National Institute for Health and Care Excellence

Final

Twin and triplet pregnancy

[K] Progesterone for preventing spontaneous preterm birth in twin and triplet pregnancy

NICE guideline NG137

Evidence review underpinning recommendations 1.5.1 to 1.5.4 and a research recommendation in the NICE guideline

April 2024

Final

These evidence reviews were developed by NICE

Disclaimer

The recommendations in this guideline represent the view of NICE, arrived at after careful consideration of the evidence available. When exercising their judgement, professionals are expected to take this guideline fully into account, alongside the individual needs, preferences and values of their patients or service users. The recommendations in this guideline are not mandatory and the guideline does not override the responsibility of healthcare professionals to make decisions appropriate to the circumstances of the individual patient, in consultation with the patient and/or their carer or guardian.

Local commissioners and/or providers have a responsibility to enable the guideline to be applied when individual health professionals and their patients or service users wish to use it. They should do so in the context of local and national priorities for funding and developing services, and in light of their duties to have due regard to the need to eliminate unlawful discrimination, to advance equality of opportunity and to reduce health inequalities. Nothing in this guideline should be interpreted in a way that would be inconsistent with compliance with those duties.

NICE guidelines cover health and care in England. Decisions on how they apply in other UK countries are made by ministers in the Welsh Government, Scottish Government, and Northern Ireland Executive. All NICE guidance is subject to regular review and may be updated or withdrawn.

Copyright

ISBN: 978-1-4731-5886-3

Contents

Progesterone for preventing spontaneous preterm birth in twin and triplet pregnancy

Review question

What is the clinical and cost-effectiveness of progesterone in preventing spontaneous preterm birth in twin and triplet pregnancies?

Introduction

Spontaneous preterm birth occurs more frequently in twin and triplet pregnancies than in singleton pregnancies. Preterm birth is associated with an increased risk of morbidity and mortality and increased use of healthcare resources, with many preterm babies requiring specialist neonatal care. The severity of complications related to preterm birth and the likelihood of long-term effects is generally proportional to the degree of prematurity, with babies born extremely preterm (at less than 28 weeks' gestation) often requiring long stays in neonatal care. Preventing preterm birth and prolonging a pregnancy to nearer 37 weeks is therefore beneficial in improving the likelihood of a healthy baby and reducing the need for neonatal care.

The aim of this review is to determine if progesterone reduces preterm birth in twin and triplet pregnancies, without causing adverse effects for the mother or baby.

Summary of the protocol

See Table 1 for a summary of the Population, Intervention, Comparison and Outcome (PICO) characteristics of this review.

Table 1:Summary of the protocol (PICO table)

Population	Women at risk of preterm birth in twin and triplet pregnancy Strata: Cervical length • Women with a short cervix (≤25 mm) • Women with a longer cervix (>25 mm) Previous preterm birth • Women with previous preterm birth • Women with no previous preterm birth
Intervention	• Vaginal progesterone • Oral progesterone • Intramuscular 17-hydroxyprogesterone caproate (17-OHPC) Progesterone use in first, second and third trimester (part) will be included as all are relevant.
Comparison	With no intervention (placebo or control or standard care) or with each other
Outcome	**Critical** • Stillbirth or neonatal death* (to report neonatal death outcome separately if reported) • Preterm birth** at 22+0 - 27+6 weeks • Preterm birth** at 28+0 - 31+6 weeks • Preterm birth** at 32+0 - 36+6 weeks • Spontaneous preterm birth <34 weeks of gestation (this will include spontaneous preterm birth <33 weeks) *Stillbirth is a baby that dies after 24 weeks of pregnancy but before they are born, and a neonatal death is death within 28 days after birth. **This includes spontaneous preterm birth and indicated preterm birth (in which a baby is delivered by early induction of labour or caesarean birth due to maternal or fetal illness). **Important** • Composite of serious neonatal complications (for example, severe necrotising enterocolitis stages 2–3, intraventricular haemorrhage grades 3–4, retinopathy of prematurity stage 3 or worse, bronchopulmonary dysplasia, confirmed sepsis, patent ductus arteriosus, and neonatal infection) • Composite of adverse maternal outcomes (for example, gestational hypertension, pre-eclampsia, gestational diabetes, and maternal infection including chorioamnionitis)

17-OHPC: 17-hydroxyprogesterone caproate

For further details see the review protocol in appendix A.

Methods and process

This evidence review was developed using the methods and process described in Developing NICE guidelines: the manual. Methods specific to this review question are described in the review protocol in appendix A.

Declarations of interest were recorded according to NICE's conflicts of interest policy.

Effectiveness evidence

Included studies

Four studies were included for this review, 2 randomised controlled trials (RCTs) (Rehal 2021, Shabaan 2018) and 2 individual participant data (IPD) reviews (Conde-Agudelo 2022, EPPPIC Group 2021).

