervical length is indicated but is not available or not acceptable. The guideline notes that if a swab for fetal fibronectin testing is anticipated, the swab should be taken before any digital vaginal examination.

When using the Rapid fFN 10Q Cassette Kit, the following equipment is required in addition to the test kit: an analyser, printer, Rapid fFN Control Kit, QCette (quality control device) and pipette. A health care professional collects a sample from the posterior fornix of the vagina in the specimen transport tube during a sterile speculum examination. The patient sample is pipetted into the sample application well once the Rapid fFN 10Q Cassette is inserted into the analyser. The analysis takes 2 to 3 minutes once the 7 minute incubation period has

been completed. Concentrations of fetal fibronectin reported by the analyser range from 0 to 500 ng/ml, where concentrations over 500ng/ml are displayed as >500ng/ml. Suggested thresholds are not provided by the manufacture in their instructions, therefore these are set by the hospital and consequently result in some variability between hospitals. With every test, an internal control is performed automatically.[5]

NICE's guideline on pre-term labour and birth recommends a result of ≥50 ng/ml indicates a positive fetal fibronectin result.

1.3 Current Guidelines

1.3.1 NICE guidelines

NICE's guideline on preterm labour and birth states that women reporting symptoms of preterm labour who have intact membranes should have a clinical assessment which includes:[11, 12]

- clinical history taking
- observations of the woman, including: the length, strength and frequency of her contractions; any pain she is experiencing; pulse, blood pressure and temperature; and urinalysis
- observations of the unborn baby, including: asking about the baby's movements in the last 24 hours; palpation of the woman's abdomen to determine the fundal height, the baby's lie, presentation, position, engagement of the presenting part, and frequency and duration of contractions; and auscultation of the fetal heart rate for a minimum of 1 minute immediately after a contraction
- a speculum examination (followed by a digital vaginal examination if the extent of cervical dilatation cannot be assessed).

If the clinical assessment suggests that the woman is in suspected preterm labour and she is 29^{+6} weeks pregnant or less, treatment for preterm labour is recommended.[11, 12]

If the clinical assessment suggests that the woman is in suspected preterm labour and she is 30^{+0} weeks pregnant or more then the following tests should be conducted:[11, 12]

- Transvaginal ultrasound measurement of cervical length (as a diagnostic test to determine likelihood of birth within 48 hours).
    - If cervical length is >15 mm, the woman is unlikely to be in preterm labour and could be discharged home with routine follow up in community and advised to return if symptoms reappear.
    - If cervical length is <15 mm, the woman is diagnosed as being in preterm labour and should be offered treatment.
- If transvaginal ultrasound measurement of cervical length is indicated but is not available or not acceptable then fetal fibronectin testing as a diagnostic test may be used for women who are 30+0 weeks pregnant or more.

- If the fetal fibronectin testing is negative (concentration 50 ng/ml or less), the woman is unlikely to be in preterm labour and could be discharged home with routine follow up in community and advised to return if symptoms reappear.
- If the fetal fibronectin testing is positive (concentration more than 50 ng/ml), the woman is diagnosed as being in preterm labour and should be offered treatment

- It is not recommended to use transvaginal ultrasound measurement of cervical length and fetal fibronectin testing in combination to diagnose preterm labour.

See Figure 1, in Appendix 1 for the NICE guidance care pathway.

2 Decision problem

2.1 Purpose of the decision to be made

To evaluate the effectiveness and cost-effectiveness of PartoSure, Actim Partus and Rapid fFN 10Q Cassette Kit (quantitative test) for the diagnosis of preterm labour (≤36^{+6} weeks of gestation) in symptomatic women with intact amniotic membranes and for whom transvaginal ultrasound is not available or acceptable.

2.2 Clear definition of the interventions

PartoSure is a point-of-care qualitative test that detects PAMG-1 in the vaginal secretions of pregnant women. It is manufactured by Parsagen Diagnostics Inc. and distributed in the UK by Advanced Global Health.

Actim Partus is a point-of-care qualitative test that detects insulin-like growth factor binding protein-1 in the cervical secretions of pregnant women. It is manufactured by Medix Biochemica and distributed in the UK by Alere.

Rapid fFN 10Q Cassette Kit (assessed at thresholds other than 50ng/ml) is a point-of-care test that detects fetal fibronectin. It is manufactured by Hologic.

