UK National Screening Committee

Screening for preterm birth in asymptomatic, low-risk women

External review against programme appraisal criteria for the UK National Screening Committee

Version: Publication Document

Author: Bazian

Date: October 2020

The UK National Screening Committee secretariat is hosted by Public Health England.

About the UK National Screening Committee (UK NSC)

The UK NSC advises ministers and the NHS in the 4 UK countries about all aspects of population screening and supports implementation of screening programmes. Conditions are reviewed against evidence review criteria according to the UK NSC's evidence review process.

Read a complete list of UK NSC recommendations.

UK NSC, Floor 5, Wellington House, 133-155 Waterloo Road, London, SE1 8UG
www.gov.uk/uknsc
Twitter: @PHE_Screening Blog: phescreening.blog.gov.uk

For queries relating to this document, please contact: phe.screeninghelpdesk@nhs.net

Published Month 20XX

Contents

Plain English summary

1 in 12 babies in the UK are born preterm. Premature babies can have serious medical problems including breathing issues. They might need to go to an intensive care unit.

This review found studies on 2 possible tests to screen low risk asymptomatic pregnant women for risk of preterm birth. One test involved measuring the length of the cervix (using an ultrasound probe). The other test involved taking a vaginal swab to look for a particular substance (called fetal fibronectin) that is released from the amniotic sac that surrounds the developing baby. If this substance is present at higher levels, there could be an increased risk of preterm birth within the next couple of weeks. Both tests can be done at the same time as the routine anomaly scan at around 20 weeks of pregnancy.

The studies showed that these are not very good tests for screening all pregnant women. This is because they found that fewer than half of women who had a high risk result went on to have preterm birth. But at the same time, the tests picked up other women who went on to have a normal, full term birth.

There is another possible screening test that looks for a type of bacterial imbalance in the vagina called bacterial vaginosis. This can sometimes increase the chances of having a preterm birth. This review did not find any recent studies looking at this test. But, a big study found that using antibiotics to treat women with bacterial vaginosis had no effect on their risk of preterm birth.

This review also looked at possible treatments to try and reduce risk of preterm birth in women found by screening to be at an increased risk. There was some evidence that a hormone tablet (progesterone) inserted into the vagina may reduce the risk of preterm birth in women found to have a short cervix on ultrasound. But we need more research to better understand which women may benefit more from this treatment.

The review found that placing a stitch in the cervix was not helpful to prevent preterm birth. It was not clear whether or not inserting a device called a pessary into the vagina, to support the cervix, might help some women. We need more research to be sure of this.

In conclusion:

- the screening tests for preterm birth in low risk are not reliable enough, and

- it is not certain that treating women identified to be at increased risk through screening will reduce their risk of preterm birth

As a result, screening low risk asymptomatic pregnant women for their risk of preterm birth is not recommended.

Executive summary

Purpose of the review

This evidence review update aimed to evaluate whether the evidence available supports the introduction of a universal screening programme to screen all pregnant women (asymptomatic and without existing risk factors) for risk of preterm birth and related neonatal and maternal outcomes.

Background

Preterm birth, defined as birth occurring before 37+0 weeks' gestation, is the single largest cause of morbidity and mortality in neonates in the UK.[1] Several pregnancy pathologies are associated with an increased risk of indicated or spontaneous preterm birth, some of these are pre-existing conditions, for example, chronic hypertension, pre-pregnancy diabetes mellitus, systemic lupus erythematosus and maternal underweight or obesity. Short inter-pregnancy interval and a family history of preterm birth can also be indicators of higher risk of spontaneous preterm birth. However, some other pathologies are pregnancy dependent such as preeclampsia or gestational diabetes mellitus.

Preterm birth can also occur in low risk asymptomatic pregnancies as the symptom or outcome of many different aetiological processes such as infection, bleeding, uterine over-distention, cervical weakness.[1] Preterm birth itself is not the negative event that has to be prevented, however it is associated with complications for the newborn, including increased risk of respiratory distress syndrome, intraventricular haemorrhage, retinopathy of prematurity and neonatal mortality, with risk of neurodevelopmental disability in the longer term. The risk of complications and mortality increases with decreasing gestational age at birth. The effects of preterm birth, including the need for the baby to spend time in special or intensive neonatal care, have considerable impact on parents and families. These effects may be wide ranging, including emotional and psychological effects such as depression, anxiety and affecting bonding, but also effects upon the family dynamic, interpersonal relationships, work commitments and finances. There are currently around 55,000 preterm births each year in England and Wales, an annual incidence of around 7.9%.[2] The DH Maternity Safety Ambition is to reduce the national preterm birth rate to 6%.[2]

Around three-quarters of preterm births are spontaneous, following onset of preterm labour or preterm prelabour rupture of membranes (P-PROM) rather than as a result of medically-indicated 'iatrogenic' preterm delivery.[1] The mechanisms behind spontaneous preterm birth may be multifactorial, including infective or inflammatory processes, cervical dysfunction,

nutritional, socioeconomic and environmental influences.[3] Various risk factors are also known to be associated including:[1, 4]

- multiple pregnancy
- history of preterm birth <34 weeks
- history of mid-trimester loss (16 to 24 weeks)
- history of P-PROM <34 weeks
- uterine anomalies
- cervical trauma/cervical surgical procedure

Certain clinical findings or biological markers have been found to be associated with a higher risk of preterm birth and are currently used for selective testing of symptomatic women or certain high-risk groups, notably cervical length measurement. In the UK pregnant women who have history of preterm birth or mid-trimester loss and a short cervical length (<25mm) in the mid-trimester may be offered prophylactic vaginal progesterone or cervical cerclage (stitch). Similarly cervical cerclage may be considered for women with short cervix and history of P-PROM or cervical trauma.[1] Cervicovaginal fetal fibronectin (fFN) is another biological marker from the placental/fetal membranes that is usually at low concentration (<50ng/ml) in mid-pregnancy, while only increasing at term. The concentration of fFN in the vaginal fluid may be measured in symptomatic women presenting in preterm labour to indicate the likelihood of birth and to triage admission or *in utero* transfer.[1] The common vaginal imbalance of bacterial vaginosis has also been associated with risk of preterm birth, and is usually treated with antibiotics if women are symptomatic or are found to have the condition incidentally in pregnancy. However, routine screening of asymptomatic women with no existing risk factors for preterm birth, by any test, is not currently performed in the UK.

Recommendation under review

The UK National Screening Committee (NSC) does not currently recommend systematic population screening of asymptomatic low risk pregnant women for risk of preterm labour and birth. This recommendation was made on the basis of the last evidence review on the topic, published in 2015.

Prior to the last evidence review, a 2009 Health Technology Assessment (HTA)[3] had investigated the accuracy of screening tests to predict risk of preterm labour and effectiveness of prophylactic/therapeutic interventions for asymptomatic women. The HTA considered a threshold for a useful test as one with a positive likelihood ratio (LR+) >5 (where a positive screening result would mean the condition is likely) and a negative likelihood ratio (LR-) <0.2 (where a negative result would give reassurance that the condition is unlikely). Optimal LR+ was found for fFN testing and cervical length

measurement, while optimal LR- was found for home monitoring of uterine contractions and amniotic fluid C-reactive protein (CRP) measurement. There was a trade-off between sensitivity and specificity, and none of the tests demonstrated a high LR+ in combination with a low LR-. Regarding prophylactic treatment, the HTA generally found poor quality of evidence, but potential for vaginal progesterone, cervical cerclage and antibiotics for bacterial vaginosis to reduce risk of preterm birth.

The 2015 UK NSC evidence review looked for evidence published up to 2013.
Most evidence was available on cervical length measurement but this was not found to be a reliable enough screening test. Questions remained over the timing of the test, and the lack of a standardised 'normal' measure making it difficult to establish what is 'abnormal'. The review also found that there was very limited evidence on screening for bacterial vaginosis, and absent evidence on fFN testing, detection of uterine contractions or amniotic fluid CRP measurement. Regarding treatment, the 2015 review found some trials indicating that vaginal progesterone and cervical cerclage may reduce risk in low-/mixed-risk populations, but the overall body of evidence was small.

Focus of the review

The current evidence review update aimed to synthesise the evidence on universal screening of pregnant women for risk of preterm labour published since the 2015 evidence review. It aimed to see whether new evidence is available on screening test performance and on the effectiveness of prophylactic interventions, which suggests that the current policy not to offer universal screening of all pregnant women for risk of preterm birth and associated neonatal and maternal outcomes should be reconsidered. It focused on the screening tests and interventions assessed by the last evidence review and previous HTA. However, the decision was made *a priori* not to assess amniotic fluid CRP measurement as a screening test, on the basis that an invasive, risk-associated procedure would not be considered a viable universal screening test for the general population of asymptomatic pregnant women.

This review update addressed 2 key questions in low-risk asymptomatic women:

1. What is the diagnostic measurement of the following tests in predicting preterm labour, birth or associated morbidity/mortality: (Criteria 4 and 5)
    - cervical length measurement
    - cervicovaginal fetal fibronectin
    - tests for bacterial vaginosis
    - uterine contraction (by home monitoring device)
2. What is the effectiveness of available treatments for the prevention of preterm labour, birth or associated morbidity/mortality: (Criterion 9)
    - progesterone

- cervical cerclage
- cervical pessary
- antibiotics for bacterial vaginosis
- probiotics

A rapid review search for these questions was conducted in September 2019 for studies published from January 2013 onwards, the search date of the last UK NSC evidence review.

Findings and gaps in the evidence of this review

This evidence review update found that the available evidence remains insufficient to support a programme to routinely screen all pregnant women for risk of preterm birth and related neonatal and maternal outcomes.

Screening tests

Fetal fibronectin (fFN) measurement

This evidence update identified one systematic review[5] and one prospective US cohort study [6] that assessed fetal fibronectin (fFN) testing and 4 prospective cohort studies[6-9] assessing cervical length measurement. None of the studies indicated that fFN would be a reliable screening test to predict risk of spontaneous preterm birth (<37 weeks) in the low-risk/general population of asymptomatic women with singleton pregnancies.

For fFN screening, both studies tested the standard ≥50ng/ml threshold measured at ≥22 weeks' gestation, giving inconsistent results. The systematic review[5] (n=1,236 women) found that a positive screen indicated a high likelihood of preterm birth (LR+ 12), but that there would be no confidence in a negative screen (LR- 0.54). Pooled sensitivity was 48%, but this was imprecise ranging from 20 to 77% across the meta-analysed studies. There is greater confidence in the findings of the single US cohort study[6] due to the larger sample size (n=9,469) and homogenous population/methods. This study found sensitivity of only 8% and PPV 11% for the same ≥50ng/ml threshold at ≥22 weeks (LR+ 2.53, LR- 0.95). Testing at other cut-offs/gestations little improved test performance, with peak sensitivity 35% (at PPV 7%) and peak PPV 14% (at sensitivity 4%). Both the systematic review and US study did have some quality and applicability limitations, including 30 to 50% of women in the studies being from non-western/Caucasian women populations.

Cervical length measurement

For cervical length screening, a US study[6] and a large Dutch study[9] (n=11,943) tested the standard ≤25mm cut-off measured in the mid-trimester (as used for selective testing of high-risk women). Both studies found this test identified fewer than 10% of women with preterm birth with PPV less than 30% (LR+ <4, LR- >0.9). Testing the same cut-off at later gestation (US study) or raising the cut-off to 35mm (Dutch study) achieved peak sensitivity <30% with peak PPV <15%. Two lower quality studies[7, 8] used receiver operating curves to identify optimal cut-offs for their populations of 37-38mm, which achieved higher sensitivity 50 to 75% but at very low PPV (6 to 7%). It is unknown whether these thresholds could be applied to other populations. There were also several quality issues with these studies, including small samples and high drop-out.

Similar to the last UK NSC evidence review and 2009 HTA, this evidence indicates that fFN testing and cervical length measurement are not useful to predict preterm birth in asymptomatic low-risk women (where a useful test is defined by LR+ >5 and LR- <0.2). A balance of high sensitivity and specificity is not achieved. Testing at different cut-offs and/or gestations to achieve optimal (though still inadequate) sensitivity results in poorer specificity with the majority of screen-positives being false.

This evidence update did not identify studies looking at screening for bacterial vaginosis or home monitoring for uterine contractions as screening tests.

Interventions

This evidence update identified a total of 8 eligible studies in applicable western populations looking at the treatment of pathologies that might increase the risk of preterm birth in women identified to have risk factors through screening:

- 1 systematic review (SR) with meta-analysis of individual patient data (IPD MA)[10] and 1 randomised controlled trial (RCT)[11] assessing vaginal progesterone;
- 1 SR[12] and 2 RCTs[13, 14] assessing cervical pessary; 1 RCT[15] comparing progesterone and pessary;
- 1 SR with IPD MA[16] assessing cervical cerclage; and 1 large RCT[17] assessing antibiotics for bacterial vaginosis.

The studies on progesterone, pessary and cerclage all assessed prophylactic treatment given from the time of screen-detection of short cervix in the mid-trimester (16 to 24 weeks) up until near term. In the antibiotic trial, treatment was given following screen-detection of bacterial vaginosis in the first trimester. This evidence update identified no studies where treatment was indicated on the basis of fetal fibronectin measurement.

The evidence was overall of good quality and applicable to a potential UK screening programme. The main limitations were that some studies included mixed-risk populations (including some with previous preterm birth), and that studies would be too small to reliably detect an effect on rarer preterm and neonatal outcomes.

Vaginal progesterone

Vaginal progesterone was associated with a modest reduction in the risk of preterm birth at all gestations <36 weeks, with numbers needed to treat (NNT) of around 12 to 14. There was no effect on overall preterm birth <37 weeks. There was also limited assessment of spontaneous preterm birth specifically (excluding medically-indicated). The single analysis conducted for the primary outcome (very preterm birth <33 weeks') on this basis indicated that the effect could be attenuated when considering spontaneous preterm births only (NNT 19 rather than 12). The main analyses were also for the mixed-risk/general antenatal population, with limited assessment specific to low-risk women with short cervix but no history of preterm birth. There was also evidence that vaginal progesterone reduced the risk of neonatal morbidity outcomes of low birthweight, admission to neonatal intensive care, respiratory distress syndrome, and the composite neonatal outcome. There was moderate-to-low quality evidence that vaginal progesterone had no effect on other neonatal outcomes. There was no new evidence on the effect of intramuscular or oral progesterone.

These findings on progesterone are essentially unchanged from the 2015 UK NSC evidence review, which was based on most of the same evidence. A further meta-analysis of IPD[18] is awaited which will compare any type and dose of progesterone, and may provide more comprehensive evidence on the effect of progesterone, including confirming whether there is an effect in otherwise-low risk women.

Cervical pessary

It is uncertain whether cervical pessary may benefit women with short cervix. Only a single trial was available at the 2015 UK NSC evidence review (which found a benefit). Four RCTs have since been published comparing with expectant management and one RCT comparing pessary with progesterone. The results are conflicting, with some finding a benefit of pessary and others not. The effect on risk of associated neonatal morbidity or mortality was also inconsistent across studies. However, even the trials finding a benefit showed little consistency in their findings or study populations, some of which included low-risk women only, while others included those with existing risk factors for preterm birth. Future IPD MA may help to understand whether variables such as cervical length, history of preterm birth or existing infection, could have an influence on effect. All trials were, however, unanimous in finding that pessary increased reports of vaginal discharge, though

the prevalence and risk increase was again inconsistent across studies. There was minimal other assessment of tolerability or acceptability which may be beneficial.

Cervical cerclage

As with the conclusions of the 2015 UK NSC review, the latest systematic review on cervical cerclage found that it had no effect on the risk of preterm birth or associated neonatal morbidity in otherwise low-risk women with short cervix. There was also no effect on any neonatal outcomes reported. Trials to date have also performed limited assessment of maternal adverse effects or acceptability of cerclage.

Bacterial vaginosis

One large, high-quality trial in otherwise low-risk women with bacterial vaginosis found that oral clindamycin (single or triple course) had no effect on risk of preterm birth or associated neonatal morbidity. It did, however, increase the risk of gastrointestinal adverse effects, though the prevalence of side effects was still low at 3% among treated women. There was no evidence available on the standard UK treatment of oral metronidazole. There were no studies on probiotics.

This evidence is consistent with the last Cochrane review[19] (included in the 2015 UK NSC evidence review) which found that antibiotics (any) for bacterial vaginosis had no effect on preterm birth risk.

None of the identified evidence on any intervention assessed whether treating women with short cervix or bacterial vaginosis who went on to have full term birth (that is false positives) was associated with any negative effects (such as psychological outcomes).

Recommendations on screening

Overall, the findings are in line with the 2015 UK NSC evidence review, finding that vaginal progesterone might have the potential to reduce risk of preterm birth in otherwise low-risk women found to have short cervix through screening in the mid-trimester. However, the poor test performance of cervical length measurement and/or cervicovaginal fetal fibronectin testing to reliably detect which asymptomatic, low-risk women are at risk of spontaneous preterm birth would appear to preclude universal screening at the current time.

The findings indicate that the current policy not to perform universal screening of otherwise low-risk asymptomatic pregnant women for risk of preterm birth and associated neonatal and maternal outcomes should not be reversed at the current time.

Evidence uncertainties

Further meta-analysis of individual patient data for progesterone and for cervical pessary may help to confirm whether or not these treatments are effective specifically in otherwise low-risk women with short cervix who have no existing risk factors for preterm birth. Future IPD may similarly help to clarify whether variables such as degree of cervical shortening, presence of infection, or method of treatment (for example dose or device) have an influence on effectiveness.

It may be beneficial to review the evidence on whether screening of asymptomatic, low-risk women (and subsequent management) reduces the risk of preterm birth and associated morbidity compared with not screening or is associated with any harms.

Future studies may also wish to explore whether other screening tests used as an alternative to, or in combination with cervical length or fFN testing (for example, measuring cervical consistency or cervical incompetence) may have potential as screening tests and demonstrate improved test performance.

Limitations

This was a rapid evidence review. The search strategy was built on a protocol developed *a priori* for each of the 2 key questions. Searching was limited to 3 literature databases and did not include grey literature resources. Studies only available in non-English language, editorials, abstracts, conference reports or poster presentations were not included. The reviewers were also unable to contact study authors or review non-published material.

Introduction and approach

Background

Preterm birth, defined as birth occurring before 37^{+0} weeks' gestation, is described to be the single largest cause of morbidity and mortality in neonates in the UK.[1] The Department of Health (DH) reported (2017) that there are around 55,000 preterm births each year in England and Wales, an annual incidence of around 7.9%.[2] The majority (85%) of preterm births are moderately preterm occurring from 32 to 37 weeks, while 11% are very preterm (28 to 32 weeks) and 5% extremely preterm (<28 weeks).[20] The rate of preterm birth varies by ethnicity. The Office for National Statistics (2016) reported that the highest preterm rate was among those of Black Caribbean origin (10.4%) and the lowest rate among the White other ethnic group (6.6%).[21] There is also marked socioeconomic variation, with those from the most deprived income groups twice as likely to have a preterm birth.[2]

Several pregnancy pathologies are associated with an increased risk of indicated or spontaneous preterm birth, some of these are pre-existent condition, for example, chronic hypertension, pre-pregnancy diabetes mellitus, systemic lupus erythematosus and maternal underweight or obesity. Short interpregnancy interval and a family history of preterm birth can also be indicators of higher risk of spontaneous preterm birth. However, some other pathologies are pregnancy dependent such as preeclampsia or gestational diabetes mellitus. Preterm birth can also occur in low risk asymptomatic pregnancies as the symptom or outcome of many different aetiological processes such as infection, bleeding, uterine over-distention, cervical weakness. Preterm birth is associated with various complications for the newborn, including increased risk of respiratory distress syndrome, intraventricular haemorrhage, retinopathy of prematurity and neonatal mortality. Neurodevelopmental disability is the major long-term complication.[1] The effects of preterm birth including the need for the baby to spend time in the special care baby unit (SCBU) or neonatal intensive care unit (NICU) have considerable impact upon parents and families. This includes emotional and psychological effects such as depression, anxiety and the restriction of physical contact that parents can have with their baby affecting bonding. There may also be wider effects upon the family dynamic, interpersonal relationships, work commitments and finances.

Globally, complications from preterm birth were the leading cause of death in children under 5 years in 2016, accounting for 16% of all deaths in under-5s and 35% of deaths among newborns.[22] The risk of complications and mortality increases with decreasing gestational age. In 2012, the National Institute for Health and Care Excellence (NICE) reported that although UK birth cohorts demonstrated improved survival for extremely preterm babies

(from 40% in 1995 to 53% in 2006), the rate of disability among survivors had remained unchanged. Similarly, as of 2012, NICE reported that the preterm birth rate (then given as 7.3% of live births) had remained constant over the past decade.[1] The DH Maternity Safety Ambition is to reduce the national preterm birth rate from 8% to 6% by 2025.[2]

Around a quarter of preterm births are medically-indicated (iatrogenic), for example due to pre-eclampsia or intrauterine growth restriction, but the remaining 75% are spontaneous, following onset of preterm labour or preterm prelabour rupture of membranes (P-PROM).[1] Various mechanisms are thought to underlie spontaneous preterm birth including, potentially, infective or inflammatory processes, cervical dysfunction, nutritional, socioeconomic and environmental influences.[3] In many cases the cause will be unknown (idiopathic). Various risk factors are also known to be associated with spontaneous preterm birth including:[1, 4]

- multiple pregnancy
- history of preterm birth <34 weeks
- history of mid-trimester loss (16 to 24 weeks)
- history of P-PROM <34 weeks
- uterine anomalies
- cervical trauma/surgical procedure

Certain clinical findings or biological markers have been found to be associated with a higher risk of preterm birth and are used for selective testing of symptomatic women or certain high-risk groups. They also have potential application for wider use as universal screening tests. Cervical length measurement (by transvaginal ultrasound scan) is perhaps the most well-established factor.[23, 24] Cervical length is described to be normally distributed and remains relatively constant until the third trimester. Mean lengths (average for all pregnant women) have been reported of 38mm at 23 weeks, 35mm at 24 weeks and 34mm at 28 weeks.[24] The rate of cervical shortening has been demonstrated to be faster in women who give birth before preterm compared with term, though this difference can be small.[24] A short cervix is generally considered to be a measure <25mm in the mid-trimester (16 to 24 weeks).[23, 24] Fetal fibronectin (fFN), a glycoprotein found between the placenta and fetal membranes, is another potential marker. It is found in cervicovaginal secretions up to around 20 weeks and then again at term, but through mid-pregnancy is usually at very low concentration (<50ng/ml).[5, 23] Increased levels after 22 weeks' gestation may be associated with increased risk of preterm birth in the coming weeks. Vaginal progesterone and cervical cerclage (stitch) are the most well established prophylactic treatments for women considered to be at high-risk of preterm birth (on the basis of obstetric history +/- such additional markers).

Bacterial vaginosis is another risk factor associated with preterm birth and related neonatal and maternal outcomes. It is the most common lower genital tract infection among women of reproductive age (studies suggesting prevalence of up to 1 in 4).[25] Diagnosis is based on characteristic vaginal discharge, pH testing of vaginal discharge (>4.5) and swab for Gram staining. Bacterial vaginosis is currently treated with antibiotics if it is detected symptomatically in pregnancy.

Current clinical practice and selective/targeted screening

Universal screening of all pregnant women for risk of preterm birth and related neonatal and maternal outcomes is not currently performed in the UK.

NICE clinical guidance on Antenatal care for uncomplicated pregnancies (published 2008)[26] gives recommendations related to screening in general, for asymptomatic bacterial vaginosis infection, and through pelvic examination. NICE CG62 states by recommendation (verbatim):[26]

- 1.8.2.1: Pregnant women should not be offered routine screening for bacterial vaginosis because the evidence suggests that the identification and treatment of asymptomatic bacterial vaginosis does not lower the risk of preterm birth and other adverse reproductive outcomes
- 1.9.3.1: Routine screening for preterm labour should not be offered
- 1.5.3.1: Routine antenatal pelvic examination does not accurately assess gestational age, nor does it accurately predict preterm birth or cephalopelvic disproportion. It is not recommended

NICE guidance on preterm labour and birth (2015)[1] indicates selective or targeted screening by cervical length measurement for (asymptomatic) women who have existing risk factors for preterm birth, with prophylactic treatment offered for those who screen below the 25mm cut-off (verbatim):

- 1.2.1: Offer a choice of prophylactic vaginal progesterone or prophylactic cervical cerclage to women who have both:
    - a history of spontaneous preterm birth (up to 34^{+0} weeks of pregnancy) or mid-trimester loss (from 16^{+0} weeks of pregnancy onwards) **and**
    - results from a transvaginal ultrasound scan carried out between 16^{+0} and 24^{+0} weeks of pregnancy that show a cervical length of 25 mm or less
- 1.2.2: Consider prophylactic vaginal progesterone for women who have either:
    - a history of spontaneous preterm birth (up to 34^{+0} weeks of pregnancy) or mid-trimester loss (from 16^{+0} weeks of pregnancy onwards) **or**
    - results from a transvaginal ultrasound scan carried out between 16^{+0} and 24^{+0} weeks of pregnancy that show a cervical length of 25 mm or less

- 1.2.4: Consider prophylactic cervical cerclage for women when results of a transvaginal ultrasound scan carried out between 16^{+0} and 24^{+0} weeks of pregnancy show a cervical length of 25 mm or less, and who have had either:
    - P-PROM in a previous pregnancy **or**
    - a history of cervical trauma

[Subsequent recommendations cover use of cervical length measurement and fFN testing in women presenting with signs and symptoms of preterm labour.]

Therefore, although routine screening is not recommended, it is notable that recommendation 1.2.2 covers the use of vaginal progesterone for women who have short cervix, apparently as an isolated finding or risk factor. This potentially reflects practice variation within the UK, and uncertainty around the contexts in which the decision may be taken to measure cervical length.

There is also international variation in recommendations. A UK systematic review (2018)[27] identified guidelines on screening and management of risk of preterm labour to see where there was consensus and variation (identifying 56 guidance documents: 49 guidelines and 7 methods papers). There was general consensus around screening by cervical length measurement for high-risk women only. Of 9 guidelines that made recommendations, 8 recommended screening in high-risk women, with only one guideline stating there was not enough evidence to screen this group. Eight guidelines gave recommendations on universal cervical length screening: 4 gave explicit recommendation not to screen, 2 stating that there was insufficient evidence, while 2 endorsed universal cervical length screening (University of Michigan and Japanese Society of Obstetrics and Gynaecology). However, the management recommendations then add further ambiguity, with 8/8 guidelines said to endorse vaginal progesterone for asymptomatic women *without* history of preterm birth who have short cervix (<20mm) at 24 weeks'. It is unclear therefore why cervical length would have been measured in these cases. Cervical cerclage meanwhile is only recommended for women with history of preterm birth: 4/4 recommending for women with >3 preterm births, and 5/5 recommending for women with short cervix (<25mm at 24 weeks') and past preterm birth.

The systematic review[27] found that asymptomatic screening for bacterial vaginosis was not recommended by 5/7 guidelines, while 2 guidelines recommended selective screening in women with prior preterm birth (Queensland Australia) or for symptomatic women or those with other risk factors for preterm birth (Society of Obstetricians and Gynaecologists of Canada). There was reported lack of consensus on fetal fibronectin screening for women considered to be at risk of preterm birth (1 yes, 3 conditional yes, 2 no, 1 insufficient evidence), but did not specify whether any guidelines had made recommendation on universal fFN screening.

Since publication of the systematic review, guidance published by the French College of Gynaecologists and Obstetricians[28] gives a 'cautious' recommendation on universal cervical length screening, stating on the one hand that a cervical length measurement ≤15mm is the best method for identifying asymptomatic low-risk women at risk of preterm birth, but that it is too early to conclude that universal screening is justified due to various evidence limitations. Meanwhile guidance jointly published by the German, Swiss and Austrian Societies for Gynaecology and Obstetrics[29] also gives some ambiguity. They give a 'moderately binding' recommendation that universal screening for short cervical length *should not be carried out* in asymptomatic women without risk factors; but then conversely give a 'non-binding' recommendation that cervical length measurement *may be carried out* in asymptomatic women without risk factors. They give an explicit recommendation that biomarkers *must not be used* to evaluate risk of preterm birth in asymptomatic women without risk factors. They also state some evidence that diagnosis and treatment of asymptomatic (and symptomatic) bacterial vaginosis (<23 weeks) reduces risk of preterm birth.

Current policy context and previous reviews

The UK NSC does not currently recommend systematic population screening of asymptomatic pregnant women for risk of preterm birth and related neonatal and maternal outcomes. This recommendation was made on the basis of the last evidence review on the topic, published in 2015 and which looked at evidence published up to 2013.

Prior to the last evidence review, the decision not to screen was primarily informed by a 2009 Health Technology Assessment (HTA) by Honest et al[3] which had investigated the accuracy of screening tests to predict risk of preterm labour/birth and effectiveness of prophylactic/therapeutic interventions for both asymptomatic and symptomatic women. The evidence was generally stronger for women with signs and symptoms of preterm labour. For asymptomatic women, the HTA found optimal positive likelihood ratios (LR+, where a positive test indicated higher likelihood of preterm birth) for fFN testing and cervical length measurement. Optimal negative likelihood ratios (LR-, where a negative test gives reassurance preterm birth is unlikely) were found for home monitoring of uterine contractions and amniotic fluid C-reactive protein (CRP) measurement. There was a trade-off between sensitivity and specificity, with no useful tests identified which demonstrated high LR+ in combination with LR-. Regarding prophylactic management, Honest et al generally found that some interventions such as vaginal progesterone, cervical cerclage and antibiotics for bacterial vaginosis had potential for reducing the risk of preterm birth. However, the evidence base was of poor quality. Further detail on the HTA findings is given in the introductions to the respective test and treatment criteria sections of this report.

The 2015 UK NSC evidence review looked at evidence (published since the 2009 HTA) on testing or prophylactic treatment for preterm birth. Of the potential screening tests, most of the available evidence was on cervical length measurement, but the review concluded that the test was not reliable enough. Also, questions remained over the timing of the test, and the lack of a standardised ‘normal’ measure making it difficult to establish what is an ‘abnormal’ cervical length. As such, the test may miss women who go onto have preterm birth, while falsely identifying and potentially exposing women who would not have preterm birth to unnecessary treatment. The review also found that there was a very limited volume of evidence relating to screening for bacterial vaginosis, and absent evidence on fFN testing, detection of uterine contractions or amniotic fluid CRP measurement.

Regarding treatment, the last UK NSC review found some trials indicating that vaginal progesterone and cervical cerclage may reduce risk in low-/mixed-risk populations, but the overall body of evidence was small.

Objectives

The current update review aims to synthesise and summarise the evidence on universal screening of pregnant women for risk of preterm labour published since the 2015 evidence review. It aims to see whether new evidence is available on screening test performance and on the effectiveness of prophylactic interventions, which suggests that the current policy not to offer universal screening of all pregnant women for risk of preterm labour and birth should be reconsidered.

Two key questions will be addressed to cover the issues identified by the last evidence review. These questions are outlined in Table 1.

Table 1. Key questions for the evidence summary, and relationship to UK NSC screening criteria

	Criterion	Key questions	Studies Included
	THE TEST		
4 **5**	There should be a simple, safe, precise and validated screening test. The distribution of test values in the target population should be known and a suitable cut-off level defined and agreed.	1. What is the diagnostic measurement of the following tests in predicting preterm labour, birth or associated morbidity/mortality: • cervical length measurement • cervicovaginal fetal fibronectin • tests for bacterial vaginosis • uterine contraction (by home monitoring device)	Total 6 publications: • fFN: 1 systematic review and 1 primary study • cervical length measurement: 4 primary studies • bacterial vaginosis: 0 studies • uterine contraction: 0 studies
	THE INTERVENTION		
9	There should be an effective intervention for patients identified through screening, with evidence that intervention at a pre-symptomatic phase leads to better outcomes for the screened individual compared with usual care. Evidence relating to wider benefits of screening, for example those relating to family members, should be taken into account where available. However, where there is no prospect of benefit for the individual screened then the screening programme shouldn't be further considered.	2. What is the effectiveness of available treatments for the prevention of preterm labour, birth or associated morbidity/mortality: • progesterone • cervical cerclage • cervical pessary • antibiotics for bacterial vaginosis • probiotics	Total 8 publications • progesterone: 1 systematic review and 1 RCT • cervical pessary: 1 systematic review and 2 RCTs • progesterone vs pessary: 1 RCT • cervical cerclage: 1 systematic review • antibiotics for bacterial vaginosis: 1 RCT • probiotics: 0 studies

Methods

The current review was conducted by Bazian (part of the Economist Intelligence Healthcare Unit), in keeping with the UK National Screening Committee evidence review process. Database searches were conducted on 17th September 2019 to identify studies relevant to the questions detailed in in

Table 1.

Eligibility for inclusion in the review

The systematic literature search of MEDLINE and Embase databases (Embase.com) and The Cochrane Library (Wiley Online) was performed for studies published between January 2013 (the search date for the last UK NSC evidence review) and September 2019. Individual searches were conducted for each of the 2 key questions according to the scopes developed *a priori* as outlined in Table 2 below. Searches for both questions retrieved a total of 3,304 citations after removal of 1,010 duplicates.

The following review process was followed:

1. Each of the 3,304 titles and abstract were reviewed against the inclusion/exclusion criteria for each question by one information specialist. Where the applicability was unclear, the article was included at this stage to ensure that all potentially relevant studies were captured. In total 306 citations were included at first sift: 142 for question 1 and 164 for question 2.
2. At second sift the main reviewer reviewed each of the 306 abstracts for potential relevance to either question. Where the article content was unclear from the abstract, full text was obtained to ensure that potentially relevant literature was not missed.
3. A total of 68 articles were acquired for the full-text review stage. Each full-text article was reviewed against the inclusion/exclusion criteria by the main reviewer, who determined whether the article was relevant to either of the review questions. All inclusion/exclusion decisions were reviewed by a second independent reviewer who provided input in cases of uncertainty. Any disagreements were resolved by discussion until a consensus was met.

The full search strategy is presented in Appendix 1.

0contains a full PRISMA flow diagram (Figure 1), along with a table of the included publications (Table 22) and a table of the publications excluded at full-text appraisal (Table 23).

Table 2. Inclusion and exclusion criteria for the key questions

Key question	Inclusion criteria							Exclusion criteria
	Population	**Target condition**	**Index test or intervention**	**Reference Standard**	**Comparator**	**Outcome**	**Study type**	
1. What is the diagnostic measurement of the following tests in predicting preterm labour, birth or associated morbidity/mortality ?	Asymptomatic women at low-risk of preterm labour Low-risk considered to be women without: multiple birth prior SPTB (<34 weeks) prior mid-trimester loss (16-24 weeks prior P-PROM uterine anomalies cervical trauma	Spontaneous preterm labour <37 weeks Spontaneous preterm birth (SPTB) <37 weeks Associated morbidity or mortality	Cervical length measurement Cervicovaginal fFN Tests for bacterial vaginosis Uterine contraction monitoring (by home device)	Spontaneous preterm labour <37 weeks Spontaneous preterm birth (SPTB) <37 weeks Associated morbidity or mortality	NA	NA	Diagnostic cohort studies in non-selected populations where the full sample received the index test and birth outcomes assessed Systematic reviews (SRs) of these studies	Case-controls, studies conducted exclusively in high-risk pregnant women, non-SRs, editorials, conference abstracts, non-English language studies, studies conducted exclusively in non-western populations with limited applicability to the UK.
2. What is the effectiveness of available treatments for the prevention of preterm labour,	Asymptomatic women otherwise at low-risk of preterm labour	Spontaneous preterm labour <37 weeks Spontaneous preterm	Progesterone Cervical cerclage Cervical pessary	NA	Placebo, no treatment or other intervention	Spontaneous preterm labour <37 weeks Spontaneous preterm	Randomised controlled trials (RCTs) prioritised followed by	Studies conducted exclusively in high-risk pregnant women,

birth or associated morbidity/mortality ?	(according to the above definition) who have been identified through the screening tests outlined in Q1	birth (SPTB) <37 weeks Associated morbidity or mortality	Antibiotics for bacterial vaginosis Probiotics	birth (SPTB) <37 weeks Associated fetal, neonatal or maternal morbidity or mortality Any harms from treatment	non-RCTs or comparative cohort studies (depending on evidence availability) SRs of these studies	non-comparative studies, non-SRs, editorials, conference abstracts, non-English language studies, studies conducted exclusively in non-western populations with limited applicability to the UK.