The included studies are summarised in Table 2.

Population

One RCT included women with a twin pregnancy, and this study included women with and without short cervix and women with and without previous preterm birth (Rehal 2021).

One RCT included women with a twin pregnancy, but this study did not report information on cervical length and previous preterm birth (Shabaan 2018)

One IPD review included women with a singleton pregnancy or a multifetal pregnancy (that is, women with a twin pregnancy in vaginal progesterone trials and women with a twin or triplet pregnancy in intramuscular 17-hydroxyprogesterone caproate [17-OHPC] trials), including women with and without short cervix (≤25 mm) and with and without previous preterm birth (EPPPIC Group 2021). Only evidence from multifetal pregnancies has been reported from this IPD in this review. The study did not report evidence for twin pregnancy and triplet pregnancy separately.

One IPD review included women with a twin pregnancy and short cervix (≤25 mm), but there was no information on previous preterm birth in this paper (Conde-Agudelo 2022).

Interventions and comparisons

All studies compared vaginal progesterone to placebo or control (no intervention) in women with a twin pregnancy (Conde-Agudelo 2022, EPPPIC Group 2021, Rehal 2021, Shabaan 2018). Among these studies, 1 IPD review also compared intramuscular 17-OHPC to control in women with twin or triplet pregnancies (EPPPIC Group 2021).

No evidence was identified for oral progesterone in twin and triplet pregnancies.

Timing of progesterone use was reported in the studies. One RCT investigated the effectiveness of progesterone use from first trimester to third trimester (from 11-14 weeks to 34 weeks) (Rehal 2021), and 1 RCT investigated the effectiveness of progesterone use in third trimester (from 28 weeks to delivery) (Shabaan 2018). The remaining studies included in IPD reviews investigated the effectiveness of progesterone use from second trimester to third trimester (from 16 weeks to 37 weeks) (Conde-Agudelo 2022, EPPPIC Group 2021).

Outcomes

Evidence was available for all outcomes stated in the protocol. Stillbirth and neonatal death was a composite outcome in the protocol. However, where data was available, stillbirth and neonatal death were reported separately due to the differences in the definitions of these outcomes. Two IPD reviews reported stillbirth and neonatal death separately (Conde-Agudelo 2022, EPPPIC Group 2021), but 2 RCTs did not report them separately (Rehal 2021, Shabaan 2018).

Three cut-offs for preterm birth and 1 cut-off for spontaneous preterm birth were stated in the protocol. However, none of the included studies used the cut-offs as in the protocol. Therefore, we reported preterm birth and spontaneous preterm birth outcomes as reported in the papers, which were assessed at relatively similar time points as in the protocol. Two IPD reviews reported preterm birth <28 weeks (Conde-Agudelo 2022, EPPPIC Group 2021), and

1 RCT reported any birth between 24 weeks and <28 weeks (spontaneous or indicated birth) (Rehal 2021). One IPD review reported preterm birth <32 weeks (Conde-Agudelo 2022), 1 RCT reported any birth between 24 weeks and <32 weeks (Rehal 2021) and 1 RCT reported preterm birth at 28-30 weeks (Shabaan 2018). One IPD review reported preterm birth <34 weeks (EPPPIC Group 2021), 1 IPD review reported preterm birth <37 weeks (Conde-Agudelo 2022), 1 RCT reported any birth between 24 and <37 weeks (Rehal 2021), and 1 RCT reported preterm birth at 32-36 weeks (Shabaan 2018). One RCT reported spontaneous birth between 24 and <34 weeks (Rehal 2021), and 1 IPD review reported spontaneous preterm birth <34 weeks (Conde-Agudelo 2022).

Two IPD reviews and 1 RCT reported composite of serious neonatal complications (Conde-Agudelo 2022, EPPPIC Group 2021, Rehal 2021), and 1 IPD review and 1 RCT reported composite of adverse maternal outcomes (EPPPIC Group 2021, Rehal 2021).

Analysis

Meta-analysis was performed where possible (for example, if there were at least 2 studies reporting the same intervention in populations with the same/similar characteristics) and where there was no significant variation between studies or very serious heterogeneity. For those where meta-analysis could not be performed, the results for each individual study have been reported in this review.

The evidence was stratified by cervical length and previous preterm birth where possible.

Evidence from 1 RCT (Rehal 2021) was stratified based on with or without previous preterm birth and by cervical length. Data for women with a long cervix (>30 mm) were used from this study but data for women with short cervix (<30 mm) were included in the IPD review by Conde-Agudelo 2022, which reported data for women with twin pregnancies and short cervix only.

Evidence from 1 RCT (Shabaan 2018) could not be stratified as this study included all participants with short cervix or long cervix and with or without previous preterm birth in their analyses.