2.3 Populations and relevant subgroups

The intervention will be applied to pregnant women who are preterm (≤ 36^{+6} weeks of pregnancy), presenting with symptoms of labour (e.g. regular uterine activity, with a contraction frequency of at least 1:10)[13] with intact amniotic membranes and for whom transvaginal ultrasound is not available or acceptable. If evidence permits, sub-groups based on gestational age groups will be assessed.

2.4 Place of the intervention in the treatment pathway(s)

Tests and other clinical information can be used to predict whether a woman presenting with symptoms of preterm labour will actually deliver or whether the symptoms will subside. At present, women presenting with symptoms of preterm labour and intact membranes will be diagnosed as being in threatened preterm labour if a transvaginal ultrasound cervical length is <15mm. When transvaginal ultrasound cervical length is unavailable or unacceptable, the fetal fibronectin test with a threshold of ≥50 ng/ml is the recommended test. It is at this point in the treatment pathway, the interventions are to be assessed.

2.5 Relevant comparators

Interventions will be compared against each other and against the following comparators:

- Fetal fibronectin used qualitatively with a threshold of 50 ng/ml, with clinical assessment of symptoms
- Clinical assessment alone

2.6 Key factors to be addressed

2.6.1 Clinical outcomes

Key Question: How effective are PartoSure, Actim Partus and Rapid fFN 10Q Cassette Kit (quantitative fetal fibronectin) at predicting preterm birth? How do the results of these tests affect clinical outcomes including hospital admissions, maternal and neonatal morbidity and mortality, long term health problems of the child and health related quality of life for both the mother and child?

2.6.2 Cost outcomes

Key Question: Is testing for preterm labour using PartoSure, Actim Partus or Rapid fFN 10Q Cassette Kit (quantitative fetal fibronectin) an effective use of NHS resources?

2.6.3 Key challenges (further considerations/problematic factors)

- Understanding how the results from the quantitative fetal fibronectin test are used and whether there are discrete bands which result in the quantitative test being used qualitatively. Interpreting this challenge is likely to be guided by how the literature has reported its findings.
- Assessing the accuracy of the tests by gestational subgroups will also be challenging. How gestational age is grouped will be determined by the literature available, unless individual patient data is made available. If no data are available on test accuracy by gestational age, subgroup analysis will be performed by gestational age ranges relevant to variation in neonatal mortality and morbidity risks contingent on diagnostic outcomes (true and false positives and true and false negatives), assuming the same diagnostic test accuracy. These analyses will also allow for variations in preterm labour rates by gestational age.
- Interpreting the false positive data will need to be considered with some caution, since treatment may be given to delay preterm delivery thus impacting the number of false positive results. This will confound the test accuracy results.

3 Report methods for assessing the outcomes arising from the use of the interventions

From our scoping work, it has been identified that there is a paucity of literature surrounding PartoSure and Actim Partus. Therefore the three systematic reviews (for diagnostic test accuracy, clinical and cost-effectiveness) of these tests will be conducted using similar methods but will differ in their analysis of the data (see section 3.6). In addition to these reviews, the health economic model (section 4, page 15) will be used to extrapolate outcomes.

3.1 Inclusion criteria

3.1.1 Population

Women with signs and symptoms of preterm labour, with intact amniotic membranes and who are not in established labour. Studies need to identify their population as 'preterm' or be clearly including only preterm women if the term is not used. The number of weeks of gestation will not be a limiting factor for inclusion. Women must only be expecting a singleton pregnancy.

3.1.1.1 Details of sub-groups to be examined, if any

Subgroups based on the number of weeks of gestation will be examined. The bracketing used to denote these groups will be led primarily by how the literature has defined the subgroups. If individual patient data is provided from included studies, this may be used in an assessment of different subgroups.