Appraisal for quality/risk of bias tool

The following tools were used to assess the quality and risk of bias of each study included in the review:

- systematic reviews: Critical Appraisal Skills Programme (CASP) Systematic Review Checklist
- diagnostic accuracy studies: Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool
- RCTs: Cochrane Collaboration's "Risk of Bias" Tool

Question level synthesis

Criterion 4 — There should be a simple, safe, precise and validated screening test

Criterion 5 — The distribution of test values in the target population should be known and a suitable cut-off level defined and agreed

Question 1 – What is the diagnostic measurement of the following tests in predicting preterm labour, birth or associated morbidity/mortality:

- *cervical length measurement*
- *cervicovaginal fetal fibronectin*
- *tests for bacterial vaginosis*
- *uterine contraction (by home monitoring device)*

Evidence to date on screening in asymptomatic, low-risk women

The Honest et al HTA (2009) assessed 22 different tests for the prediction of spontaneous preterm birth in women with singleton pregnancies. The quality and strength of evidence tended to be higher for studies in women with signs and symptoms of preterm labour rather than for asymptomatic women. Honest et al considered a threshold for a useful test as one with positive likelihood ratio (LR+) >5 (where a positive screening result would mean the condition was likely), and one with a negative likelihood ratio (LR-) <0.2 (where a negative result would give reassurance that the condition is unlikely). In asymptomatic women, Honest et al found optimal LR+ for cervicovaginal fetal fibronectin (fFN) testing and cervical length measurement. For fFN testing, LR+ for predicting preterm birth <37 weeks ranged from 0.43 to 26.38 across 15 studies with summary LR+ 3.40 (with corresponding LR- ranging from 0.28 to 1.20, summary 0.87). LR+ was higher for predicting birth <34 weeks (summary LR+ 7.65, LR- 0.80). For cervical length, the standard threshold of ≤25mm measured before 20 weeks' gestation gave summary LR+ 13.38 and LR- 0.80 to predict very preterm birth <34 weeks (from 3 high quality studies). Only a single study assessed prediction of preterm birth <37 weeks finding LR+ 3.99 and LR- 0.33 (for measure ≤32.5mm at 20 to 24 weeks). Therefore, despite good LR+ there was a trade-off against test sensitivity, with neither fFN nor cervical length measurement demonstrating an optimal balance of high LR+ in combination with low LR-. Consequently, women with preterm birth would be missed.

The 2 tests demonstrating the best LR- were home monitoring of uterine contractions and amniotic fluid CRP measurement. However, again the best evidence was for symptomatic women. Only 2 studies assessed uterine activity monitoring in asymptomatic women (using

different methods), one found detecting 4 significant contractions within a one-hour period had LR+ 4.90 and LR- 0.15 for predicting birth <37, though the other found only LR+ 0.51 and LR- 1.01 for day-time monitoring to predict birth <34 weeks.

Bacterial vaginosis has also been widely studied but did not appear to be a useful screening test to predict preterm birth <37 weeks in asymptomatic women: 12 studies gave summary LR+ 1.77 and LR- 0.80 for a single second trimester test (using Nugent's criteria), and 3 studies gave summary LR+1.38 and LR-0.94 for serial testing.

Honest et al noted quality and reporting issues across studies that may have affected the validity of results. They report how studies may have included women from across the clinical spectrum (for example, including or exclusively representing those with existing risk factors for preterm birth), making it difficult to know whether the results were directly applicable to low-risk women. The authors emphasised that studies needed to adequately report the study population, design and execution of tests in order to be considered as scientific evidence. There was also the potential that screen-positives may have received treatment or different antenatal care from screen-negatives which could have reduced risk of spontaneous preterm birth and so reduced the apparent accuracy of the screening test. This was described by Honest et al as the 'treatment paradox'.

The last 2015 UK NSC evidence review looked for evidence published between 2007 and 2013. Most of evidence identified was for cervical length measurement, and the results were conflicting. The reviewers identified one systematic review which found optimal test performance using the cut-off ≤20mm to predict preterm delivery at <35 weeks (sensitivity 22.1%, specificity 98.2%, LR+ 12.4 and LR- 0.74). Ten additional primary studies measured cervical length at variable gestations, from 10 to 28 weeks, and for the prediction of preterm birth from <30 weeks to <37 weeks. There was wide variation in test accuracy with only 2 studies fulfilling the HTA criteria for a useful test (LR+ >5 and LR- <0.2).

The 2015 evidence review also identified a single study of bacterial vaginosis testing, which found that vaginal pH ≥5 had high specificity (98.9) and LR+ (46.7) but low sensitivity (47.9). No studies were identified on screening by fFN testing, detection of uterine contractions or amniotic fluid CRP measurement.

The current evidence summary reviewed the evidence published since 2013 on the 4 potentially most promising screening tests (as indicated by the 2009 HTA and last NSC evidence review) to predict preterm labour/birth or associated morbidity and mortality:

- cervical length measurement
- cervicovaginal fetal fibronectin
- tests for bacterial vaginosis

- uterine contraction (by home monitoring device)

The decision was made *a priori* not to assess amniotic fluid CRP measurement on the basis that an invasive, risk-associated procedure would not be considered a viable universal screening test for the general population of asymptomatic pregnant women.

Eligibility for inclusion in the review

The evidence review update aimed to identify prospective cohort studies evaluating any of these 4 tests in consecutively-enrolled or randomly-selected asymptomatic pregnant women. Women could be either from the general pregnant population (mixed-risk) or specifically low-risk, with no known risk factors for preterm birth. The following were considered to be risk factors, which may be indicators for selective or targeted screening, or for alternative care pathways, as part of current antenatal care in the UK:

- multiple pregnancy
- history of spontaneous preterm labour (<34 weeks)
- history of P-PROM (<34 weeks)
- history of mid-trimester loss (from 16 to 24 weeks)
- cervical trauma/cervical surgical procedure
- uterine anomaly

Studies evaluating screening tests exclusively in these high-risk populations would be excluded.

Studies from the UK would be prioritised but studies from other representative western populations would also be eligible. The index test could be performed at any gestation and using any threshold/cut-off to predict risk of preterm labour or birth or associated morbidity. The outcome of interest was spontaneous preterm birth (SPTB) or labour, excluding iatrogenic preterm birth which is indicated by medical factors would not be preventable through such screening. The primary outcome of interest was overall SPTB (<37 weeks) but evidence on prediction of SPTB at earlier gestation was also reviewed.

Systematic reviews of eligible diagnostic cohort studies would also be eligible for inclusion.

Further detail on the inclusion and exclusion process following appraisal of the available evidence is discussed below.

Description of the evidence

Database searches yielded a total 3,304 results across both questions. A total 142 were judged to be potentially relevant to this question at first sift by the information specialist, 32 of which were selected for full text appraisal by the main reviewer.

A hierarchical approach was taken to full text appraisal, considering firstly any systematic reviews of diagnostic cohort studies in the eligible population, followed by individual diagnostic cohort studies.

A total of 5 studies in asymptomatic mixed-/low-risk women (one publication covering 2 tests) met inclusion criteria for this question:

- cervicovaginal fetal fibronectin: 1 systematic review[5] and 1 prospective cohort study[6]
- cervical length measurement: 4 prospective cohort studies[6-9]
- tests for bacterial vaginosis: 0 studies
- uterine contraction (by home monitoring device): 0 studies

The 27studies not selected for inclusion are listed in Appendix 2 (Table 23) alongside the reason for exclusion. There are a couple of excludes that are worthy of further mention. The search retrieved a number of prospective screening cohort studies where prophylactic treatment (such as progesterone) was routinely offered to screen-positives, or where they were offered participation in a treatment trial. This introduces the issue of the ‘treatment paradox’ highlighted by Honest et al, where the different management approach for screen-positives and negatives invalidates test performance data. Such studies were only considered eligible for inclusion if treated-women were excluded from the analysis, or if the proportion of screen-positives treated was less than 20% (this accounted for 2 included studies, which both treated around 10% of screen-positives).

The search additionally retrieved 4 Cochrane reviews (published post-2013) evaluating fFN testing, cervical length measurement, bacterial vaginosis testing and uterine contraction monitoring for prevention of preterm labour/birth. The inclusion criteria for these Cochranes were randomised controlled trials that compared the risk of preterm labour/birth in populations assigned to screening (plus knowledge of results and subsequent treatment) vs no screening (and no treatment). These studies are essentially reviewing whether or not screening is effective in reducing morbidity or mortality (UK NSC Criterion 11). This may be

a question and Criterion relevant for further study, but the evidence is not an applicable to the diagnostic accuracy question.*

This evidence update also excluded systematic reviews where all studies included by the review were published prior to 2013 and the review provided only narrative synthesis of the results. In such cases the individual study results would have been covered by the 2015 UK NSC evidence review (or by the Honest et al HTA) and in the absence of meta-analysis the review was not considered to be providing new evidence.

Other studies excluded from this question either at abstract level (if study design and method were clearly non-applicable) or at full text appraisal is as follows:

- studies conducted exclusively in women with risk factors (multiple pregnancy, previous preterm birth, P-PROM, mid-trimester loss, cervical trauma or uterine anomaly)
- tests in symptomatic women with signs of preterm labour
- case controls in women with and without signs of preterm labour
- studies in women already receiving prophylactic treatment (for example, cerclage in situ)
- studies assessing tests not defined *a priori* in the PICO (for example, biomarkers in cervicovaginal fluid, uterocervical angle, pelvic muscle function)
- studies validating new instruments or tests (for example, cervical length measure by transabdominal vs standard transvaginal ultrasound measure, or use of cervicometer)
- studies evaluating test combinations where there was no individual assessment of the test of interest (for example, evaluating cervical length measurement in combination with measure of inflammatory cytokines)
- subtests in women identified to have short cervix (for example, cervical volume or vascularisation)
- studies purely evaluating the inter-rater reliability of measures
- secondary analysis of treatment RCTs
- tests for predicting prolonged pregnancy or labour in full-term women
- risk factors associated with short cervix
- incidence of short cervix by ethnicity
- studies in conducted exclusively in non-representative populations (for example, non-western countries, or a study of African American women alone)
- non-applicable scenarios for universal screening, such as tests to predict the onset of preterm labour within the following 48 hours
- non-systematic reviews or editorials

* The Cochrane reviews were considered for eligibility for question 2 on treatment but did not meet inclusion criteria for other reasons as outlined in the table of excluded studies. Much of the evidence identified by these Cochranes was in high-risk women or symptomatic populations.

The findings from the included studies are discussed below by screening test. Tables 3 to 7 present the key data from each of the studies alongside summarised quality appraisal. Complete data extraction and quality appraisal for each of the 5 included publications is presented in Appendix 3.

Presentation and discussion of findings

Fetal fibronectin measurement

Table 3. Test accuracy of fetal fibronectin (fFN) measurement to predict spontaneous preterm birth <37 weeks

Study	Population	Preterm birth incidence	Time of screen	Cut-off	Accuracy for prediction of spontaneous preterm birth <37 weeks (95% CI)					
					Sensitivity	Specificity	PPV	NPV	LR+	LR-
Dos Santos et al 2018[5] Systematic review (SR) with meta-analysis (MA) Studies from Italy, Belgium, US, Japan Singapore and Mexico; all published 1995 to 1999	6 prospective cohort studies in n=1,236 low-risk women with singleton pregnancy: No previous preterm birth No other risk factors as defined by individual studies Population size range across studies: n=60 to n=438	Study range: 4.1% to 10.3% NB not specified if spontaneous only	22-37 weeks gestation: 3 studies 22-34 weeks 3 studies 24-37 weeks	≥50ng/ml	48% (20 to 77) Study range: 6 to 81%	96% (CI 86 to 99) Study range: 84 to 100%	NR Study range: 23.8 to 71.0%	NR Study range: 93.4 to 97.8%	12.01 (4.70 to 30.68) By study: NR	0.54 (0.30 to 0.97) By study: NR
Esplin et al 2017[6] Prospective cohort study Multicentre US, 2010 to 2014	n=9,469 nulliparous women with singleton pregnancy	5.0%	6^{+0} to 14^{+6} weeks	≥10ng/ml	**34.5%** (30.0 to 39.1)	74.1% (73.2 to 75.1)	6.7% (5.6 to 7.7)	95.5% (95.0 to 96.0)	1.34 (1.15 to 1.52)	0.88 (0.82 to 0.95)
				≥50ng/ml	21.2% (17.2 to 25.1)	87.6% (86.9 to 88.3)	8.4 % (6.7 to 10.0)	95.4% (94.9 to 95.9)	1.71 (1.37 to 2.04)	0.90 (0.85 to 0.95)
				≥200 ng/ml	9.5% (6.7 to 12.3)	94.6 (94.1 to 95.1)	8.6 (6.0 to 11.1)	95.1% (94.7 to 95.6)	1.75 (1.20 to 2.30)	0.96 (0.93 to 0.99)
			16^{+0} to 22^{+6} weeks	≥10ng/ml	15.1% (11.7 to 18.6)	88.5% (87.8 to 89.2)	6.4% (4.9 to 8.0)	95.2% (94.8 to 95.7)	1.32 (1.01 to 1.63)	0.96 (0.92 to 1.00)
				≥50ng/ml	7.3%	96.0%	8.8%	95.2%	1.85	0.97

					(4.8 to 9.8)	(95.6 to 96.5)	(5.8 to 11.8)	(94.7 to 95.7)	(1.18 to 2.51)	(0.94 to 0.99)
				≥200 ng/ml	2.9% (1.5 to 5.1)	98.3% (98.0 to 98.6)	8.3% (3.8 to12.8)	95.1% (94.6 to 95.6)	1.73 (0.72 to 2.74)	0.99 (0.97 to 1.00)
			22^{+0} to 30^{+6} weeks	≥10ng/ml	21.9% (17.7 to 26.0)	91.8% (91.2 to 92.4)	11.2% (9.0 to 13.5)	96.1% (95.7 to 96.5)	2.66 (2.12 to 3.20)	**0.85** (0.81 to 0.90)
				≥50ng/ml	8.1% (5.3 to 10.8)	96.8% (96.4 to 97.2)	10.7% (7.2 to 14.3)	**95.7%** (95.2 to 96.1)	2.53 (1.62 to 3.44)	0.95 (0.92 to 0.98)
				≥200 ng/ml	3.9% (2.2 to 6.4)	**98.9%** (98.6 to 99.1)	**14.0%** (7.4 to 20.6)	95.6% (95.2 to 96.0)	**3.44** (1.59 to 5.28)	0.97 (0.95 to 0.99)

NR=not reported; Bold=peak value

Table 4: CASP assessment of systematic review

Study	Are the review findings valid?	Are the results clear and precise?	Will the results help locally?	QUADAS-2 risk of bias for n=6 included studies (by study authors)
Dos Santos et al 2018[5]	Yes (focused question, relevant papers) Unclear (detail on quality assessment, appropriate meta-analysis)	Unclear (clear interpretation) No (precision of results)	Unclear (applicability, outcomes)	Patient selection: n=3 low, n=2 unclear, n=1 high Index test: n=5 low, n=1 high risk Reference standard: n=5 low, n=1 high Flow and timing: n=3 low, n=2 unclear, n=1 high

Table 5: QUADAS-2 assessment of diagnostic accuracy studies

Study	Risk of bias by domain				Applicability to review question
	Patient selection	Index test	Reference Standard	Flow and timing	
Esplin et al 2017[6]	Low	Unclear (conduct)	Low	Low	Unclear (population, index test) Low (reference standard)

The 2 identified studies find no evidence to suggest that fFN measurement would be an accurate screening test for asymptomatic low-risk women with singleton pregnancies.

The Dos Stantos[5] systematic review identified 6 small cohort studies in low-risk pregnant women (excluding women with prior preterm birth and other risk factors) that measured fFN at ≥22 weeks' gestation, using the standard threshold of fFN ≥50ng/ml to indicate risk of preterm birth. Nearly all individual studies had LR+ >5 resulting in a high pooled LR+ of 12 combined with a consistently high specificity across studies (mean 96%, all studies >84%). This suggests that the false positive rate from fFN screening would be low, and that a positive screen would indicate a high likelihood of preterm birth. Conversely the pooled sensitivity (48%) and LR- (0.54) are very poor, indicating that around half of women who have preterm birth would not have raised fFN and would screen negative. Sensitivity also varied widely across studies from 6 to 81%, decreasing confidence in the pooled estimate. This may be a result of the small sample size and low number of preterm births across these studies (total sample 1,236; study range 60 to 438).

The single prospective cohort study by Esplin et al[6] included 9,469 nulliparous women (thereby inherently excluding most risk factors for preterm birth) with the larger sample size giving greater precision in the results. The standard ≥50ng/ml threshold at ≥22 weeks was this time not demonstrated as a useful test with LR+ 2.53 and LR- 0.95 (compared with LR+ 12 and LR-0.54 in the Dos Stantos[5] review). It identified only 8% of women who had preterm birth, while 89% of screen-positives were false (PPV 11%). The peak PPV was only 14% and peak LR+ 3.44 using a much higher cut-off of ≥200g/ml at ≥22 weeks. Sensitivity was extremely poor across all thresholds and gestations tested. Peak sensitivity of 34.5% was achieved using cut-off >10ng/ml at 6 to 14^{+6} weeks' at the expensive of a low PPV 6.7% and specificity 74%. This is unlikely to be useful as a test considering that fFN is known to be raised in early pregnancy. Even the combination of 3 serial fFN measures only gave an area under the curve (AUC) of 0.59 for predicting preterm birth, which is little better than chance.†

There are quality and applicability issues to note for both studies. It is not possible to know whether the findings of the Esplin cohort study would be applicable to low-risk, multiparous women (that is those without previous preterm birth, P-PROM or mid-trimester loss). It is also not possible to be sure that the findings are applicable to the UK in terms of the screened population, as one third of included women were of Black or Hispanic ethnicity. Similarly, 3 of 6 studies included in the Dos Santos review were conducted in non-Western countries (Japan, Singapore and Mexico). The Dos Santos review was considered on balance to be applicable evidence to this evidence update, but such studies conducted exclusively in non-Western/Caucasian populations would not have been selected for inclusion on an individual basis.

With both studies there are also outstanding questions regarding the external validity/applicability of the index tests. The Dos Santos review tested the accepted standard of the 50ng/ml cut-off at gestational age ≥22weeks, but the included studies tested at a broad age range from 22 to 37 weeks' gestation. Furthermore, 3 studies used single testing while used 3 serial testing. In the latter cases it was unclear whether the uthors extracted test performance data for single measurements or for the overall serial measure. The Esplin study meanwhile used self-obtained fFN samples, and it is unknown whether these samples would be as reliable as those taken by trained health professionals. With such variability in testing strategy, even had better test performance been achieved, it would be difficult to apply this evidence to inform the optimal testing strategy to use in a screening programme.

---

† This evidence review update focused on the reference standard of all spontaneous preterm birth <37 weeks. For prediction of SPTB <32 weeks Esplin et al found improved test performance but this was still not indicated as a reliable screening test: maximum sensitivity 50% (≥10ng/ml at 22 to 30^{+6} weeks) and maximum PPV 5.6% (≥200ng/ml at same gestation).

There are also points to note regarding the reference standard. This evidence update aimed to look at spontaneous preterm birth as the outcome, excluding iatrogenic preterm birth which would not be preventable through screening. Spontaneous birth was clearly defined in the US study, but the Dos Santos review included studies looking at 'preterm birth'. It is unclear whether all of these studies restricted their analyses to spontaneous preterm birth. It is also unknown whether any meta-analysed studies may have given prophylactic treatment to screen-positives (as reported, this evidence update would have excluded individual studies where >20% of screen-positives were treated).

This latter point on prophylactic treatment raises a final notable point on the reference standard in general, which is applicable to both these studies looking at fFN testing and the following studies on cervical length measurement. All studies have looked at preterm birth, rather than preterm labour. If a woman presents in preterm labour, treatment such as rescue cerclage or tocolytic drugs may be given to try and halt labour and prevent birth. In current clinical practice, fFN or cervical length measurement may be used to guide the use of such treatments. It is assumed that if a woman presents in preterm labour, new measures would be taken at the time of symptomatic presentation, regardless of earlier screening results. That is, the approach to assessment and management of preterm labour would be equivalent for screen-positive and screen-negative women. In that case the performance of screening tests to predict preterm birth could be considered compatible with the performance of screening tests to predict preterm labour. However, if screening results were used to guide management decisions in symptomatic women (for example, screen-positives being more like to receive tocolysis) then this could affect the validity of screening test performance to predict preterm birth.

Cervical length measurement

As a result, screening. Test accuracy of cervical length measurement to predict spontaneous preterm birth <37 weeks

Study	Population	Preterm birth incidence	Time of screen	Cut-off	Accuracy for prediction of spontaneous preterm birth <37 weeks (95% CI)					
					Sensitivity	Specificity	PPV	NPV	LR+	LR-
Esplin et al 2017[6] Prospective cohort study Multicentre US, 2010 to 2014	n=9,469 nulliparous women with singleton pregnancy	5.0%	16^{+0} to 22^{+6} weeks	≤25mm	8.0% (5.4 to 10.5)	97.8% (97.5 to 98.1)	16.2% (11.3 to 21.1)	95.3% (94.8 to 95.7)	3.67 (2.39 to 4.95)	0.94 (0.91 to 0.97)
				≤20mm	4.1% (2.4 to 6.4)	98.8% (98.6 to 99.1)	15.5% (8.9 to 22.1)	95.1% (94.7 to 95.6)	3.49 (1.77 to 5.21)	0.97 (0.95 to 0.99)
			22^{+0} to 30^{+6} weeks	≤25mm	**23.3%** (19.2 to 27.5)	93.6% (93.1 to 94.1)	15.1% (12.3 to 17.9)	96.2% (95.8 to 96.6)	3.65 (2.94 to 4.37)	0.82 (0.77 to 0.86)
				≤20mm	17.4% (13.7 to 21.1)	96.8% (96.4 to 97.2)	**20.8%** (16.5 to 25.2)	96.0% (95.6 to 96.4)	**5.42** (4.10 to 6.74)	0.85 (0.82 to 0.89)
van der Ven et al 2015[9] Prospective cohort study Multicentre The Netherlands, 2009 to 2013	n=11,943 women with singleton pregnancy and no history of preterm birth <34 weeks divided into: n=5710 nulliparous n=6233 low-risk multiparous	3.9% overall: 5.3% nulliparous and 2.6% multiparous	16^{+0} to 21^{+6} weeks	≤35mm* *see sub-note	nllp. **29.0%** mltp. **17.7%**	nllp. 86.9% mltp. 90.4%	Nllp. 10.9% Mltp. 4.8%	nllp. 95.7% mltp. 97.6%	31 to 35mm: nllp. 2.0 (1.6 to 2.5) mltp. 1.5 (0.97 to 2.4)	NR
				≤30mm* *see sub-note	nllp. 6.3% mltp. 5.5%	nllp. 98.0% mltp. 98.7%	nllp. 15.2% mltp.10.3%	nllp. 95.0% mltp.97.5%	26 to 30mm: nllp. 1.8 (0.86 to 3.6) mltp. 3.9 (1.7 to 8.9)	NR
				≤25mm* *see sub-note	nllp. 3.7% mltp.1.8%	nllp. 99.6% mltp.99.7%	nllp. 31.4% mltp.12.5%	nllp. 94.9% mltp.97.4%	21 to 25mm	NR

Study	Population	Preterm birth incidence	Time of screen	Cut-off	Accuracy for prediction of spontaneous preterm birth <37 weeks (95% CI)					
					Sensitivity	**Specificity**	**PPV**	**NPV**	**LR+**	**LR-**
									nllp. 4.5 (1.7 to 12) mltp. 0 cases	
				≤20mm* *see sub-note	nllp. 2.0% mltp.1.8%	nllp. **99.6%** mltp. **100**%	nllp. **31.4%** mltp**.50.0%**	nllp. 94.9% mltp.97.4%	≤20mm: **nllp. 27 (7.7 to 95)** **mltp. 37 (7.5 to 182)**	NR
Banos et al 2018[7] Prospective cohort study Single centre, Spain, 2014 to 2015	n=532 low-risk women with singleton pregnancy Exclusions: history of preterm birth <34 weeks P-PROM late miscarriage >16 weeks cervical/uterine trauma or malformation known short cervical length (<25mm) receipt of current prophylactic treatment	4.1%	19^{+0} to 24^{+6} weeks	≤37.9mm (40^{th} centile)	**72.7%** (NR)	61.2% (NR)	7.5% (NR)	**98.1%** (NR)	1.9 (1.4 to 2.5)	**0.4** (0.2 to 0.9)
				≤33.0mm (10^{th} centile)	31.8% (NR)	89.6% (NR)	11.7% (NR)	96.8% (NR)	3.1 (1.6 to 5.9)	0.8 (0.6 to 1.0)
				≤30.0mm (5^{th} centile)	18.2% (NR)	96.5% (NR)	18.2% (NR)	96.5% (NR)	5.2 (1.9 to 13.9)	0.8 (0.7 to 1.0)
				≤25mm (1^{st} centile)	13.6% (NR)	99.6% (NR)	**60.0%** (NR)	96.4% (NR)	**34.8** (95% CI 6.1 to 197.6)	0.9 (0.7 to 1.0)
Kuusela et al 2015[8] Prospective cohort study 2 centres, Sweden, 2012 to 2015	n=2,061 women with singleton pregnancy (no risk-based criteria) NB exclusion of 7/11 women with measure <25mm opting to participate in treatment trial	4.2%	16^{+0} to 24^{+0}	≤37mm (35^{th} centile)	**53%** (48 to 59)	65% (64 to 67)	6% (6 to 6)	**97%** (96 to 98)	1.52 (1.24 to 1.87)	**0.72** (0.58 to 0.90)
				≤35mm (25^{th} centile)	36% (32 to 39)	75% (74 to 76)	6% (6 to 6)	96% (95 to 97)	1.42 (1.06 to 1.91)	0.86 (0.73 to 1.01)
				≤33mm (15^{th} centile)	22% (20 to 24)	85% (83 to 86)	6% (6 to 6)	96% (95 to 97)	1.44 (0.95 to 2.17)	0.92

Study	Population	Preterm birth incidence	Time of screen	Cut-off	Accuracy for prediction of spontaneous preterm birth <37 weeks (95% CI)					
					Sensitivity	Specificity	PPV	NPV	LR+	LR-
										(0.82 to 1.03)
				≤31mm (5th centile)	11% (11 to 12)	95% (94 to 96)	9% (8 to 9)	96% (95 to 97)	2.20 (1.19 to 4.07)	0.93 (0.87 to 1.01)
				≤28mm (1st centile)	3% (3 to 4)	99% (98 to 99)	**10%** (9 to 11)	96% (95 to 97)	**2.52** (0.78 to 8.15)	0.98 (0.94 to 1.02)

***van der Ven reported only +LRs for 4 distinct categories. Absolute numbers with SPTB and term birth (including iatrogenic) were reported which allowed test performance data for continuous categories to be calculated by the reviewer. See appendix 3 for contingency tables.**

Table 7: QUADAS-2 assessment of diagnostic accuracy studies

Study	Risk of bias by domain				Applicability to review question
	Patient selection	Index test	Reference Standard	Flow and timing	
Esplin et al 2017[6]	Low	Low	Low	Low	Unclear (population) Low (index test, reference standard)
van der Ven et al 2015[9]	Low	Low	Low	Low	Low
Banos et al 2018[7]	Low	High (thresholds)	Low	Unclear (drop-out)	Low (population, reference standard) High (Index test)
Kuusela et al 2015[8]	Unclear (recruitment)	High (thresholds)	Unclear (definition)	High (drop-out)	High (population, index test) Unclear (reference standard)

The evidence available does not indicate that cervical length measurement would be a reliable screening test to predict risk of preterm birth in low-risk women.

Of the 4 studies, the Esplin[6] and van der Ven[9] cohort studies provide the best quality evidence. Both are large studies that included nulliparous women who are inherently at low-risk having no history of preterm birth. The van der Ven study also benefited from including and separately analysing low-risk multiparous women. Both studies measured cervical length in the mid-trimester (16 to 22/23 weeks) at around the time of the routine anomaly scan when high-risk women may be selectively tested in current practice, and assessed the standard ≤25mm threshold. The Esplin study demonstrated extremely poor sensitivity of 8% using this threshold, while 84% of screen-positives would be false (PPV 16%). This was not a useful test according to the definition of LR+ >5 and LR- <0.2, with LR+ 3.67 and LR- 0.94. Lowering the threshold to ≤20mm did not improve performance. The van der Ven study found similarly poor performance for the ≤25mm cut-off. They did find that lowering the cut-off to ≤20mm resulted in fewer false positives (PPV 31% and LR+27 for nulliparous women; PPV 50% and LR+ 37 for multiparous). However, these results are based on only 10 nulliparous and 6 multiparous women who had cervical length ≤20mm in this study, and are very likely to be inaccurate, as evident by the extremely wide confidence intervals (see table 6). van der Ven also tested higher cut-offs of ≤35mm and ≤30mm, finding that both gave LR+ less than 5. The 35mm cut-off gave the best sensitivity, but this still identified only around 1 in 3 nulliparous and 1 and 6 low-risk multiparous women who gave birth preterm. The overall AUC for cervical length measure in the van der Ven study was 0.61 for nulliparous and 0.56 for multiparous women, which is little better than chance.‡

The Esplin study[6] also measured cervical length at the later 22 to 30^{+6} gestation when a greater number of women would be expected to demonstrate cervical shortening, as was demonstrated in this study (n=216 in the mid-trimester increased to n=624 at >22 weeks'). Testing at this later gestation gave better test performance, with LR+ of 5.42 for measure ≤20mm. This is at the threshold for a useful test, but the PPV still shows only 1 in 5 screen-positives experienced preterm birth. LR- and sensitivity were also very poor, identifying only 1 in 6 women who had preterm birth. Maximum sensitivity was only 23% using the 25mm cut-off, with the trade-off of lower LR+ and specificity.§ The AUC using the combination of both mid-trimester and later gestational measures was 0.67** in the Esplin study, so only

‡ For prediction of preterm birth <34 weeks van der Ven et al found improved but still poor sensitivity for nulliparous (33.1 for ≤35mm,10.8 for ≤30mm, AUC 0.63) and multiparous women (23.6 for ≤35mm, 9.1 for ≤30mm, AUC 0.58).
§ For prediction of preterm birth <32 weeks, Esplin et al found improved sensitivity (maximum 52% with ≤25mm at 22 to 30^{+6} weeks) but still very poor PPV (maximum 8.6% with ≤20mm at 16 to 22^{+6} weeks).
** The combination of both cervical length and fFN measures in the Esplin study did not improve upon cervical length measure alone with the same AUC 0.67.

marginally better than seen by van der Ven for an equivalent population of nulliparous women.

Both studies had a low risk of bias aside from the potential applicability issue of the Esplin study (nulliparous only and 30% Black/Hispanic women). Ten percent of screen-positive women in both studies (those with length ≤25mm in Esplin and ≤30mm in van der Ven) received prophylactic progesterone. However, this is expected to have had minimal effect on test performance results, as demonstrated by Esplin et al who performed sensitivity analysis on this basis (AUC increased only to 70 for non-treated women rather than 0.67 when including all women).

Two further studies[7, 8] also assessed cervical length measurement in the mid-trimester. However, both were of poor quality as diagnostic accuracy studies. The researchers did not pre-specify cut-offs, but instead constructed receiver operating curves (ROC) to identify the optimal cut-off. The 2 studies found optimal test performance using higher cut-offs at the 40th centile (≤37.9mm) in the Banos cohort study[7] and 35th centile (≤37mm) in the Kuusela cohort study[8]. At these cut-offs sensitivity was improved but was still inadequate for a screening test, at 73% and 53%, respectively. Meanwhile specificity (61%, 65%), PPV (7.5%, 6%) and LR+ (both <2) were very poor indicating there would be a high number of false positives.†† It is difficult to know whether these results would apply to other populations.

Neither of these 2 studies could give reliable assessment of the standard ≤25mm threshold due to small numbers and high drop-out. The Banos study included only 532 women (70% of those recruited‡‡) with only 5 women having cervical length ≤25mm. Three of these 5 had preterm birth giving a very high LR+ 35 for this cut-off, but like the van der Ven study, the extremely wide confidence intervals demonstrate the imprecision in this estimate. The Kuusela study had larger initial sample size but had to exclude 7/11 women with measure ≤25mm because they participated in a treatment trial, therefore ≤28mm was the lowest measure tested. This study had further risk of bias regarding the reference standard used (spontaneous birth was not defined) and had uncertain representation and applicability (the study recruited only one quarter of the eligible antenatal population, and also included women with prior preterm birth).

---

†† For prediction of preterm birth <34 weeks, both studies found improved sensitivities of 86% and 59%, though at the same low specificities (61% and 65%).

‡‡ Exclusions were because the study was additionally assessing cervical ~~contingency~~ consistency index (CCI) alongside cervical length, and included only those with adequate imaging quality. CCI was not an index test being assessed by this evidence review update, but it demonstrated slightly better test performance than cervical length. Maximum sensitivity was 77 % and specificity 83% at cut-off 64.6% (AUC 0.84). The combination of cervical length<37.9mm *and* CCI <64.6% gave sensitivity 55% and specificity 90%, with the figures reversed for either measure (*or*, Sn 91%, Sp 54%).

Summary of Findings Relevant to Criteria 4 and 5: Criteria not met[§§]

Fetal fibronectin (fFN) measurement: criterion not met

This evidence update identified one systematic review (6 cohort studies, n=1,236) and one larger US cohort study (n=9,469) assessing fetal fibronectin (fFN) testing in asymptomatic low-risk women. Both studies measured fFN at ≥22 weeks using the standard threshold of ≥50ng/ml. Their results were inconsistent. The systematic review indicated that this test gave a high likelihood of preterm birth (LR+ 12), but that there would be no confidence in a negative screen (LR- 0.54). Pooled sensitivity was 48% but this was imprecise, ranging from 20 to 77% across the meta-analysed studies. There is greater confidence in the findings of the US study due to the larger sample size and homogenous population/methods. This study found sensitivity of only 8% and PPV 11% for the same ≥50ng/ml threshold at ≥22 weeks (LR+ 2.53, LR- 0.95). Testing at other cut-offs/gestations little improved test performance, with peak sensitivity 35% (at PPV 7%) and peak PPV 14% (at sensitivity 4%). Both the systematic review and US study had some quality and applicability limitations, including 30 to 50% of women in the studies being from non-western/Caucasian women populations.

Cervical length measurement: criterion not met

The same US study discussed above and one large Dutch cohort (n=11,943) assessed cervical length, using the standard ≤25mm cut-off measured in the mid-trimester (as used for selective testing of high-risk women). Both studies found this test identified fewer than 10% of women with preterm birth with PPV less than 30% (LR+ <4, LR- >0.9). Testing the same cut-off at later gestation (US study) or raising the cut-off to 35mm (Dutch study) achieved peak sensitivity <30% with peak PPV <15%. Two small, lower quality studies used receiver operating curves to identify optimal cut-offs for their populations of 37-38mm, which achieved higher sensitivity 50 to 75% but with very low PPV (6 to 7%). It is unknown whether these thresholds could be applied to other populations. There were also several quality issues with these studies, including small samples and high drop-out.

Similar to the last UK NSC evidence review and 2009 HTA, this evidence indicates that fFN testing and cervical length measurement are not useful to predict preterm birth in asymptomatic low-risk women (where a useful test is defined by LR+ >5 and LR- <0.2). A

[§§] **Met** -for example, this should be applied in circumstances in which there is a sufficient volume of evidence of sufficient quality to judge an outcome or effect which is unlikely to be changed by further research or systematic review.
Not Met - for example, this should be applied in circumstances where there is insufficient evidence to clearly judge an outcome or effect or where there is sufficient evidence of poor performance.
Uncertain -for example, this should be applied in circumstances in which the constraints of an evidence summary prevent a reliable answer to the question. An example of this may be when the need for a systematic review and meta-analysis is identified by the rapid review.

balance of high sensitivity and specificity is not achieved. Testing at different cut-offs and/or gestations to achieve optimal (though still inadequate) sensitivity results in poorer specificity with the majority of screen-positives being false.

No studies were identified on tests for bacterial vaginosis or home monitoring for uterine contractions.

Criterion 9 — There should be an effective intervention for patients identified through screening, with evidence that intervention at a pre-symptomatic phase leads to better outcomes for the screened individual compared with usual care.

Question 2 – What is the effectiveness of available treatments for the prevention of preterm labour, birth or associated morbidity/mortality:

- *progesterone*
- *cervical cerclage*
- *cervical pessary*
- *antibiotics for bacterial vaginosis*
- *probiotics*

Evidence to date on prophylactic treatment of asymptomatic women

The Honest et al HTA (2009)[3] reviewed the effectiveness of interventions to prevent preterm birth in asymptomatic women with singleton pregnancies. Two studies indicated that vaginal or intramuscular progesterone reduced risk of spontaneous preterm birth compared with placebo. One meta-analysis of individual patient data (IPD) also indicated that cervical cerclage may reduce risk in women with short cervix, though this evidence was mostly applicable to women with additional risk factors like prior preterm birth or mid-trimester loss. Antibiotics demonstrated potential to reduce risk of spontaneous preterm birth in women with bacterial vaginosis and intermediate flora, though there was little evidence to support the use of prophylactic antibiotics in low-risk women in general. Most non-pharmacological measures such as bed rest, education or supplements did not demonstrate benefit. The available evidence across all tests was generally of poor quality, with small sample sizes precluding reliable assessment of most outcomes and a variable population mix.