One IPD review (EPPPIC Group 2021) reported overall data for all women, and although they reported subgroup analyses for participants with short cervix, long cervix, and with previous preterm birth and without previous preterm birth, these data were not available for those with multifetal pregnancies (EPPPIC Group 2021).

See the literature search strategy in appendix B and study selection flow chart in appendix C.

Excluded studies

Studies not included in this review are listed, and reasons for their exclusion are provided in appendix J.

Summary of included studies

Summaries of the studies that were included in this review are presented in Table 2.

Table 2: Summary of included studies.

Study	Population	Intervention	Comparison	Outcomes	Comments
Conde-Agudelo 2022 IPD review International	N=95 Vaginal progesterone: N=52 Placebo:	Vaginal progesterone (100-600 mg per day)	Placebo	• Stillbirth (fetal death) • Neonatal death • Preterm birth <28 weeks	This is an IPD meta-analysis of women with twin pregnancy and short

	N=43 N=6 RCTs investigating vaginal progesterone for the prevention of preterm birth in women with a twin pregnancy and short cervix (≤25 mm) or women with an unselected twin pregnancy and short cervix Participants characteristics not reported			• Preterm birth <32 weeks • Preterm birth <37 weeks • Spontaneous preterm birth <34 weeks • Composite of serious neonatal complications	cervix. This paper did not report data on previous preterm birth.
EPPPIC Group 2021 IPD review International	N=11,644 (Whole sample size; N=31 RCTs) RCTs that compared progesterone with placebo or standard care or other forms of progesterone **Twin pregnancies (8 RCTs)** Vaginal progesterone: N=2068 Control: NR Age in years, mean (SD): Vaginal progesterone: 31.3 (5.4) Control: NR Cervical length, N: Vaginal progesterone: Shorter cervix (≤25 mm): 72	Vaginal progesterone (90-400 mg per day) IM 17-OHPC (250 mg weekly or 500 mg twice weekly)	Control: Participants received placebo or no intervention	• Stillbirth • Neonatal death • Preterm birth <28 weeks • Preterm birth <34 weeks • Composite of serious neonatal complications • Composite of adverse maternal outcomes	Subgroup analysis included mainly singleton pregnancies with only one study with both singleton and twin pregnancies. This study included >80% singleton pregnancies. Hence, no subgroup analyses according to cervical length and previous preterm birth were available for women with twin or triplet pregnancy

	Longer cervix (>25 mm): 1155 Unknown: 841 Control: NR Previous preterm birth and parity, N: Vaginal progesterone: Previous preterm birth: 61 No previous preterm birth: 611 Nulliparous: 1058 Unknown: 338 Control: NR **Twin and triplet pregnancies (8 RCTs)** IM 17-OHPC: N=2270 Control: NR Age in years, mean (SD): IM 17-OHPC: 31.2 (5.9) Control: NR Cervical length, N: IM 17-OHPC: Shorter cervix (≤25 mm): 227 Longer cervix (>25 mm): 1121 Unknown: 922 Control: NR Previous preterm birth and parity, N: IM 17-OHPC: Previous preterm birth: 202 No previous preterm birth: 872				

	Nulliparous: 1191 Unknown: 5 Control: NR				
Rehal 2021 RCT International	N=1,194 Vaginal progesterone: N=596 Placebo: N:598 Pregnant women aged over 18 years with twin pregnancy and two live fetuses, who were fluent in the local language Age in years, median (IQR): Vaginal progesterone: 34.1 (30.3-37.7) Placebo: 34.0 (30.0-37.6) Cervical length in mm, median (IQR): Vaginal progesterone: 34.4 (31.0-38.0) Placebo: 34.6 (31.5-38.0) Nulliparous, N: Vaginal progesterone: 317 Placebo: 326 Parous with preterm birth <37 weeks, N: Vaginal progesterone: 23 Placebo: 33 Parous without preterm birth <37 weeks, N:	Vaginal progesterone (600 mg per day): Participants received 300 mg vaginal progesterone capsule twice a day from 11-14 weeks to 34 weeks of gestation	Placebo: Participants received vaginal placebo capsules, which were identical to vaginal progesterone, twice a day from 11-14 weeks to 34 weeks of gestation	• Stillbirth or neonatal death • Any birth between 24 weeks and <28 weeks • Any birth between 24 weeks and <32 weeks • Any birth between 24 and <37 weeks • Spontaneous birth between 24 and <34 weeks • Composite of serious neonatal complications • Composite of adverse maternal outcomes	Subgroup analyses according to cervical length and previous preterm birth were available. However, only study level data on previous preterm birth and long cervix were used because IPD on short cervix was reported by an included IPD review (Conde-Agudelo 2022)