3.1.2 Interventions

- PartoSure (with or without an assessment of clinical symptoms)
- Actim Partus (with or without an assessment of clinical symptoms)
- Rapid fFN 10Q Cassette Kit (or quantitative version of fetal fibronectin), used with a threshold other than 50 ng/ml, (with or without an assessment of clinical symptoms)

3.1.3 Comparators

- Fetal fibronectin (or a qualitative version of fetal fibronectin) used with a threshold of 50 ng/ml, (with or without an assessment of clinical symptoms)
- Clinical assessment of symptoms alone
- One of the other interventions as in 3.1.2 above
- For the review of diagnostic test accuracy studies, the interventions and comparators may be assessed against each other, and/or against a reference standard (e.g. the onset/occurrence of preterm labour)

3.1.4 Outcomes

For the review of diagnostic test accuracy, the outcomes to be assessed for interventions and comparators are:

- Diagnostic accuracy (sensitivity and specificity), for preterm birth within 48 hours and 7 days
- Time to test result
- Test failure rate

For the review of clinical effectiveness, the outcomes to be assessed for interventions and comparators are:

- Perinatal mortality
- Neonatal morbidity and mortality
- Long-term health problems in the child
- Maternal morbidity and mortality
- Health-related quality of life
- Anxiety associated with confidence in the test results
- Number of women admitted to hospital
- Number of re-presentations to hospital within 48 hours and 7 days
- Number of women who have tocolytics / corticosteroids
- Length of inpatient hospital stay
- Number of transfers of pregnant women and neonates between hospitals
- Time to delivery from presentation
- Number of women treated with maternal corticosteroids appropriately (that is, they deliver within 7 days following treatment)
- Number of women treated with maternal corticosteroids inappropriately (that is, they do not deliver within 7 days following treatment)
- Impact on neonatal intensive care resource planning
- Gestational age at birth

For the review of cost effectiveness, the outcomes to be assessed for interventions and comparators are:

- The cost and/or cost-effectiveness of an intervention or comparator (also including studies where health outcomes are valued in natural units, e.g. cost per preterm delivery, or monetary terms, i.e. cost-benefit analysis).

3.2 Study design

For the test accuracy review, studies will compare the outcomes of an index test (an intervention or comparator in the review) to the outcomes of a reference standard in patients receiving both the index tests and the reference standard. The timing of the index test will be in line with current or potential clinical practice, while the reference standard may incorporate information not known at the time of administering the index test (e.g., the gestational age at birth). Studies will be eligible for inclusion whether or not the index test results were used in the clinical management of patients. (The role and impact of index test results on the validity of test accuracy outcomes will be assessed as part of the quality assessment, section 3.5, p.13).

For the review of clinical effectiveness, studies will report clinical outcomes for patients who have received one or more of the interventions or comparators in line with current or potential practice. The review will include controlled studies with and without a randomised design, and will include studies with prospectively and retrospectively identified populations. Studies with historical controls will be eligible for inclusion. If there is insufficient evidence within these study designs, we may widen our inclusions.

For the review of economic evaluations, studies can be based on trials, other data sources (e.g., registries and patient records), decision models, or systematic reviews of existing economic evaluations. If set in the NHS, studies must report costs and/or resource use. If not set in the NHS, studies must report incremental costs and/or resource use, as well as incremental effectiveness outcomes (which may be presented in: natural units, e.g., preterm births avoided; utility-based measures, e.g., QALYs; or, monetary value). Studies not reporting incremental outcomes but reporting sufficient information for these to be calculated will also be included.

3.3 Search strategy

The search strategy will not be limited by date, language, study design literature search filters or by human population. A sample search strategy is as follows:

PartoSure Search

(PartoSure or Parto Sure or PartoSureR or Parto SureR).ti,ab,kw.

(Placental alpha adj5 test$).ti,ab,kw.

Actim Partus Search

MEDLINE	(Actim Partus or Actim PartusR phIGFBP-1).ti,ab,kw.
EMBASE	(Actim Partus or Actim PartusR phIGFBP-1).ti,ab,kw. or *actim partus test/

Fetal Fibronectin Search

((Fetal adj3 (fibronectin$)) or fFN).ti,ab,kw.

The following bibliographic databases will be searched:

- MEDLINE via OVID;
- EMBASE via OVID;
- The Cochrane Library (DARE, CENTRAL, CDSR, HTA and NHS EEDs);
- BIOSIS via Thomson Reuters;
- Web of Science via Thomson Reuters; and
- CINAHL via EBSCO Host.

The following supplementary search methods will be undertaken:

- Studies identified by Specialist Committee Members or included in company submissions;
- The Cochrane Register of Diagnostic Test Accuracy Studies (CRDTAS) will be searched;
- Studies included at full-text will be citation chased and study authors will be contacted to identify unpublished studies; and
- Web-searching will be undertaken to identify unpublished studies.