The 2015 UK NSC evidence review drew from the conclusions of the 2009 HTA and searched for further evidence on the effect of these interventions to prevent preterm birth in the same population of asymptomatic, low-risk women. The 2015 UK NSC evidence review identified 2 randomised controlled trials which found that vaginal progesterone reduced risk of very preterm birth in low-/mixed-risk asymptomatic women with short cervix. One trial found that treating 14 women with length 10-20mm would prevent one birth <33 weeks, while the other found that treating 7 women with length ≤15mm could prevent one birth <34 weeks. There was also inconsistent evidence that progesterone reduced the risk of morbidity/mortality associated with prematurity. However, there was no evidence that vaginal progesterone reduced risk of overall preterm birth <37 weeks. The optimal cervical length cut-off, treatment protocol or formulation of vaginal progesterone was unclear. One randomised controlled trial of intramuscular progesterone (17-alpha-hydroxyprogesterone

caproate, 17OHP-C) was identified, which found that it was of no benefit in reducing risk of preterm birth.

Cervical pessary was also found to be of benefit in one RCT which found that treating 5 women with cervical length ≤25mm could prevent one preterm birth <34 weeks. Conversely, 3 systematic reviews did not find that cervical cerclage reduced risk of preterm birth in low-/mixed-risk populations.

The 2015 UK NSC evidence review found conflicting evidence on the effect of antibiotics in women with abnormal vaginal flora or confirmed bacterial vaginosis. One 2011 systematic review found that clindamycin reduced risk of preterm birth <37 weeks in women with abnormal flora. A later 2013 Cochrane systematic review concluded that antibiotics did not affect risk of preterm birth and associated morbidity in women with bacterial vaginosis. Similarly, when broadening the criteria to include women with abnormal flora, 2 studies identified by the Cochrane suggested antibiotics may be of benefit, but these findings required validation.

This evidence review update aimed to see whether there was new evidence on the effectiveness of treatments to reduce risk of spontaneous preterm birth and/or associated neonatal and maternal morbidity and mortality in asymptomatic low-risk pregnant women. The focus was upon these treatments reviewed by the 2015 evidence review:

- progesterone (vaginal or intramuscular)
- cervical pessary
- cervical cerclage
- antibiotics for bacterial vaginosis
- probiotics

Eligibility for inclusion in the review

This evidence update aimed to preferentially identify randomised controlled trials comparing any of these interventions with placebo or alternative treatment in asymptomatic pregnant women. Women would ideally be selected on the basis of risk factors that may be screen-detected (for example, short cervical length, raised fFN or positive bacterial vaginosis) but otherwise have no existing risk factors for preterm birth. The following were considered to be risk factors, which may be indicators for selective or targeted screening and prophylactic management, or for alternative care pathways, as part of current antenatal care in the UK:

- multiple pregnancy
- history of spontaneous preterm labour (<34 weeks)
- history of P-PROM (<34 weeks)
- history of mid-trimester loss (from 16 to 24 weeks)

- cervical trauma
- uterine anomaly

Trials would be excluded if they evaluated interventions exclusively in such high-risk populations, even if these women had additional screen-detected risk factors (for example, trials recruiting women with history of preterm birth and short cervix).

Studies from the UK would be prioritised but studies from other representative western populations would also be eligible. Interventions could be administered at any gestation and using any formulation or dose of treatment. The outcomes of interest were spontaneous preterm birth or labour (excluding iatrogenic), maternal or neonatal morbidity/mortality associated with preterm birth, and adverse events.

In the absence of randomised controlled trials, the plan was to move down through the evidence hierarchy, secondly reviewing comparative cohort studies in eligible populations. Systematic reviews of randomised controlled trials (or comparative cohorts in their absence) would also be eligible for inclusion.

Further detail on the inclusion and exclusion process following appraisal of the available evidence is discussed below.

Description of the evidence

Database searches yielded a total 3,304 results across both questions. A total 164 were judged to be potentially relevant to this question at first sift by the information specialist, 36 of which were selected for full text appraisal by the main reviewer.

A hierarchical approach was taken to full text appraisal, considering firstly any systematic reviews of randomised controlled trials in eligible populations, followed by individual randomised controlled trials. Due to the availability of RCT evidence, the reviewers did not move further through the evidence hierarchy to include comparative cohort studies.

A total of 8 studies met inclusion criteria that were conducted in asymptomatic, otherwise low-risk/general pregnant women with risk factors that may be identified through screening. Seven of the studies recruited women with short cervical length, while one included women positive for bacterial vaginosis. Seven studies compared the intervention with placebo/no treatment, while one trial had active comparators (progesterone vs pessary). The included studies were as follows:

- vaginal progesterone: 1 systematic review[10] and 1 RCT[11]
- cervical pessary: 1 systematic review[12] and 2 RCTs[13, 14]

- vaginal progesterone vs pessary: 1 RCT[15]
- cervical cerclage: 1 systematic review[16]
- antibiotics for bacterial vaginosis: 1 RCT[17]

No randomised controlled trials of intramuscular or oral progesterone or probiotics met eligibility criteria.

No RCTs were identified that evaluated interventions in otherwise low-risk, asymptomatic women where treatment was indicated on the basis of raised fFN or uterine contractions detected through home-monitoring.

The 28 studies not selected for inclusion are listed in Appendix 2 (Table 23) alongside the reason for exclusion. A couple of exclusions are worthy of note. The search retrieved 4 randomised controlled trials that were included by the selected systematic reviews evaluating progesterone (Norman 2016[30]), pessary (Nicolaides 2016[31] and Hui 2013[32]) and cervical cerclage (Otsuki 2016[33]). In these cases the systematic reviews, which meta-analysed their data alongside other trial data in equivalent populations, were considered to provide the most comprehensive evidence on these interventions and were prioritised for inclusion.

A number of systematic reviews were also identified for each intervention. Unless there was reason for inclusion of more than one systemic review per intervention (for example, different population covered or method of administration) a single systematic review was prioritised for inclusion. Selection was typically on the basis of systematic review quality, population applicability and search date. Like question 1, systematic reviews were also excluded if all included trials were published prior to 2013 and the review provided narrative synthesis of the results only. In such cases the individual study results would have been covered by the 2015 UK NSC evidence review (or by the Honest et al HTA) and in the absence of meta-analysis the review was not considered to be providing new evidence.

Other studies excluded from this question either at abstract level (if study design and method were clearly non-applicable) or at full text appraisal are as follows:

- studies conducted exclusively in women with other risk factors (multiple pregnancy, previous preterm birth, P-PROM, mid-trimester loss, cervical trauma or uterine anomaly)
- evaluating interventions in symptomatic women
- covering only outcomes outside of the PICO (for example, effect of probiotics on restoration of vaginal flora, maternal diabetes or cholesterol, or child atopy)
- evaluating administration of corticosteroids or tocolytics to asymptomatic women with short cervix (on the basis that treatments are usually reserved for symptomatic women

with threatened preterm birth and are not expected to be used in the context of a universal screening programme)
- studies of unlicensed drugs (for example, ulinastatin)
- studies evaluating technical aspects of treatment (for example, the best approach for cervical cerclage)
- secondary treatment of women who have not responded to primary treatment (for example, women with progressive cervical shortening)
- predictors of treatment failure
- population comparisons (for example, the treatment response among White compared with African American women)
- quasi-randomised controlled trials, cohorts, case series or case studies
- non-systematic reviews

The findings from the 8 included studies are discussed below by intervention. Tables 8 to 17 present the key data from each of the studies alongside summarised quality appraisal. Complete data extraction and quality appraisal for each of the 8 included publications is presented in Appendix 3.

Presentation and discussion of findings

Vaginal progesterone

Table 8. Effect of progesterone on risk of spontaneous preterm birth and associated neonatal/maternal morbidity

Study	Eligibility	Population	Intervention	Comparator	Additional treatments (both groups)	Outcomes	
						Spontaneous preterm birth (SPTB)	Associated morbidity/mortality and adverse effects
Romero et al (2018)[10] SR and MA of individual patient data (IPD)	RCTs in asymptomatic women with singleton pregnancy and cervical length (CL) ≤25mm IPD collected for this relevant subgroup only in trials with mixed eligibility.	5 RCTs n=974 women 38% White 39% Black 19% Asian 45% nulliparous 30% prior spontaneous preterm birth (SPTB) 76% with CL 10-20mm Fonesca (2007) International, 5 centres, n=250 with CL ≤15mm: IPD for n=226 O'Brien (2007) International, 53 centres, n=659 women with prior SPTB: IPD for n=31 Cetingoz (2011) Turkey, single centre, n=160 with prior SPTB or uterine malformation: IPD for n=8	Vaginal progesterone (n=498) Dose range 90 to 200mg daily From mean 22^{+6} weeks (range 18 to 25) up to 34 to 37 weeks.	Placebo (n=476)	None reported.	Progesterone reduced the risk preterm birth* at all gestations <36 weeks but not 37 weeks (all graded as high quality evidence) *NB not specified as SPTB; which was only collected for the primary outcome and <34 weeks Preterm <33 weeks (primary outcome) 14% vs 22% Relative risk (RR) 0.62 (95% CI 0.47 to 0.81); p=0.0006; I^2 0%; Number needed to treat (NNT) 12 (SPTB <33 weeks 12% vs 17% RR 0.70 (0.51 to 0.95); p=0.02 I^2 0%; NNT 19) Preterm <34 weeks 17% vs 26% RR 0.65 (0.51 to 0.83); p=0.0006; I^2 0%; NNT 11 (SPTB <34 weeks 15% vs 20% RR 0.72 (0.55 to 0.95); p=0.02; I^2 0%; NNT 18) Preterm <32 weeks 12% vs 19%	**Neonatal** Progesterone reduced the risk of (all graded as high quality evidence): Low birthweight (<2,500g) 29% vs 36% RR 0.82 (95% CI 0.68 to 0.98); p=0.03; I^2 0%; NNT16 Very low birthweight (<1,500g) 10% vs 16% RR 0.62 (0.44 to 0.86); p=0.004; I^2 0%; NNT16 Neonatal Intensive Care Unit (NICU) admission 17% vs 25% RR 0.68 (0.53 to 0.88); p=0.003; I^2 0%; NNT 13 Respiratory distress syndrome (RDS): 5% vs 10% RR 0.47 (0.27 to 0.81); p=0.007; I^2 0%; NNT 18 Composite neonatal morbidity/mortality* 8% vs 14% RR 0.59 (0.38 to 0.91); p=0.02; I^2 0%; NNT18 * RDS, IVH, NEC, sepsis, neonatal death

Study	Eligibility	Population	Intervention	Comparator	Additional treatments (both groups)	Outcomes	
						Spontaneous preterm birth (SPTB)	**Associated morbidity/mortality and adverse effects**
		Hassan (2011) International, 44 centres, n=465 women with CL 10-20mm: IPD for n=458. Norman (2016) UK and Sweden, 66 centres, n=1,228 with prior SPTB; or CL≤25mm; or positive fFN plus other risk factors: IPD for n=251				RR 0.64 (0.48 to 0.86); p=0.003; I^2 0%; NNT 14 <28 weeks 8% vs 11% RR 0.67 (0.45 to 0.99); p=0.04; I^2 0%; NNT 27 <36 weeks: 28% vs 35% RR 0.80 (0.67 to 0.97); p=0.02; I^2 0%; NNT14 No effect on preterm birth <37 weeks: 38% vs 42% RR 0.90 (0.77 to 1.05); p=0.19; I^2 0% Progesterone reduced risk of preterm birth <33 weeks in subgroup analysis for women with no history of SPTB (n=686): RR 0.65 (95% CI 0.45 to 0.94) (no further detail).	No significant effect on (moderate quality evidence): Sepsis (proven) Perinatal mortality Apgar score <7 at 5 minutes Mechanical ventilation No significant effect on (low quality evidence): Necrotizing enterocolitis (NEC) Intraventricular haemorrhage (IVH) Bronchopulmonary dysplasia Retinopathy of Prematurity (RoP) Neonatal mortality Fetal mortality Congenital anomaly Child neurodevelopment at 2 years (see appendix for data) **Maternal** No difference in any maternal events (not specified) (moderate quality evidence): 12% vs 11% RR 1.21 (0.87 to 1.69); p=0.26; ; I^2 5% (moderate quality evidence)
Van Os et al 2015[11] Multicentre RCT, The Netherlands (treatment trial of van der Ven[9])	Asymptomatic, nulliparous, women with singleton pregnancy, no history of SPTB (<34 weeks) and cervical length (CL) ≤30mm when screened	n=80 69% White ethnicity 11% prior cervical surgery or uterine anomaly 6% bacterial vaginosis	Vaginal progesterone (n=41) Dose 200mg daily From 22 to 34 weeks'	Placebo (n=39)	None reported.	Progesterone had no effect on SPTB (secondary outcomes) <37 weeks 15% (6/41) vs 13% (5/39) RR 1.17 (95% CI 0.39 to 3.52)	**Neonatal** Progesterone had no effect on the composite neonatal outcome* (primary outcome): 5% (2/41) vs 11% (4/39) RR 0.47 (95% CI 0.09 to 2.4) *defined as NEC, IVH, RDS, bronchopulmonary dysplasia, sepsis (proven), neonatal death

Study	Eligibility	Population	Intervention	Comparator	Additional treatments (both groups)	Outcomes	
						Spontaneous preterm birth (SPTB)	**Associated morbidity/mortality and adverse effects**
	at routine anomaly scan	mean CL: 26mm From n=20,234 screened • n=375 with one measure ≤30mm • n=151 confirmed at 2 weeks • n=80 agreed participation				<34 weeks 7% (3) vs 10% (4) RR 0.73 (0.17 to 3.06) <32 weeks 2% (1) vs 8% (3) RR 0.33 (0.04 to 2.99)	No effect on NICU admission: 7% (3) vs 13% (5) RR 0.53 (0.12 to 2.25) **Maternal** No difference in reported adverse effects 12% (4) vs 23% (7) RR 0.51 (0.16 to 1.6)

Table 9: CASP assessment of systematic review

Study	Are the review findings valid?	Are the results clear and precise?	Will the results help locally?	Cochrane risk of bias assessment for 5 included trials (by review authors)
Romero et al 2018[10]	Yes	Yes	Some uncertainty (mixed population, limited analysis of SPTB outcomes)	Low risk for randomisation, allocation concealment, blinding, incomplete outcome data, selective reporting. 1 trial (Norman) with 'other' risk of bias (related to compliance 63% vs 69% placebo and attrition for childhood outcomes).

Table 10: Cochrane risk of bias assessment of randomised controlled trials

Study	Random sequence (selection bias)	Allocation concealment (selection bias)	Blinding of participant and personnel (performance bias)	Blinding of outcome assessment (detection bias)	Incomplete outcome assessment (attrition bias)	Selective reporting (reporting bias)	Other bias
Van Os et al 2015[11]	+	+	+	+	+	+	-

+Low risk of bias; - High risk; ? Unclear risk

The 2015 UK NSC evidence review identified 3 systematic reviews of vaginal progesterone in the management of low-/mixed-risk asymptomatic women with short cervix, which were predominantly based on the findings of the Hassan et al (2011) and Fonesca et al (2007) trials. The Romero meta-analysis of IPD[10] identified by this review updates this evidence, incorporating IPD from these trials in addition to data from the large OPPTIMUM trial (Norman et al 2016[30)]. The Romeo review and its included studies were of high methodological quality. There was low heterogeneity for nearly all outcomes giving confidence in the pooled findings. The results are applicable to a potential screening scenario where short cervix is detected at the time of the routine anomaly scan (mean 22 weeks') and progesterone is given until around term.

The review found high quality evidence (as graded by the study authors) that vaginal progesterone reduced the risk of preterm birth at all gestations less than 36 weeks' compared with placebo. The primary outcome of the review and included trials was the rate of very preterm birth <33 weeks, with absolute rates of 14% in the progesterone group vs 22% in the placebo group (RR 0.62, 95% CI 0.47 to 0.81, I^2 0%[***]). The number of women needing treatment (NNT) to prevent one very preterm birth was 12. However, this was for

[***] Although the meta-analysis for risk of preterm birth <33 weeks was significant with low heterogeneity, the risk reductions crossed the threshold of statistical significance in only 1 of the 3 large studies: Hassan, RR 0.55 (95% CI 0.33 to 0.92); Fonseca, RR 0.60 (95% CI 0.36 to 1.00); Norman 0.74 (95% CI 0.48 to 1.12)

the risk of very preterm birth in general, including iatrogenic preterm births. The effect of progesterone specifically on spontaneous preterm birth (SPTB) <33 weeks was still statistically significant, but slightly attenuated with only a 5% absolute risk reduction (12% vs 17%) and higher NNT of 19. Progesterone had statistically significant effect on the rate of preterm birth (in general) at later gestations of 35 and 36 weeks with NNT of 12 and 14, but there is no analysis for SPTB specifically for these gestations. Therefore, it is unclear whether they would remain statistically significant. There was also no effect on the rate of preterm birth (in general) by the standard definition of <37 weeks (rate 38% with progesterone vs 42% with placebo). This lack of benefit for overall preterm birth is the same as had been found in the 2015 NSC evidence review. Therefore, although it can generally be concluded from these findings that progesterone reduces the risk of preterm birth in low-/mixed-risk, asymptomatic women with short cervix, the effect appears to be modest, particularly if restricting this to spontaneous births only, or considering late preterm births.

There are a few other limitations to the Romero review, including population applicability. The meta-analysed population were all women with short cervix, but a third had history of preterm birth. Romero et al performed subgroup analysis for the primary outcome of preterm birth <33 weeks according to history of preterm birth. They found this had minimal influence on the effectiveness of progesterone, with 35% relative risk reduction for women with short cervix and no history of preterm birth (RR 0.65, 95% CI 0.45 to 0.94) compared with similar 41% risk reduction for women with short cervix and history preterm birth (RR 0.59, 95% CI 0.40 to 0.88). However, the review authors conducted no further analyses for low-risk women without history of preterm birth. Absolute risk reduction or NNT are not given, there is no analysis of the effect on moderate-to-late preterm birth (>33 weeks), and no analysis for SPTB specifically, rather than overall preterm birth.

Another potential applicability issue of this review is that only 38% of the analysed population were of White/Caucasian ethnicity. The review authors performed subgroup analysis on this basis, which found that progesterone was effective in White women but not in those of Black, Asian or another ethnicity. It is difficult to be sure from this analysis alone whether ethnicity may be a true mediator of progesterone effectiveness. Romero et al also analysed whether other factors such as age, BMI, cervical length or progesterone dose may mediate effectiveness. They found a non-significant interaction for all subgroups tested, suggesting that none of these factors were associated with progesterone efficacy. However, there was different effectiveness within some of the individual subgroups. For example, by cervical length, progesterone only demonstrated benefit in women with very short cervix of length10 to 20mm. However, as these women accounted for three-quarters of the sample, there is greater uncertainty around the effect estimates for women with length <10mm or 20 to 25mm. As with ethnicity, this makes it difficult to be sure whether cervical length is a true mediator of progesterone effectiveness and there is an optimal length to treat. Notably,

despite the variation in dose used across the studies, progesterone doses 90-100mg and 200mg were both equally effective in preventing preterm birth <33 weeks.

Aside from rate of preterm birth, Romero et al also found high quality evidence that vaginal progesterone reduced the risk of neonatal morbidity outcomes of low birthweight, admission to neonatal intensive care, respiratory distress syndrome, and the composite neonatal outcome. There was moderate evidence that progesterone was not associated with maternal side effects, and moderate-to-low quality evidence that it had no effect on neonatal mortality and other neonatal outcomes (for example, intracranial haemorrhage or retinopathy of prematurity). The evidence for individual neonatal morbidity and mortality outcomes is graded low quality primarily due to the small numbers of cases. This may have reduced statistical power to estimate the true effect of progesterone. This point is the central limitation to the following RCT identified by this evidence update.

van Os et al (2015)[11] was a placebo-controlled RCT of vaginal progesterone for low-risk women with short cervix, that was an extension to the Dutch screening cohort[9] covered by question 1. It was excluded by Romero et al because their review only included RCTs where progesterone was given to women with cervical length ≤25mm and where the primary outcome was effect on preterm birth. The van Os RCT included women with length ≤30mm and assessed neonatal morbidity as the primary outcome. The trial was included in this evidence update for completeness, but it provides very low quality evidence on the effect of progesterone. The trial was well-conducted and included an applicable, low-risk, screen-detected, western population. However, although the screened population was large (n=11,943), the prevalence of short cervix was lower than expected (n=375). This sample size was further reduced because women were only eligible for the trial if they had measure ≤30mm confirmed 2 weeks later (n=151) only half of whom (n=80) chose to participate. Consequently, the trial was highly underpowered to detect a difference in neonatal morbidity, recruiting only 4% of the required sample size of 1,920. Furthermore, treatment compliance was very poor with only half of the sample taking over 80% of the prescription. Therefore, although this trial found that progesterone had no effect on neonatal outcomes or preterm birth (secondary outcomes) there is little confidence in the findings.

The direct-comparison trial of progesterone vs pessary (Cruz-Melguizo 2018[15]) is covered below.

Cervical pessary

Table 11. Effect of pessary on risk of spontaneous preterm birth and associated neonatal/maternal morbidity

Study	Eligibility	Population	Intervention	Comparator	Additional treatments (both groups)	Outcomes	
						Spontaneous preterm birth (SPTB)	Associated morbidity/mortality and adverse effects
Saccone et al 2017[12] SR and MA	RCTs in asymptomatic women with singleton pregnancy and cervical length (CL) ≤25mm	3 RCTs n=1,420 Goya (2012), Spain, n=380, 11% prior preterm birth Hui (2013), China, n=108, 8% prior preterm birth Nicolaides (2016), Multicentre, n=932, 17% prior preterm birth	Arabian pessary From 20 to 24^{+6} weeks (mean 22) up to 37 weeks	Expectant management	Vaginal progesterone (200mg to week 33^{+6}): 25% of review population used in Nicolaides for women with CL ≤15mm (n=359, 39% of study population) use not reported by Goya or Hui	Pessary had no effect on the risk of SPTB at any gestation. <34 weeks (primary outcome): 10.2% vs 14.6% RR 0.71 (95% CI 0.21 to 2.42); I^2 90% <37 weeks: 20.2% vs 50.2% RR 0.50 (95% CI 0.23 to 1.09); I^2 0% Goya and Hui only (n=488) <32 weeks: 9.9% vs 7.5% RR 1.32 (95% CI 0.87 to 2.01) Nicolaides only (n=932) <28 weeks: 4.4% vs 4.8% RR 0.70 (95% CI 0.18 to 2.67); I^2 72% P-PROM: 3.7% vs 10.2% RR 0.39 (95% CI 0.09 to 1.71); I^2 72% Goya and Hui only (n=488)	**Neonatal** Pessary had no effect on any outcome: Low birthweight Average birthweight Necrotizing enterocolitis (NEC) Respiratory distress syndrome (RDS) Intraventricular haemorrhage (IVH) Neonatal intensive care (NICU) Neonatal mortality Perinatal mortality (see appendix for data) **Maternal** Pessary was associated with increased vaginal discharge: 37.3% vs 18.0% RR 2.12 (95% CI 1.84 to 2.44); I^2 0% No difference in bacterial vaginosis 25.8% vs 22.8% RR 1.14 (95% CI 0.95 to 1.36); I^2 0%
Saccone et al 2017[14] Single centre RCT, Italy	Asymptomatic women with singleton pregnancy and cervical length (CL) ≤25mm at routine anomaly scan Exclusions: prior preterm birth prior mid-trimester loss	n=300 89% White 70% nulliparous 4% prior cervical surgery mean CL: 12mm	Arabian pessary From 18 to 23^{+6} weeks (mean 22) up to 37 weeks	Expectant management	Vaginal progesterone (200mg to week 37): 89% pessary and 83% of controls recommended or women with CL ≤20mm Antibiotics (not specified):	Pessary reduced the risk of SPTB at 34 and 37 weeks. <34 weeks (primary outcome): 7.3% vs 15.3% RR 0.48 (95% CI 0.24 to 0.95); p=0.04 <37 weeks: 20.0% vs 32.7% RR 0.61 (95% CI 0.41 to 0.91); p=0.02	**Neonatal** Pessary was associated with difference in: Average birthweight: 2889.9 vs 2644.6 grams MD 245.3 (95% CI 69.2 to 421.4); p=0.006 NICU admission rate: 10.0% vs 18.7% RR 0.54 (95% CI 0.30 to 0.96); p=0.04

Study	Eligibility	Population	Intervention	Comparator	Additional treatments (both groups)	Outcomes	
						Spontaneous preterm birth (SPTB)	Associated morbidity/mortality and adverse effects
	prior P-PROM				22% pessary and 25% of controls for bacterial infection evident on vaginal swab taken at randomisation	No effect at earlier gestations. <32 weeks: 6.7% vs 9.3% RR 0.71 (95% CI 0.33 to 1.56); p=0.52 <28 weeks: 4.0% vs 6.0% RR 0.67 (95% CI 0.24 to 1.83); p=0.60	Rate of the composite perinatal outcome: 14.7% vs 32.0% RR 0.46 (95% CI 0.29 to 0.72); p=0.01 defined as ≥1 of NEC, IVH, RDS, bronchopulmonary dysplasia, retinopathy of prematurity (ROP), sepsis, neonatal death No effect on neonatal or perinatal mortality. **Maternal** Pessary was associated with increased vaginal discharge: 86.7% vs 46.0% RR 1.88 (95% CI 1.57 to 2.27); p<0.001 No difference in rate of chorioamnionitis or pelvic discomfort.
Dugoff et al 2018 5 centre RCT, US	Asymptomatic women with singleton pregnancy and cervical length (CL) ≤25mm following screening at routine anomaly scan Exclusions: prior preterm birth prior mid-trimester loss	n=118 61% Black 66% nulliparous mean CL: 18mm	Bioteque pessary From 18 to 23^{+6} weeks (mean 21) up to 37 weeks	Expectant management	Vaginal progesterone (200mg to week 37): 84% pessary and 91% of controls recommended for women with CL ≤20mm	Pessary had no effect on the risk of SPTB at any gestation. <37 weeks (primary outcome): 38.3% vs 32.8% RR 1.17 (95% CI 0.72 to 1.91); p=0.59 <34 weeks: 31.7% vs 25.9% RR 1.22 (95% CI 0.69 to 2.17); p=0.55 <28 weeks: 18.3% vs 20.7% RR 0.89 (95% CI 0.43 to 1.85); p=0.82	**Neonatal** Pessary had no effect on any outcome: Average birthweight Sepsis NEC IVH ROP RDS Bronchopulmonary dysplasia NICU admission Neonatal mortality Intrauterine mortality Composite of above (see appendix for data) **Maternal** Pessary was associated with increased vaginal discharge: 73.3% vs 48.3% RR 1.48 (95% CI 1.15 to 1.89); p=0.002 No difference in genitourinary infection or chorioamnionitis.

Study	Eligibility	Population	Intervention	Comparator	Additional treatments (both groups)	Outcomes	
						Spontaneous preterm birth (SPTB)	**Associated morbidity/mortality and adverse effects**
Cruz-Melguizo et al 2018[15] Multicentre non-inferiority RCT, Spain	Asymptomatic women with singleton pregnancy and cervical length (CL) ≤25mm following screening at routine anomaly scan Exclusions: ≥3 prior preterm births uterine abnormality prior cervical biopsy/excision	n=246 77% White 46% nulliparous 11% prior preterm birth mean CL: 21mm	Arabian pessary From 20^{+1} to 23^{+6} weeks (mean 21) up to between 34^{+4} and 37 weeks	Vaginal progesterone 200mg From 20^{+1} to 23^{+6} weeks (mean 21) up to between 34^{+4} and 37 weeks	Antibiotics (not specified): 29% of pessary and 24% of progesterone groups For women with positive bacterial culture at time of randomisation	Pessary had no effect on the risk of SPTB at any gestation. <34 weeks (primary outcome): 14% vs 14% Risk difference (RD): 0.11% (95% CI -8.85% to 8.62%) <37 weeks: 22% vs 21% RD 0.41 (-9.90 to 10.73) <28 weeks 8% vs 8% RD 0.37 (-6.38 to 7.12) P-PROM <37 weeks 10% vs 9% RD 0.28 (-7.08 to 7.64) P-PROM <34 weeks 6% vs 6% RD -0.33 (-6.20 to 5.53)	**Neonatal** Pessary had no effect on any outcome: Average/low birthweight Sepsis NEC IVH ROP RDS NICU admission Neonatal mortality Intrauterine mortality **Maternal** No difference in overall reporting of adverse effects: 16% vs 11%; p=0.27 (no further detail reported) Pessary was associated with increased: vaginal discharge: 87% vs 71%; p=0.002 vaginal discomfort: 27% vs 3%; p<0.001 increased emergency department visits (first month only): 25% vs 15%; p<0.05 No difference in reported infections, pain or sexual activity. Only 3% of pessary group reported to require removal for tolerability.

Table 12: CASP assessment of systematic review

Study	Are the review findings valid?	Are the results clear and precise?	Will the results help locally?	Cochrane risk of bias assessment for 3 included trials (by review authors)
Saccone et al 2017[12]	Yes	Yes	Some uncertainty related to mixed populations	Low risk for randomisation, allocation concealment, incomplete outcome data, selective reporting. No studies blinded, though not considered feasible. One trial (Hui) rated to have 'other' risk of bias.

Table 13: Cochrane risk of bias assessment of randomised controlled trials

Study	Random sequence (selection bias)	Allocation concealment (selection bias)	Blinding of participant and personnel (performance bias)	Blinding of outcome assessment (detection bias)	Incomplete outcome assessment (attrition bias)	Selective reporting (reporting bias)	Other bias
Saccone et al 2017[14]	+	+	-	+	+	+	+
Dugoff et al 2018[13]	+	?	-	-	+	+	-
Cruz-Melguizo et al 2018[15]	+	+	-	-	+	+	+

+Low risk of bias; - High risk; ? Unclear risk

The Saccone systematic review (pooling 3 RCTs)[12] and 2 subsequent RCTs[13, 14] have assessed the use of pessary in low-/mixed-risk asymptomatic women with singleton pregnancies with short cervical length (≤25mm) detected by screening at the time of the routine anomaly scan. All trials assessed the use of pessary from around 18 and 24 weeks' gestation continued to near term (compared with expectant management). As such, the findings are wholly applicable to a potential screening programme. The review and all included trials were also of high methodological quality. Lack of blinding of study participants or personnel to treatment allocation was the only consistent limitation across trials. However, with the potential exception of maternal side effects, it is not expected that this would have introduced much risk of bias with such objective outcomes as preterm birth and serious neonatal morbidity. Two of the RCTs (Dugoff[13] and Hui in the meta-analysis) had other risk of bias due to small sample size (n=118 and n=108) which may have reduced their statistical power to detect a difference in outcomes.

The evidence is overall conflicting. The Saccone meta-analysis[12] and Dugoff RCT[13] found that pessary had no effect on the risk of preterm birth. However, the Saccone RCT[14] found that pessary reduced the relative risk of preterm birth <37 weeks by 39% compared with expectant management (absolute risk reduction 12.7%). Risk of birth preterm birth <34

weeks was reduced by 52% (absolute risk reduction 8%). One of the 3 trials included in the Saccone meta-analysis[12] (Goya 2012) had also found a benefit of pessary, with 76% relative risk reduction for births <34 weeks at a large absolute risk reduction of 21% (preterm rate 6% rate vs 27%).†††

It is difficult to find explanation for the inconsistent findings in terms of the included trial populations or use of additional treatment. The findings of the Saccone meta-analysis[12] are driven primarily by the large Nicolaides trial (n=932), which found no benefit of pessary. Nicolaides included a mixed-risk population (17% with history of preterm birth) and used additional progesterone for 39% (all of whom had very short cervix <15mm). The rate of preterm birth <34 weeks was 12% in the pessary vs 11% in the control group. The Saccone RCT,[14] which conversely found a benefit of pessary, included low-risk women (no history of preterm birth) but most had very short cervix (mean length 12mm) and 86% were prescribed progesterone. The rate of preterm <34 week was 7% in the pessary compared with 15% among controls. Saccone et al[14] considered that one potential reason why they found a benefit of pessary when the Nicolaides trial did not, may be because pessary is beneficial only for women who have very short cervix but no history of prior preterm birth.

However, the Goya trial included in the meta-analysis, which like the Saccone RCT[14] found a benefit of pessary, does not fit with this explanation. This trial included a mixed population (11% with history of preterm birth) and did not report use of progesterone or high prevalence of short cervix. Goya did though have an unusually high rate of preterm birth: 27% of women in the control group having preterm birth <34 weeks (6% with pessary) and 60% having preterm birth <37 weeks (22% with pessary). The reasons for the high prevalence of preterm birth in this trial population are not explained. The small Dugoff RCT[13] was the only trial with preterm rates similar to this. However, despite having compatible baseline population to the Saccone RCT (women with no history of preterm birth but very short cervix and prescribed additional progesterone) it did not find a benefit of pessary. Therefore, neither does this trial fit with the proposed suggestion that pessary is only beneficial for otherwise low-risk women with very short cervix. Some additional variables that may have also influenced the effect of pessary, include screening for concomitant infection and use of antibiotics (Saccone RCT), rate of nulliparity, extent of training in pessary insertion and follow-up assessments, type of pessary (Bioteque in Dugoff, standardly used for uterine prolapse), or ethnic representation (majority African-American in Dugoff and Chinese in the Hui trial of the Saccone review). It is not possible to know what effect these factors may have had.

††† At the time of the 2015 evidence review, only this single Goya et al trial (2012) had been available; hence the last review had concluded the potential benefit of pessary in reducing preterm birth risk.

One additional RCT[15] adds further to the mixed pattern of results. Cruz-Melguizo et al directly compared pessary with vaginal progesterone in a similar population to the trials comparing against expectant management: women with cervical length ≤25mm identified at the time of the routine anomaly scan (11% with preterm birth). The rate of preterm birth <34 was equivalent among women given pessary or progesterone: 14% in both treatment arms. One interpretation from this trial could be that the trial provides evidence in support of the findings of the Saccone and Goya RCTs and imply that pessary is effective in preventing preterm birth (and as good as progesterone). However, without an untreated control group for comparison, it could also be the case that neither treatment made a difference to what would have been the preterm rate in this population. It is not possible to know which may be the case. Overall with the current level of evidence it is not possible to conclude whether or not pessary may reduce the risk of preterm birth for some or all low-risk women identified to have short cervical length in the mid-trimester.

It is also not possible to know whether pessary may affect the risk of associated neonatal morbidity. The Saccone trial[14] found that pessary was associated with reduced risk of neonatal complications. None of the other trials demonstrated a benefit on neonatal outcomes. However, with the small sample sizes and multiple outcomes assessed it is expected that trials were underpowered to reliably detect difference in rarer neonatal outcomes. A consistent finding across all trials, though, was that pessary was associated with increased reports of vaginal discharge. There was no evidence that this was associated with increased infection rates. However, there appears to have been minimal assessment (or reporting) of tolerability or acceptability of pessary across trials, which may be worthy of further assessment.