	Vaginal progesterone: 242 Placebo: 228				
Shabaan 2018 RCT Egypt	N=140 Vaginal progesterone: N=70 Control: N=70 Pregnant women with uncomplicated twin pregnancy and 28 weeks of gestation and without major fetal anomalies Age in years, mean (SD): Vaginal progesterone: 29.1 (4.1) Control: 28.6 (3.5) Cervical length: NR Previous preterm birth: NR	Vaginal progesterone (400 mg per day): Participants received progesterone vaginal pessaries 400 mg per day at bedtime from 28 weeks of gestation to delivery	Control: Participants received the normal tonics during pregnancy	• Stillbirth or neonatal death • Preterm birth at 28-30 weeks • Preterm birth at 32-36 weeks	No data on cervical length and previous preterm birth reported

IPD: Individual participant data; IM: intramuscular; IQR: interquartile range; NR: not reported; RCT: randomised controlled trial; SD: standard deviation; 17-OHPC: 17-hydroxyprogesterone caproate

See the full evidence tables in appendix D and the forest plots in appendix E.

Summary of the evidence

This section is a narrative summary of the findings of the review, as presented in the GRADE tables in appendix F. For details of the committee's confidence in the evidence and how this affected recommendations, see 'The committee's discussion and interpretation of the evidence' section.

The evidence was assessed as being very low to high quality. Downgrading the evidence was due to risk of bias arising from missing outcome data (loss to follow-up ~ 16%) and imprecision (95% confidence intervals crossing decision making thresholds or low event rates). See the GRADE tables in appendix F for the certainty of the evidence for each individual outcome.

For stillbirth or neonatal death, statistical significance was used for assessing clinical importance. Default minimally important differences (MIDs) (0.8 and 1.25) were used for all other dichotomous outcomes to assess clinical importance.

MIDs, the line of no effect, and 95% confidence interval (CI) were used to assess whether there were important differences in outcomes between groups. Outcomes were considered to have an important benefit, an important harm, no evidence of important difference, or no important difference using the following approach:

- Where the point estimate is greater than the upper MID and the 95% CI do not cross line of no effect, an intervention was described as having an important harm
- Where the point estimate is greater than the upper MID or lower than the lower MID, and the 95% CI cross the line of no effect, the result was described as no evidence of an important difference
- Where the point estimate is between two MIDs, the result was described as no important difference
- Where the point estimate is lower than the lower MID and the 95% CI do not cross line of no effect, an intervention was described as having an important benefit.

Imprecision was assessed using default MIDs (0.8 and 1.25) for all dichotomous outcomes (except stillbirth/neonatal death). The evidence was downgraded for imprecision by one level when the confidence interval around the point estimate crossed one of the thresholds for minimally important difference (0.8 and 1.25 for dichotomous outcomes) and downgraded by two levels when the confidence interval around the point estimate crossed both thresholds. For stillbirth or neonatal death, statistical significance was used for assessing clinical importance, hence the following cut-offs for event rates were used to assess imprecision: (1) >300 events – no imprecision, (2) 150-300 events – serious imprecision, and (3) <150 events – very serious imprecision. For stillbirth or neonatal death, the evidence was downgraded by one level when event rate was 150-300 and downgraded by two levels when event rate was less than 150.

Vaginal progesterone (400 or 600 mg per day) versus placebo/control in all women with twin pregnancies (participants with short cervix/long cervix and with/without previous preterm birth) (study level analysis) (2 studies)

Meta-analysis was conducted for stillbirth or neonatal death; however, the data were not pooled for the remaining outcomes due to variations between studies. In women with a twin pregnancy, evidence showed that there were no important differences between vaginal progesterone and placebo/control for the outcomes of stillbirth or neonatal death (low quality), any birth (spontaneous or indicated birth) between 24 weeks and <28 weeks (low quality), any birth between 24 weeks and <32 weeks (low quality), any birth between 24 and <37 weeks (high quality), preterm birth at 28-30 weeks (very low quality), the 2 composites of serious neonatal complications (moderate and high quality) and the 2 composites of adverse maternal outcomes (low and moderate quality). There was no evidence of important differences between vaginal progesterone and placebo/control for the outcomes of preterm birth at 32-36 weeks (very low quality) and spontaneous preterm birth between 24 and <34 weeks (moderate quality).

Vaginal progesterone (600 mg per day) versus placebo in twin pregnancies (participants with long cervix, ≥30 mm) (study level analysis) (1 study)

In women with a twin pregnancy with long cervix (≥30 mm), evidence from one study showed that vaginal progesterone (600 mg per day) had important harm in terms of spontaneous birth between 24 and <34 weeks when compared with placebo. The quality of evidence was moderate.

Vaginal progesterone (600 mg per day) versus placebo in twin pregnancies (participants with short cervix/long cervix and with previous preterm birth) (study level analysis) (1 study)

In women with a twin pregnancy with previous preterm birth, evidence showed that there was no important difference between vaginal progesterone (600 mg per day) and placebo for spontaneous birth between 24 and <34 weeks. The quality of evidence was low.