Computer assisted deduplication will be performed, following which studies will be screened according to title and abstract for potential inclusion. For the systematic review of diagnostic test accuracy, the screening will be done independently by at least two reviewers. Any disagreements will be resolved by discussion, with the involvement of a third reviewer if necessary.

Full texts will be sought for any studies not excluded by title and abstract screening. These full texts will be screened independently by at least two reviewers for ultimate inclusion in the review. Any disagreements will be resolved by discussion, with the involvement of a third reviewer if necessary.

3.4 Data extraction strategy

Data will be extracted by one reviewer and checked by a second reviewer. Any disagreements will be resolved by discussion, with the involvement of a third reviewer if necessary. Standardised data extraction tables will be developed for all three reviews.

3.5 Quality assessment strategy

Diagnostic test accuracy studies will be quality assessed using the QUADAS-2 tool.[14] Quality appraisal of the clinical effectiveness evidence will be assessed using the Cochrane Risk of Bias (if randomised controlled trials) or alternative bias tool. The quality appraisal of the cost effectiveness literature will be conducted using the Drummond checklist.[15]

All quality appraisal assessments will be conducted by one reviewer and checked by another, with disagreements resolved by discussion, with the involvement of a third reviewer if necessary.

3.6 Methods of analysis/synthesis

Where outcomes of interest are not directly reported in studies but can be reliably imputed from information provided, imputed values will be calculated and presented.

Test accuracy review:

For the review of test accuracy, studies which provide estimates of both sensitivity and specificity will have their point estimates plotted in ROC space. If appropriate, these estimates may be synthesised following the general principles described in the Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy.[16]

Clinical effectiveness review:

Narrative synthesis will be performed, supported by tabulation of study characteristics and results. Meta-analysis will be considered as appropriate. If any studies are known to have overlapping participants (e.g., because they recruit from the same registry) this will be highlighted and accounted for where possible in any data synthesis.

Cost effectiveness review:

Where studies do not conduct a fully incremental cost-effectiveness analysis (e.g., if they perform a cost–consequences analysis), but it is possible to conduct such an analysis based on reported results, this will be done.

Currency conversion will not be performed, but an indication will be given of purchasing-power-parity exchange rates, and if currency- or country-specific cost-effectiveness thresholds are supplied by the authors these will also be reported (in the original currency).

Narrative synthesis will be performed, supported by tabulation of study characteristics and results.

4 Report methods for synthesising evidence of cost-effectiveness

Evidence pertaining to the cost-effectiveness of PartoSure, Actim Partus and Rapid fFN 10Q Cassette Kit (quantitative fetal fibronectin) will include a systematic review of economic evaluations and the development of an economic model. The systematic review of economic evaluations will be broad and capable of identifying studies answering the present decision problem.

An economic model will be developed, including all diagnostic strategies relevant to the decision problem and current or potential clinical practice. It will be populated with data on diagnostic test performance based on the review of test accuracy, linked to data on clinical outcomes and the costs of tests, procedures, treatments and other relevant healthcare costs.

4.1 Identifying and systematically reviewing published cost-effectiveness studies

As mentioned in section 3, scoping of the literature has indicated that evidence on this topic area is limited. With this in mind, we will not run a separate search for cost-effectiveness literature, but studies will be identified from screening the search results from the diagnostic test accuracy and clinical effectiveness review.

4.1.1 Population, intervention, comparators, outcomes and study designs

As described in sections 3.1 and 3.2.

4.1.2 Search strategy

As described in section 3.3

4.1.3 Data extraction strategy

As described in section 3.4

4.1.4 Quality assessment strategy

As described in section 3.5

4.1.5 Methods of analysis/synthesis

As described in section 3.6.

4.2 Development of a health economic model

The health economic model previously developed by NICE to inform the guidance on diagnosing preterm labour for women with intact membranes, NICE Guideline [NG25][11] is likely to form the basis of the new economic evaluation. This model was based on a comprehensive review of the evidence on diagnostic accuracy and treatment effectiveness, and modern methods of evidence synthesis including data from indirect comparisons. We will seek to validate this model using individual mother and infant data from the South West of England.