Cervical cerclage

Table 14. Effect of cervical cerclage on risk of spontaneous preterm birth and associated neonatal/maternal morbidity

Study	Eligibility	Population	Intervention	Comparator	Additional treatments (both groups)	Outcomes	
						Spontaneous preterm birth (SPTB)	Associated morbidity/mortality and adverse effects
Berghella et al 2017[16] SR and MA of IPD	RCTs in asymptomatic women with singleton pregnancy and cervical length (CL) <25mm and no prior history of SPTB	5 RCTs n=419 Rust (2001), US, n=105, CL <25mm at 16 to 24 weeks Althuisius (2001) Netherlands, n=9, CL <25mm at 14 to 27 weeks To (2004), Multicentre, n=209, CL ≤15mm at 22 to 24 weeks Berghella (2004), US, n=21, CL <25mm at 14 to 24 weeks Otsuki (2016), Japan, n=75, CL <25mm at 16 to 26 weeks All excluding prior history SPTB including mid-trimester loss.	Cerclage McDonald 3 studies Shirodkar 1 study Either 1 study From 14 to 27 weeks (mean 22) up to 36 to 37 weeks	No cerclage	Varied across trials including antibiotics, anti-inflammatories (indomethacin) and bed rest, with tocolytics given to the cerclage group in 1 study. Progesterone not used in any trial.	Cerclage had no effect on the risk of SPTB at any gestation. <35 weeks (primary outcome): 21.9% (49/224) vs 27.7% (54/195) RR 0.88 (95% CI 0.63 to 1.23); I^2 0% <37 weeks: 36.2% (81/224) vs 41.0% (80/195) RR 0.93 (0.73 to 1.18); I^2 57% <34 weeks: 20.1% (45/224) vs 25.1% (49/195) RR 0.89 (0.63 to 1.27); I^2 0% <28 weeks: 11.6% (26/224) vs 11.3% (22/195) RR 1.15 (0.68 to 1.93); I^2 0% P-PROM: 20.5% (34/166) vs 13.6% (23/169) RR 1.52 (0.94 to 2.46); I^2 0% Subgroup analysis found significant effect of cerclage in: Women with CL <10mm 39.5% (30/76) vs 58.0% (29/50); RR 0.68 (0.47 to 0.98); I^2 0% Women given tocolytics:	**Neonatal** Cerclage had no effect on any outcome: Low birthweight Average birthweight Neonatal intensive care (NICU) Neonatal mortality NEC* RDS* IVH* Sepsis* (* rates assessed/reported for only n=30 newborns; see appendix for data) **Maternal** None assessed/reported

Study	Eligibility	Population	Intervention	Comparator	Additional treatments (both groups)	Outcomes	
						Spontaneous preterm birth (SPTB)	Associated morbidity/mortality and adverse effects
						17.5% (20/114) vs 32.7% (18/55); RR 0.54 (0.31 to 0.93) Women given antibiotics: 18.3% (20/109) vs 31.5% (17/54); RR 0.58 (0.33 to 0.98)	

Table 15: CASP assessment of systematic review

Study	Are the review findings valid?	Are the results clear and precise?	Will the results help locally?	Cochrane risk of bias assessment for 5 included trials (by review authors)
Berghella et al 2017[16]	Yes	No – imprecise and overall low quality evidence	Yes – applicable population Unclear – lack of maternal outcomes	Low risk for randomisation, allocation concealment, incomplete outcome data, selective reporting. No studies blinded, though not considered feasible. 2 trials (Althuisius and Otsuki) with 'other' risk of bias (not described).

This Berghella 2017 systematic review[16] pools the evidence available to date on the effect of cervical cerclage in asymptomatic women with singleton pregnancies, who have short cervical length (<25mm) but who are otherwise low risk, with no history or prior preterm birth or mid-trimester loss. The trials were also mostly of representative Western countries, applicable population to a potential UK screening programme.

The review adds one additional trial (Otsuki 2016) to the 4 trials available at the time of the authors' 2005 systematic review and IPD meta-analysis on the effect of cerclage in women with short cervix[34] (included in the Honest et al HTA), and their 2010 analysis of the same 4 trials[35] which had investigated in more depth whether the degree of cervical shortening influences effectiveness (included in the 2015 UK NSC review). The additional trial has not altered the findings. In line with the conclusions of both the HTA and last UK NSC review, the 2017 updated IPD found that cervical cerclage had no effect on the risk of spontaneous preterm birth or of associated neonatal morbidity in women with short cervix.

The 5 RCTs included in the 2017 update were rated by Berghella et al[16] to be at low risk of bias, again with the exception that blinding of participants and personnel was not feasible. Two of the trials were rated to have other risk of bias which was not described by the review. The studies varied in certain aspects, including the gestation at screening, cervical length of participants, type of cerclage and stitch used, and additional care given alongside. Despite this variation, there was low heterogeneity in the results, with none of the individual trials finding that cerclage had an effect on preterm birth.

This should add confidence to the results, though the total number of participants was relatively small at 419. Some of the pooled risk estimates are imprecise, particularly for certain neonatal outcomes, which were only assessed for a sample of 30. Therefore, the overall strength of evidence on cerclage remains low and there could still be potential for type 2 error, where a benefit of cerclage upon preterm birth or associated morbidity has been missed. The IPD meta-analysis also indicated there could be a potential benefit in

women with very short cervix (<10mm), and where antibiotics and tocolytics had been used. At this stage, these findings require validation and need to be interpreted with caution. However, it may be that future study may find a benefit of cerclage in certain subgroups. Type of cerclage used was not found to be associated with effect in this review.

There has also been an apparent lack of assessment (or reporting) of maternal adverse effects or acceptability of cerclage in trials to date, which may be worthy of further study.

Antibiotics for bacterial vaginosis

Table 16. Effect of antibiotics for bacterial vaginosis on risk of spontaneous preterm birth and associated neonatal/maternal morbidity

Study	Eligibility	Population	Intervention	Comparator	Additional treatments (both groups)	Outcomes	
						Spontaneous preterm birth (SPTB)	Associated morbidity/mortality and adverse effects
Subtil et al 2018[17] Multicentre RCT, France	Screen-positive asymptomatic women with bacterial vaginosis (Nugent score ≥7) detected at ≤14 weeks' gestation Exclusions: prior preterm birth prior mid-trimester loss	n=2,869 52% nulliparous 2% with multiple pregnancy 55% of n=5,246 screen-positives with no history of SPTB of n=84,530 women screened	1. Single course oral clindamycin (4 days of 300mg twice daily): n=943 2. Triple course oral clindamycin (4 days of 300mg twice daily, once a month for 3 months): n=968 Treatment from mean 12^{+4} weeks' gestation.	3. Identical placebo: n=958	Other antibiotics if indicated (not specified): 17.5% (no difference between groups)	Clindamycin had no effect on SPTB at any gestation. Primary outcome: late miscarriage to very preterm delivery (16 to 32 weeks): 1.2%* clindamycin vs 1.0% placebo RR 1.10 (95% CI 0.53 to 2.32); p=0.82 SPTB at <37 weeks: 4.8%* vs 4.1% RR 1.17 (0.81 to 1.69); p=0.40 P-PROM <37 weeks: 2.2%* vs 1.9% RR 1.18 (0.65 to 2.13); p=0.57 *Data for the single and triple course clindamycin groups combined; intervention groups were equivalent (see appendix for data)	**Neonatal** Clindamycin had no effect on any outcome: Low birthweight Average birthweight NICU admission Sepsis Need for ventilation Neonatal mortality Intrauterine death **Maternal** Clindamycin increased the risk of: Overall adverse effects: 3.0% vs 1.3%; p=0.0035 Diarrhoea: 1.6% vs 0.4%; p=0.0071 Abdominal pain: 0.5% vs 0; p=0.034 Incomplete treatment/non-compliance: 19.6% vs 16.3%; p=0.031 Clindamycin had no effect on: Chorioamnionitis Fever during or after labour Need for antibiotics after delivery

Table 17: Cochrane risk of bias assessment

Study	Random sequence (selection bias)	Allocation concealment (selection bias)	Blinding of participant and personnel (performance bias)	Blinding of outcome assessment (detection bias)	Incomplete outcome assessment (attrition bias)	Selective reporting (reporting bias)	Other bias
Subtil et al 2018[17]	+	+	+	+	+	+	? (compliance and power)

+Low risk of bias; - High risk; ? Unclear risk

This evidence update identified only a single trial[17] of antibiotic treatment for bacterial vaginosis. This was a large, high quality trial of a representative western population of asymptomatic pregnant women who screened positive for bacterial vaginosis during the first trimester but were otherwise low risk with no history of preterm birth. This trial found that clindamycin, given as either a single or triple course from the start of the second trimester, made no difference to the risk of spontaneous preterm birth. Neither was there an effect on associated neonatal morbidity. Antibiotics were, however, associated with significantly increased risk of adverse effects, specifically diarrhoea and abdominal pain, and with a higher rate of non-compliance. Although notably, side effects of antibiotics were still fairly rare at 3%.

The trial had low risk of bias across domains with the exception of a couple of uncertainties. The prevalence of the primary outcome of mid-trimester loss or very preterm birth <32 weeks was lower than expected at 1% (prior research had indicated 4% prevalence among untreated women with bacterial vaginosis). As such it is unclear whether the trial may have been underpowered and could have missed a true effect of clindamycin upon the primary outcome (type 2 error). The authors do not comment upon this. However, given the large sample size large and the lack of effect for overall preterm birth <37 weeks (prevalence 4.8% with antibiotics vs 4.1% placebo) the chance that a true effect has been missed seems less likely.

Compliance could be another potential issue. This was around 80% by self-report at each follow-up visit, but only 49% based on returned pill packs. It is uncertain whether compliance may have reduced effectiveness of the intervention. However, per protocol analysis using the most conservative estimate of 49% still found no effect of clindamycin. Additionally, this level of compliance may be representative of what may be achieved in standard clinical practice

As the Subtil et al acknowledge, their results conflict with the results of the 2011 systematic review[36] (included by the 2015 UK NSC review) which found that clindamycin given prior to

22 weeks' gestation for women with abnormal vaginal flora reduced the risk of late miscarriage and preterm birth in women. However, for diagnosed bacterial vaginosis specifically, these findings are consistent with the latest Cochrane systematic review[19] (also included by the 2015 evidence review) which found no effect of antibacterial treatment (any) upon risk of preterm birth. No trials have been published since 2013 assessing the standard UK treatment for bacterial vaginosis of oral metronidazole, or of either metronidazole or clindamycin administered vaginally. Therefore, these treatments could potentially be studied further to confirm they have no effect on the risk of preterm birth in asymptomatic women.

Summary of Findings Relevant to Criterion 9‡‡‡

This evidence update identified a total of 8 eligible studies looking at treatments to prevent risk of preterm birth in women identified to have risk factors through screening: 1 SR with IPD MA and 1 RCT assessing vaginal progesterone; 1 SR and 2 RCTs assessing cervical pessary; 1 RCT comparing progesterone and pessary; 1 SR with IPD MA assessing cervical cerclage; and 1 large RCT assessing antibiotics for bacterial vaginosis.

The studies on progesterone, pessary and cerclage all assessed prophylactic treatment given from the time of screen-detection of short cervix in the mid-trimester (16 to 24 weeks) up until near term. In the antibiotic trial, treatment was given following screen-detection of bacterial vaginosis in the first trimester. No studies were identified where treatment was indicated on the basis of raised fetal fibronectin.

The evidence was overall of good quality and applicable to a potential UK screening programme. The main limitations were that some studies included mixed-risk populations, and that studies would be too small to reliably detect an effect on rarer preterm and neonatal outcomes.

‡‡‡ **Met** -for example, this should be applied in circumstances in which there is a sufficient volume of evidence of sufficient quality to judge an outcome or effect which is unlikely to be changed by further research or systematic review.
Not Met - for example, this should be applied in circumstances where there is insufficient evidence to clearly judge an outcome or effect or where there is sufficient evidence of poor performance.
Uncertain -for example, this should be applied in circumstances in which the constraints of an evidence summary prevent a reliable answer to the question. An example of this may be when the need for a systematic review and meta-analysis is identified by the rapid review.

Vaginal progesterone: criterion uncertain

One SR and IPD found high quality evidence that vaginal progesterone was associated with a modest reduction in the risk of preterm birth at all gestations <36 weeks (NNT of around 12 to 14), but had no effect on overall preterm birth <37 weeks. There was limited assessment of the outcome of spontaneous preterm birth specifically, or in the subgroup of women with no history of preterm birth. There was also high quality evidence that vaginal progesterone reduced the risk of neonatal morbidity outcomes of low birthweight, admission to neonatal intensive care, respiratory distress syndrome, and the composite neonatal outcome. There was moderate-to-low quality evidence that vaginal progesterone had no effect on other neonatal outcomes. There was no new evidence on the effect of intramuscular or oral progesterone. These findings are essentially unchanged from the 2015 UK NSC evidence review, which was based on most of the same evidence. A further meta-analysis of IPD[18] is awaited which will compare any type and dose of progesterone, and may provide more comprehensive evidence on the effect of progesterone, including by population subgroup.

Cervical pessary: criterion uncertain

It is uncertain whether cervical pessary may benefit women with short cervix. Only a single trial was available at the 2015 UK NSC evidence review (which found a benefit). Four trials have since been published comparing with expectant management, and one trial comparing pessary with progesterone. The results are conflicting, with some finding a benefit of pessary on risk of preterm birth and others not. The effect on risk of associated neonatal morbidity or mortality was also inconsistent across studies. Future IPD MA may help to understand whether variables such as cervical length, history of preterm birth or existing infection, could have an influence on effect. All trials were, however, unanimous in finding pessary increased reports of vaginal discharge, though there was minimal further assessment of maternal tolerability or acceptability.

Cervical cerclage: criterion not met

As with the conclusions of the 2015 UK NSC review, the latest systematic review on cervical cerclage found that it had no effect on the risk of preterm birth for otherwise low-risk women with short cervix. There was also no effect on any neonatal outcomes reported. Trials to date have also performed limited assessment of maternal adverse effects or acceptability of cerclage.

Bacterial vaginosis: criterion not met

One large, high-quality trial in otherwise low-risk women with bacterial vaginosis found that oral clindamycin (single or triple course) had no effect on risk of preterm birth or any neonatal outcomes assessed. It did, however, increase the risk of gastrointestinal adverse effects. This evidence is consistent with the 2013 Cochrane review[19] (included in

the 2015 UK NSC evidence review) which found that antibiotics (any) for bacterial vaginosis had no effect on preterm birth risk. There were no studies on probiotics.

None of the identified evidence on any intervention assessed whether treating women with short cervix or bacterial vaginosis who went on to have full term birth (that is false positives) was associated with any negative effects (such as psychological outcomes).

Overall the findings are compatible with the last UK NSC evidence review, finding that vaginal progesterone may benefit women with short cervix. Further meta-analysis of patient data may help to understand whether maternal characteristics, formulation or dose are associated with effect. Similarly, future meta-analysis may help to resolve whether there is an effect of pessary in any subgroup. However, the evidence seems to suggest no benefit of cerclage or antibiotics for low-risk women.

Review summary

Conclusions and implications for policy

The evidence to support a universal screening programme to routinely screen all pregnant women for risk of preterm birth and associated neonatal and maternal morbidity is not currently available. As such, the findings do not indicate that a change to the current policy should be made and systematic population screening should not be recommended.

Screening tests

Fetal fibronectin (fFN) measurement

This evidence update identified one systematic review[5] and one prospective US cohort study [6] that assessed fetal fibronectin (fFN) testing and 4 prospective cohort studies[6-9] assessing cervical length measurement. None of the studies indicated that fFN would be a reliable screening test to predict risk of spontaneous preterm birth (<37 weeks) in the low-risk/general population of asymptomatic women with singleton pregnancies.

For fFN screening, both studies tested the standard ≥50ng/ml threshold measured at ≥22 weeks' gestation, giving inconsistent results. The systematic review[5] (n=1,236 women) found that a positive screen indicated a high likelihood of preterm birth (LR+ 12), but that there would be no confidence in a negative screen (LR- 0.54). Pooled sensitivity was 48%, but this was imprecise ranging from 20 to 77% across the meta-analysed studies. There is greater confidence in the findings of the single US cohort study[6] due to the larger sample size (n=9,469) and homogenous population/methods. This study found sensitivity of only 8% and PPV 11% for the same ≥50ng/ml threshold at ≥22 weeks (LR+ 2.53, LR- 0.95). Testing at other cut-offs/gestations little improved test performance, with peak sensitivity 35% (at PPV 7%) and peak PPV 14% (at sensitivity 4%). Both the systematic review and US study did have some quality and applicability limitations, including 30 to 50% of women in the studies being from non-western/Caucasian women populations.

Cervical length measurement

For cervical length screening, a US study[6] and a large Dutch study[9] (n=11,943) tested the standard ≤25mm cut-off measured in the mid-trimester (as used for selective testing of high-risk women). Both studies found this test identified fewer than 10% of women with preterm birth with PPV less than 30% (LR+ <4, LR- >0.9). Testing the same cut-off at later gestation (US study) or raising the cut-off to 35mm (Dutch study) achieved peak sensitivity

<30% with peak PPV <15%. Two lower quality studies[7, 8] used receiver operating curves to identify optimal cut-offs for their populations of 37-38mm, which achieved higher sensitivity 50 to 75% but at very low PPV (6 to 7%). It is unknown whether these thresholds could be applied to other populations. There were also several quality issues with these studies, including small samples and high drop-out.

Similar to the last UK NSC evidence review and 2009 HTA, this evidence indicates that fFN testing and cervical length measurement are not useful to predict preterm birth in asymptomatic low-risk women (where a useful test is defined by LR+ >5 and LR- <0.2). A balance of high sensitivity and specificity is not achieved. Testing at different cut-offs and/or gestations to achieve optimal (though still inadequate) sensitivity results in poorer specificity with the majority of screen-positives being false.

This evidence update did not identify studies looking at screening for bacterial vaginosis or home monitoring for uterine contractions as screening tests.

Interventions

This evidence update identified a total of 8 eligible studies in applicable western populations looking at the treatment of pathologies that might increase the risk of preterm birth in women identified to have risk factors through screening:

- 1 systematic review (SR) with meta-analysis of individual patient data (IPD MA)[10] and 1 randomised controlled trial (RCT)[11] assessing vaginal progesterone;
- 1 SR[12] and 2 RCTs[13, 14] assessing cervical pessary; 1 RCT[15] comparing progesterone and pessary;
- 1 SR with IPD MA[16] assessing cervical cerclage; and 1 large RCT[17] assessing antibiotics for bacterial vaginosis.

The studies on progesterone, pessary and cerclage all assessed prophylactic treatment given from the time of screen-detection of short cervix in the mid-trimester (16 to 24 weeks) up until near term. In the antibiotic trial, treatment was given following screen-detection of bacterial vaginosis in the first trimester. This evidence update identified no studies were treated was indicated on the basis of fetal fibronectin measurement.

The evidence was overall of good quality and applicable to a potential UK screening programme. The main limitations were that some studies included mixed-risk populations (including some with previous preterm birth), and that studies would be too small to reliably detect an effect on rarer preterm and neonatal outcomes.

Vaginal progesterone

Vaginal progesterone was associated with a modest reduction in the risk of preterm birth at all gestations <36 weeks, with numbers needed to treat (NNT) of around 12 to 14. There was no effect on overall preterm birth <37 weeks. There was also limited assessment of spontaneous preterm birth specifically (excluding medically-indicated). The single analysis conducted for the primary outcome (very preterm birth <33 weeks') on this basis indicated that the effect could be attenuated when considering spontaneous preterm births only (NNT 19 rather than 12). The main analyses were also for the mixed-risk/general antenatal population, with limited assessment specific to low-risk women with short cervix but no history of preterm birth. There was also evidence that vaginal progesterone reduced the risk of neonatal morbidity outcomes of low birthweight, admission to neonatal intensive care, respiratory distress syndrome, and the composite neonatal outcome. There was moderate-to-low quality evidence that vaginal progesterone had no effect on other neonatal outcomes. There was no new evidence on the effect of intramuscular or oral progesterone.

These findings on progesterone are essentially unchanged from the 2015 UK NSC evidence review, which was based on most of the same evidence. A further meta-analysis of IPD[18] is awaited which will compare any type and dose of progesterone, and may provide more comprehensive evidence on the effect of progesterone, including confirming whether there is an effect in otherwise-low risk women.

Cervical pessary

It is uncertain whether cervical pessary may benefit women with short cervix. Only a single trial was available at the 2015 UK NSC evidence review (which found a benefit). Four RCTs have since been published comparing with expectant management and one RCT comparing pessary with progesterone. The results are conflicting, with some finding a benefit of pessary and others not. The effect on risk of associated neonatal morbidity or mortality was also inconsistent across studies. However, even the trials finding a benefit showed little consistency in their findings or study populations, some of which included low-risk women only, while others included those with existing risk factors for preterm birth. Future IPD MA may help to understand whether variables such as cervical length, history of preterm birth or existing infection, could have an influence on effect. All trials were, however, unanimous in finding that pessary increased reports of vaginal discharge, though the prevalence and risk increase was again inconsistent across studies. There was minimal other assessment of tolerability or acceptability which may be beneficial.

Cervical cerclage

As with the conclusions of the 2015 UK NSC review, the latest systematic review on cervical cerclage found that it had no effect on the risk of preterm birth or associated neonatal morbidity in otherwise low-risk women with short cervix. There was also no effect

on any neonatal outcomes reported. Trials to date have also performed limited assessment of maternal adverse effects or acceptability of cerclage.

Bacterial vaginosis

One large, high-quality trial in otherwise low-risk women with bacterial vaginosis found that oral clindamycin (single or triple course) had no effect on risk of preterm birth or associated neonatal morbidity. It did, however, increase the risk of gastrointestinal adverse effects, though the prevalence of side effects was still low at 3% among treated women. There was no evidence available on the standard UK treatment of oral metronidazole. There were no studies on probiotics. This evidence is consistent with the last Cochrane review[19] (included in the 2015 UK NSC evidence review) which found that antibiotics (any) for bacterial vaginosis had no effect on preterm birth risk.

None of the identified evidence on any intervention assessed whether treating women with short cervix or bacterial vaginosis who went on to have full term birth (that is false positives) was associated with any negative effects (such as psychological outcomes).

Overall, the findings are in line with the 2015 UK NSC evidence review, finding that vaginal progesterone might have the potential to reduce risk of preterm birth in otherwise low-risk women found to have short cervix through screening in the mid-trimester. However, the poor test performance of cervical length measurement and/or cervicovaginal fetal fibronectin testing to reliably detect which asymptomatic, low-risk women are at risk of spontaneous preterm birth would appear to preclude universal screening at the current time.

Evidence uncertainties

Further meta-analysis of individual patient data for progesterone and for cervical pessary may help to confirm whether or not these treatments are effective specifically in otherwise low-risk women with short cervix who have no existing risk factors for preterm birth. Future IPD may similarly help to clarify whether variables such as degree of cervical shortening, presence of infection, or method of treatment (for example dose or device) have an influence on effectiveness.

It may be beneficial to review the evidence on whether screening of asymptomatic, low-risk women (and subsequent management) reduces risk of preterm birth and associated morbidity compared with not screening, or is associated with any harms.

Future studies may also wish to explore whether other screening tests used as an alternative to, or in combination with cervical length or fFN testing (for example, measuring cervical consistency or cervical incompetence) may have potential as screening tests and demonstrate improved test performance.

Limitations

This was a rapid evidence review process. The search strategy was built on a protocol developed *a priori* for each of the 2 key questions. Searching was limited to 3 literature databases and did not include grey literature resources. Studies only available in non-English language, editorials, abstracts, conference reports or poster presentations were not included. The reviewers were also unable to contact study authors or review non-published material.

Appendix 1 — Search strategy

Electronic databases

The search strategy included searches of the databases shown in Table 18.

Table 18. Summary of electronic database searches and dates

Database	Platform	Searched on date	Date range of search
PubMed	PubMed.gov	17 September 2019	1946 to search date
Embase	Embase.com	17 September 2019	1974 to search date
The Cochrane Library, including: Cochrane Database of Systematic Reviews (CDSR) Cochrane Central Register of Controlled Trials (CENTRAL)	Wiley Online	17 September 2019	CDSR: Issue 9 of 12, September 2019

Search Terms

Search terms included combinations of free text and subject headings (Medical Subject Headings [MeSH] for PubMed, and Emtree terms for Embase), grouped into the following categories:

- Condition: Preterm labour
- Index tests: Cervical length measurement, Cervicovaginal fetal fibronectin, Uterine contraction, Tests for bacterial vaginosis)
- Interventions: Antibiotics for bacterial vaginosis, Probiotics, Vaginal or intramuscular progesterone, Cervical pessary, Cervical cerclage
- Diagnostic accuracy: sensitivity and specificity

Search terms for PubMed are shown in Table 19, for Embase in Table 20, and search terms for the Cochrane Library databases are shown in Table 21.

Table 19. Search strategy for PubMed

Term Group	#	Search terms	Results

Condition	1	(((((preterm labor[Text Word] OR preterm labour[Text Word] OR preterm birth*[Text Word] OR preterm deliver*[Text Word])) OR (pre-term labor[Text Word] OR pre-term labour[Text Word] OR pre-term birth*[Text Word] OR pre-term deliver*[Text Word])) OR (spontaneous labor[Text Word] OR spontaneous labour[Text Word] OR spontaneous birth*[Text Word] OR spontaneous deliver*[Text Word])) OR (premature labor[Text Word] OR premature labour[Text Word] OR premature birth*[Text Word] OR premature deliver*[Text Word])) OR (pre-mature labor[Text Word] OR pre-mature labour[Text Word] OR pre-mature birth*[Text Word] OR pre-mature deliver*[Text Word])	45030
Condition	2	(morbid*[Text Word] OR mortal*[Text Word])	1276220
Condition	3	((("Obstetric Labor, Premature"[Mesh]) OR "Maternal Mortality"[Mesh]) OR "Fetal Mortality"[Mesh]) OR "Morbidity"[Mesh]	558254
Condition	4	(#1 or #2 or #3)	1729242
Index tests	5	((((((cervi* length[Text Word]) OR cervico* fibronectin*[Text Word]) OR cervico* secretion*[Text Word]) OR (fetal fibronectin*[Text Word] OR foetal fibronectin*[Text Word])) OR ffn protein[Text Word]) OR uterine contraction*[Text Word]) OR bacterial vagin*[Text Word]	15262

Index tests	6	"FFN protein, human" [Supplementary Concept] OR "Fibronectins/analysis"[Mesh] OR "Vaginosis, Bacterial"[Mesh] OR "Uterine Contraction"[Mesh] OR "Cervical Length Measurement"[Mesh]	15980
Index tests	7	(#5 or #6)	20400
Diagnostic accuracy	8	(((accuracy[Title/Abstract] OR sensitiv*[Title/Abstract] OR specificity[Title/Abstract] OR diagnos*[Title/Abstract])) OR prognostic value[Title/Abstract]) OR test performance*[Title/Abstract]	3930054
Diagnostic accuracy	9	"Diagnosis"[Mesh] OR "Diagnostic Screening Programs"[Mesh] OR "Diagnostic Techniques, Obstetrical and Gynecological"[Mesh] OR "Diagnostic Imaging"[Mesh] OR "Sensitivity and Specificity"[Mesh] OR "diagnosis" [Subheading]	9258587
Diagnostic accuracy	10	(#8 or #9)	11056714
Condition AND Index tests AND Diagnostic accuracy Date/Language limit	11	(#4 and #7 and #10) Filters: Publication date from 2013/01/01 to 2020/12/31; English	828
Intervention	12	((prevent*[Text Word]) OR treat*[Text Word]) OR intervention*[Text Word]	7629528
Intervention	13	((pessar*[Text Word] OR probiotic*[Text Word] OR cerclage[Text Word] OR antibiotic*[Text Word] OR anti-biotic*[Text Word] OR progesterone[Text Word])) OR (anti-bacterial*[Text Word] OR antibacterial*[Text Word])	695961

Intervention	14	"Progesterone"[Mesh] OR "Probiotics"[Mesh] OR "Anti-Bacterial Agents"[Mesh] OR "Cerclage, Cervical"[Mesh] OR "Pessaries"[Mesh]	447510
Index tests	15	#12 or #13 or #14	8029104
Condition AND Index tests AND Intervention	16	#4 and #7 and #15	2493
Diagnostic Accuracy OR Intervention	17	#11 or #16	3838
Animal studies	18	"Animals"[Mesh] NOT ("Animals"[Mesh] AND "Humans"[Mesh])	4629995
Remove animal studies	19	#17 NOT #18	3647
Date and language limit	20	#19 Filters: Publication date from 2013/01/01 to 2020/12/31; English	1007

Table 20. Search strategy for Embase

Term Group	#	Search terms	Results
Condition	#1	((preterm OR 'pre term' OR premature OR 'pre mature' OR spontaneous*) NEAR/3 (birth* OR labor* OR labour* OR deliver*)):ti,ab,kw	75.421
Condition	#2	morbid*:ti,ab,kw OR mortal*:ti,ab,kw	1,327,291
Condition	#3	'immature and premature labor'/exp	147,347
Condition	#4	'fetus mortality'/exp	3,300
Condition	#5	'maternal mortality'/exp	21,737
Condition	#6	‘maternal morbidity'/exp	8,142
Condition	#7	'fetal morbidity'/exp	14
Condition	#8	#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7	1,482,329
Index tests	#9	(cervi* NEAR/1 length):ti,ab,kw	2,863
Index tests	#10	cervico* NEAR/4 (fibronectin* OR secretion*)):ti,ab,kw	580
Index tests	#11	(ffn NEAR/1 protein):ti,ab,kw	3
Index tests	#12	(uterine NEAR/1 contraction*):ti,ab,kw	5,382

Index tests	#13	(bacterial NEAR/1 vagin*):ti,ab,kw	5,409
Index tests	#14	'fibronectin'/exp	46,730
Index tests	#15	cervical length measurement'/exp	2,342
Index tests	#16	'uterus contraction'/exp	9,919
Index tests	#17	'bacterial vaginosis'/exp	94
Index tests	#18	'vaginitis'/exp	15,982
Index tests	#19	#9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18	78,311
Diagnostic accuracy	#20	accuracy:ti,ab OR sensitiv*:ti,ab OR specificity:ti,ab OR diagnos*:ti,ab	5,195,088
Diagnostic accuracy	#21	(prognostic NEAR/2 value):ti,ab	72,904
Diagnostic accuracy	#22	(test NEAR/2 performance*):ti,ab	19,392
Diagnostic accuracy	#23	'diagnosis'/de	1,356,463
Diagnostic accuracy	#24	'diagnostic imaging'/de	172,385
Diagnostic accuracy	#25	'gynecological examination'/exp	98,006
Diagnostic accuracy	#26	sensitivity and specificity'/exp	333,589
Diagnostic accuracy	#27	'prognostic value'/exp	76
Diagnostic accuracy	#28	#20 OR #21 OR #22 OR #23 OR #24 OR #25 OR #26 OR #27	6,047,063
Condition AND Index test AND Diagnostic accuracy search Date/Language limit	#29	#8 AND #19 AND #28 AND [2013-2019]/py AND [english]/lim	1,328
Intervention	#30	prevent*:ti,ab,kw	1,778,099
Intervention	#31	treat*:ti,ab,kw	7,232,524
Intervention	#32	pessar*:ti,ab,kw OR probiotic*:ti,ab,kw OR cerclage:ti,ab,kw OR antibiotic*:ti,ab,kw OR 'anti biotic*':ti,ab,kw OR progesterone:ti,ab,kw	566,567
Intervention	#33	'anti bacterial*':ti,ab,kw OR antibacterial*:ti,ab,kw	101,684
Intervention	#34	progesterone'/de	95,227
Intervention	#35	'probiotic agent'/de	32,734
Intervention	#36	'antiinfective agent'/exp	3,494,077

Intervention	#37	'uterine cervix cerclage'/exp	2,118
Intervention	#38	'vagina pessary'/exp	2,687
Intervention	#39	'antibiotic agent'/exp	1,467,866
Intervention	#40	'anti-biotic*':ti,ab,kw OR antibiotic*:ti,ab,kw	435,983
Intervention	#41	#30 OR #31 OR #32 OR #33 OR #34 OR #35 OR #36 OR #37 OR #38 OR #39 OR #40	10,552,989
Condition AND Index test AND Intervention Date/Language limit	#42	#8 AND #19 AND #41 AND [2013-2019]/py AND [english]/lim	2,243
Diagnostic accuracy OR Intervention	#43	#29 OR #42	2,841

Table 21. Search strategy for the Cochrane Library Databases (Searched via the Wiley Online platform)

Term Group	#	Search terms	Results
Condition	#1	((preterm OR pre-term OR premature OR pre-mature OR spontaneous*) NEAR/3 (birth* OR labor* OR labour* OR deliver*)):ti,ab,kw	7755
Condition	#2	morbid*:ti,ab,kw OR mortal*:ti,ab,kw	108380
Condition	#3	MeSH descriptor: [Obstetric Labor, Premature] explode all trees	1849
Condition	#4	MeSH descriptor: [Fetal Mortality] explode all trees	2
Condition	#5	MeSH descriptor: [Maternal Mortality] explode all trees	111
Condition	#6	MeSH descriptor: [Morbidity] explode all trees	14729
Condition	#7	(or #1-#6)	125341
Index tests	#8	(cervi* NEAR/1 length):ti,ab,kw	447
Index tests	#9	(cervico* NEAR/4 (fibronectin* OR secretion*)):ti,ab,kw	40
Index tests	#10	ffn protein*:ti,ab,kw	4
Index tests	#11	(uter* NEAR/1 contraction*):ti,ab,kw	1338

Index tests	#12	(bacterial NEAR/1 vagin*):ti,ab,kw	896
Index tests	#13	MeSH descriptor: [Fibronectins] explode all trees	161
Index tests	#14	MeSH descriptor: [Cervical Length Measurement] explode all trees	59
Index tests	#15	MeSH descriptor: [Uterine Contraction] explode all trees	374
Index tests	#16	MeSH descriptor: [Vaginosis, Bacterial] explode all trees	373
Index tests	#17	[or #8-#16]	2815
Diagnostic accuracy	#18	accuracy:ti,ab OR sensitiv*:ti,ab OR specificity:ti,ab OR diagnos*:ti,ab	199838
Diagnostic accuracy	#19	(prognostic NEAR/2 value):ti,ab	2714
Diagnostic accuracy	#20	(test NEAR/2 performance*):ti,ab	2388
Diagnostic accuracy	#21	MeSH descriptor: [Diagnosis] explode all trees	320246
Diagnostic accuracy	#22	MeSH descriptor: [] explode all trees and with qualifier(s): [diagnosis - DI]	51045
Diagnostic accuracy	#23	MeSH descriptor: [Diagnostic Imaging] explode all trees	47308
Diagnostic accuracy	#24	MeSH descriptor: [Diagnostic Techniques, Obstetrical and Gynecological] explode all trees	2648
Diagnostic accuracy	#25	MeSH descriptor: [Sensitivity and Specificity] explode all trees	16207
Diagnostic accuracy	#26	[or #18-#25]	487859
Condition AND Index tests AND Diagnostic accuracy	#27	7 and #17 and #26	419
Intervention	#28	(prevent* or treat*):ti,ab,kw	885823
Intervention	#29	pessar*:ti,ab,kw OR probiotic*:ti,ab,kw OR cerclage:ti,ab,kw OR	42249

		antibiotic*:ti,ab,kw OR 'anti biotic*':ti,ab,kw OR progesterone:ti,ab,kw	
Intervention	#30	anti-bacterial*:ti,ab,kw OR antibacterial*:ti,ab,kw	12623
Intervention	#31	MeSH descriptor: [Progesterone] explode all trees	2987
Intervention	#32	MeSH descriptor: [Probiotics] explode all trees	1836
Intervention	#33	MeSH descriptor: [Anti-Bacterial Agents] explode all trees	11384
Intervention	#34	MeSH descriptor: [Cerclage, Cervical] explode all trees	53
Intervention	#35	MeSH descriptor: [Pessaries] explode all trees	176
Intervention	#36	MeSH descriptor: [] explode all trees and with qualifier(s): [prevention & control - PC]	89183
Intervention	#37	[or #28-#36]	897148
Condition AND Index tests AND Intervention	#38	7 and #17 and #37	831
Diagnostic accuracy or Intervention	#39	#27 or #38	911
CDSR only	#40	#39 with Cochrane Library publication date Between Jan 2013 and Dec 2019, in Cochrane Reviews	39
CENTRAL only	#41	#39 with Publication Year from 2013 to 2019, in Trials	430

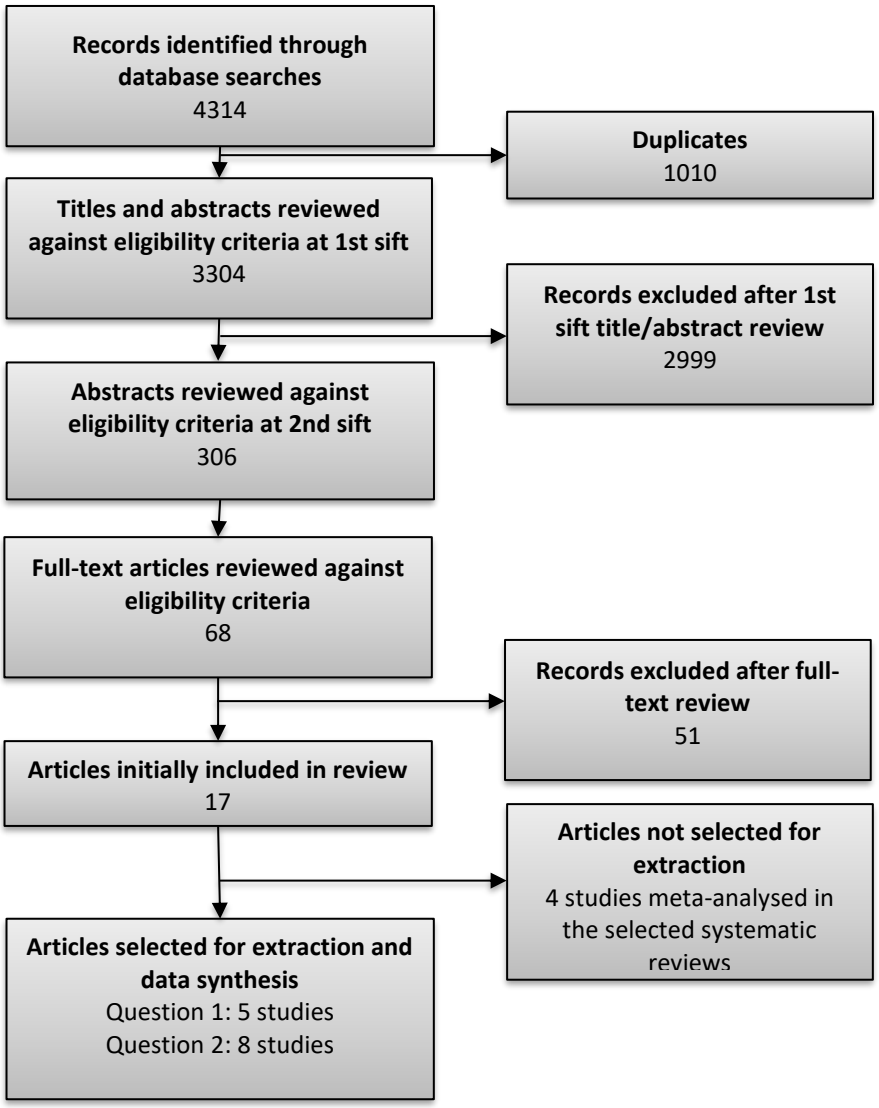
Results were imported into EndNote and de-duplicated (removing 1,010 references).