Vaginal progesterone (600 mg per day) versus placebo in twin pregnancies (participants with short cervix/long cervix and with no previous preterm birth) (study level analysis) (1 study)

In women with a twin pregnancy with no previous preterm birth, evidence from one study showed that there was no evidence of important difference between vaginal progesterone (600 mg per day) and placebo for spontaneous birth between 24 and <34 weeks. The quality of evidence was moderate.

Vaginal progesterone (90-400 mg per day) versus control in all women with twin pregnancies (participants with short cervix/long cervix and with/without previous preterm birth) (1 IPD)

In women with a twin pregnancy, evidence showed that there were no important differences between vaginal progesterone (90-400 mg per day) and control for the outcomes of stillbirth (moderate quality), neonatal death (moderate quality), preterm birth <28 weeks (low quality), preterm birth <34 weeks (high quality), composite of serious neonatal complications (moderate quality), and composite of adverse maternal outcomes (moderate quality).

Vaginal progesterone (100-600 mg per day) versus placebo in twin pregnancies (participants with short cervix ≤25 mm) (1 IPD)

In women with a twin pregnancy with short cervix (≤25 mm), moderate quality evidence showed that vaginal progesterone (100-600 mg per day) had important benefits in terms of reducing preterm birth <28 weeks, preterm birth <32 weeks, spontaneous preterm birth <34 weeks, and composite of serious neonatal complications when compared with placebo. However, there were no evidence of important differences for stillbirth (low quality) and neonatal death (low quality), and no important difference for preterm birth <37 weeks (moderate quality) between the two groups.

Intramuscular 17-hydroxyprogesterone caproate (17-OHPC; 250 mg weekly or 500 mg twice weekly) versus control in all women with twin and triplet pregnancies (participants with short cervix/long cervix and with/without previous preterm birth) (1 IPD)

In women with twin and triplet pregnancies, the evidence showed that there were no important differences between intramuscular 17-OHPC (250 mg weekly or 500 mg twice weekly) and control for the outcomes of stillbirth (moderate quality), neonatal death (moderate quality), preterm birth <28 weeks (low quality), preterm birth <34 weeks (high quality), composite of serious neonatal complications (low quality), and composite of adverse maternal outcomes (moderate quality).

See appendix F for full GRADE tables.

Economic evidence

Included studies

A systematic review of the economic literature was conducted but no economic studies were identified which were applicable to this review question.

Excluded studies

Economic studies not included in this review are listed and reasons for their exclusion are provided in appendix J.

Summary of included economic evidence

See Table 3 for the economic evidence profile of the economic model developed for this guideline.

Table 3: Economic evidence profile of a systematic review of economic evaluations of screening to predict the risk of preterm birth in twin pregnancies and treatment to delay or prevent spontaneous preterm birth in those pregnancies identified as being at higher risk of preterm birth

Study	Limitations	Applicability	Other comments	Incremental			Uncertainty
				Costs	Effect	Cost effectiveness	
NICE guideline model 2023	Minor limitations[1]	Directly applicable	Cost-utility analysis	-£518	0.05 QALYs	Screening and treatment dominates Incremental NMB = £1,521	Probabilistic sensitivity analysis suggested there was a 99% probability that screening to predict the risk of preterm birth in twin pregnancies and treatment to delay or prevent spontaneous preterm birth in those pregnancies identified as being at higher risk of preterm birth was cost-effective Mean incremental NMB = £1,499 (95% CrInt: £1,483 to £1,514)

CrInt = Credible Intervals; NMB = Net monetary benefit; QALY = Quality adjusted life-year
[1] Health state utilities were obtained from published literature, but they were not derived using NICE's preferred method

Economic model

An original cost utility analysis was developed to reflect the new clinical evidence identified in this review. The model compared the cost effectiveness of the following 2 strategies:

i. screening to predict the risk of preterm birth, undertaken by measurement of cervical length using transvaginal ultrasound, and a daily dose of vaginal progesterone to delay or prevent spontaneous preterm birth in those pregnancies identified as being at higher risk of preterm birth by screening (intervention)
ii. no screening for short cervix and no treatment (comparator)