The model will be updated with the following key changes:

- Interventions
  - Evaluation of PartoSure
  - Evaluation of Actim Partus
  - Evaluation of quantitative fetal fibronectin test (i.e. at varying diagnostic thresholds)
- Patient population
  - Expanded to include pregnant women between 20^{+0} and 36^{+6} weeks suspected of preterm labour with intact membranes
- Costs
  - Allow variability in costs of neonatal care by level of care using national data (BadgerNET Neonatal Electronic Patient Records)[17]
- Outcomes
  - Account for the effect on neonatal mortality due to in utero transfer to tertiary centres at <28 weeks gestational age group

4.2.1 Model characteristics

The economic model will adhere to the NICE Diagnostic Assessment Programme Manual,[18] specifically:

- Cost-effectiveness results will be presented as incremental cost-effectiveness ratios of incremental costs to incremental QALYs;
- An infant lifetime time horizon will be used (to age 100 years);
- Costs will be included from a NHS and Personal Social Services perspective;
- Health effects on infants (incidence of preterm birth, impact of inappropriate corticosteroid treatment, neonatal morbidity and mortality and long term consequences of preterm birth) and mothers (anxiety associated with diagnostic test results or in utero transfers) will be included;
- A discount rate of 3.5% will be used for costs and QALYs.

4.2.2 Data sources

Diagnostic performance aspects of the model will be parameterised using the results of the systematic review of diagnostic test accuracy or clinical opinion if no suitable published estimates are obtained. Key parameters will also be identified through structured reviews, to ensure that they are not identified serendipitously, opportunistically or preferentially.[19] If these structured reviews identify existing recent systematic reviews, then these may be utilised without further searching.

4.2.3 Model structure

The model will comprise of two components:

- Diagnostic submodel (a decision tree);
- Treatment submodel (a decision tree).

Diagnostic submodel

The diagnostic submodel will produce the expected diagnostic yield of the different diagnostic strategies in terms of false positive (FP), false negative (FN), true positive (TP) and true negative (TN) outcomes, given a preterm labour prevalence rate typically observed in routine practice, which increases with gestational age. Diagnostic test costs will be measured, including those of test kits, staff time to perform test and undergo training. The strategies considered are as follows:

- Rapid fFN 10Q Cassette Kit (quantitative fetal fibronectin, using thresholds other than 50 ng/ml) with clinical assessment
- PartoSure test with clinical assessment
- Actim Partus test with clinical assessment
- Qualitative fetal fibronectin (set at 50 ng/ml threshold) with clinical assessment
- Treat all after clinical assessment alone

Treatment submodel

Conditional on the outcomes of the diagnostic model phase, the treatments and their clinical and cost consequences will be determined, by gestational age. Treatments will include:

- Corticosteroid
- Tocolysis
- Magnesium sulphate

Among the cost consequences to be measured are those of in utero transfers from local maternity units to hospitals with specialised neonatal units, hospital stays and treatments received by mothers, and the costs of neonatal care received, including inpatient hospitalisation, and long term healthcare costs of infants.

The effects of treatments will be measured in terms of mortality (stillbirth and neonatal) and morbidity due to respiratory distress syndrome (RDS) and intraventricular haemorrhage (IVH). Mortality will be divided between that related to RDS and IVH in order to estimate the long term morbidity impact for those infants who survive the neonatal hospitalisation period. Data from the literature on long term costs and QALYs of long term infant morbidity will be applied to these outcomes and added to diagnostic test, treatment and neonatal hospitalisation costs and maternal costs and utilities to estimate total costs and QALYs of diagnostic strategies and the incremental cost-effectiveness ratio (ICER) of PartoSure, Actim Partus and quantitative fetal fibronectin. A systematic search of the literature of utility values will be undertaken to assess the long-term effects of neonatal outcomes.

4.2.4 Exploration of uncertainty

Uncertainty in the cost-effectiveness of different diagnostic strategies will be explored through one-way sensitivity analyses and scenario analyses (in which alternative parameter values for one or more parameters are substituted).

Due to the inherently high level of uncertainty associated with predicting long term outcomes from neonatal outcomes, a limited probabilistic sensitivity analysis considering only uncertainty in diagnostic performance and immediate neonatal treatment outcomes will be conducted, using recent methodology.[20]

5 Handling information from the companies

All data submitted by the company(s) will be considered if received by the EAG no later than 1st October 2017. Data arriving after this date may not be considered. If the data meet the inclusion criteria for the review, they will be extracted and quality assessed in accordance with the procedures outlined in this protocol.