Appendix 2 — Included and excluded studies

PRISMA flowchart

Figure 2 summarises the volume of publications included and excluded at each stage of the review. Eleven publications were ultimately judged to be relevant to one or more review questions and were considered for extraction. Publications that were included or excluded after the review of full-text articles are detailed below.

Figure 1. Summary of publications included and excluded at each stage of the review



Publications included after review of full-text articles

The publications included after review of full-texts are summarised in Table 22 below.

Studies were prioritised for extraction and data synthesis. It was planned *a priori* that the following approach would be taken to prioritise studies for extraction:

1. Systematic reviews and meta-analyses were considered the highest quality of evidence and reviewed initially for applicability to each key question.
2. Following this, the reviewers moved down through the evidence hierarchy, prioritising prospective diagnostic cohort studies in applicable randomly-selected/consecutive populations for question 1, and randomised controlled trials for question 2.
3. Due to the availability of randomised controlled trials for question 2, the decision was made not to move further through the evidence hierarchy to include comparative cohort studies.
4. Studies in UK populations were prioritised if identified, followed by studies from Western populations analogous to the UK.

Table 22. Summary of publications included after review of full-text articles, and the question each publication was identified as being relevant to

Study	The condition	The test	The intervention	The screening programme	Implementation criteria	Comments
Dos Santos 2018[5]	-	Q1	-	-	4, 5	-
Esplin 2017[6]	-	Q1	-	-	4, 5	2 tests
van der Ven 2015[9]	-	Q1	-	-	4, 5	-
Banos 2018[7]	-	Q1	-	-	4, 5	-
Kuusela 2015[8]	-	Q1	-	-	4, 5	-
Romero 2018[10]	-	-	Q2	-	9	-
van Os 2015[11]	-	-	Q2	-	9	-
Saccone 2017[12]	-	-	Q2	-	9	-
Saccone 2017[14]	-	-	Q2	-	9	-
Dugoff 2018	-	-	Q2	-	9	-
Cruz-Melguizo 2018[15]	-	-	Q2	-	9	-
Berghella 2017[16]	-	-	Q2	-	9	-
Subtil 2018[17]	-	-	Q2	-	9	-

Publications excluded after review of full-text articles

Of the 68 publications reviewed at full text, 55 were not selected for inclusion: 27 identified at first sift as being potentially applicable to the diagnostic accuracy question, and 28 identified at first sift as being potentially applicable to the treatment question. This latter group included 4 randomised controlled trials, the data from which had been included by selected systematic reviews and meta-analyses. These 55 publications, along with reasons for exclusion, are listed in Table 23. The order of presentation is by question and individual test, with systematic reviews listed first, in most recent date order, followed by primary studies.

Table 23. Publications excluded after review of full-text articles

Reference	Reason for exclusion
Fetal fibronectin (fFN) testing	
Berghella V, Saccone G. Fetal fibronectin testing for reducing the risk of preterm birth. Cochrane Database Syst Rev. 2019;7:Cd006843.	Cochrane review with search date Sept 2018 including RCTs where women were screened by fFN and then subsequently randomised to a) knowledge of/disclosure of the result and intervention or b) no knowledge/intervention. Relative risk of preterm labour compared between the groups. The review could not provide evidence applicable to assessment of diagnostic accuracy. For question 2, all of the 6 included RCTs were conducted in symptomatic women in preterm labour so there were not applicable by population.
Faron G, Balepa L, Parra J, et al. The fetal fibronectin test: 25 years after its development, what is the evidence regarding its clinical utility? A systematic review and meta-analysis. J Matern Fetal Neonatal Med. 2018:1-31.	Systematic review (SR) with more recent search date (Feb 2018) than Dos Santos et al. However, the search was for fFN testing for predicting preterm birth in symptomatic or asymptomatic women. Pooled likelihood ratios are given (no other test performance data) for 6 meta-analyses in asymptomatic women, only one of which specifies low risk women (6 studies). The paper does not clarify which studies are included in these meta-analyses; the supplementary table only lists the data extracted for the 193 studies included in the whole review. Without going through all studies and assuming which may be the asymptomatic studies it is difficult to ensure these are low risk women. The Dos Santos review was therefore selected in preference despite the earlier search date (2017) having a focused search for asymptomatic women; giving greater clarity

	on which studies included women without existing risk factors and giving pooled sensitivity and specificity data for these groups.
Hezelgrave NL, Shennan AH. Quantitative fetal fibronectin to predict spontaneous preterm birth: a review. Womens Health (Lond). 2016;12(1):121-8.	Non-systematic review (uncertain design from the abstract only)
Gao L, Zhang JP, Chen H, et al. Fetal fibronectin detection for preterm birth prediction. Genet Mol Res. 2014;13(1):1323-8.	Cohort, China, n=124 women tested for fFN between 20-34 weeks including symptomatic and asymptomatic. Analysis of full group assessing test performance for birth within 7-14 days, or preterm. Excluded on basis of small, non-representative country and sample, including unclear numbers with signs of preterm labour.
Cervical length (CL) measurement	
Lim K, Butt K, Crane JM. No. 257-Ultrasonographic Cervical Length Assessment in Predicting Preterm Birth in Singleton Pregnancies. Journal of Obstetrics and Gynaecology Canada. 2018;40(2):e151-e64.	Canadian guideline recommendations based on systematic literature search to 2009. Reports data from individual pre-2013 studies but no new meta-analysis. However, the document does provide background data on practice.
Berghella V, Baxter JK, Hendrix NW. Cervical assessment by ultrasound for preventing preterm delivery. Cochrane Database Syst Rev. 2013(1):Cd007235.	Cochrane review search Aug 2012 for RCTs where women were randomised to cervical length screening or no screening; or to knowledge of the result and intervention or no knowledge/intervention. No applicability to the diagnostic accuracy question. 5 RCTs were included but none were conducted in asymptomatic women with singleton pregnancies (either multiple pregnancy, or symptomatic women with singletons). Therefore neither could the studies meet population eligibility for the treatment question. NB. At second sift it was found that the 2019 update of this Cochrane was now available (published 25 Sept just after the 17 Sept literature search date for this rapid review). The updated version could not therefore have met eligibility criteria for inclusion, but it was reviewed for interest. A single RCT (Mishra 2018, India) had been identified in 296 asymptomatic singletons without risk factors. The review gives relative risks for preterm birth and neonatal outcomes for women randomised to knowledge (and treatment) or no knowledge of results. The individual study would not have met inclusion criteria as a

	diagnostic cohort and would be excluded from the treatment question due to non-western population representation.
Saccone G, Simonetti B, Berghella V. Transvaginal ultrasound cervical length for prediction of spontaneous labour at term: a systematic review and meta-analysis. Bjog. 2016;123(1):16-22.	Inapplicable population SR search date Oct 2014 for studies assessing cervical length for predicting spontaneous labour onset within 7 days in women at term (37 weeks plus).
Barros-Silva J, Pedrosa AC, Matias A. Sonographic measurement of cervical length as a predictor of preterm delivery: a systematic review. J Perinat Med. 2014;42(3):281-93.	SR search date Dec 2012 for cohorts assessing cervical length by TVUS at 18-24 weeks (any cut-off; single or serial measure) among low or high risk asymptomatic women. Identified n=12 studies in the general pregnant population or low risk women with single pregnancies. The review reports the individual test performance results by cut-off and by gestation. All studies were conducted prior to 2013 (search date of last rapid review) and without meta-analysis the individual study results were not considered to be contributing new evidence.
Conde-Agudelo A, Romero R. Predictive accuracy of changes in transvaginal sonographic cervical length over time for preterm birth: a systematic review and metaanalysis. Am J Obstet Gynecol. 2015;213(6):789-801	SR search date June 2015 for cohorts assessing serial cervical length measures among asymptomatic women. N=14 studies met inclusion criteria but only 2 were conducted in non-high risk women with single pregnancies: one in general population (1996) and one in specifically low risk women (2007). The review reports the individual test performance results, but as the studies were conducted prior to 2013 with no meta-analysis, the individual study results were not considered to be contributing new evidence.
Rosenbloom JI, Raghuraman N, Temming LA, et al. Predictive Value of Midtrimester Universal Cervical Length Screening Based on Parity. Journal of ultrasound in medicine: official journal of the American Institute of Ultrasound in Medicine. 2019.	Secondary analysis of a retrospective, single centre cohort, US. N=13,508 women with single pregnancy and without history of preterm labour were offered CL measurement at 17-23 weeks. Women with CL <20mm treated; those measuring 20-24mm were asked to return for further measures. Gives complete test performance data separately for nulliparous and multiparous women at 4 different cut-offs beneath <25mm. Would meet eligibility criteria, except that of n=122 with length <20mm, treatment was known for n=100, 89% of whom received prophylactic treatment. Therefore treatment may be influencing the prediction of the measure for preterm birth. Sensitivity analysis was performed but only excluding the n=11 untreated, and then women who received treatment other than progesterone.

Wulff CB, Rode L, Rosthoj S, et al. Transvaginal sonographic cervical length in first and second trimesters in a low-risk population: a prospective study. Ultrasound Obstet Gynecol. 2018;51(5):604-13.	Prospective cohort in 3 centres, Denmark. n=3302 of the general population of pregnant women with single pregnancies receiving serial cervical length measures at 11-14, 19-21 and 23-24 weeks (method not specified). Gives range of LR+ for different gestations of preterm birth, for four measurement categories (all <25mm); and at four measurement times (1, 2, 3, [2 or 3]). However, of n=67 women with length <25mm, n=42 (63%) received treatment.
Hermans FJR, Koullali B, van Os MA, et al. Repeated cervical length measurements for the verification of short cervical length. Int J Gynaecol Obstet. 2017;139(3):318-23.	Secondary analysis of van der Ven. Women with length <30mm were asked to attend for a second measure. Analysed odds for preterm birth with either or both measures <30mm. No test performance data is available and there is insufficient information to construct contingency tables with certainty (for example, whether screen-negatives would also include those with first measure >30mm in addition to those with first measure <30mm who were not screened or measured >30mm on repeat).
Radhouane A, Nadia BJ, Imen K, et al. Ultrasound cervical length in predicting preterm birth: Prospective study. Australasian Medical Journal. 2017;10(8):647-55.	Single centre cohort in Tunisia measuring cervical length at 11-13 weeks in n=117 women of the general population. Does give contingency tables for preterm birth with CL<35mm. Primarily excluded on the basis of being a non-western-representative country, but also very small sample size.
Kokanali MK, Celik H, Kokanali D, et al. Predictive role of transvaginal ultrasonographic measurement of cervical length at 34 weeks for late pre-term and late-term deliveries in nulliparous women. J Matern Fetal Neonatal Med. 2016;29(11):1789-94.	Prospective cohort, Turkey. N=362 low risk women with single pregnancies screened by cervical length measure at 34 weeks (by TVUS) to predict risk of late preterm delivery (34-37 weeks). Gives test performance data but excluded for relevance as a nationwide screening programme would be unlikely to screen specifically for late-preterm deliveries, which are lower risk and may be less likely to be treated/benefit from treatment compared with moderate to extreme preterm. deliveries <34 weeks.
Papastefanou I, Pilalis A, Eleftheriades M, et al. Prediction of Preterm Delivery by Late Cervical Length Measurement after 24 Weeks. Fetal Diagn Ther. 2015;38(3):200-4.	Cross-sectional study, Greece. Cervical length measurement is routinely performed at 20-24 weeks when women with length <15mm are advised progesterone prophylaxis. n=1180 women received repeat measurement at 24-30 weeks (reportedly excluding those being treated). Assesses the median difference in cervical length measure for those delivering preterm or not at <34 or <37 weeks and gives the overall AUC. It also lists the proportion of preterm deliveries detected at different false positive

	rate. Excluded due to overall lack of clarity over absolute numbers, cut-off measures and associated test performance.
Son M, Grobman WA, Ayala NK, et al. A universal mid-trimester transvaginal cervical length screening program and its associated reduced preterm birth rate. Am J Obstet Gynecol. 2016;214(3):365.e1-5.	Inapplicable study design for diagnostic accuracy. Before-after study comparing the number of preterm births among low-risk women before (n=46,598) and after (n=17,609) implementation of cervical length measurement screening at 18-24 weeks with treatment recommended for those with short cervix (<25mm).
Cho HJ, Roh HJ. Correlation Between Cervical Lengths Measured by Transabdominal and Transvaginal Sonography for Predicting Preterm Birth. J Ultrasound Med. 2016;35(3):537-44.	Prospective cohort, Korea. N=771 low risk women with single pregnancies assessed at 20-29 weeks by transabdominal or TVUS. The primary analysis is the accuracy of transabdominal for predicting short length on TVUS. Test performance is given for both methods at two cut-offs for predicting preterm birth <34 weeks. However, there is high loss to follow-up and the analysis is based on only one third of the sample (n=241) who gave birth at this tertiary centre, which may affect representation. There is also uncertain population applicability.
Souka AP, Papastefanou I, Papadopoulos G, et al. Cervical length in late second and third trimesters: a mixture model for predicting delivery. Ultrasound Obstet Gynecol. 2015;45(3):308-12.	Cross sectional study, Greece, of n=647 with single pregnancies who had cervical length measurement 20-24 weeks. Gaussian distribution models used to describe the distribution of cervical lengths and probability of labour or birth at any given age. No test performance data
Banicevic AC, Popovic M, Ceric A. Cervical length measured by transvaginal ultrasonography and cervicovaginal infection as predictor of preterm birth risk. Acta Informatica Medica. 2014;22(2):128-32.	Small cohort, Bosnia, in only n=100 high- and n=100 low-risk women who received cervical length measurement and pathogen smear at 16 weeks. Reports the frequency of pathogens and cervical length among those with preterm delivery in the high-risk group. Small study, no data for the low risk group with which to calculate test performance, additionally uncertain representation of western countries.
Facco FL, Simhan HN. Short ultrasonographic cervical length in women with low-risk obstetric history. Obstet Gynecol. 2013;122(4):858-62.	Secondary analysis of 1996 US prospective cohort with the purpose to identify predictors of preterm birth. N=1284 low risk women with single pregnancies had serial measurements 22-24 weeks. Only gives the incidence of preterm birth among those with length <20 or <15mm, with no data to calculate test performance.
Tests of bacterial vaginosis	

Sangkomkamhang US, Lumbiganon P, Prasertcharoensuk W, et al. Antenatal lower genital tract infection screening and treatment programs for preventing preterm delivery. Cochrane Database Syst Rev. 2015(2):Cd006178.	Cochrane review (search Nov 2014) for RCTs where asymptomatic women were randomised to infection screening and treatment or no screening. A single 2004 trial in a large sample of the general pregnant population was identified. This study could not meet criteria for the test performance question and was excluded from the treatment question as it is a pre-2013 study, the individual results of which were not considered to be contributing new evidence.
Nelson DB, Hanlon A, Nachamkin I, et al. Early pregnancy changes in bacterial vaginosis-associated bacteria and preterm delivery. Paediatr Perinat Epidemiol. 2014;28(2):88-96.	US study where n=1890 women received vaginal swabs at recruitment <16 weeks and at 20-24 weeks. Introduced as a prospective cohort but the analysis used is a case-control design including all women with preterm birth and a 30% random sample of the rest of the cohort. Would also be considered to have limited applicability to the UK as 70% African American population.
Uterine contraction monitoring (home device)	
Urquhart C, Currell R, Harlow F, Callow L. Home uterine monitoring for detecting preterm labour. Cochrane Database of Syst Rev 2017 (2). CD006172.	Cochrane review (search June 2016) for RCTs where women identified to be at risk for preterm birth were assigned to home uterine monitoring or routine care. Includes 15 RCTs all pre-2013. No description of population characteristics (for example, how they were considered to be at risk) or interventions (for example, what was the definition of increased contractions in a screening context). Results given are as relative risks for those with home-monitoring vs not. Would not meet inclusion criteria as a screening study, and home contraction monitoring to detect preterm labour was not being considered as an intervention.
General reviews/mixed tests	
Glover AV, Manuck TA. Screening for spontaneous preterm birth and resultant therapies to reduce neonatal morbidity and mortality: A review. Semin Fetal Neonatal Med. 2018;23(2):126-32.	Non-systematic review (uncertain design from the abstract only)
Lucaroni F, Morciano L, Rizzo G, et al. Biomarkers for predicting spontaneous preterm birth: an umbrella systematic review. J Matern Fetal Neonatal Med. 2018;31(6):726-34.	Review of reviews assessing maternal and fetal biomarkers. The two reviews on fFN from 2012 onwards did not meet inclusion criteria for this rapid review (assessing, respectively, symptomatic women, and value for short-term prediction of birth within 48 hours).

Sananès N, Langer B, Gaudineau A, et al. Prediction of spontaneous preterm delivery in singleton pregnancies: Where are we and where are we going? A review of literature. Journal of Obstetrics and Gynaecology. 2014;34(6):457-61.	Systematic review search 2012. Presents the results of individual studies on different tests including cervical length and fFN. As all studies were pre-2013, without meta-analysis this review was not considered to be contributing new evidence.
Jwala S, Tran TL, Terenna C, et al. Evaluation of additive effect of quantitative fetal fibronectin to cervical length for prediction of spontaneous preterm birth among asymptomatic low-risk women. Acta Obstet Gynecol Scand. 2016;95(8):948-55.	Prospective cohort, US single centre. N=528 low-risk women with single pregnancies (risk factors not specifically reported) receiving cervical length and fetal fibronectin measurement at 18-23 weeks. Gives test performance for each test (length <20mm or fFN >5ng/ml) or each individually for predicting preterm birth. This study would be eligible for inclusion; however, it is reported that progesterone is routinely offered to all women with cervical length <20mm at this centre. The study does not state how many were treated; therefore it is unknown whether treatment could be influencing test performance results.
Question 2	
Progesterone treatment	
Kuon RJ, Voss P, Rath W. Progesterone for the Prevention of Preterm Birth - an Update of Evidence-Based Indications. Geburtshilfe Frauenheilkd. 2019;79(8):844-53.	Systematic review with search date Sept 2018 (one database). Inclusion of studies published in English or German that looked at either progesterone in women with short cervix, or after previous preterm birth. Those after preterm birth would be excluded as exclusively a high risk population. Search for studies in women with short cervix included only the Romero et al 2018 meta-analysis therefore the primary publication of the Romero review was prioritised. Studies on safety are also narratively reported but include either those published <2013 or exclusively in high risk pregnancies (for example multiples).
Romero R, Conde-Agudelo A, Da Fonseca E, et al. Vaginal progesterone for preventing preterm birth and adverse perinatal outcomes in singleton gestations with a short cervix: a meta-analysis of individual patient data. American Journal of Obstetrics and Gynecology. 2018;218(2):161-80.	The Romero et al 2018 meta-analysis of individual patient data (search 2017) was prioritised. This earlier review (search 2016) included the same 5 trials but the subsequent analysis had greater individual data available for the OPPTIMUM trial.
Ahn KH, Bae NY, Hong SC, et al. The safety of progestogen in the prevention of preterm birth: Meta-analysis of neonatal mortality. Journal of Perinatal Medicine. 2017;45(1):11-20.	Systematic review (search Nov 2015) for any RCTs that assessed the effect of progesterone on neonatal death, with separate data for singletons and multiples (some uncertainties around eligibility and exclusions). Only 3 relevant studies of progesterone for singleton pregnancies in meta-analysis: O'Brien (women with prior preterm birth), Hassan and van Os (both

	based on cervical length) (no effect on neonatal death). The Romero review was selected in preference which included applicable data on women with short cervix from Hassan and O'Brien (along with 3 other studies) to give a larger sample size with wider assessment of neonatal outcomes. The van Os study was included separately, though notably this was a highly underpowered study for this outcome.
Conde-Agudelo A, Romero R. Vaginal progesterone to prevent preterm birth in pregnant women with a sonographic short cervix: clinical and public health implications. Am J Obstet Gynecol. 2016;214(2):235-42.	Non-systematic review (uncertain design from the abstract only)
O'Brien JM, Lewis DF. Prevention of preterm birth with vaginal progesterone or 17-alpha-hydroxyprogesterone caproate: A critical examination of efficacy and safety. American Journal of Obstetrics and Gynecology. 2016;214(1):45-56.	Non-systematic review (uncertain design from the abstract only)
Schmouder VM, Prescott GM, Franco A, et al. The rebirth of progesterone in the prevention of preterm labor. Annals of Pharmacotherapy. 2013;47(4):527-36.	Systematic review (search date Sept 2012) for double-blind, placebo-controlled trials of progesterone with primary outcome of preterm birth or adverse neonatal outcomes. Two trials included women with short cervix, both of which were included in the Romero et al review which was prioritised for inclusion. Other inclusions were studies in women with multiple gestation or prior spontaneous birth which would be excluded as exclusively in women with risk factors.
Cervical pessary	
Jin Z, Chen L, Qiao D, et al. Cervical pessary for preventing preterm birth: a meta-analysis. Journal of Maternal-Fetal and Neonatal Medicine. 2019;32(7):1148-54.	Systematic review with search Dec 2016 for 'case-controls' (though included studies are RCTs) comparing pessary with expectant management or other treatment to prevent preterm birth (published in English or Chinese). Included n=8 studies, but only n=3 RCTs in women with singletons. Provides relative risk for preterm birth <28 and <34 weeks in singletons. Same study includes as the Saccone et al (and Jin et al) reviews, but Saccone was prioritised as the search was specific to RCTs in the population of interest and contained more information on the individual studies and more detailed analysis.
Zheng L, Dong J, Dai Y, et al. Cervical pessaries for the prevention of preterm birth: a systematic review and meta-analysis. J Matern Fetal Neonatal Med. 2019;32(10):1654-63.	Systematic review with search July 2016 for RCTs or cohorts comparing pessary with no pessary in asymptomatic women in second trimester. Included n=11 studies, but only n=3 RCTs and

	1 cohort including women with singletons. Provides effect estimates for preterm birth and adverse neonatal outcomes, overall and by singles or multiples (studies included in each analysis not specified). Same study includes as the Saccone et al (and Jin et al) reviews, but Saccone was prioritised as the search was specific to RCTs in the population of interest and contained more information on the individual studies.
Abdel-Aleem H, Shaaban OM, Abdel-Aleem MA. Cervical pessary for preventing preterm birth. Cochrane Database Syst Rev. 2013 (5): CD007873.	Cochrane review of pessary for preventing preterm birth in women with cervical incompetence (search 2012). High-risk included women with short cervix but the analysis was not exclusive to this. The Saccone 2017 review (search Feb 2016) was prioritised as this included subsequent RCTs and analysis of women selected only on the basis of short cervix.
Mendoza M, Goya M, Gascon A, et al. Modification of cervical length after cervical pessary insertion: correlation weeks of gestation. J Matern Fetal Neonatal Med. 2017;30(13):1596-601.	Secondary analysis of the PECEP trial (Goya et al 2012) which assessed the effect of pessary vs control in women with short cervix upon preterm birth.
Cervical cerclage	
Alfirevic Z, Stampalija T, Medley N. Cervical stitch (cerclage) for preventing preterm birth in singleton pregnancy. Cochrane Database Syst Rev. 2017(6): CD008991.	Cochrane review (search June 2016) of RCTs of cerclage for prevention in women with single pregnancies and high risk factors for preterm birth. High risk included women with short cervix but the analysis was not exclusive to this. The Berghella 2017 review (search Feb 2017) was prioritised as this included subsequent RCTs and analysis of women selected only on the basis of short cervix.
Parrish MR, Salpekar M, Lee G. Pregnancy outcomes after cerclage placement in nulliparous women with a short cervix on transvaginal ultrasonography. J Matern Fetal Neonatal Med. 2016;29(20):3281-5.	Retrospective cohort of n=70 nulliparous women with single pregnancies and with no other risk factors and cervical length <25mm. N=11 had cerclage, n=27 received progesterone and n=32 managed expectantly. Comparison of methods for preterm delivery. Excluded on the basis of small sample size and RCTs available.
Antibiotics for bacterial vaginosis	
Rebouças KF, José Eleutério JE, Peixoto RC, et al. Treatment of bacterial vaginosis before 28 weeks of pregnancy to reduce the incidence of preterm labor. International Journal of Gynecology and Obstetrics. 2019;146(3):271-6.	Systematic review and meta-analysis (search Dec 2017) for RCTs of low-risk women with bacterial vaginosis (diagnosed by demonstration of bacteria) prescribed oral metronidazole or vaginal clindamycin. N=6 RCTs in clindamycin meta-analysis (n=2200) and n=2 RCTs in metronidazole meta-analysis

	(n=2275). All trials are pre-2006 and were included by the 2013 Cochrane systematic review (Brocklehurst et al) included by the last evidence review, which also analysed by antibiotic and mode of administration (both reviews finding no effect). Therefore the review was not considered to contributing new evidence.
Haahr T, Ersboll AS, Karlsen MA, et al. Treatment of bacterial vaginosis in pregnancy in order to reduce the risk of spontaneous preterm delivery - a clinical recommendation. Acta Obstet Gynecol Scand. 2016;95(8):850-60.	Systematic review and meta-analysis with GRADE recommendations (search Oct 2014) around 14 questions including whether metronidazole, clindamycin or probiotics reduce risk of spontaneous preterm birth for low-risk women. Includes guidelines, reviews, RCTs and observational studies. N=5 studies of metronidazole (all pre-2013) n=10 trials of clindamycin, 2 of which are post-2013: Subtil 2014 (abstract only of Subtil 2016 RCT) and Gupta 2014 (Indian RCT, which would be an individual exclude as a non-western population). There is additional narrative report of n=2 probiotics studies, both pre-2013. The publication and supplement give little information on the included studies, including whether any studies in the meta-analysis of antibiotics were observational. There is also limited quality assessment. Therefore the Subtil 2016 RCT alone was selected for inclusion; this review was not considered to update evidence from the 2013 Cochrane review.
Yudin MH, Money DM. No. 211-Screening and Management of Bacterial Vaginosis in Pregnancy. J Obstet Gynaecol Can. 2017;39(8):e184-e91.	Canadian guideline recommendations based on systematic literature search to 2007. Reports data from individual pre-2013 studies but no new meta-analysis. Useful background.
Thinkhamrop J, Hofmeyr GJ, Adetoro O, et al. Antibiotic prophylaxis during the second and third trimester to reduce adverse pregnancy outcomes and morbidity. Cochrane Database Syst Rev. 2015(6):Cd002250.	Cochrane review. Inclusion criteria routine prophylaxis for all women considered to be at high risk of preterm birth including prior preterm birth or low birthweight baby. ‘Bacterial vaginosis in the current pregnancy’ is listed as a high risk factor; however, it is not possible to separately analyse studies where women have been detected to have infection. Additionally three studies are listed as excluded because ‘Antibiotic usage was for treatment after having identified the infection, not for prophylaxis.’ Therefore these apparently contradictory inclusions/criteria give some uncertainty but the review was excluded as it appears to be routine prophylaxis rather than treatment indicated following detection of infection.

Brocklehurst P, Gordon A, Heatley E, et al. Antibiotics for treating bacterial vaginosis in pregnancy. Cochrane Database Syst Rev. 2013(1):Cd000262.	2013 Cochrane review included by the last 2015 UK NSC evidence review (captured by both searches covering 2013).
Shimaoka M, Yo Y, Doh K, et al. Association between preterm delivery and bacterial vaginosis with or without treatment. Sci Rep. 2019;9(1):509.	Retrospective cohort, single centre Japan. Women screened for bacterial vaginosis. Comparison of n=867 who received observational care and n=628 who received metronidazole treatment. Large comparative cohort but excluded as the decision was made to prioritise RCT evidence.
Probiotics	
Cooper NA, Moores R. A review of the literature regarding nutritional supplements and their effect on vaginal flora and preterm birth. Curr Opin Obstet Gynecol. 2014;26(6):487-92.	Systematic review of reviews on nutritional supplements to prevent preterm birth (search January 2014) with narrative synthesis. Identified 5 reviews including the 2012 updated Cochrane on probiotics to prevent preterm labour, which would have been available at the time of the last evidence review. Two reviews are available post-2012 but evaluate supplements to prevent or treat bacterial vaginosis rather than prevent preterm birth.
General/mixed reviews	
Matei A, Saccone G, Vogel JP, et al. Primary and secondary prevention of preterm birth: a review of systematic reviews and ongoing randomized controlled trials. Eur J Obstet Gynecol Reprod Biol. 2019;236:224-39.	Review of any published systematic review of randomised controlled trials or individual patient data for the primary or secondary prevention of preterm birth (search Nov 2016). Regardless of preterm birth risk. 49 Cochranes and 63 non-Cochranes identified, 60 in primary prevention (not defined; presumed to imply no prior history of preterm birth). Includes: drug, device, psychosocial, procedural, nutritional supplements, lifestyle or behaviour, health system change, and screening with or without treatment. Presents several pages of results for the individual reviews, with relevant post-2013 reviews identified for the listed treatments (the only relevant review [Sangkomkamhang] was identified by this search). The primary reviews were therefore selected for inclusion.
Medley N, Poljak B, Mammarella S, et al. Clinical guidelines for prevention and management of preterm birth: a systematic review. BJOG. 2018;125(11):1361-9.	Systematic search May 2017 for clinical practice guidelines on the prevention or management of preterm birth. Based around 27 questions including whether individual strategies should be used in women with different risk factors. Covers NICE, RCOG and other international recommendations. Useful background but does not provide evidence for question.

Medley N, Vogel JP, Care A, et al. Interventions during pregnancy to prevent preterm birth: An overview of Cochrane systematic reviews. Cochrane Database of Systematic Reviews. 2018;2018(11).	Overview of other Cochrane systematic reviews, relevant reviews of which were identified and the primary reviews analysed.
Jarde A, Lutsiv O, Beyene J, et al. Vaginal progesterone, oral progesterone, 17-OHPC, cerclage, and pessary for preventing preterm birth in at-risk singleton pregnancies: an updated systematic review and network meta-analysis. Bjog. 2019;126(5):556-67.	Systematic review and network meta-analysis (search 2018) to compare the risk of progesterone, oral progesterone, 17-OHPC, cerclage, and pessary for preventing preterm birth in at-risk singleton pregnancies. At-risk was defined as women with history of preterm birth, by cervical length or other factors defined by the study authors. The overall network MA for women overall, women with preterm birth and by type/route of administration was thought less applicable as it could not be applied to all women. The separate subgroup analysis specific to the risk factor of short cervical length included standard rather than network MA: 4 trials of pessary (Goya 2012, Hui 13, Nicolaides 16 and Saccone 17), 2 of cerclage (Althuisius 2001 and Otsuki 2016), one of vaginal progesterone (Fonesca 2007) and none of oral. Therefore all trials were included in the systematic reviews of these interventions, which also included additional trials. The only added value of this review is that the Saccone 2017 RCT is included in the meta-analysis (currently included as a separate study include alongside the Saccone 2017 review). However, it was decided to prioritise the specific systematic reviews due to the limited applicability of this network meta-analysis which was overall applicable to at-risk women.
Sentilhes L, Senat MV, Ancel PY, et al. Prevention of spontaneous preterm birth: Guidelines for clinical practice from the French College of Gynaecologists and Obstetricians (CNGOF). Eur J Obstet Gynecol Reprod Biol. 2017;210:217-24.	French guidance with graded recommendations. Methodology unclear from the publication.
Trials included by the selected systematic reviews	
Progesterone	
Norman JE, Marlow N, Messow CM, et al. Vaginal progesterone prophylaxis for preterm birth (the OPPTIMUM study): a multicentre, randomised, double-blind trial. Lancet. 2016;387(10033):2106-16.	Large, multicentre UK and Swedish RCT in n=1228 women with single pregnancies and risk factors for preterm birth. Initial eligibility was women with positive fetal fibronectin test at 22-24 weeks plus additional risk factors (previous preterm birth, second trimester loss, P-PROM, or cervical procedure). Eligibility was later extended (after recruitment of n=84) to include women with negative fibronectin but with history of preterm birth, or with short

	cervical length. All the main outcomes are for the full study population, 80% of whom had history of prior previous preterm birth. Therefore this is majority high-risk population. Subgroup analysis is performed for the women with risk factor of short cervical length though the applicable individual patient data was included by the Romero 2018 systematic review.
Pessary	
Nicolaides KH, Syngelaki A, Poon LC, et al. A randomized trial of a cervical pessary to prevent preterm singleton birth. New England Journal of Medicine. 2016;374(11):1044-52.	Multicentre international RCT. N=935 women with single pregnancy and cervical length <25mm detected at anomaly scan randomised to pessary or control (some receiving additional treatment). Primary outcome of preterm birth <34 weeks with secondary outcomes of other gestations or neonatal morbidity. Majority of participants UK. Meta-analysed in the Saccone 2017 systematic review.
Hui SA, Chor CM, Lau TK, et al. Cerclage pessary for preventing preterm birth in women with a singleton pregnancy and a short cervix at 20 to 24 weeks: A randomized controlled trial. American Journal of Perinatology. 2013;30(4):283-8.	Single centre RCT, China. N=103 women with single pregnancy and cervical length <25mm detected at anomaly scan randomised to pessary or control. Primary outcome of preterm birth <34 weeks with secondary outcomes of other gestations or neonatal morbidity. Included by the Saccone 2017 systematic review (though note would not have been eligible as a single study include due to being a non-western population).
Cervical cerclage	
Otsuki K, Nakai A, Matsuda Y, et al. Randomized trial of ultrasound-indicated cerclage in singleton women without lower genital tract inflammation. J Obstet Gynaecol Res. 2016;42(2):148-57.	Single centre RCT, Japan. N=106 women with single pregnancies who received cervical length screening at 16-25 weeks and found to have length <25mm. Randomised to 2 different cerclage procedures or bedrest. Primary outcome of preterm delivery. Included by the Berghella 2017 meta-analysis (note non-'western' though representation may be considered on an OECD status).