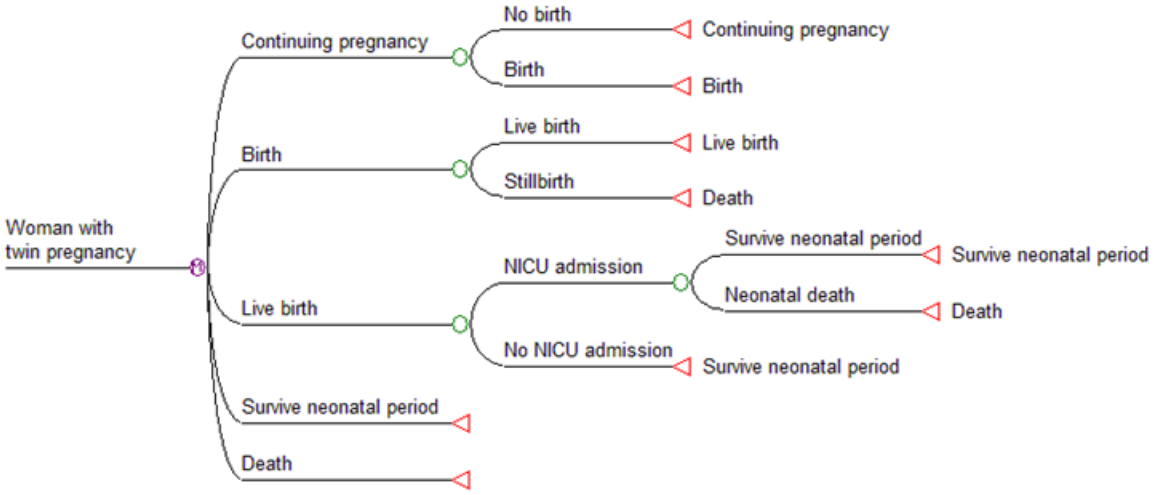
The model is summarised below with full details available in appendix I.

If a pregnancy was identified as being at higher risk of spontaneous preterm birth by the screening strategy, the woman would be treated with daily vaginal progesterone until a gestational age of 34 weeks, or birth if earlier.

A Markov approach was used to model the pregnancy from a gestational age of 24 weeks to a maximum of 37 weeks. Pregnant women with twins enter the model in the state of 'continuing pregnancy' but for each week of gestational age they can transition to the state of 'birth'. This model structure is illustrated in Figure 1 below. This Markov process serves as the 'birth engine' in the model with the transition probabilities dependant on gestational age, the distribution of cervical length across the model population, the probability of preterm birth at each gestational age by cervical length, and the effectiveness of treatment to prevent preterm birth in the women identified for treatment by screening. Figure 1 also highlights the health state transitions from 'birth' which are used to quantify the probability of various

adverse neonatal outcomes, with these probabilities being determined by gestational age at birth.

Figure 1: Schematic to illustrate Markov approach across pregnancy and the neonatal period



In order to estimate the proportion of pregnancies that would be identified as being at higher risk of spontaneous preterm birth, the model factored in a distribution of cervical length at the time of screening. In the base case analysis this was estimated from personal communication undertaken for the previous NICE guideline (Liem, 2018). Data from Kindinger (2016) was then used to estimate the baseline risk of spontaneous preterm birth by gestational age for twin pregnancies according to their cervical length at the time of screening. Data from an individual patient data meta-analysis (Conde-Agudelo, 2022) was then used to modify these baseline risk for pregnancies identified by screening as being at higher risk of preterm birth and treated with vaginal progesterone, using the relative risks reported for <28 weeks, <32 weeks and <36 weeks.

To evaluate the impact of intervention on health-related quality of life and “downstream” costs related to neonatal mortality and morbidity, the model included the following clinical outcomes for babies related to preterm birth:

• Neonatal death

• Post neonatal death

• Neonatal unit admission

• Cerebral palsy

• Intraventricular haemorrhage

• Respiratory distress syndrome

For each of these outcomes the analysis modelled a relationship between the risk and gestational age at birth. Depending on the outcome, costs and quality adjusted life years (QALYs) were assigned to these outcomes.

The results of the analysis suggested that it was cost-effective to screen for the risk of spontaneous preterm birth using a cervical length threshold of 25 mm and to treat those pregnancies identified at being at higher risk of preterm birth due to the presence of a short

cervix. Probabilistic sensitivity analysis showed that intervention had an incremental net monetary benefit (iNMB) of £1,489 when compared to the no screening strategy, with a 99.1% probability of being the most cost-effective strategy. It also indicated that "downstream" costs averted as a result of fewer preterm births and its associated morbidity would more than offset the costs of intervention.

The costs of a daily dose of vaginal progesterone accounted for an insignificant part of the overall costs of intervention. This is because vaginal progesterone is inexpensive and because the number of pregnancies identified as being at higher risk of spontaneous preterm is a small proportion of all twin pregnancies.

Sensitivity analysis indicated that the cost-effectiveness of screening for spontaneous preterm birth using a cervical length threshold of 25 mm and treatment of those pregnancies identified at being at higher risk of preterm birth was not particularly sensitive to changes in model input parameters. Therefore, the committee considered that a recommendation to offer daily vaginal progesterone to women whose twin pregnancy had been identified as being at higher risk of preterm birth would be cost-effective to the NHS.