Any 'commercial in confidence' data provided by a company and specified as such will be highlighted in blue and underlined in the assessment report (followed by an indication of the relevant company name e.g. in brackets). Any 'academic in confidence' data provided and specified as such will similarly be highlighted in yellow and underlined.

6 Competing interests of authors

All authors confirm that they have no potential competing interests.

7 Timetable/milestones

Milestone	Date to be completed
Draft protocol	21st June 2017
Final protocol	17th July 2017
Progress report	13th October 2017
Draft assessment report	8th December 2017
Final assessment report	10th January 2018

References

1. Office for National Statistics. Pregnancy and ethnic factors influencing births and infant mortality. 2015.

2. Ross M, Ridout, A., Shennan, A. Promising recent advances in preterm birth management. Frontiers in Women's Health. 2016;1(2):27-33.

3. March of Dimes P, Save the Children, WHO. Born Too Soon: The Global Action Report on Preterm Birth. Geneva: 2012.

4. World Health Organization. Preterm Birth [updated November 2016]. Available from: http://www.who.int/mediacentre/factsheets/fs363/en/.

5. NICE. DAP40: Biomarker tests to help diagnose preterm labour in women with intact membranes - Final Scope. 2017.

6. Phillips C, Velji Z, Hanly C, Metcalfe A. Risk of recurrent spontaneous preterm birth: a systematic review and meta-analysis. BMJ Open. 2017;7(6).

7. NICE. DAP40: The PartoSure Test to aid diagnosis of preterm labour in women with intact membranes - Briefing Note. 2017.

8. Royal Devon and Exeter NHS Foundation. Clinical Guideline: Management of pre-term labour. 2017.

9. Peaceman AM, Andrews WW, Thorp JM, Cliver SP, Lukes A, Iams JD, et al. Fetal fibronectin as a predictor of preterm birth in patients with symptoms: A multicenter trial. American Journal of Obstetrics and Gynecology. 1997;177(1):13-8.

10. Lee SM, Romero R, Park JW, Kim SM, Park C-W, Korzeniewski SJ, et al. The Clinical Significance Of A Positive Amnisure Test(™) In Women With Preterm Labor And Intact Membranes. The Journal of Maternal-Fetal and Neonatal Medicine. 2012;25(9):1690-8.

11. NICE. Preterm labour and birth, NICE guideline [NG25]. 2015.

12. NICE. Preterm labour and birth (full guideline), NICE guideline [NG25]. 2015.

13. Shennan A, McClaren J, Carter J. Threatened Preterm Labour: Guideline for Management at Guy's and St Thomas' NHS Foundation Trust. 2017.

14. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Annals of internal medicine. 2011;155(8):529-36.

15. Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. BMJ (Clinical research ed). 1996;313(7052):275-83.

16. Macaskill P GC, Deeks JJ, Harbord RM, Takwoingi Y. Chapter 10: Analysing and Presenting Results. In: Deeks JJ BP, Gatsonis C,, editor. Cochrane Handbook for

Systematic Reviews of Diagnostic Test Accuracy Version 10: The Cochrane Collaboration 2010.

17. CleverMed. BadgerNet Neonatal Electronic Patient Record 2017. Available from: https://www.clevermed.com/badgernet/badgernet-neonatal/.

18. NICE. Diagnostics Assessment Programme manual. 2011.

19. Kaltenthaler E, Tappenden P, Paisley S, Squires H. Identifying and reviewing evidence to inform the conceptualisation and population of cost-effectiveness models [available from http://www.nicedsu.org.uk/]. NICE DSU Technical Support Document. 2011;13.