Appendix 3 — Summary and appraisal of individual studies

Data Extraction

Table 24. Studies relevant to criteria 4 and 5

Study reference	Study design	Eligibility	Population/included studies	Index test and reference standard	Test accuracy
Dos Santos et al 2018[5]	Systematic review and met-analysis to determine the accuracy of fetal fibronectin (fFN) in cervico-vaginal secretions for identifying the risk of preterm birth in asymptomatic pregnant women	Longitudinal, cross-sectional or case-control studies meeting criteria: • asymptomatic pregnant women with either singleton or multiple pregnancies • fFN sampling undertaken after 22 weeks' gestation using a validated method • using threshold cut-off ≥50 ng/mL for a positive test • reference standard of preterm birth <37 weeks' gestation Exclusions: • women with signs or symptoms of preterm birth Published 2005 to 2015 (studies <2005 identified from Honest et al 2009	n=15 studies met inclusion criteria: • n=10 women with singletons • n=5 women with multiple birth Of n=10 studies in singletons, n=6 were in 1,236 women without previous preterm birth or other risk factors (as defined by individual studies): • Arinami et al 1999 (n=438). Prospective, Japan, preterm prevalence 4.1% • Chang et al 1997 (n=234). Prospective, Singapore, preterm prevalence 7.7% • DiStefano et al 1999 (n=60). Prospective, Italy, preterm prevalence 10% • Garcia et al 1999 (n=263). Prospective,	**Index test** All n=6 studies in low-risk population described as standard fFN collection (with speculum and visualisation of the posterior fornix as appose to blind sampling) n=3 studies sampled at 22-34 weeks, n=3 at 24-37 weeks n=3 used serial testing (n=3 presumably single collection) n=5 studies analysed using the Adeza Biomedical model (not specified for n=1) All n=6 used threshold >50ng/ml **Reference standard** n=5 assessed preterm birth as <37 weeks, n=1 used <36 weeks Gestational age confirmed by ultrasound in n=5 studies (not specified for n=1)	Meta-analysis of n=6 studies in women without risk factors. Preterm birth <36/37 weeks: • sensitivity (Sn): 0.48 (95% CI 0.20 to 0.77) • specificity (Sp): 0.96 (95% CI 0.86 to 0.99) • likelihood ratio (LR) positive: 12.01 (95% CI 4.70 to 30.68) • LR negative: 0.54 (95% CI 0.30 to 0.97) Individual studies show high heterogeneity for sensitivity: • Arinami: Sn 0.06 (0.00 to 0.30), Sp 1.00 (0.99 to 1.00), PPV 50.0%, NPV (calculated) 96.5% • Chang: Sn 0.17 (0.04 to 0.41), Sp (0.99 (0.97 to 1.00), PPV 60.0%, NPV (calculated) 93.4%

Study reference	Study design	Eligibility	Population/included studies	Index test and reference standard	Test accuracy
		review). Update search February 2017. Search databases: EMBASE, MEDLINE, CINHAL, AMED and BNI. No language restrictions.	Mexico, preterm prevalence 10.3% • Greenhagen et al 1996 (n=108). Prospective, US, preterm prevalence 7.4% • Hellemans et al 1995 (n=133). Prospective, Belgium, preterm prevalence 8.1% (described as low-risk but n=7 had risk factors) All included non-smokers only. Quality appraisal QUADAS-2 with risk of bias: • patient selection: n=3 low risk, n=2 unclear risk, n=1 high risk • index test: n=5 low risk, n=1 high risk • reference standard: n=5 low risk, n=1 high risk • flow and timing: n=3 low risk, n=2 unclear risk, n=1 high risk Applicability (to review question): low for the index test and reference standard;	NB. Not specified whether this is spontaneous preterm birth only (excluding iatrogenic) or whether any treatment following screen detection was permitted.	• DiStefano: Sn 0.67 (0.22 to 0.96), Sp 0.85 (0.73 to 0.93), PPV 33.3%, NPV (calculated) 95.8% • Garcia: Sn 0.81 (0.62 to 0.94), Sp 0.96 (0.93 to 0.98), PPV 71.0%, NPV (calculated) 97.8% • Greenhagen: Sn 0.63 (0.24 to 0.91), Sp 0.84 (0.75 to 0.91), PPV 23.8%, NPV (calculated) 96.6% • Hellemans: Sn 0.60 (0.26 to 0.88), Sp 0.85 (0.78 to 0.91), PPV 25.0%, NPV (calculated) 96.3%

Study reference	Study design	Eligibility	Population/included studies	Index test and reference standard	Test accuracy
			patient selection, n=5 low and n=1 high risk		
Esplin et al 2017[6] Monitoring Mothers-to-Be (nuMoM2b) study	Prospective cohort to assess the accuracy of universal screening using serial transvaginal cervical length and quantitative measurement of fFN to predict spontaneous preterm birth in nulliparous women. 8 centres, US. Oct 2010 to May 2014.	Nulliparous women with singleton pregnancies prior to 14 weeks' gestation who completed at least one of the fFN or cervical length measurements. Nulliparous was defined as women with no previous self-reported pregnancy extending beyond 20 weeks' gestation.	n=9469 women (median age 27 years, 59% white, 15% black, 16% Hispanic) n=477 spontaneous preterm births (5% prevalence): • <32 weeks: n=76 • 32 to 37 weeks: n=401 n=474 preterm births analysed: • n=451 with both measurement data • n=12 cervical length only • n=11 fFN only N=8992 term births or iatrogenic preterm births (controls) • n=8936 included in the analysis, excluding n=56 with neither measure Total exclusions <1% of n=9469, but initial recruitment was n=10,038: n=110 were excluded due to miscarriage <20 weeks and n=459	**Index tests** <u>fFN</u> 3 vaginal swabs (self-obtained) at: • visit 1: 6^{+0} to 14^{+6} weeks (median 12^{+4}) • visit 2: 16^{+0} to 22^{+6} weeks (median 19.0) • visit 3: 22^{+0} to 30^{+6} weeks (median 28^{+0}) Analysed by Hologic assay. Cut-offs ≥10, ≥50, and ≥200 ng/mL. <u>Cervical length measurement</u> taken by transvaginal ultrasound (TVUS) at: • visit 2 • visit 3 Cut-offs ≤20 mm or 25mm. **Reference standard** Primary outcome: Spontaneous preterm birth <37 weeks' gestation occurring after spontaneous onset of labour or preterm prelabour rupture of the membranes (P-PROM), regardless of subsequent	Women with spontaneous preterm birth had shorter cervical length than those who gave birth at term at visit 2 (median 36mm cases vs 39mm controls) and visit 3 (median 32mm vs 37mm) (p<0.001). For fFN, use of the commonly accepted threshold of >50ng/mL identified 87/411women (21.2%) with spontaneous preterm birth at visit 1, 30/410 (7.3%) at visit 2, and 31/384 (8.1%) at visit 3. For prediction of spontaneous preterm birth <37 weeks all thresholds had poor test performance. Specificity was generally high but sensitivity very poor. The best sensitivity was at 34.5 for fFN and 23.3 for cervical length with the lowest specificity. Even with good specificity, peak PPV was only 14.0% for fFN and 20.8% for cervical length. The positive likelihood ratio was poor across thresholds. **fFN** *****peak values <u>Visit 1 (6^{+0} to 14^{+6} weeks)</u>

Study reference	Study design	Eligibility	Population/included studies	Index test and reference standard	Test accuracy
			excluded due to lack of data on pregnancy outcomes. Characteristics significantly associated with preterm birth: • younger maternal age • smoking history • lower educational level • diabetes	labour augmentation or caesarean. Iatrogenic preterm birth was included in the control group. Secondary outcome: Preterm birth <32 weeks' gestation. <u>Treated women:</u> Women were informed of cervical length measures less than 15mm and their clinician could have treated with progesterone. Of n=742 women (8% of the cohort) with cervical length <25mm, n=66 (8.9%) received progesterone. The authors performed sensitivity analysis where these women were considered to have preterm birth, which made little difference to test performance (optimal AUC 0.70 [95% CI 0.67 to 0.73] vs 0.67 [0.64 to 0.70]).	≥10ng/mL • **Sn 34.5 (95% CI 30.0 to 39.1)*** • Sp 74.1 (73.2 to 75.1) • PPV 6.7 (5.6 to 7.7) • NPV 95.5 (95.0 to 96.0) • LR+ 1.34 (1.15 to 1.52) • LR- 0.88 (0.82 to 0.95) • AUC 0.54 (0.52 to 0.57) ≥50ng/mL • Sn 21.2 (17.2 to 25.1) • Sp 87.6 (86.9 to 88.3) • PPV 8.4 (6.7 to 10.0) NPV 95.4 (94.9 to 95.9) • LR+ 1.71 (1.37 to 2.04) • LR- 0.90 (0.85 to 0.95) • AUC 0.54 (0.52 to 0.56) ≥200ng/mL • Sn 9.5 (6.7 to 12.3) • Sp 94.6 (94.1 to 95.1) • PPV 8.6 (6.0 to 11.1) • NPV 95.1(94.7 to 95.6) • LR+ 1.75 (1.20 to 2.30) • LR- 0.96 (0.93 to 0.99)

Study reference	Study design	Eligibility	Population/included studies	Index test and reference standard	Test accuracy
					• AUC 0.52 (0.51 to 0.53) Visit 2 (16^{+0} to 22^{+6} weeks) ≥10ng/mL • Sn 15.1 (11.7 to 18.6) • Sp 88.5 (87.8 to 89.2) • PPV 6.4 (4.9 to 8.0) • NPV 95.2 (94.8 to 95.7) • LR+1.32 (1.01 to 1.63) • LR- 0.96 (0.92 to 1.00) • AUC 0.52 (0.50 to 0.54) ≥50ng/mL • Sn 7.3 (4.8 to 9.8) • Sp 96.0 (95.6 to 96.5) • PPV 8.8 (5.8 to 11.8) • NPV 95.2 (94.7 to 95.7) • LR+ 1.85 (1.18 to 2.51) • LR- 0.97 (0.94 to 0.99) • AUC 0.52 (0.50 to 0.53) ≥200ng/mL • Sn 2.9 (1.5 to 5.1) • Sp 98.3 (98.0 to 98.6) • PPV 8.3 (3.8 to12.8) • NPV 95.1 (94.6 to 95.6) • LR+ 1.73 (0.72 to 2.74)

Study reference	Study design	Eligibility	Population/included studies	Index test and reference standard	Test accuracy
					• LR- 0.99 (0.97 to 1.00) • AUC 0.51 (0.50 to 0.51) <u>Visit 3 (22^{+0} to 30^{+6} weeks)</u> ≥10ng/mL • Sn 21.9 (17.7 to 26.0) • Sp 91.8 (91.2 to 92.4) • PPV 11.2 (9.0 to 13.5) • NPV 96.1 (95.7 to 96.5) • LR+ 2.66 (2.12 to 3.20) • LR- 0.85 (0.81 to 0.90) • AUC 0.57 (0.55 to 0.59) **≥50ng/mL (common threshold – used by SR and MA)** • Sn 8.1 (5.3 to 10.8) • Sp 96.8 (96.4 to 97.2) • PPV 10.7 (7.2 to 14.3) • NPV 95.7 (95.2 to 96.1) • LR+ 2.53 (1.62 to 3.44) • LR- 0.95 (0.92 to 0.98) • AUC 0.52 (0.51 to 0.54) ≥200ng/mL • Sn 3.9 (2.2 to 6.4) • Sp 98.9 (98.6 to 99.1) • **PPV 14.0 (7.4 to 20.6)***

Study reference	Study design	Eligibility	Population/included studies	Index test and reference standard	Test accuracy
					• NPV 95.6 (95.2 to 96.0) • LR+ 3.44 (1.59 to 5.28) • LR- 0.97 (0.95 to 0.99) • AUC 0.51 (0.50 to 0.52) Combination of serial fFN measurements at visit 3 gave AUC 0.59 (0.56 to 0.62) **Cervical length** ***peak values <u>Visit 2 (16^{+0} to 22^{+6} weeks)</u> ≤25 mm: • Sn 8.0 (5.4 to 10.5) • Sp 97.8 (97.5 to 98.1) • PPV 16.2 (11.3 to 21.1) • NPV 95.3 (94.8 to 95.7) • LR+ 3.67 (2.39 to 4.95) • LR- 0.94 (0.91 to 0.97) • AUC 0.53 (0.52 to 0.54) ≤20 mm: • Sn 4.1 (2.4 to 6.4) • Sp 98.8 (98.6 to 99.1) • PPV 15.5 (8.9 to 22.1) • NPV 95.1 (94.7 to 95.6) • LR+ 3.49 (1.77 to 5.21)

Study reference	Study design	Eligibility	Population/included studies	Index test and reference standard	Test accuracy
					• LR- 0.97 (0.95 to 0.99) • AUC 0.51 (0.51 to 0.52) <u>Visit 3 (22^{+0} to 30^{+6} weeks)</u> ≤25 mm: • **Sn 23.3 (19.2 to 27.5)*** • Sp 93.6 (93.1 to 94.1) • PPV 15.1 (12.3 to 17.9) • NPV 96.2 (95.8 to 96.6) • LR+ 3.65 (2.94 to 4.37) • LR- 0.82 (0.77 to 0.86) • AUC 0.58 (0.56 to 0.61) ≤20 mm • Sn 17.4 (13.7 to 21.1) • Sp 96.8 (96.4 to 97.2) • **PPV 20.8 (16.5 to 25.2)*** • NPV 96.0 (95.6 to 96.4) • LR+ 5.42 (4.10 to 6.74) • LR- 0.85 (0.82 to 0.89) • AUC 0.57 (0.55 to 0.59) Combination of cervical length measurements at visit 3 gave AUC 0.67 (0.64 to 0.70) Secondary outcome: Prediction of preterm birth <32 weeks:

<table>
<tr><th>Study reference</th><th>Study design</th><th>Eligibility</th><th>Population/included studies</th><th>Index test and reference standard</th><th>Test accuracy</th></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>Gave slightly higher, but still poor, sensitivity values with a peak of 50% for fFN and 52% for cervical length. Peak PPV was 5.6 for fFN and 8.6 for cervical length.</td></tr>
<tr><td>van der Ven et al 2015[9]

The Triple P screening study</td><td>Prospective cohort to assess the accuracy of cervical length measurement to predict preterm birth in women with no history of preterm birth.

Multicentre, The Netherlands. Nov 2009 to July 2013.</td><td>Nulliparous and multiparous women (>18 years) with singleton pregnancies and without history of SPTB <34 weeks.

Other exclusions:
• regular uterine contractions
• ruptured membranes
• cervical cerclage in situ
• fetal anomaly</td><td>n=16,204 women screened

n=12,360 (74.4%) could be linked to Dutch Perinatal Registry to obtain outcome data (reasons typically incorrect digit entry or change of residence).

n=11,943 following exclusion criteria:

n=5710 nulliparous (mean 30 years, 88% white)

n=6233 low-risk multiparous (mean 32 years, 87% white)

Mean CL was shorter in nulliparous women: 43.1mm vs. 45.1mm, p<0.0001).

More nulliparous women had short cervix ≤30 mm: 2.2% (n=125) vs 1.4% (n=87) multiparous, p=0.001, and ≤35mm: 14.0% vs 9.8%, p<0.0001.

No difference between women with data linkage and not for measured characteristics of:
• maternal age: 31 both groups</td><td>Index test

Cervical length measurement taken by TVUS at 16^{+0} to 21^{+6} at the time of the routine anomaly scan (mean 20^{+2}).

Women with measure ≤30mm invited to participate in an RCT comparing progesterone with placebo (assumed to be the primary cut-off of interest).

The study also tested likelihood ratio (not specified positive or negative) for cut-offs <20, 21-25, 26-30 and 31-35.

Reference standard

SPTB and iatrogenic preterm birth <37, <34 and <37 weeks were assessed separately for nulliparous and multiparous women.

Iatrogenic was defined as elective caesarean or induction of labour for fetal or maternal reasons.</td><td>The rate of SPTB was 3.9% overall but higher in nulliparous women: 5.3% (n=300) vs 2.6% (n=164), p<0.001

The study reports only LR for distinct categories <20, 21-25, 26-30 and 31-35 rather than continuous; and separately for SPTB, iatrogenic and term births.

Test performance data has been calculated by the reviewer, combining iatrogenic preterm births in the term comparison group, as the authors describe doing for calculation of LRs.

Nulliparous: SPTB <37 weeks
<35mm threshold
<table><tr><th>Screen</th><th>SPTB</th><th>Term</th><th>Total</th></tr><tr><td><35 (+)</td><td>87</td><td>710</td><td>797</td></tr><tr><td>>35 (-)</td><td>213</td><td>4700</td><td>4913</td></tr><tr><td>Total</td><td>300</td><td>5410</td><td>5710</td></tr></table>• Sn (87/300)=29.0%
• Sp (4700/5410)=86.9%
• PPV (87/797)=10.9%
• NPV (4700/4913)=95.7%</td></tr>
</table>

<table>
<tr><th>Study reference</th><th>Study design</th><th>Eligibility</th><th>Population/included studies</th><th>Index test and reference standard</th><th>Test accuracy</th></tr>
<tr><td></td><td></td><td></td><td>

median CL: 43mm both groups
CL ≤30 mm 1.8% vs 2.1%, p=0.21).

</td><td>
Data was obtained through linkage with the Dutch Perinatal Registry Database.

<u>Treated women:</u>

n=375 women had measure <30mm (1.8%) of whom n=80 (21.3%) agreed to participate in the treatment trial (van Os[11] below, where eligibility also required a 2nd repeat measure <30mm).

Of n=80, n=41 received progesterone (10.9% of screen-positives) with the remainder receiving placebo.
</td><td>
(on ROC curve authors report Sn 28.2 and Sp 87.3: AUC 0.61)

<30mm threshold
<table>
<tr><th>Screen</th><th>SP TB</th><th>Term</th><th>Total</th></tr>
<tr><td><30 (+)</td><td>19</td><td>106</td><td>125</td></tr>
<tr><td>>30 (-)</td><td>281</td><td>5304</td><td>5585</td></tr>
<tr><td>Total</td><td>300</td><td>5410</td><td>5710</td></tr>
</table>

Sn=6.3%
Sp=98.0%
PPV=15.2%
NPV =95.0%

(on ROC curve authors report Sn 5.7 and Sp 98.1)

<25mm threshold
<table>
<tr><th>Screen</th><th>SP TB</th><th>Term</th><th>Total</th></tr>
<tr><td><25 (+)</td><td>11</td><td>24</td><td>35</td></tr>
<tr><td>>25 (-)</td><td>289</td><td>5386</td><td>5675</td></tr>
<tr><td>Total</td><td>300</td><td>5410</td><td>5710</td></tr>
</table>

Sn=3.7%
Sp=99.6%
PPV=31.4%
NPV =94.9%

<20mm threshold
<table>
<tr><th>Screen</th><th>SP TB</th><th>Term</th><th>Total</th></tr>
<tr><td><20 (+)</td><td>6</td><td>4</td><td>10</td></tr>
<tr><td>>20 (-)</td><td>294</td><td>5406</td><td>5700</td></tr>
<tr><td>Total</td><td>300</td><td>5410</td><td>5710</td></tr>
</table>

Sn=2.0%

</td></tr>
</table>

<table>
<tr><th>Study reference</th><th>Study design</th><th>Eligibility</th><th>Population/included studies</th><th>Index test and reference standard</th><th>Test accuracy</th></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>
Sp=99.6%
PPV=31.4%
NPV =94.9%
Study authors reported LRs for SPTB in nulliparous women:
>35mm: 0.81 (0.75 to 0.87)
31-35mm: 2.0 (1.6 to 2.5)
26-30mm: 1.8 (0.86 to 3.6)
21-25mm: 4.5 (1.7 to 12)
<20mm: 27 (7.7 to 95)
Multiparous: SPTB <37 weeks
<35mm threshold
<table>
<tr><td>Screen</td><td>SP TB</td><td>Term</td><td>Total</td></tr>
<tr><td><35 (+)</td><td>29</td><td>581</td><td>610</td></tr>
<tr><td>>35 (-)</td><td>135</td><td>5488</td><td>5623</td></tr>
<tr><td>Total</td><td>164</td><td>6069</td><td>6233</td></tr>
</table>
Sn=17.7%
Sp=90.4%
PPV=4.8%
NPV=97.6%
(on ROC curve authors report Sn 16.0 and Sp 90.4: AUC 0.56)
<30mm threshold
<table>
<tr><td>Screen</td><td>SP TB</td><td>Term</td><td>Total</td></tr>
<tr><td><30 (+)</td><td>9</td><td>78</td><td>87</td></tr>
<tr><td>>30 (-)</td><td>155</td><td>5991</td><td>6146</td></tr>
<tr><td>Total</td><td>164</td><td>6069</td><td>6233</td></tr>
</table>
Sn=5.5%
Sp=98.7%
PPV=10.3%
</td></tr>
</table>

<table>
<tr><th>Study reference</th><th>Study design</th><th>Eligibility</th><th>Population/included studies</th><th>Index test and reference standard</th><th>Test accuracy</th></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>
NPV=97.5%
(on ROC curve authors report Sn 4.4 and Sp 98.7)

<25mm threshold
<table>
<tr><td>Screen</td><td>SP TB</td><td>Term</td><td>Total</td></tr>
<tr><td><25 (+)</td><td>3</td><td>21</td><td>24</td></tr>
<tr><td>>25 (-)</td><td>161</td><td>6048</td><td>6209</td></tr>
<tr><td>Total</td><td>164</td><td>6069</td><td>6233</td></tr>
</table>

Sn=1.8%
Sp=99.7%
PPV=12.5%
NPV=97.4%

<20mm threshold
<table>
<tr><td>Screen</td><td>SP TB</td><td>Term</td><td>Total</td></tr>
<tr><td><20 (+)</td><td>3</td><td>3</td><td>6</td></tr>
<tr><td>>20 (-)</td><td>161</td><td>6066</td><td>6227</td></tr>
<tr><td>Total</td><td>164</td><td>6069</td><td>6233</td></tr>
</table>

Sn=1.8%
Sp=100.0%
PPV=50.0%
NPV=97.4%

Study authors reported LRs for SPTB in multiparous women:

>35mm: 0.92 (0.86 to 0.99)
31-35mm: 1.5 (0.97 to 2.4)
26-30mm: 3.9 (1.7 to 8.9)
21-25mm: 0
<20mm: 37 (7.5 to 182)

<u>For preterm birth <34 weeks</u>
</td></tr>
</table>

Study reference	Study design	Eligibility	Population/included studies	Index test and reference standard	Test accuracy
					Test performance was little changed: Nulliparous women: AUC 0.63 <35mm Sn 33.1 and Sp 86.5 <30mm Sn 10.8 and Sp 98.0 Multiparous women: AUC 0.58 <35mm Sn 23.6 and Sp 90.2 <30mm Sn 9.1 and Sp 98.7 The study authors further calculated the number needed to screen and treat to prevent one SPTB <37 weeks using assumed risk reductions with treatment of 20%, 40% and 60%.
Banos et al 2018[7]	Prospective cohort to investigate the effectiveness of mid-trimester cervical ~~contingency~~ consistency index (CCI) measurement for the prediction of spontaneous preterm birth in a selected low-risk pregnant	Low-risk women attending a routine ultrasound scan from 19^{+0} to 24^{+6} weeks. Exclusions: multiple pregnancy, history of preterm birth <34 weeks, P-PROM, late miscarriage >16 weeks, cervical/uterine trauma/malformation, known short cervical length (<25mm) or current treatment to prevent preterm birth (progesterone, cervical cerclage or pessary)	n=532 women (median age 32 years, 72% white) n=22 spontaneous preterm births (4.1% prevalence): • <34 weeks: n=7 n=510 term births (controls) Total sample representative of n=749 recruited (71.0%) after exclusion of n=10 with iatrogenic preterm birth or termination, n=54 lost to follow-up and n=153 with poor image quality. No other characteristics (apart from the index tests)	**Index tests** Cervical length measurement taken by transvaginal ultrasound (TVUS). Cut-offs ≤25mm (1st centile), 30.0mm (5th), 33.0mm (10th), 37.9mm (40th centile). (Also CCI percentage, not being assessed as an index test in this review) Intra- and inter-observer agreement between images was assessed blinded to outcomes.	The cervix was significantly shorter in women who had spontaneous preterm birth (median 39.8mm vs 36.2 mm; p=0.004) Cervical length for predicting spontaneous preterm birth <37 weeks **≤25mm*** • Sn 13.6 (n=3/22) • **Sp 99.6** (508/510) • **PPV 60.0** (3/5) • NPV 96.4 (508/527)

Study reference	Study design	Eligibility	Population/included studies	Index test and reference standard	Test accuracy
	population and to compare it with that of ultrasound cervical length measurement. Single centre, Barcelona, Spain. March 2014 to Nov 2015.		were significantly associated with preterm birth.	**Reference standard** Primary outcome: Spontaneous preterm birth <37 weeks' gestation defined as spontaneous preterm delivery or induction following P-PROM. Iatrogenic preterm birth excluded. Secondary outcome: Spontaneous preterm birth <34 weeks' gestation	• **LR+ 34.8** (95% CI 6.1 to 197.6) • LR- 0.9 (0.7 to 1.0) *Standard cut-off and peak predictive ability with positive test but very low sensitivity ≤30mm: • Sn 18.2 (4/22) • Sp 96.5 (492/510) • PPV 18.2 (4/22) • NPV 96.5 (492/510) • LR+ 5.2 (1.9 to 13.9) • LR- 0.8 (0.7 to 1.0) ≤33mm: • Sn 31.8 (7/22) • Sp 89.6 (457/510) • PPV 11.7 (7/60) • NPV 96.8 (457/472) • LR+ 3.1 (1.6 to 5.9) • LR- 0.8 (0.6 to 1.0) ≤37.9mm: • **Sn 72.7 (16/22)** • Sp 61.2 (312/510) • PPV 7.5 (16/214) • NPV 98.1 (312/318)

Study reference	Study design	Eligibility	Population/included studies	Index test and reference standard	Test accuracy
					• LR+ 1.9 (1.4 to 2.5) • LR- 0.4 (0.2 to 0.9) AUC 0.68 (95% CI 0.56 to 0.81) with optimal cut-off 37.9mm For the secondary outcome of prediction of spontaneous preterm birth <34 weeks, 37.9 was again optimum with slightly higher Sn 85.7 and Sp 61.3. Optimal PPV and LR+ was again at the lower threshold of <25mm. (No evidence of intra- or inter-observer bias in any measures). CCI is not assessed by this review but was significantly lower in women with spontaneous preterm birth (median 73.0% vs 58.1%; p<0.001). CCI was considered to be optimal with AUC 0.84 (95% CI 0.75 to 0.93) and optimal cut-off 64.6% (Sn 77.3%, Sp 82.7%). The combination of both length 37.9mm *or* CCI 63.6% gave Sn 90.9 but Sp 53.9, with the reverse of Sn 54.5 and Sp 90.2 for the combination. (No evidence of intra- or inter-observer bias in any measures).

Study reference	Study design	Eligibility	Population/included studies	Index test and reference standard	Test accuracy
Kuusela et al 2015[8]	Prospective cohort to measure 2nd trimester cervical length (by TVUS) in asymptomatic women with singleton pregnancies and to examine the relation between these measurements and spontaneous preterm delivery <34 weeks as the primary outcome, <37 weeks as the secondary outcome. 2 centres, Sweden. Aug 2012 to May 2015.	Asymptomatic general pregnant population of women attending a routine ultrasound scan from 16^{+0} to 24^{+0} weeks. Exclusions: multiple pregnancy, fetal anomalies, age <18 years, signs of preterm labour/miscarriage, women with cervical length <25mm electing to participate in the OPPTIMUM trial (progesterone vs placebo). NB women with prior preterm labour, P-PROM, mid-trimester loss or cervical trauma would have been eligible.	n=2061 women (median age 31 years, 9% with ≥1 prior preterm delivery) n=22 spontaneous preterm births <34 weeks (1.1%) • n=2039 >34 weeks n=87 spontaneous preterm births <37 weeks (4.2%) • n=1974 term (calculated) Representative of n=2122 analysed excluding n=61: • n=7/11 women with cervical length <25mm participating in OPPTIMUM trial • n=35 with iatrogenic preterm birth • n=19 with missing data (no difference in cervical length between women with prior preterm birth or not) n=2122 represent only 22.7% of n=9338 eligible women; n=7216 pregnant women did not participate. Low uptake reported to be a result of non-consent, busy workloads at ultrasound departments or language barriers.	**Index tests** <u>Cervical length measurement</u> taken by transvaginal ultrasound (TVUS). Cut-offs ≤28mm (1st centile), 31mm (5th), 33mm (15th), 35mm (25th), 37mm (35th centile). NB n=7/11 women with cervical length <25mm were excluded due to trial participation which may affect representation (2/7 delivered at 36 weeks; 5/7 at term) **Reference standard** Primary outcome: Spontaneous preterm birth <34 weeks' gestation. Secondary outcome: Spontaneous preterm birth <37 weeks' gestation (not further defined)	Univariate logistic regression analysis showed a significant association between cervical length and spontaneous preterm birth <34 weeks (odds ratio [OR] 1.78, 95% CI 1.19 to 2.65, p=0.005, for a 5mm decrease in cervical length). This was unaffected by adjustment for smoking, parity and maternal height which were also associated (adjusted OR 1.70, p<0.012). There was no association between cervical length and spontaneous preterm birth <37 weeks in either univariate (OR 1.19, 95% CI 0.99 to 1.42, p=0.059 for a 5mm decrease in cervical length) or adjusted analysis (adjusted OR 1.14, 95% CI 0.95 to 1.38, p=0.16). <u>Prediction of spontaneous preterm birth <37 weeks (secondary outcome)</u> ≤28mm (1st centile) • Sn 0.03 (95% CI 0.03 to 0.04) • Sp 0.99 (0.98 to 0.99) • **PPV 0.10 (0.09 to 0.11)** • NPV 0.96 (0.95 to 0.97)

Study reference	Study design	Eligibility	Population/included studies	Index test and reference standard	Test accuracy
			Higher rate of nulliparous women among the screened vs non-screened group (49.3% vs 43.2%; p<0.0001). Comparable preterm birth rates between screened and not screened.		• **LR+ 2.52 (0.78 to 8.15)** • LR- 0.98 (0.94 to 1.02) ≤31mm (5th centile) • Sn 0.11 (0.11 to 0.12) • Sp 0.95 (0.94 to 0.96) • PPV 0.09 (0.08 to 0.09) • NPV 0.96 (0.95 to 0.97) • LR+ 2.20 (1.19 to 4.07) • LR- 0.93 (0.87 to 1.01) ≤33mm (15th centile) • Sn 0.22 (0.20 to 0.24) • Sp 0.85 (0.83 to 0.86) • PPV 0.06 (0.06 to 0.06) • NPV 0.96 (0.95 to 0.97) • LR+ 1.44 (0.95 to 2.17) • LR- 0.92 (0.82 to 1.03) ≤35mm (25th centile) • **Sn 0.36 (0.32 to 0.39)** • **Sp 0.75 (0.74 to 0.76)** • PPV 0.06 (0.06 to 0.06) • NPV 0.96 (0.95 to 0.97) • LR+ 1.42 (1.06 to 1.91) • LR- 0.86 (0.73 to 1.01) ≤37mm (35th centile) • Sn 0.53 (0.48 to 0.59) • Sp 0.65 (0.64 to 0.67)

Study reference	Study design	Eligibility	Population/included studies	Index test and reference standard	Test accuracy
					• PPV 0.06 (0.06 to 0.06) • NPV 0.97 (0.96 to 0.98) • LR+ 1.52 (1.24 to 1.87) • LR- 0.72 (0.58 to 0.90) The best sensitivity was obtained with cervical length 37mm with Sn 53% at low Sp 65% (AUC 0.582). Length ≤28mm gave the best PPV and LR+ which was still very poor, with the key limitation of exclusion of 7/11 women with length ≤25mm. <u>The primary outcome of preterm birth <34 weeks</u> 37mm was also assessed to be optimum with slightly improved Sn 59% and Sp 65% (AUC 0.689).

Table 25. Studies relevant to criterion 9

Study reference	Study design	Eligibility	Population/included studies	Intervention and Comparator	Preterm birth/labour risk	Other outcomes
Progesterone						
Romero et al 2018[10]	Systematic review with meta-analysis	RCTs comparing vaginal	5 RCTs in n=974 women with singleton pregnancy	Vaginal progesterone (n=498) vs placebo (n=476).	Pessary reduced the risk of preterm birth at all	<u>Neonatal</u>

Study reference	Study design	Eligibility	Population/included studies	Intervention and Comparator	Preterm birth/labour risk	Other outcomes
	of individual patient data (IPD). Aim: to evaluate whether vaginal progesterone prevents preterm birth and improves perinatal outcomes in asymptomatic women with singleton gestation and a mid-trimester sonographic short cervix.	progesterone with placebo/no treatment in singleton pregnancies where the primary aim was to prevent preterm birth (not specified spontaneous) in women with cervical length ≤25mm (or to prevent preterm birth in women with other risk factors but where data was available for those with short cervix). Exclusions: • quasi-randomised trials • symptomatic women • first trimester administration to prevent miscarriage	(all included in ITT analysis). Studies: • Fonesca 2007, International, 5 centres. n=250 women with CL ≤15mm (single or multiple pregnancy): IPD available for n=226. Progesterone dose 200mg daily from 24 to 33^{+6} weeks; 92% compliance (94% placebo) • O'Brien 2007, International, 53 centres. n=659 women with single pregnancy and previous spontaneous preterm birth (SPTB): IPD available for n=31. Progesterone dose 90mg daily from 18-22 to 37 weeks; 100% compliance (95% placebo) • Cetingoz 2011, Turkey, single centre. n=160 with multiple or twin and prior SPTB or uterine malformation: IPD available for n=8.	Progesterone dose 90-200mg daily. Mean gestation at randomisation: 22^{+6} weeks (18-25) to 34-37 weeks.	gestations <36 weeks (all high quality evidence). <33 weeks (primary outcome) • 14% (70/498) progesterone vs 22% placebo (107/476); relative risk (RR) 0.62 (95% CI 0.47 to 0.81); p=0.0006; I^2 0%; number needed to treat (NNT) 12 SPTB <33 weeks • 12% (60/498) vs 17% (82/476); RR 0.70 (0.51 to 0.95); p=0.02 I^2 0%; NNT 19 <34 weeks • 17% (86/498) vs 26% (126/476); RR 0.65 (0.51 to 0.83); p=0.0006; I^2 0%; NNT 11 SPTB <34 weeks • 15% (73/498) vs 20% (97/476); RR 0.72 (0.55 to 0.95); p=0.02; I^2 0%; NNT 18 <32 weeks • 12% (62/498) vs 19% (92/476); RR 0.64 (0.48 to 0.86); p=0.003; I^2 0%; NNT 14 <28 weeks	Progesterone reduced risk of (all high quality evidence): Low birthweight (<2,500g): • 29% (144/497) vs 36% (168/473); RR 0.82 (95% CI 0.68 to 0.98); p=0.03; I^2 0%; NNT16 Very low birthweight (<1,500g) • 10% (50/497) vs 16% (77/473); RR 0.62 (0.44 to 0.86); p=0.004; I^2 0%; NNT16 NICU admission: • 17% (83/496) vs 25% (117/474); RR 0.68 (0.53 to 0.88); p=0.003; I^2 0%; NNT 13 Respiratory distress syndrome (RDS): • 5% (17/365) vs 10% (37/358); RR 0.47 (0.27 to 0.81); p=0.007; I^2 0%; NNT 18 Composite neonatal morbidity/mortality*: • 8% (29/365) vs 14% (49/358); RR 0.59 (0.38 to 0.91); p=0.02; I^2 0%; NNT18 * RDS, IVH, NEC, sepsis, neonatal death No effect on: Necrotizing enterocolitis (NEC)

Study reference	Study design	Eligibility	Population/included studies	Intervention and Comparator	Preterm birth/labour risk	Other outcomes
		Subsequent exclusion of 7/12 identified RCTs: • otherwise high-risk women where data on cervical length (CL) could not be collected (5 RCTs) • women also received cerclage (1 RCT) • where IPD could not be obtained Searches: MEDLINE, EMBASE, LILACS, CINAHL, the Cochrane Central Register of Controlled Trials, Google Scholar. No language restrictions, supplemented by hand-searching.	Progesterone dose 100mg daily from 24 to 34 weeks; 100% compliance (both groups) • Hassan 2011, International, 44 centres. n=465 women with single pregnancy and CL 10-20mm: IPD available for n=458. Progesterone dose 90mg daily from 20-23^{+6} to 36^{+6} weeks; 89% compliance (93% placebo) • Norman 2016, 66 centres in UK and Sweden. n=1228 with single pregnancy and prior SPTB; or CL≤25mm; or positive fFN plus other risk factors: IPD available for n=251. Progesterone dose 200mg daily from 22-24 to 34 weeks; 63% compliance (69% placebo) 3 studies in women with short cervix provided 96% of the data.		• 8% (38/498) vs 11% (54/476); RR 0.67 (0.45 to 0.99); p=0.04; I^2 0%; NNT 27 No effect on overall preterm birth at <37 weeks: • 38% (187/498) vs 42% (199/476); RR 0.90 (0.77 to 1.05); p=0.19; I^2 0% Effect at <36 weeks: • 28% (139/498) vs 35% (166/476); RR 0.80 (0.67 to 0.97); p=0.02; I^2 0%; NNT14 and all earlier gestations (including <35 and <30) For the primary outcome <33 weeks, Norman (RR 0.74, 0.48 to 1.12) Centingoz (0.33, 0.02 to 6.37) and O'Brien (0.4, 0.05 to 3.13) showed no association, but the latter 2 studies based on n=8 and n=31 participants only. Hassan showed an effect (0.55, 0.33 to 0.92) while the effect in Fonesca was borderline (0.60, 0.36 to 1.00).	• 2% (11/495) vs 3% (12/475); RR 0.89 (0.41 to1.93); p=0.77; I^2 0% (low quality evidence) Intraventricular haemorrhage (IVH): • 1% (5/494) vs 2% (10/475); RR 0.50 (0.18 to 1.38); p=0.18; I^2 0% (low quality evidence) Proven sepsis: • 4% (18/494) vs 6% (28/470); RR 0.61 (0.34 to 1.08); p=0.09; I^2 0% (moderate quality evidence) Bronchopulmonary dysplasia • 3% (11/367) vs 4% (13/340); RR 0.77 (0.35 to 1.68); p=0.51; I^2 0% (low quality evidence) Retinopathy of prematurity (RoP) • 2% (6/365) vs 1% (3/358); RR 1.78 (0.49 to 6.47); p=0.38; I^2 29% (low quality evidence) Fetal death • 2% (9/498) vs 2% (8/476); RR 1.06 (0.41 to 2.72); p=0.91; I^2 0% (low quality evidence) Neonatal death • 1% (7/498) vs 3% (15/476); RR 0.44 (0.18