The committee's discussion and interpretation of the evidence

The outcomes that matter most

The aim of this review was to determine the effectiveness of progesterone in preventing preterm birth in women with twin and triplet pregnancies. Preterm birth is associated with both stillbirth and neonatal death, and the risk is greater at earlier gestations. Therefore, stillbirth or neonatal death, and preterm birth at 22+0 to 27+6 weeks, preterm birth at 28+0 to 31+6 weeks, preterm birth at 32+0 to 36+6 weeks, and spontaneous preterm birth <34 weeks of gestation (including preterm birth <33 weeks) were prioritised as critical outcomes.

As preterm birth or spontaneous preterm birth can be associated with a range of health problems or adverse events for both the mother and baby, a composite of serious neonatal complications (for example, severe necrotising enterocolitis stages 2–3, intraventricular haemorrhage grades 3–4, retinopathy of prematurity stage 3 or worse, bronchopulmonary dysplasia, confirmed sepsis, patent ductus arteriosus, and neonatal infection) and a composite of adverse maternal outcomes (for example, gestational hypertension, pre-eclampsia, gestational diabetes, and maternal infection including chorioamnionitis) were considered as important outcomes.

Evidence was identified for all critical and important outcomes.

The quality of the evidence

The quality of the evidence was assessed using GRADE methodology. The quality of the evidence ranged from very low to high quality, and most evidence was moderate quality. The main reasons for downgrading were risk of bias and imprecision. Risk of bias was most commonly due to missing outcome data. The committee took into account the quality of the evidence, including the uncertainty in their interpretation of the evidence.

Benefits and harms

The committee discussed the evidence on vaginal progesterone and intramuscular 17-hydroxyprogesterone caproate (17-OHPC) in women or pregnant people at risk of preterm birth in twin and triplet pregnancies.

The committee discussed that there was no evidence of benefit (from both study level data and IPD meta-analysis data) for vaginal progesterone for women or pregnant people with twin pregnancies for the outcomes of stillbirth/neonatal death, preterm birth (<28 weeks, <32

weeks, <34 weeks, <37 weeks), spontaneous preterm birth (<34 weeks), and composite neonatal and maternal complications. This group included women or pregnant people with short cervix/long cervix and with/without previous preterm birth. Hence due to lack of evidence of benefit for vaginal progesterone in all women or pregnant people with twin or triplet pregnancies, the committee agreed not to make a recommendation for this group.

The evidence also suggested that there was no benefit (from IPD meta-analysis data) for intramuscular 17-hydroxyprogesterone caproate (17-OHPC) for a mixed group of women or pregnant people with twin or triplet pregnancies and short/long cervix and with/without previous preterm birth, for the outcomes of stillbirth/neonatal death, preterm birth <28 weeks and <34 weeks, composite of serious neonatal complications and composite of adverse maternal outcomes. The committee agreed that the evidence base for 17-OHPC was limited and showed no benefit and therefore they did not make any changes to the existing recommendations on 17-OHPC which advise against its use.

The committee next discussed the evidence for vaginal progesterone in groups of women or pregnant people with or without a previous preterm birth and in women or pregnant people with different cervical lengths.

The committee discussed that there was no evidence of benefit for the outcome of spontaneous birth before 34 weeks for vaginal progesterone in women or pregnant people with twin pregnancies who had a history of previous preterm birth. For women or pregnant people without a history of previous preterm birth there was an increase in spontaneous birth before 34 weeks, but it was not statistically significant. The committee discussed the potential reasons for this increased spontaneous preterm birth such as high dose of progesterone used (600 mg per day), the early onset of therapy (11-14 weeks of gestation) or the duration of therapy (from 11-14 weeks till 34 weeks of pregnancy) in this study. Overall, given the lack of clear evidence in women or pregnant people with and without previous preterm birth, the committee did not make a recommendation for this group. However, the committee agreed that in singleton pregnancies, previous preterm birth was a known risk factor and they agreed more research was required to determine whether it affected risk in the same way in twin or triplet pregnancies. They therefore made a research recommendation for women or pregnant people with twin and triplet pregnancies with a history of preterm birth to help inform future guidelines (see appendix K for full details of the research recommendation). The committee discussed that due to the low numbers of twin and triplet pregnancies compared to singleton pregnancies this research may not be possible using a randomised controlled trial but may need to consider cohort methods or the use of registry or NHS data.