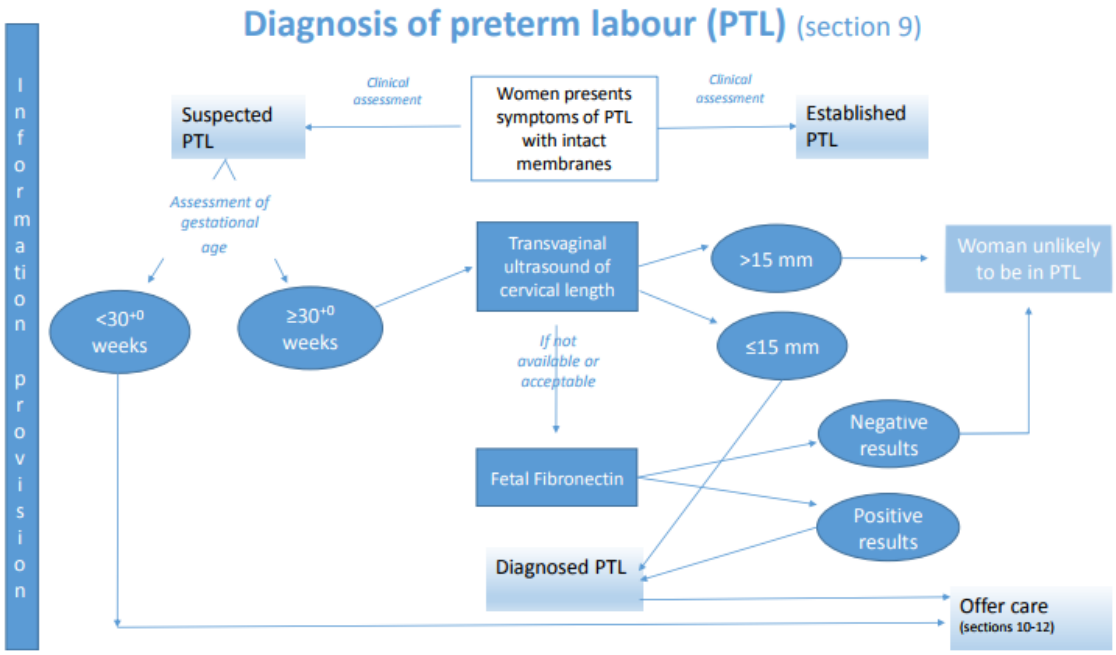
20. Briggs A. SM, Claxton K. Decision Modelling for health economic evaluation. : Oxford University Press; 2006.

Abbreviations

BMFMS	British Maternal and Fetal Medicine Society
CRDTAS	The Cochrane Register of Diagnostic Test Accuracy Studies
EAG	External Assessment Group
ELISA	enzyme-linked immunosorbent assay
fFn	fetal fibronectin
FN	false negative
FP	false positive
ICER	incremental cost-effectiveness ratio
IGFBP-1	insulin-like growth factor binding protein-1
IVH	intraventricular haemorrhage
NICE DSU	National Institute for Health and Care Research Decision Support Unit
PAMG-1	placental alpha microglobulin-1
PenTAG	Peninsula Technology Assessment Group
QALYs	quality-adjusted life-years
RDS	respiratory distress syndrome
TN	true negative
TP	true positive
WHO	World Health Organisation

Appendix 1.

Figure 1. Diagnosis of Preterm Labour from the 2015 NICE Guidelines on Preterm Labour and Birth (full guideline)[12]



Appendix 2. Details of EAG and clinical advisors

Name	Institution	Role/expertise
External Assessment Group		
Jo Varley-Campbell	PenTAG	Research Fellow; lead for systematic review of test accuracy; project lead
Ruben Mujica Mota	PenTAG	Senior Lecturer; lead for cost-effectiveness (economic modelling and review of economic evaluations)
Chris Cooper	PenTAG	Senior Research Fellow; Information Specialist
Helen Coelho	PenTAG	Research Fellow; systematic reviewer
Irina Tikhonova	PenTAG	Research Fellow; economic modelling
Max Barnish	PenTAG	Associate Research Fellow; systematic reviewer
Sue Whiffin	PenTAG	Senior administrator
Jenny Lowe	ESMI	Administrator; information officer
Martin Hoyle	PenTAG	Associate Professor of Health Technology Assessment; Director of PenTAG; economic modelling
Chris Hyde	PenTAG, Exeter Test Group, PenCLAHRC	Professor of Public Health and Clinical Epidemiology; public health physician; project guarantor
Clinical advisors		
Andrew Shennan	Kings College London / Guys and St Thomas	Professor of Obstetrics
Tracey Kay	RD&E	Consultant Obstetrician & Gynaecologist and Clinical Lead for Labour Ward
Neil Liversedge	RD&E	Consultant Obstetrician & Gynaecologist
Bridget Knight	RD&E	Research Midwife/Nurse and RDETB Nurse Manager