Study reference	Study design	Eligibility	Population/included studies	Intervention and Comparator	Preterm birth/labour risk	Other outcomes
		Search: September 2017. IPD collected from study authors.	Characteristics: mean age 28 years; 38% White, 39% Black, 19% Asian ethnicity; 54% European; 45% nulliparous; 30% with history of ≥1 SPTB; 76% with length 10-20 mm (13% 21-25mm, 11% <10mm); randomised at mean 22^{+6} weeks. 4 studies were assessed to be at low risk of bias across domains. Norman was at risk of bias for attrition related to the main childhood outcome (Bayley-III cognitive score) but not for obstetric or neonatal, and at risk of compliance bias making the trial potentially underpowered to detect an outcome.		Subgroup analysis found an effect in women with no prior SPTB (n=686): RR 0.65 (95% CI 0.45 to 0.94) (also in women with past SPTB By cervical length there was effect only for those with CL 10-20mm but they made up the majority population (n=741): RR 0.59 (0.42 to 0.81) Both 90-100mg and 200mg doses effective, By gestational age at randomisation 22-25 weeks (n=703; RR 0.58, 0.42 to 0.78) was effective rather than 18-21 weeks though again the majority population. By ethnicity, significant effect for White women. However, test for interaction across all subgroups was not significant.	to 1.07); p=0.07; I^2 0% (low quality evidence) Perinatal death • 3% (16/498) vs 5% (23/476); RR 0.66 (0.35 to 1.22); p=0.19; I^2 0% (moderate quality evidence) Apgar score <7 at 5 mins • 8% (38/491) vs 9% (43/469); RR 0.83 (0.55 to 1.26); p=0.39; I^2 0% (moderate quality evidence) Congenital anomaly • 1% (4/491) vs 1% (6/469); RR 0.72 (0.23 to 2.26); p=0.57; I^2 0% (low quality evidence) Mechanical ventilation • 8% (28/365) vs 12% (43/358); RR 0.65 (0.41 to 1.01); p=0.06; I^2 0% (moderate quality evidence) No effect on childhood outcome at 2 years (one study; all low quality evidence): Bayley-III cognitive composite score • 95.5 vs 97.7; mean difference -2.17 (-7.16 to +2.83); p=0.40

Study reference	Study design	Eligibility	Population/included studies	Intervention and Comparator	Preterm birth/labour risk	Other outcomes
						Moderate/severe neurodevelopmental impairment • 12% (10/81) vs 9% (7/77); RR 1.36 (0.54 to 3.39); p=0.51 Visual or hearing impairment • 0% (0/100) vs 2% (2/87); RR 0.17 (0.01 to 3.58); p=0.26 Disability in renal, gastrointestinal, or respiratory function • 1% (1/91) vs 1% (1/84); RR 0.92 (0.06 to 14.52); p=0.95 <u>Maternal adverse effects</u> No difference in any maternal events (not specified): • 12% (51/424) vs 11% (47/422); RR 1.21 (0.87 to 1.69); p=0.26; ; I^2 5% (moderate quality evidence)
Van Os 2015	RCT Multicentre, The Netherlands (Nov 2009 to Aug 2013)	Nulliparous, low-risk women (≥18 years) with singleton pregnancy and short cervix (≤30mm*) at routine anomaly	n=80 women n=80 included in ITT analysis Mean age 30 years, mean CL 26mm, 69% White, n=7 in the progesterone group with cervical	Vaginal progesterone (n=41) (200mg daily) vs identical placebo (n=39) From 22 to 34 weeks	No difference in SPTB (secondary outcomes) <37 weeks: • 15% (6/41) vs 13% (5/39); RR 1.17 (95% CI 0.39 to 3.52)	<u>Neonatal</u> Progesterone made no difference to the primary neonatal outcome: Composite of RDS, IVH (>grade II), NEC (>stage I), bronchopulmonary

Study reference	Study design	Eligibility	Population/included studies	Intervention and Comparator	Preterm birth/labour risk	Other outcomes
	Aim: to assess the effectiveness of vaginal progesterone in reducing adverse neonatal outcome due to preterm birth in low risk women with a short cervical length.	scan (18 to 22 weeks) Exclusions: • multiple pregnancy • history of preterm birth(<34 weeks) • cervical cerclage • congenital anomaly • symptoms of labour *measure repeated 2 weeks later and only those with 2 CL measurements ≤30mm were eligible.	procedure/uterine anomaly and n=2 in the placebo group (unclear whether significant); n=3 and n=2 in each group with bacterial vaginosis. n=20,234 screened n=375 women with CL ≤30mm (1.8%; expected prevalence had been 10%) n=151 with confirmed CL >30mm at second measure: n=80 agreed participation Of remaining 224/375 who did not have confirmed short cervix: n=121 (32%) had CL >30mm at second measure; n=103 (28%) refused to undergo a second measure. Low prevalence of women with preterm birth so termination was ended early after enrolling only 80 of the planned 1,920 over 4 years.		<34 weeks: • 7% (3) vs 10% (4); RR 0.73 (0.17 to 3.06) <32 weeks: • 2% (1) vs 8% (3); RR 0.33 (0.04 to 2.99)	dysplasia, proven sepsis, and death before discharge: • 5% (2/41) vs 11% (4/39); RR 0.47 (95% CI 0.09 to 2.4) • comprising n=2 vs n=2 with RDS; n=1 vs n=2 who died before discharge; and n=1 cases of bronchopulmonary dysplasia in the placebo group NICU admission • 7% (3) vs 13% (5); RR 0.53 (0.12 to 2.25) Days in NICU • 3 (1.5 to 5.5) vs 8 (7 to 31); mean difference -5.0 (-27 to 0.15) <u>Maternal</u> No difference in adverse effects: • 12% (4) vs 23% (7); RR 0.51 (0.16 to 1.6) Used >80% of tablets: • 57% (23) vs 50% (18) (risk difference not reported) Used >50%:

Study reference	Study design	Eligibility	Population/included studies	Intervention and Comparator	Preterm birth/labour risk	Other outcomes
						• 80% (32) vs 75% (27) (risk difference not reported)
Cervical pessary						
Saccone et al 2017[12]	Systematic review with meta-analysis Aim: to evaluate the effectiveness of cervical pessary for preventing spontaneous preterm birth (SPTB) in singleton pregnancies with a second trimester short cervix.	RCTs comparing cervical pessary with expectant management for prevention of SPTB in singleton pregnancies with cervical length ≤25mm (as measured by transvaginal ultrasound). Exclusions (for review): • quasi-randomised trials • multiple pregnancies Exclusions (across studies): • symptomatic women • placenta praevia • previews cone biopsy	3 RCTs in n=1,420 women with singleton pregnancy. Studies: • Goya 2012. Spain (5 centres), n=380 women, 11% with prior preterm birth. Study duration 36 months. • Hui 2013. China (1 centre), n=108 women, prior preterm birth: 6% intervention vs 11% control group. Study duration 29 months. • Nicolaides 2016. Multicentre (16 centres), n=932 women, prior preterm birth: 15% intervention vs 18% control group. Study duration 53 months. All studies were assessed to be of high quality with low risk of bias for	Arabian pessary (n=708, 49.8%) vs expectant management (n=712, 50.2%) in all studies. Mean gestation at randomisation: 22 weeks across all studies (range 20 to 24^{+6}). All studies removed at 37 weeks or earlier if rupture of membranes, vaginal bleeding or contractions. Progesterone: Nicolaides, n=359 of women with cervical length ≤15mm (38.5%; 25% of review population) were assigned to vaginal progesterone (200mg to week 33^{+6}). Unclear for Goya and Hui.	Pessary had no effect on risk of SPTB at any gestation. <34 weeks (primary outcome of review and studies): • 10.2% pessary (72/708) vs 14.6% control (104/712); RR 0.71 (95% CI 0.21 to 2.42); I^2 90% • NB. Significant in Goya which had rate 6.3% vs 26.8% controls (RR 0.24 [0.13 to 0.43]); rates in other trials comparable and lower (9.4 vs 5.5 Hui and 11.8 vs 10.7 Nicolaides) Secondary outcomes: <37 weeks: • 20.2% (49/243) vs 50.2% (123/245); RR 0.50 (0.23 to 1.09); I^2 0%	<u>Neonatal</u> No difference in: Average birthweight (Hui and Nicolaides only): • MD -113.00 grams (-364.95 to +138.95) Low birthweight (Nicolaides only): • 96/465 (20.6%) vs 84/497 (18.4%); RR 1.15 (0.88 to 1.49) NEC (Goya and Nicolaides only): • 6/655 (0.9%) vs 5/657 (0.8%); RR 0.95 (0.11 to 8.07); I^2 50% RDS: • 38/708 (5.4%) vs 49/712 (6.9%); RR 0.80 (0.22 to 3.00); I^2 83% IVH: • 9/708 (1.3%) vs 6/712 (0.8%); RR 45% (0.15 to 6.04); I^2 0.94 NICU admission (Hui and Nicolaides only): • 73/518 (14.1%) vs 75/522 (14.4%); RR

Study reference	Study design	Eligibility	Population/included studies	Intervention and Comparator	Preterm birth/labour risk	Other outcomes
		• major fetal abnormality • cerclage in situ Searches: MEDLINE, Scopus, ClinicalTrials, the PROSPERO International Prospective Register of Systematic Reviews, EMBASE, and the Cochrane Central Register of Controlled Trials Search: February 2016.	randomisation and allocation, incomplete outcome data, and selective reporting. No studies were blinded, which was not considered feasible. Publication bias could not be assessed due to small numbers of studies.		• NB. Rates 21.6% vs 59.5% in Goya and 15.1% vs 18.2% in Hui (not assessed by Nicolaides) <32 (Nicolaides only): • 9.9% (46/465) vs 7.5% (35/467); RR 1.32 (0.87 to 2.01) <28 weeks: • 4.4% (31/708) vs 4.8% (34/712); RR 0.70 (0.18 to 2.67); I^2 78% PPROM (Goya and Hui only) • 9/243 (3.7%) vs 25/245 (10.2%); RR 0.39 (0.09 to 1.71); I^2 72% Mean difference (MD) in gestational age at delivery: • 1.63 weeks (-0.82 to 4.07) The researchers consider the significant benefit in Goya may be due to greater training on pessary insertion at this site, and that position was confirmed by ultrasound.	1.02 (0.73 to 1.42); I^2 20% Neonatal mortality: • 8/708 (1.1%) vs 6/712 (0.8%); RR 1.32 (0.48 to 3.65); I^2 0% Perinatal mortality: • 16/708 (2.3%) vs 11/712 (1.5%); RR 1.44 (0.69 to 3.04); I^2 0% <u>Maternal adverse effects</u> Bacterial vaginosis: • 183/708 (25.8%) vs 162/712 (22.8%); RR 1.14 (0.95 to 1.36); I^2 0% Difference only in vaginal discharge: • 264/708 (37.3%) vs 128/712 (18.0%); RR 2.12 (1.84 to 2.44); I^2 0% No difference in rate of Caesarean (Goya only): • 41/190 (21.6%) vs 40/190 (21.1%); RR1.02 (0.70 to 1.51)

Study reference	Study design	Eligibility	Population/included studies	Intervention and Comparator	Preterm birth/labour risk	Other outcomes
					The multiple sites in Nicolaides may introduce potential for less experienced management in certain centres, while training was not mentioned in Hui. However, preterm rates were also notably high in Goya.	
Saccone et al 2017[14]	RCT Single centre, Italy (March 2016 to May 2017) Aim: to test whether cervical pessary reduces risk of preterm birth <34 weeks in asymptomatic women with singleton pregnancies and no prior SPTB but with short cervical length (on transvaginal ultrasound).	Women (18 to 50 years) with singleton pregnancy referred following detection of short cervix (≤25mm) at routine anomaly scan (18^{+0} to 23^{+6} weeks) Exclusions: • multiple pregnancy • history of SPTB (including mid-trimester loss from 16 weeks plus) • fetal abnormality	n=300 women n=300 included in ITT analysis Mean age 28 years, 89% white ethnicity, 70% nulliparous, 4% with prior cervical surgery. Mean cervical length 11.5mm pessary and 12.5mm control groups Sample representative of n=503 referred for short cervix; others excluded due to not meeting eligibility criteria	Arabian pessary (n=150) vs no pessary (n=150) Mean gestation at randomisation: 22 weeks Insertion to 37 weeks or earlier removal if indicated by signs of labour. Randomisation stratified by length <20mm or 20 to 25mm Position rechecked at monthly follow-up to delivery. At time of randomisation vaginal swabs were taken and any bacterial infections (24% pessary and 27% controls) were treated (22% and 25%).	Pessary reduced risk of SPTB at 34 weeks and 37 weeks: <34 weeks (primary outcome) • 7.3% pessary (11/150) vs 15.3% control (23/150); MD -8.0 (-15.7 to -0.4); RR 0.48 (95% CI 0.24 to 0.95); p=0.04 <37 weeks (secondary) • 20.0% (30/150) vs 32.7 (49/150); MD -12.7 (-22.9 to -2.3); RR 0.61 (0.41 to 0.91); p=0.02 No difference for earlier gestations: <32 weeks • 6.7% (10) vs 9.3% (14); RR 0.71 (0.33 to 1.56)	<u>Neonatal</u> Significant difference in: Birthweight • 2889.9 grams (2770.3 to 3009.6) vs 2644.6 grams (2513.5 to 2775.7); MD 245.3 (69.2 to 421.4); p=0.006 NICU admission • 10.0% (15) vs 18.7 (28); MD -8.7 (-17.1 to -0.3); RR 0.54 (0.30 to 0.96); p=0.04 Composite perinatal outcome (≥1 of NEC, IVH, RDS, ROP bronchopulmonary dysplasia, sepsis, neonatal death): • 14.7% (22) vs 32.0 (48); MD -17.3 (-27.0 to -7.3); RR 0.46 (0.29 to 0.72); p=0.01 No difference in:

Study reference	Study design	Eligibility	Population/included studies	Intervention and Comparator	Preterm birth/labour risk	Other outcomes
		• P-PROM • vaginal bleeding • bulging membrane • uterine contractions • suspected chorioamnionitis • placenta praevia •		Progesterone: Women with length ≤20mm were given vaginal progesterone (200mg) to 37 weeks: • 133/150 of pessary group (89%) • 125/150 of controls (83%) (No significant effect of treatment in post-hoc analysis)	<28 weeks • 4.0% (6) vs 6.0% (9); RR 0.67 (0.24 to 1.83) Pessary increased mean age at delivery • 37.6 weeks (37.1 to 38.2) vs 36.2 weeks (35.5 to 36.9); MD 1.4 (0.6 to 2.3); p=0.001 (SPTB defined as spontaneous labour or P-PROM. Iatrogenic preterm births excluded.)	Neonatal death • 0.7% (1) vs 2.0% (3) 0.33 (0.04 to 3.17) Perinatal death (composite of fetal death >20 weeks or neonatal death) • 1.3% (2) vs 2.7% (4); RR 0.50 (0.09 to 2.69) <u>Maternal adverse effects</u> Pessary increased vaginal discharge: • 86.7% (130) vs 46.0% (69); MD 40.7 (30.1 to 50.3); RR 1.88 (1.57 to 2.27); p<0.001 No difference in: Pelvic discomfort • 3.3% (5) vs 0.6% (1); RR 5.00 (0.59 to 42.29) Chorioamnionitis • 3.3% (5) vs 4.7% (7); RR 0.71 (0.23 to 2.20) Caesarean • 30.0% (45) vs 38.0 (57%); RR 0.79 (0.57 to 1.09) Operative vaginal delivery • 3.3% (5) vs 6.7 (10); RR 0.50 (0.18 to 1.43)

Study reference	Study design	Eligibility	Population/included studies	Intervention and Comparator	Preterm birth/labour risk	Other outcomes
Dugoff et al 2018[13] PoPPS study	RCT 5 sites, US (March 2014 to July 2016) Aim: to see whether pessary prevents preterm birth in women with short cervical length (measured on transvaginal ultrasound) and without previous SPTB	Women (18 to 50) with singleton pregnancy with short cervix (≤25mm) following screening at routine anomaly scan (18^{+0} to 23^{+6} weeks) Exclusions: • multiple pregnancy • history of SPTB (including mid-trimester loss from 16 weeks plus) • fetal abnormality • present or planned cerclage • abnormal smear • vaginal bleeding • bulging membrane	n=122 women n=118 included in ITT analysis excluding n=1 from the pessary group with history prior SPTB (discovered 6 days after enrolment), n=3 in control group lost to follow-up Mean 28 years, majority Black (57% pessary, 66% controls; only 28% white), nulliparous (63% and 69%) Median cervical length 17.6mm pessary group and 19.0mm controls; 36.7% and 41.4% of with length 20 to 25mm. Initial eligible sample: n=17,383 screened n=422 (2.4%) with length ≤25mm N=391 (92.7%) met inclusion criteria, n=213 decline and n=56 not offered due to no trial coordinator. NB. The trial was underpowered. The planned recruitment was n=242 women with n=121 in each trial arm; recruitment had to be	Bioteque pessary (n=61) vs no pessary (n=61) Described to be similar to the Arabian pessary but designed for the treatment of mild uterine prolapse. States 'The smaller diameter of the pessary should encompass the cervix, and the side of the pessary with the larger diameter should face the introitus' Mean gestation at randomisation: 21 weeks Insertion to 37 weeks or earlier removal if indicated by signs of labour. Randomisation stratified by length <20mm or 20 to 25mm. Reported that subjects were contacted monthly to get information about hospital or other	Pessary had no effect on risk of SPTB: <37 weeks (primary outcome*) • 38.3% (23/60) vs 32.8% (19/58); RR 1.17 (95% CI 0.72 to 1.91) <34 weeks • 31.7% (19) vs 25.9% (15); RR 1.22 (0.69 to 2.17) <28 weeks • 18.3% (11) vs 20.7% (12); RR 0.89 (0.43 to 1.85) *NB overall preterm birth was the main outcome; this review has reported only those for SPTB (defined as spontaneous labour or P-PROM). P-PROM <37 weeks • 31.7% (19) vs 25.9% (15); RR 1.20 (0.68 to 2.13) <34 weeks • 30.0% (18) vs 17.2% (10); RR 1.71 (0.86 to 3.38)	No effect on any neonatal outcomes: Birthweight • 2788 grams (1285 to 3188) vs 2843 grams (1035 to 3329) (p=0.58) Sepsis • 11.6% (7) vs 10.3% (6); RR 1.08 (0.39 to 3.03) IVH • 6.7% (4) vs 3.4% (2); RR 1.83 (0.35 to 9.60) NEC • (2) vs 1.7% (1); RR 1.83 (0.17 to 19.6) ROP • 3.3% (2) vs 6.9% (4); RR 0.47 (0.09 to 2.48) Bronchopulmonary dysplasia • 8.3% (5) vs 8.6% (5); RR 0.87 (0.27 to 2.83) RDS • 16.7% (10) vs 15.5% (9); RR 1.01 (0.45 to 2.3 Neonatal death • 5.0% (3) vs 10.3% (6); RR 0.48 (0.13 to 1.84) Intrauterine death • 3.3% (2) vs 8.6% (5); RR 0.39 (0.07 to 1.91)

Study reference	Study design	Eligibility	Population/included studies	Intervention and Comparator	Preterm birth/labour risk	Other outcomes
		• uterine contractions • suspected chorioamnionitis • placenta praevia	terminated in June 2016 due to unable to continue to enrol subjects, given precedence of a competing National Institute of Child Health and Human Development (NICHD) Maternal–Fetal Medicine Unit Network (MFMU) pessary trial after their entrance into the NICHD MFMU Network.	appointments but position not reported as rechecked Progesterone: Women with length ≤20mm were recommended vaginal progesterone (200mg) to 37 weeks: • 84% of pessary group • 91% of controls Cerclage was also used for 3% of pessary group and 5% of controls	Gestational age at delivery • 37.2 weeks (30.0 to 39.1) vs 38.1 (27.8 to 39.4) Pessary removal due to preterm labour/P-PROM was reported for 38.3% (23 women)	Composite adverse neonatal outcome (NEC, IVH, RDS, sepsis, bronchopulmonary dysplasia, neonatal death): • 20.0% (12) vs 17.2% (10); RR 1.16 (0.54 to 2.47) <u>Maternal outcomes</u> Pessary increased vaginal discharge • 73.3% (44) vs 48.3% (28); RR 1.48 (1.15 to 1.89); p=0.002 No difference for: Genitourinary infection • 25.0% (15) vs 24.1% (15); RR 1.09 (0.58 to 2.06) Chorioamnionitis • 11.6% (7) vs 6.9% (4); RR 1.63 (0.50 to 5.28) Caesarean • 16.7 (10) vs 17.2 (10%); RR 0.97 (0.43 to 2.15)
Cruz-Melguizo et al 2018[15]	Non-inferiority RCT Multicentre, 27 sites in Spain (August 2012 to April 2016).	Women (≥18 years) with singleton pregnancy and cervical length ≤25mm measured at the	n=254 randomised. n=246 included in ITT analysis after exclusion of n=8: n=1 from the pessary group where it was not inserted; and from the	Arabian pessary (n=128; ITT n=127; PP=125) vs vaginal progesterone (200mg daily) (n=126; ITT n=119; PP n=118)	No difference in the rate of SPTB <34 weeks (primary outcome, ITT): • 14% (18/127) pessary vs 14% (17/119)	No effect on any neonatal outcomes (PP analysis): Birthweight • 2,855 grams vs 2,921; p=0.52

Study reference	Study design	Eligibility	Population/included studies	Intervention and Comparator	Preterm birth/labour risk	Other outcomes
	Aim: To compare the effectiveness of cervical pessary and vaginal progesterone to prevent SPTB in pregnant women with cervical length ≤25m (measured by transvaginal ultrasound)	anomaly scan (19 to 22 weeks). Exclusions: • ≥3 prior preterm births • uterine abnormality • prior cone biopsy or loop excision • fetal abnormality • uterine contractions • vaginal bleeding • bulging membrane • P-PROM	progesterone group n=4 not receiving a single dose of treatment and n=3 not meeting criteria. All secondary outcome results are reported only for the per protocol (PP) analysis including n=243 after exclusion of 2 women with medically induced deliveries and 1 with major protocol deviation (non-receipt of progesterone for >7 days). Of analysed women n=6 in the pessary group and n=5 of the progesterone group received double treatment at doctor's discretion and, while n=1 and n=2 of each respective group received additional cerclage. Characteristics of n=243: median age 33 years, 77% white, 46% nulliparous, previous preterm birth: 9% pessary, 13% progesterone groups. Median cervical length 21mm.	Treatment given from between 20^{+1} and 23^{+6} (mean 21 weeks) Treatment to between 34^{+4} and 37 weeks (or earlier if symptoms of labour). Cervical length re-measured at monthly assessment for women with pessary. Bacterial culture performed prior to randomisation with treatment if indicated (given to 29% pessary 24% of progesterone groups).	progesterone; risk difference (RD) -0.11%; 95% CI -8.85% to 8.62% • PP analysis: -0.01 (-8.84 to 8.83) • (non-inferiority margin set at 4%) No difference in secondary outcomes (PP analysis only): SPTB <37 weeks • 22% (27/125) vs 21% (25/118); RD 0.41 (-9.90 to 10.73) SPTB <28 weeks • 8% (10) vs 8% (9); RD 0.37 (-6.38 to 7.12) P-PROM <37 weeks • 10% (12) vs 9% (11); RD 0.28 (-7.08 to 7.64) P-PROM <34 weeks • 6% (7) both groups; RD -0.33 (-6.20 to 5.53) Mean age at delivery • 37 weeks both groups	Birthweight <2,500 grams • 26% (32/125) vs 21% (25/118); p=0.38 Neonatal/fetal death • 5% (6) vs 3% (3); p=0.35 NICU • 12% (14) both groups RDS • 6% (7) vs 5% (6); p=0.81 IVH • 0 vs 0.9% (1); p=0.31 NEC • 1.7% (2) vs 0; p=0.16 ROP • 1.7% (2) vs 0.9% (1); p=0.58 <u>Maternal</u> No difference in overall reporting of adverse effects: • 16% pessary vs 11% progesterone (numbers and detail not reported); p=0.27 However, increased reporting in the pessary group of: • discharge: 87% vs 71%; p=0.002

Study reference	Study design	Eligibility	Population/included studies	Intervention and Comparator	Preterm birth/labour risk	Other outcomes
			9% of controls and 11% of progesterone group with length ≤15mm.		Planned subgroup analysis: primary outcome not influenced by cervical length, parity, prior preterm birth or bacterial cultures.	• vaginal discomfort: 27% vs 3%; p<0.001 No differences reported to infections, pain or sexual activity. Pessary insertion described as unpleasant by 22% (28/127) though only 3% (4/127) described as unbearably painful. More emergency department visits were reported for the pessary group during the first month (25% vs 15%; p<0.05) but there was no difference afterwards. Early pessary removal required by 3% (4/127) for tolerability.
Cervical cerclage						
Berghella et al 2017[16]	Systematic review with IPD meta-analysis Aim: to evaluate the effectiveness of cervical cerclage for preventing SPTB in singleton pregnancies	RCTs comparing cervical cerclage with no cerclage in asymptomatic, singleton pregnancies with cervical length <25mm (as measured by transvaginal ultrasound).	5 RCTs in n=419 asymptomatic women with cervix <25mm (mean 12mm) singleton pregnancy and no history of SPTB. Characteristics: mean age 29.7 years; 53% White, 36% Black; mean CL 12.7mm; 23% with prior cone biopsy. Studies:	Cerclage (n=224) vs no cerclage (n=195) Mean 22 weeks gestation to 36^{+0} to 37^{+6} weeks or earlier if indicated. 3 studies used McDonald cerclage (one with permanent filament, 2 with braided tape), To used Shirodkar (with braided tape); and	Pessary had no effect on risk of SPTB at any gestation. <35 weeks (primary outcome of review): • 21.9% pessary (49/224) vs 27.7% control (54/195); RR 0.88 (95% CI 0.63 to 1.23); I^2 0% • (all individual studies had non-significant results)	<u>Neonatal</u> No difference in: Average birthweight: • 2635 grams vs 2540 grams; RD 94.65 (–146.23 to 335.53); I^2 0% Low birthweight: • 18.8% (42/224) vs 25.1% (49/195); RR 0.88 (0.44 to 1.74); I^2 52% Very low birthweight:

Study reference	Study design	Eligibility	Population/included studies	Intervention and Comparator	Preterm birth/labour risk	Other outcomes
	with a second trimester short cervix and no history of prior SPTB.	Exclusions (for review): • quasi-randomised trials • multiple pregnancies • symptomatic women • previous SPTB • physical examination-indicated (e.g. cervical dilation on examination) Searches: MEDLINE, ClinicalTrials, PROSPERO, and the Cochrane Central Register of Controlled Trials. No language restrictions Search: February 2017.	• Rust 2001. US, n=105 women with cervix <25mm at 16-24 weeks; excluding those with prior SPTB 16-36 weeks; primary outcome SPTB <34 weeks • Althuisius 2001. Netherlands, n=9 women with cervix <25mm at 14-27 weeks; excluding those with prior SPTB 17-33 weeks; primary outcome SPTB <34 weeks • To 2004. Multicentre, n=209 women with cervix ≤15mm at 22-24 weeks; excluding those with prior SPTB 16-32 weeks; primary outcome SPTB <33 weeks • Berghella 2004. US, n=21 women with cervix <25mm at 14-24 weeks; excluding those with prior SPTB 16-34 weeks; primary outcome SPTB <35 weeks • Otsuki 2016. Japan, n=75 women with cervix <25mm at 16-	Otsuki half McDonald and half Shirodkar (with braided tape). <u>Additional treatments</u> No studies used progesterone. • Rust, inpatient bed rest for 48–72 for all women followed by amniocentesis, urogenital cultures, antibiotics and indomethacin • Althuisius, cerclage group: perioperative antibiotics up to 6 days plus indomethacin; both groups: similar activity restriction • To, no interventions • Bergella, indomethacin at discretion of physician • Otsuki, cerclage group: tocolytic drugs and antibiotics up to 2	Secondary outcomes: <37 weeks: • 36.2% (81/224) vs 41.0% (80/195); RR 0.93 (0.73 to 1.18); I^2 57% <34 weeks: • 20.1% (45/224) vs 25.1% (49/195); RR 0.89 (0.63 to 1.27)l I^2 0% <32 weeks: • 17.0% (38/224) vs 20.0% (39/195); RR 0.96 (0.64 to 1.42); I^2 0% <28 weeks: • 11.6% (26/224) vs 11.3% (22/195); RR 1.15 (0.68 to 1.93); I^2 0% PPROM • 20.5% (34/166) vs 13.6% (23/169); RR 1.52 (0.94 to 2.46); I^2 0% <u>Subgroup analysis</u> Found significant effect upon SPTB <35 weeks in:	• 9.8% (22/224) vs 10.8 (21/195); RR 0.97 (0.57 to 1.68); I^2 0% NEC: • 0/14 vs 0/16 RDS: • 14.3% (2/14) vs 12.5% (2/16); RR 1.33 (0.23 to 7.74); I^2 0% IVH: • 7.1% (1/14) vs 0/16; RR 3.90 (0.18 to 85.93); I^2 0% Sepsis: • 14.3% (2/14) vs 12.5% (2/16); RR 1.33 (0.23 to 7.74); I^2 0% NICU admission: • 4.5% (3/67) vs 10.5% (4/38); RR 0.80 (0.26 to 2.47); I^2 31% Neonatal mortality: • 5.9% (7/118) vs 6.5% (6/92); RR 1.08 (0.41 to 2.86); I^2 0% • <u>Maternal adverse effects</u> Bacterial vaginosis: • 183/708 (25.8%) vs 162/712 (22.8%); RR 1.14 (0.95 to 1.36); I^2 0% Difference only in vaginal discharge: • 264/708 (37.3%) vs 128/712 (18.0%); RR

Study reference	Study design	Eligibility	Population/included studies	Intervention and Comparator	Preterm birth/labour risk	Other outcomes
		Contact with study authors to access IPD.	26 weeks; excluding those with prior SPTB 16-36 weeks; gestational age at delivery. All trials judged to be at low risk of bias, except for participant blinding. Otsuki had uncertain outcome blinding, and both Otsuki and Althuisius had uncertain risk of other bias. Little statistical heterogeneity between studies. No evidence of publication bias. Quality of evidence overall graded as low because of imprecision in results.	days; both groups: bed rest for 7 days	• women with CL <10mm: 39.5% (30/76) vs 58.0% (29/50); RR 0.68 (0.47 to 0.98); I^2 0% • White women: 22.1% (21/95) vs 37.5% (33/88); RR 0.59 (0.37 to 0.94); I^2 0% • Tocolytics and cerclage vs no tocolytics and no cerclage: 17.5% (20/114) vs 28.6% (40/140); RR 0.61 (0.38 to 0.98) • Tocolytics and cerclage vs tocolytics and no cerclage: 17.5% (20/114) vs 32.7% (18/55); RR 0.54 (0.31 to 0.93) – benefit of tocolysis • Antibiotics and cerclage vs antibiotics and no cerclage: 18.3% (20/109) vs 31.5% (17/54); RR 0.58 (0.33 to 0.98) – cerclage benefit if antibiotics are needed • (no difference vs no antibiotics and no cerclage)	2.12 (1.84 to 2.44); I^2 0% No difference in rate of Caesarean (Goya only): • 41/190 (21.6%) vs 40/190 (21.1%); RR1.02 (0.70 to 1.51)

Study reference	Study design	Eligibility	Population/included studies	Intervention and Comparator	Preterm birth/labour risk	Other outcomes
					No effect by type of cerclage or in women of Black ethnicity.	
Antibiotics for bacterial vaginosis						
Subtil et al 2018[17] PREMEVA (Prevention of Very PREterM Delivery by Testing for and Treatment of Bacterial VAginosis)	Double-blind, placebo-controlled RCT 40 centres, France, April 2006 to June 2011.	All pregnant women in the study region were offered screening for bacterial vaginosis during the first trimester (self-collected samples). Bacterial vaginosis was defined by Nugent score ≥7. Screen-positive women (aged ≥18 years) at ≤14 weeks' gestation and with no history of SPTB or mid-trimester loss (from ≥16 weeks) offered participation. Women with known allergy to clindamycin or with vaginal	n=84,530 women screened for bacterial vaginosis. n=5,630 with Nugent score ≥7 n=5,246 women with no history of SPTB n=2,869 included n=2860 included in ITT analysis: n=2, 5 and 2 lost to follow-up from the 2 respective groups Mean age 28 years, gestational age 12[+4] weeks, 52% nulliparous, 2% multiple pregnancy (Exclusions of multiple pregnancies had no effect on results)	1. Single course oral clindamycin (4 days of 300mg twice daily): n=943 2. Triple course oral clindamycin (4 days of 300mg twice daily, once a month for 3 months): n=968 3. Placebo: n=958 Equivalent number of capsules given in each arm: one capsule twice daily for 4 days each month, for 3 months. Treatment given from mean 12 weeks' gestation. Other antibiotics were given if indicated (received by n=502, 17.5%, with no difference between study groups).	Primary outcome: Antibiotics had no effect on late miscarriage (16 to 21 weeks) or spontaneous very preterm delivery (22 to 32 weeks): • 0.8% (8/941) single and 1.5% (14/963) triple course: combined 1.2% clindamycin (22/1904) vs placebo 1.0% (10/956); RR 1.10 (95% CI 0.53 to 2.32); p=0.82 No effect on secondary outcomes: SPTB at <37 weeks: • 4.6% (43) and 5.0% (48): total 4.8% (91) vs 4.1% (39); RR 1.17 (0.81 to 1.69); p=0.40 P-PROM <37 weeks: • 2.3% (21) and 2.2% (21): total 2.2% (42) vs 1.9% (18); RR	<u>Neonatal</u> No effect on any outcome Average birthweight • 3260 grams single and 3250 grams triple course: combined 3250 grams vs placebo 3260 grams; p=0.93 Low birthweight <2500g: • 8.5% (80) and (8.4% (80): total 8.4% (160) vs 7.9% (75); p=0.62 NICU admission • 7.5% (71) and 7.3% (70): 7.4% (141) vs 6.3% (59); RR 1.20 (95% CI 0.89 to 1.60); p=0.23 Sepsis (suspected or proven) • 2.2% (21) and 2.8% (27): 2.5% (48) vs 3.3% (31); RR 0.77 (0.49 to 1.22); p=0.27 Need for ventilation >24 hours • 1.6% (15) and 1.7% (16): 1.6% (31) vs 2.1%

Study reference	Study design	Eligibility	Population/included studies	Intervention and Comparator	Preterm birth/labour risk	Other outcomes
		bleeding in the week before screening were excluded. (Women with history of late miscarriage or SPTB were offered participation in a sub-trial, the results of which are not reported here).		Nugent score not reassessed. n=227, 223 and 189 violations in the 3 respectively, including incomplete treatment (182, 192 and 156) and those found to be outside of inclusion criteria after randomisation. n=1,409 (49%) included in per protocol analysis by pill count which had no effect on results.	1.18 (0.65 to 2.13); p=0.57 (Also measured hospitalisation for threatened preterm delivery or PROM; number of days hospitalised for both)	(20); RR 0.78 (0.43 to 1.42); p=0.38 Neonatal death • 0.2% (2) and 0.1% (1): 0.2% (3) vs 0.2% (2); RR 0.75 (0.10 to 6.44); p>0.99 Intrauterine death • 0.4% (4) and 0.5% (5); 0.5% (9) vs 0.6% (6); RR 0.75 (0.27 to 2.11); p=0.59 <u>Maternal</u> No difference in: Prenatal chorioamnionitis • 1.5% (14) and 1.0% (10): total 1.3% (24) vs 0.8% (8) placebo; RR 1.51 (0.65 to 3.67); p=0.31 Need for antibiotics >24 hours after delivery • 11.9% (112) vs 11.2% (108): 11.6% (220) vs 11.8% (113); RR 0.98 (0.79 to 1.21); p=0.83 Fever during labour • 2.3% (22) and 3.8% (37): 3.1% (59) vs 3.2% (31); RR 0.95 (0.61 to 1.48); p=0.83 Fever postpartum

Study reference	Study design	Eligibility	Population/included studies	Intervention and Comparator	Preterm birth/labour risk	Other outcomes
						• 2.8% (26) and 3.2% (31): 3.0% (57) vs 2.5% (24); RR 1.20 (0.74 to 1.94); p=0.46 Antibiotics were associated with increased risk of: Overall side effects • 2.7% (25/941) and 3.4% (33/963): 3.0% (58/1904) vs 1.3% (12/956); p=0.0035 Diarrhoea • 1.5% (14/941) and 1.7% (16/963): 1.6% (30/1904) vs 0.4% (4/956); p=0.0071 Abdominal pain • 0.5% (5/941) and 0.4% (4/963): 0.5% (9/1904) vs 0; p=0.034 Incomplete treatment • 19.3% (182/941) and 19.9% (192/963): 19.6% (374/1904) vs 16.3% (156/956); p=0.031 No severe adverse events reported.

Appraisal for quality and risk of bias

Quality assessments of included studies are reported below.

Table 26. CASP assessment of Dos Santos et al[5] systematic review

Assessment	Yes, no, unclear	Comment
Are the results of the review valid?		
Did the review address a clearly focused question?	Yes	Asymptomatic pregnant women, clear test and outcome
Did the authors look for the right type of papers?	No – overall Yes – papers relevant to our population	Inclusion criteria included case-control designs which are not optimal for assessing diagnostic accuracy, though all relevant studies in low-risk women were prospective and non-selective
Do you think all the important, relevant studies were included?	Yes	Yes - for the systematic review research question: 5 databases, no language restrictions, study authors contacted. There is no inclusion of fFN testing below 22 weeks' pregnancy or testing or of other thresholds (though this is rightly said to be in accordance with manufacturer's guidance).
Did the review's authors do enough to assess quality of the included studies?	Unclear	QUADAS-2 assessment is performed giving overall results, but detail is not given on the individual assessments per study and specific reason for low/high risk of bias in each domain.
If the results of the review have been combined, was it reasonable to do so?	Unclear	Individual results are given but there is marked variation in sensitivity between studies and heterogeneity is not statistically reported. The potential reasons for this are unclear but may reflect the variable methods.
What are the results?		
Were the overall results clear?	Unclear	The result is clear, though notably this is for a wide age range of testing and serial/single testing with lack of clarity whether these are combined or single measures.

Are the results precise?	No	Limited variation for specificity, but high variability for sensitivity. May be due to the variation in testing protocol.
Applicability/Will the results help locally?		
Can the results be applied to the local population?	Unclear	Asymptomatic low-risk women without risk factors have been analysed separately, though studies are all pre-2000 with none from the UK and some western populations. It is unclear how to apply the results given the wide gestational age of sampling (22-37 weeks) and the range of serial and single testing. The review is only of the standard threshold and age at testing; studies using other testing methods post-2013 would have been identified by this update search but this precludes this from being a full SR on fFN testing in pregnancy.
Were all important outcomes considered?	Unclear	It is not specified whether this is spontaneous preterm birth.
Are the benefits worth the harms?	Unclear	Harms or acceptability are not assessed by the review.
Overall the review has a clear question and is expected to have identified all eligible studies of fFN screening at >22 weeks in low-risk women with singleton pregnancies. However, there is considerable visual heterogeneity in the results for sensitivity and variation in the gestational age and method of testing (single or serial). Further detail on the outcome definition may have been beneficial. The general interpretation seems to be that raised fFN is specific and increases the likelihood of preterm birth but it may be difficult to directly apply the results.		

Table 27. QUADAS-2 assessment of Esplin et al 2017[6]

Domain	Signal:Yes/no/unclear Bias:High/low/unclear	Notes
Domain I: Patient selection		
Was a consecutive or random sample of the population enrolled?	Yes	No exclusions excepting criteria of gestational age and nulliparous with singleton
Was a case-control design avoided?	Yes	
Inappropriate exclusions avoided?	Yes	
Could patient selection have introduced bias?	Low	Likely to be representative of this population
Domain II: Index test		

Index test results interpreted without knowledge of reference standard results?	NA	Screening for the reference standard of spontaneous preterm birth. Both index tests interpreted without knowledge of the other.
Threshold pre-specified?	Yes	Pre-specified thresholds tested at different gestation
Could the conduct or interpretation of the index test have introduced bias?	High: fFN Low: cervical length	Thresholds pre-specified but self-obtained fFN sample. No clear indication of cervical length measurement bias; performed by specifically trained personnel.
Domain III: Reference standard		
Reference standard likely to correctly classify condition?	Yes	Excluding iatrogenic preterm births.
Reference standard results interpreted without knowledge of index test results?	NA*	*NB The study assesses preterm birth rather than preterm labour. It has been assumed that management decisions for symptomatic women in preterm labour (for example use of tocolysis) would be informed by current measures taken while symptomatic, regardless of screening measures (that is, screen performance to predict preterm birth could be considered compatible with that for preterm labour)
Could the conduct or interpretation of the reference standard have introduced bias?	Low	Clearly defined
Domain IV: Test strategy flow and timing		
Was there an appropriate interval between the index test and reference standard?	NA*	NB. It is assumed that ongoing antenatal management for screen-positive and screen-negative women would be equivalent and any prophylactic treatment would be specified. 9% of women with CL <25mm received prophylactic treatment though this was accounted for in sensitivity analysis
Did all participants receive the same reference standard?	Yes	Same definition of spontaneous preterm birth
Were all participants included in analysis?	Yes	Inclusion of all women with at least one test measure and birth outcome data available (exclusion of less than 5%)
Could the participant flow have introduced bias?	Low	

Domain V: Applicability		
Is there concern that the included participants do not match the review question?	Unclear	Applicable to the general population of nulliparous women with singletons; would inherently exclude prior risk factors excepting cervical/uterine abnormality. Expected this could be broadly compatible to low risk women though cannot be sure could be applied to women with previous pregnancy. The ethnic distribution of the population (for example, high proportion of black and Hispanic) may also not be applicable to the UK.
Is there concern that the index test, its conduct or interpretation differ from the review question?	Unclear: (fFN) Low: cervical length	The testing strategies for cervical length and thresholds used are likely to be applicable. Thresholds also applicable for fFN but the sample was self-obtained which may introduce bias.
Is there concern that the target condition as defined by the reference standard does not match the review question?	Low	Spontaneous preterm birth is applicable and clearly defined.

Table 28. QUADAS-2 assessment of van der Ven 2015[9]

Domain	Signal:Yes/no/unclear Bias:High/low/unclear	Notes
Domain I: Patient selection		
Was a consecutive or random sample of the population enrolled?	Yes	No exclusions excepting criteria of women with prior SPTB, symptomatic women and those with fetal anomalies
Was a case-control design avoided?	Yes	
Inappropriate exclusions avoided?	Yes	
Could patient selection have introduced bias?	Low	Likely to be representative of this population
Domain II: Index test		
Index test results interpreted without knowledge of reference standard results?	NA	Screening for the reference standard of spontaneous preterm birth.
Threshold pre-specified?	Yes	

Could the conduct or interpretation of the index test have introduced bias?	Low	No clear indication of cervical length measurement bias; performed by specifically trained personnel.
Domain III: Reference standard		
Reference standard likely to correctly classify condition?	Yes	Excluding iatrogenic preterm births.
Reference standard results interpreted without knowledge of index test results?	NA*	*NB The study assesses preterm birth rather than preterm labour. It has been assumed that management decisions for symptomatic women in preterm labour (for example use of tocolysis) would be informed by current measures taken while symptomatic, regardless of screening measures (that is, screen performance to predict preterm birth could be considered compatible with that for preterm labour)
Could the conduct or interpretation of the reference standard have introduced bias?	Low	Iatrogenic preterm birth was clearly defined and excluded from SPTB
Domain IV: Test strategy flow and timing		
Was there an appropriate interval between the index test and reference standard?	NA*	NB. It is assumed that ongoing antenatal management for screen-positive and screen-negative women would be equivalent and any prophylactic treatment would be specified. 10% of women with CL <30mm received prophylactic treatment
Did all participants receive the same reference standard?	Yes	Same definition of spontaneous preterm birth
Were all participants included in analysis?	No	Only 75% could be linked with the registry to obtain outcome data, though characteristics, including median cervical length and the proportion with short cervix, was comparable between those linked and not. Additionally the preterm birth rate was described to be equivalent to the national rate.
Could the participant flow have introduced bias?	Low	Given the large sample and the above factors, this exclusion is not expected to have introduced bias.
Domain V: Applicability		
Is there concern that the included participants do not match the review question?	Low	Applicable to the general population of nulliparous and multiparous women with singletons and no history of preterm birth. This would inherently make

		these women low-risk (excepting cervical/uterine abnormality).
Is there concern that the index test, its conduct or interpretation differ from the review question?	Low	The testing strategies for cervical length are likely to be applicable. A higher threshold was used to indicate risk, though data was available to allow assessment of lower 25 and 20mm thresholds
Is there concern that the target condition as defined by the reference standard does not match the review question?	Low	Spontaneous preterm birth is applicable and clearly defined.

Table 29. QUADAS-2 assessment of Banos et al 2018[7]

Domain	Signal:Yes/no/unclear Bias:High/low/unclear	Notes
Domain I: Patient selection		
Was a consecutive or random sample of the population enrolled?	Yes	No exclusions excepting high risk pregnancies
Was a case-control design avoided?	Yes	
Inappropriate exclusions avoided?	Yes	
Could patient selection have introduced bias?	Low	
Domain II: Index test		
Index test results interpreted without knowledge of reference standard results?	NA	Screening for the reference standard of spontaneous preterm birth
Threshold pre-specified?	No	Thresholds are not pre-specified; the study constructs ROC curves to find the optimal cut-off and for the 1st 5th and 10th centiles which may not be applicable to other populations.
Could the conduct or interpretation of the index test have introduced bias?	High	No indication of bias to measurement (intra- and inter-observer variation assessed), but the thresholds were not pre-specified.
Domain III: Reference standard		
Reference standard likely to correctly classify condition?	Yes	Excluding iatrogenic preterm births.
Reference standard results interpreted without knowledge of index test results?	NA*	*NB The study assesses preterm birth rather than preterm labour. It has been assumed that

		management decisions for symptomatic women in preterm labour (for example use of tocolysis) would be informed by current measures taken while symptomatic, regardless of screening measures (that is, screen performance to predict preterm birth could be considered compatible with that for preterm labour)
Could the conduct or interpretation of the reference standard have introduced bias?	Low	Clearly defined
Domain IV: Test strategy flow and timing		
Was there an appropriate interval between the index test and reference standard?	NA*	NB. It is assumed that ongoing antenatal management for screen-positive and screen-negative women would be equivalent and any prophylactic treatment would be specified.
Did all participants receive the same reference standard?	Yes	Same definition of spontaneous preterm birth
Were all participants included in analysis?	No	Non-inclusion of 28% of potential recruited sample (207/749) comprising n=153 with poor image quality and n=54 lost to follow-up. The image quality is predominantly expected to relate to the cervical ~~contingency~~ consistency index measure, though women had to have received this measure to be included in the analysis.
Could the participant flow have introduced bias?	Unclear	The effect of participants lost to follow-up is unclear. The sample size is also very low as a result.
Domain V: Applicability		
Is there concern that the included participants do not match the review question?	Low	Low-risk pregnant population with exclusion of risk factors.
Is there concern that the index test, its conduct or interpretation differ from the review question?	High	The index test thresholds may not be applicable.
Is there concern that the target condition as defined by the reference standard does not match the review question?	Low	Spontaneous preterm birth is applicable and clearly defined.

Table 30. QUADAS-2 assessment of Kuusela et al 2015[8]

Domain	Signal:Yes/no/unclear Bias:High/low/unclear	Notes
Domain I: Patient selection		
Was a consecutive or random sample of the population enrolled?	Unclear	Apparent non-selective, consecutive enrolment but very low recruitment rate of 22.7% of those eligible and unclear whether there may have been some representation issues.
Was a case-control design avoided?	Yes	
Inappropriate exclusions avoided?	Unclear	As above very low recruitment. Of the assessed factors only the rate of nulliparous women was significantly different (higher among screened) but past medical history, such as the proportion of with previous preterm delivery, is unclear.
Could patient selection have introduced bias?	Unclear	
Domain II: Index test		
Index test results interpreted without knowledge of reference standard results?	NA	Screening for the reference standard of spontaneous preterm birth
Threshold pre-specified?	No	Thresholds are not pre-specified; the study constructed ROC curves to find the optimal cut-off
Could the conduct or interpretation of the index test have introduced bias?	High	No indication of bias to assessment as performed by trained personnel and using the standard measure of internal to external os. However, as above thresholds were not pre-specified.
Domain III: Reference standard		
Reference standard likely to correctly classify condition?	Unclear	Spontaneous preterm birth is not further defined. The number of iatrogenic preterm births is given suggesting there may have been low risk but it is difficult to be certain there was consistency in the births considered spontaneous or not.
Reference standard results interpreted without knowledge of index test results?	NA*	*NB The study assesses preterm birth rather than preterm labour. It has been assumed that management decisions for symptomatic women in preterm labour (for example use of tocolysis) would be informed by current measures taken while symptomatic, regardless of screening measures (that is, screen performance to predict preterm birth

		could be considered compatible with that for preterm labour)
Could the conduct or interpretation of the reference standard have introduced bias?	Unclear	As above.
Domain IV: Test strategy flow and timing		
Was there an appropriate interval between the index test and reference standard?	NA*	NB. It is assumed that ongoing antenatal management for screen-positive and screen-negative women would be equivalent and any prophylactic treatment would be specified.
Did all participants receive the same reference standard?	Yes	
Were all participants included in analysis?	No	7/11 women with cervical length measure <25mm were entered into a treatment trial and excluded which may affect the analysis.
Could the participant flow have introduced bias?	High	As above, exclusion of 64% of women with measures below the standardly accepted 25mm threshold.
Domain V: Applicability		
Is there concern that the included participants do not match the review question?	High	General pregnant population including some at risk but low inclusion rate of only 22.7% of those eligible.
Is there concern that the index test, its conduct or interpretation differ from the review question?	High	No indication of bias in conduct but higher thresholds than standardly used were tested due to the low number remaining in the study with measure <25mm and uncertain applicability/.
Is there concern that the target condition as defined by the reference standard does not match the review question?	Unclear	As above, spontaneous preterm birth was not specifically defined.

Q2/Criterion 9

Table 31. CASP assessment of Romero et al 2018[10] systematic review

Assessment	Yes, no, unclear	Comment
Are the results of the review valid?		
Did the review address a clearly focused question?	Yes	Asymptomatic women, singleton pregnancy, cervical length ≤25mm, clear intervention and outcome
Did the authors look for the right type of papers?	Yes	RCTs only

Do you think all the important, relevant studies were included?	Yes	Several databases no whether language restrictions, hand-searching and contact with authors to obtain IPD.
Did the review's authors do enough to assess quality of the included studies?	Yes	Assessments are clear
If the results of the review have been combined, was it reasonable to do so?	Yes	Heterogeneity was low data analysis is clear.
What are the results?		
Were the overall results clear?	Yes	Results appear clear with absolute numbers, risk and NNT.
Are the results precise?	Yes	Evidence is graded and high quality for preterm outcomes with lower quality for rarer neonatal outcomes.
Will the results help locally?		
Can the results be applied to the local population?	Unclear	These are asymptomatic women with singleton pregnancy and short cervical length at a gestation relevant to screening. Overall mixed risk population sample including 30% of women with history of preterm birth, though subgroup analysis was available for the primary outcome. Also mixed ethnic population with subgroup analysis finding effect in White women only.
Were all important outcomes considered?	Yes	Extensive neonatal outcomes are covered. Overall maternal adverse effects demonstrate no difference. Slight limitation that most outcomes are for preterm birth overall rather than spontaneous but this was available for the primary outcome, so in general the evidence was thought to be applicable.
Are the benefits worth the harms?	Yes	Benefit is demonstrated with no evidence of harm identified.

Table 32. Cochrane risk of bias assessment of van Os et al 2015[11] RCT

Bias domain	Low, unclear, high risk	Comment

Random sequence generation (selection bias)	Low	Web-based randomisation.
Allocation concealment (selection bias)	Low	Disclosure of codes only after last data collected, 10 weeks after delivery of last participant.
Blinding of participants and personnel (performance bias)	Low	Identical blister pack placebo
Blinding of outcome assessment (detection bias)	Low	Assessors blind
Incomplete outcome assessment (attrition bias)	Low	All analyses by ITT analysis
Selective reporting (reporting bias)	Low	All pre-specified outcomes reported
Other bias	High	The trial was underpowered. The power calculation estimated that 960 women would be needed in each arm, based on composite neonatal outcome expected in 14% of women with CL <15mm, and in 3% with CL 15-30mm. Expectation of 1.7% of women with CL< 15 and 8.3% 15-30mm. The probability of composite outcome in women with CL <30 mm was assumed to be 5.0% in the control group with 2.5% rate in progesterone group. Plan to screen 40,000 with expectation that 10% would have CL <30mm and 50% would participate in the trial.
Other notes: applicable study population following screening but requiring 2 sequential measures. Also using higher cut-off 30mm.		

Table 33. CASP assessment of Saccone et al 2017[12] systematic review

Assessment	Yes, no, unclear	Comment
Are the results of the review valid?		
Did the review address a clearly focused question?	Yes	Asymptomatic women, singleton pregnancy, cervical length ≤25mm, clear intervention and outcome
Did the authors look for the right type of papers?	Yes	RCTs only
Do you think all the important, relevant studies were included?	Unclear	Several databases but published studies only, unclear whether language restrictions, no report of hand-searching or contact with authors.
Did the review's authors do enough to assess quality of the included studies?	Yes	Assessments are clear
If the results of the review have been combined, was it reasonable to do so?	Yes	Individual results are given, heterogeneity is assessed and data analysis is clear. The studies were compatible and potential reasons for any variation in results are discussed.

What are the results?		
Were the overall results clear?	Yes	Results appear clear with both absolute numbers and risk.
Are the results precise?	Yes	Confidence intervals are not excessively wide
Will the results help locally?		
Can the results be applied to the local population?	Unclear	These are asymptomatic women with singleton pregnancy and short cervical length at a gestation relevant to screening. However, it does include women with prior preterm birth and subanalysis according to group was not possible.
Were all important outcomes considered?	Yes	Extensive neonatal and maternal outcomes are covered (minus acceptability/psychological)
Are the benefits worth the harms?	NA	Benefit is not demonstrated, though no evidence of harm identified.
Additional notes: only 3 studies with the findings driven by the Nicolaides trial. Subanalysis by risk group was not possible. Goya had much higher preterm birth rate in the control group (26.8%) than the other 2 trials (5.5%and 11.3%) for unclear reasons.		

Table 34. Cochrane risk of bias assessment of Sacccone et al 2017[14] RCT

Bias domain	Low, unclear, high risk	Comment
Random sequence generation (selection bias)	Low	Central web-based randomisation stratified by cervical length
Allocation concealment (selection bias)	Low	No access to randomisation sequence and clinicians did not have prior access
Blinding of participants and personnel (performance bias)	High	Non-blinded but not feasible given the intervention and uncertain whether this could have introduced risk of bias as outcomes are mostly objective
Blinding of outcome assessment (detection bias)	Low	Assessors blind
Incomplete outcome assessment (attrition bias)	Low	All analyses by ITT analysis
Selective reporting (reporting bias)	Low	All pre-specified outcomes reported
Other bias	Low	Adequately powered for primary outcome, balanced groups. Except there is risk of error for secondary outcomes due to multiple analyses and small samples.
Other notes: applicable study population (excluding prior preterm birth/mid-trimester loss), though high rate of short cervical length with 79% of women with cervical length ≤15mm and 89% of intervention and 83% of control group receiving vaginal progesterone (close to a trial assessing progesterone with/without pessary). Also 22% intervention and 25% controls receiving antibiotics.		

Table 35. Cochrane risk of bias assessment of Dugoff et al 2018[13] RCT

Bias domain	Low, unclear, high risk	Comment
Random sequence generation (selection bias)	Low	Central web-based randomisation stratified by cervical length
Allocation concealment (selection bias)	Unclear	Not further reported.
Blinding of participants and personnel (performance bias)	High	Non-blinded but not feasible given the intervention and uncertain whether this could have introduced risk of bias as outcomes are mostly objective
Blinding of outcome assessment (detection bias)	High	Not blinded and feasible but uncertain whether could have influenced the outcome.
Incomplete outcome assessment (attrition bias)	Low	Analysis by ITT analysis though does excluded n=3 from the control group lost to follow-up.
Selective reporting (reporting bias)	Low	All pre-specified outcomes reported
Other bias	High	Underpowered for primary outcome due to early termination with risk of 2 error, similar for secondary outcomes.
Other notes: applicable study population (excluding prior preterm birth/mid-trimester loss), though 61% Black ethnicity. Median cervical length18mm. 84% of intervention and 91% of control group receiving vaginal progesterone (similar to progesterone with/without pessary). High preterm rate similar to Goya (Saccone review) for unclear reasons. Different pessary from other trials.		

Table 36. Cochrane risk of bias assessment of Cruz-Melguizo et al 2018[15] RCT

Bias domain	Low, unclear, high risk	Comment
Random sequence generation (selection bias)	Low	Computer-generated randomisation sequence managed by a central hospital not involved with recruitment.
Allocation concealment (selection bias)	Low	Investigators received the woman's identification number and assigned treatment by phone after written informed consent was obtained.
Blinding of participants and personnel (performance bias)	High	Non-blinded but not feasible given the intervention and uncertain whether this could have introduced risk of bias as outcomes are mostly objective
Blinding of outcome assessment (detection bias)	High	Not blinded and feasible but uncertain whether could have influenced the outcome.
Incomplete outcome assessment (attrition bias)	Low	For the primary outcome ITT analysis included n=246 of n=254 randomised after excluding n=8 receiving no treatment/not meeting criteria. Results for all secondary outcomes are given by per protocol analysis only including n=243 after excluding a further n=3, n=2 with iatrogenic preterm birth and n=1 with protocol deviation. It is unclear why ITT was not presented for all outcomes but as these are small numbers and none of the results close to statistical significance it is not expected to have influenced outcomes.
Selective reporting (reporting bias)	Low	All pre-specified outcomes reported

Other bias	Low	Adequately powered for the primary outcome where the non-inferiority margin was set at 4% based on preterm <34 week prevalence 12.9% progesterone group (14% this trial). Balanced groups at baseline.
Other notes: mixed population including those with prior preterm birth (though excluding cervical trauma/uterine abnormality) but otherwise applicable. Median cervical length 21mm and preterm birth rate not excessively high. 24% of pessary and 19% of progesterone group receiving antibiotics.		

Table 37. CASP assessment of Berghella al 2017[16] systematic review

Assessment	Yes, no, unclear	Comment
Are the results of the review valid?		
Did the review address a clearly focused question?	Yes	Asymptomatic women, singleton pregnancy, cervical length <25mm, no prior SPTB; clear intervention and outcome
Did the authors look for the right type of papers?	Yes	RCTs only
Do you think all the important, relevant studies were included?	Yes	Several databases, no language restrictions, contact with authors.
Did the review's authors do enough to assess quality of the included studies?	Yes	Assessments are clear
If the results of the review have been combined, was it reasonable to do so?	Yes	Individual study results are given only for the main outcome but heterogeneity is assessed and the proportions used in IPD meta-analysis are clear for each outcome.
What are the results?		
Were the overall results clear?	Yes	Results appear clear with both absolute numbers and risk.
Are the results precise?	No	Evidence is low quality and based on small numbers
Will the results help locally?		
Can the results be applied to the local population?	Yes	This is a relevant majority western population of otherwise low-risk women.
Were all important outcomes considered?	Unclear/Partial	Maternal outcomes are not covered.
Are the benefits worth the harms?	Unclear	No neonatal outcomes, though would benefit from maternal assessment.

Table 38.Cochrane risk of bias assessment of Subtil et al 2018[17] RCT

Bias domain	Low, unclear, high risk	Comment
Random sequence generation (selection bias)	Low	Computer-generated randomisation sequence with randomisation stratified by centre.

Allocation concealment (selection bias)	Low	Participants received a numbered box containing blister packs with participants/personnel having no knowledge of contents
Blinding of participants and personnel (performance bias)	Low	Double blind
Blinding of outcome assessment (detection bias)	Low	Double blind
Incomplete outcome assessment (attrition bias)	Low	ITT analysis including all but 9/2869
Selective reporting (reporting bias)	Low	All pre-specified outcomes reported
Other bias	Unclear	The sample size was based on an expected prevalence of the primary outcome (late miscarriage/very preterm birth) of 2% in the general population and doubled in those with BV. The prevalence was lower than this and it is unclear whether the trial could have been underpowered for this outcome as a result, but as this is a large trial it would have had sufficient number to detect difference at later gestations. Compliance was low at around 80% by self-report and 49% by pill count, but using this most conservative estimate in per protocol analysis still did not alter results.
Other notes: high quality trial with applicable large screen-detected population. Included multiple pregnancies but only 2% and exclusion had no effect on outcomes.		

Appendix 4 – UK NSC reporting checklist for evidence summaries

All items on the UK NSC Reporting Checklist for Evidence Summaries have been addressed in this report. A summary of the checklist, along with the page or pages where each item can be found in this report, is presented in Table 39.

Table 39. UK NSC reporting checklist for evidence summaries

	Section	Item	Starting page numbers
1.	TITLE AND SUMMARIES		
1.1	Title sheet	Identify the review as a UK NSC evidence summary.	Title page
1.2	Plain English summary	Plain English description of the executive summary.	5
1.3	Executive summary	Structured overview of the whole report. To include: the purpose/aim of the review; background; previous recommendations; findings and gaps in the evidence; recommendations on the screening that can or cannot be made on the basis of the review.	7
2.	INTRODUCTION AND APPROACH		
2.1	Background and objectives	Background – Current policy context and rationale for the current review – for example, reference to details of previous reviews, basis for current recommendation, recommendations made, gaps	15

		identified, drivers for new reviews	
		Objectives – What are the questions the current evidence summary intends to answer? – statement of the key questions for the current evidence summary, criteria they address, and number of studies included per question, description of the overall results of the literature search.	20
		Method – briefly outline the rapid review methods used.	22
2.2	Eligibility for inclusion in the review	State all criteria for inclusion and exclusion of studies to the review clearly (PICO, dates, language, study type, publication type, publication status etc.) To be decided *a priori*.	22
2.3	Appraisal for quality/risk of bias tool	Details of tool/checklist used to assess quality, e.g. QUADAS 2, CASP, SIGN, AMSTAR.	25
3.	SEARCH STRATEGY AND STUDY SELECTION (FOR EACH KEY QUESTION)		
3.1	Databases/ sources searched	Give details of all databases searched (including platform/interface and coverage dates) and date of final search.	22
3.2	Search strategy and results	Present the full search strategy for at least one database (usually a version of Medline), including limits and search filters if used.	75

		Provide details of the total number of (results from each database searched), number of duplicates removed, and the final number of unique records to consider for inclusion.	
3.3	Study selection	State the process for selecting studies – inclusion and exclusion criteria, number of studies screened by title/abstract and full text, number of reviewers, any cross checking carried out.	22
4.	STUDY LEVEL REPORTING OF RESULTS (FOR EACH KEY QUESTION)		
4.1	Study level reporting, results and risk of bias assessment	For each study, produce a table that includes the full citation and a summary of the data relevant to the question (for example, study size, PICO, follow-up period, outcomes reported, statistical analyses etc.). Provide a simple summary of key measures, effect estimates and confidence intervals for each study where available. For each study, present the results of any assessment of quality/risk of bias.	Study level reporting: 101 Quality assessment: 137
4.2	Additional analyses	Describe additional analyses (for example, sensitivity, specificity, PPV, etc.) carried out by the reviewer.	109
5.	QUESTION LEVEL SYNTHESIS		
5.1	Description of the evidence	For each question, give numbers of studies screened,	26 and 45

		assessed for eligibility, and included in the review, with summary reasons for exclusion.	
5.2	Combining and presenting the findings	Provide a balanced discussion of the body of evidence which avoids over reliance on one study or set of studies. Consideration of four components should inform the reviewer's judgement on whether the criterion is 'met', 'not met' or 'uncertain': quantity; quality; applicability and consistency.	32 and 50
5.3	Summary of findings	Provide a description of the evidence reviewed and included for each question, with reference to their eligibility for inclusion. Summarise the main findings including the quality/risk of bias issues for each question. Have the criteria addressed been 'met', 'not met' or 'uncertain'?	43 and 70
6.	REVIEW SUMMARY		
6.1	Conclusions and implications for policy	Do findings indicate whether screening should be recommended? Is further work warranted? Are there gaps in the evidence highlighted by the review?	72

6.2	Limitations	Discuss limitations of the available evidence and of the review methodology if relevant.	76

References

1. NICE. Preterm Labour and Birth. NICE guideline [NG25] [Internet]. London: National Institute for Health and Care Excellance; [updated August 2019]. Available from: https://www.nice.org.uk/guidance/ng25/chapter/Recommendations.
2. DH. Safer Maternity Care - The National Maternity Safety Strategy - Progress and Next Steps. London: Department of Health and Social Care, 2017. Available from: https://www.gov.uk/government/publications/safer-maternity-care-progress-and-next-steps.
3. Honest H, Forbes C, Durée K, et al. Screening to prevent spontaneous preterm birth: systematic reviews of accuracy and effectiveness literature with economic modelling. Health Technol Assess. 2009;13(43):1-627.
4. Story L, Simpson NA, David AL, et al. Reducing the impact of preterm birth: Preterm birth commissioning in the United Kingdom. European journal of obstetrics & gynecology and reproductive biology: X. 2019.
5. Dos Santos F, Daru J, Rogozinska E, et al. Accuracy of fetal fibronectin for assessing preterm birth risk in asymptomatic pregnant women: a systematic review and meta-analysis. Acta Obstet Gynecol Scand. 2018;97(6):657-67.
6. Esplin MS, Elovitz MA, Iams JD, et al. Predictive Accuracy of Serial Transvaginal Cervical Lengths and Quantitative Vaginal Fetal Fibronectin Levels for Spontaneous Preterm Birth Among Nulliparous Women. Jama. 2017;317(10):1047-56.
7. Banos N, Murillo-Bravo C, Julia C, et al. Mid-trimester sonographic cervical consistency index to predict spontaneous preterm birth in a low-risk population. Ultrasound Obstet Gynecol. 2018;51(5):629-36.
8. Kuusela P, Jacobsson B, Söderlund M, et al. Transvaginal sonographic evaluation of cervical length in the second trimester of asymptomatic singleton pregnancies, and the risk of preterm delivery. Acta Obstetricia et Gynecologica Scandinavica. 2015;94(6):598-607.
9. van der Ven J, van Os MA, Kazemier BM, et al. The capacity of mid-pregnancy cervical length to predict preterm birth in low-risk women: a national cohort study. Acta Obstet Gynecol Scand. 2015;94(11):1223-34.
10. Romero R, Conde-Agudelo A, Da Fonseca E, et al. Vaginal progesterone for preventing preterm birth and adverse perinatal outcomes in singleton gestations with a short cervix: a meta-analysis of individual patient data. American Journal of Obstetrics and Gynecology. 2018;218(2):161-80.
11. van Os MA, van der Ven AJ, Kleinrouweler CE, et al. Preventing Preterm Birth with Progesterone in Women with a Short Cervical Length from a Low-Risk Population: A Multicenter Double-Blind Placebo-Controlled Randomized Trial. Am J Perinatol. 2015;32(10):993-1000.
12. Saccone G, Ciardulli A, Xodo S, et al. Cervical Pessary for Preventing Preterm Birth in Singleton Pregnancies With Short Cervical Length: A Systematic Review and Meta-analysis. J Ultrasound Med. 2017;36(8):1535-43.
13. Dugoff L, Berghella V, Sehdev H, et al. Prevention of preterm birth with pessary in singletons (PoPPS): randomized controlled trial. Ultrasound Obstet Gynecol. 2018;51(5):573-9.
14. Saccone G, Maruotti GM, Giudicepietro A, et al. Effect of Cervical Pessary on Spontaneous Preterm Birth in Women With Singleton Pregnancies and Short Cervical Length: A Randomized Clinical Trial. Jama. 2017;318(23):2317-24.

15. Cruz-Melguizo S, San-Frutos L, Martínez-Payo C, et al. Cervical pessary compared with vaginal progesterone for preventing early preterm birth a randomized controlled trial. Obstetrics and Gynecology. 2018;132(4):907-15.
16. Berghella V, Ciardulli A, Rust OA, et al. Cerclage for sonographic short cervix in singleton gestations without prior spontaneous preterm birth: systematic review and meta-analysis of randomized controlled trials using individual patient-level data. Ultrasound Obstet Gynecol. 2017;50(5):569-77.
17. Subtil D, Brabant G, Tilloy E, et al. Early clindamycin for bacterial vaginosis in pregnancy (PREMEVA): a multicentre, double-blind, randomised controlled trial. Lancet. 2018;392(10160):2171-9.
18. Stewart LA, Simmonds M, Duley L, et al. Evaluating progestogens for prevention of preterm birth international collaborative (EPPPIC) individual participant data (IPD) meta-analysis: protocol. Systematic reviews. 2017;6(1):235.
19. Brocklehurst P, Gordon A, Heatley E, et al. Antibiotics for treating bacterial vaginosis in pregnancy. Cochrane Database Syst Rev. 2013(1):Cd000262.
20. Tommy's. Premature Birth Statistics [Internet]. London: Tommy's. Available from: https://www.tommys.org/our-organisation/why-we-exist/premature-birth-statistics.
21. ONS. Birth characteristics in England and Wales: 2016 [Internet]. London: Office for National Statistics. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/livebirths/bulletins/birthcharacteristicsinenglandandwales/2016.
22. Matei A, Saccone G, Vogel JP, et al. Primary and secondary prevention of preterm birth: a review of systematic reviews and ongoing randomized controlled trials. Eur J Obstet Gynecol Reprod Biol. 2019;236:224-39.
23. Glover AV, Manuck TA. Screening for spontaneous preterm birth and resultant therapies to reduce neonatal morbidity and mortality: A review. Semin Fetal Neonatal Med. 2018;23(2):126-32.
24. Lim K, Butt K, Crane JM. No. 257-Ultrasonographic Cervical Length Assessment in Predicting Preterm Birth in Singleton Pregnancies. Journal of Obstetrics and Gynaecology Canada. 2018;40(2):e151-e64.
25. Yudin MH, Money DM. No. 211-Screening and Management of Bacterial Vaginosis in Pregnancy. J Obstet Gynaecol Can. 2017;39(8):e184-e91.
26. NICE. Antenatal care for uncomplicated pregnancies. Clinical Guideline [CG62] [Internet]. London: National Institute for Health and Care Excellance; [updated February 2019]. Available from: https://www.nice.org.uk/guidance/cg62/chapter/1-Guidance.
27. Medley N, Poljak B, Mammarella S, et al. Clinical guidelines for prevention and management of preterm birth: a systematic review. Bjog. 2018;125(11):1361-9.
28. Sentilhes L, Senat MV, Ancel PY, et al. Prevention of spontaneous preterm birth: Guidelines for clinical practice from the French College of Gynaecologists and Obstetricians (CNGOF). Eur J Obstet Gynecol Reprod Biol. 2017;210:217-24.
29. Berger R, Abele H, Bahlmann F, et al. Prevention and Therapy of Preterm Birth. Guideline of the DGGG, OEGGG and SGGG (S2k Level, AWMF Registry Number 015/025, February 2019)–Part 2 with Recommendations on the Tertiary Prevention of Preterm Birth and the Management of Preterm Premature Rupture of Membranes. Geburtshilfe und Frauenheilkunde. 2019;79(08):813-33.
30. Norman JE, Marlow N, Messow CM, et al. Vaginal progesterone prophylaxis for preterm birth (the OPPTIMUM study): a multicentre, randomised, double-blind trial. Lancet. 2016;387(10033):2106-16.

31. Nicolaides KH, Syngelaki A, Poon LC, et al. A randomized trial of a cervical pessary to prevent preterm singleton birth. New England Journal of Medicine. 2016;374(11):1044-52.
32. Hui SA, Chor CM, Lau TK, et al. Cerclage pessary for preventing preterm birth in women with a singleton pregnancy and a short cervix at 20 to 24 weeks: A randomized controlled trial. American Journal of Perinatology. 2013;30(4):283-8.
33. Otsuki K, Nakai A, Matsuda Y, et al. Randomized trial of ultrasound-indicated cerclage in singleton women without lower genital tract inflammation. J Obstet Gynaecol Res. 2016;42(2):148-57.
34. Berghella V, Odibo AO, To MS, et al. Cerclage for short cervix on ultrasonography: meta-analysis of trials using individual patient-level data. Obstetrics & Gynecology. 2005;106(1):181-9.
35. Berghella V, Keeler S, To M, et al. Effectiveness of cerclage according to severity of cervical length shortening: a meta-analysis. Ultrasound in Obstetrics and Gynecology: The Official Journal of the International Society of Ultrasound in Obstetrics and Gynecology. 2010;35(4):468-73.
36. Lamont RF, Nhan-Chang C-L, Sobel JD, et al. Treatment of abnormal vaginal flora in early pregnancy with clindamycin for the prevention of spontaneous preterm birth: a systematic review and metaanalysis. American Journal of Obstetrics and Gynecology. 2011;205(3):177-90.