The committee discussed that in women or pregnant people with twin pregnancies and with a short cervix (≤25 mm), the evidence (from IPD meta-analysis) showed that vaginal progesterone reduced preterm birth (<28 weeks and <32 weeks), spontaneous preterm birth (<34 weeks), and reduced the composite of serious neonatal complications. There was no evidence of important difference for stillbirth and neonatal death and no important difference for preterm birth <37 weeks between the two groups. The committee considered the benefit of vaginal progesterone in reducing preterm birth with no evidence of negative effects when compared to placebo group to justify a recommendation. The timing of initiation of vaginal progesterone in most studies was in second trimester (after 18 weeks of gestation) and continued until 34 weeks of gestation. The committee acknowledged that the evidence was based on a small sample size but noted that the population of interest – women or pregnant people with a twin or triplet pregnancy AND who have a short cervix – is a small cohort and so a difficult group on whom to obtain large sets of data. However, the recommendations were based on the overall evidence from an individual patient data (IPD) meta-analysis. IPD meta-analysis is considered to be gold standard for meta-analysis, and this offers the most robust approach to answer the research question, and therefore the committee agreed the evidence was robust enough to make recommendations. The committee recommended the use of vaginal progesterone to reduce the risk of preterm birth in women or pregnant people

with a twin or triplet pregnancy and a short cervix, and that treatment be continued until 34 weeks (or birth if sooner).

In women or pregnant people with twin pregnancy and with a long cervix (≥30 mm), single study evidence for 1 outcome showed that vaginal progesterone was associated with an increase in spontaneous birth before 34 weeks. The finding was seriously imprecise (possibly because of the small sample size) and of moderate quality. Therefore, the committee agreed that, based on this evidence for 1 outcome, there was insufficient evidence to make a recommendation specifically for women or pregnant people with long cervix but agreed that this evidence of harm in women or pregnant people with a long cervix reinforced their decision to recommend the use of vaginal progesterone only to women or pregnant people with a short cervix and not all those with a twin or triplet pregnancy.

There was no evidence available for vaginal progesterone in women or pregnant people with triplet pregnancies, however the committee agreed that it was reasonable to extrapolate on a physiological basis that benefits in twin pregnancies would also be seen in triplet pregnancies and so included both twin and triplet pregnancies in their recommendation. The committee discussed whether it would be feasible to make a research recommendation for the use of progesterone in triplet pregnancies but agreed that the rarity of triplet pregnancies meant that it would be very difficult to collect enough data.

The committee discussed that as the evidence had shown benefit only in women or pregnant people with a short cervix, a transvaginal ultrasound cervical length scan (the gold standard method of assessing cervical length) would be needed to identify this population. The committee discussed that in practice a cervical length scan can be performed between 16 and 20 weeks of gestation, although it is not routinely carried out in all women or pregnant people with a twin and triplet pregnancy. Most of the included studies initiated treatment with progesterone after 18 weeks of gestation and cervical length was measured pre-randomisation so the timing of cervical length scan in the studies agreed with the committee's view of when it would be measured in practice. Therefore, the committee agreed, based on their knowledge and experience, that a single cervical length scan should be offered between 16 and 20 weeks of gestation to identify women or pregnant people with short cervix (≤25 mm). The committee agreed this range would give units the flexibility to coordinate scan with antenatal appointments and other planned scans. However, the committee noted that if women or pregnant people were not proactively scanned between 16 and 20 weeks but were found incidentally after this time to have a short cervix of 25 mm or less (for example if they presented in threatened preterm labour, or were late booking), then they should not be denied vaginal progesterone. The committee noted that in the evidence the studies had initiated vaginal progesterone at a variety of different time periods, but always by 24 weeks. The committee therefore agreed there was evidence to recommend vaginal progesterone be initiated at any time from 16 to 24 weeks. The committee recommended that reason for carrying out cervical length measurement and the reason for offering progesterone should be discussed with women and pregnant people.

The IPD meta-analysis did not report results from individual studies, hence it was not possible to assess if the beneficial effects of vaginal progesterone were impacted by the different doses of vaginal progesterone used in the studies. The committee noted that the majority of the studies in the IPD used a dose of vaginal progesterone of 200 mg, and that this was the dose approved for use in singleton pregnancies, although one study had used 100 mg and one had used 600 mg. The committee therefore recommended a dose of 200 mg vaginal progesterone daily and were aware that this was an off-label use for twin and triplet pregnancies.

Cost effectiveness and resource use

An original model was developed for the guideline which jointly assessed the cost-effectiveness of both screening to predict the risk of spontaneous preterm birth, undertaken

by measurement of cervical length using transvaginal ultrasound, and intervention with vaginal progesterone to delay or prevent preterm birth. The committee considered this analysis when making recommendations on screening to predict the risk of preterm birth and the use of vaginal interventions to prevent or delay spontaneous preterm birth.

The results of the economic analysis suggested that screening to predict the risk of spontaneous preterm birth using a cervical length threshold of 25 mm, measured by transvaginal ultrasound, and a daily dose of vaginal progesterone for women or pregnant people whose pregnancies were identified as being at higher risk of preterm birth would represent a cost-effective use of NHS resources. Furthermore, the analyses also suggested that screening for a short cervix and the treatment of women or pregnant people identified would dominate, being cheaper as well as more effective, because “downstream” savings from reduced and delayed preterm birth would more than offset the costs of screening and treatment.

The committee noted that most of the economic analyses suggested that their recommendation to screen for short cervix in twin pregnancies and treat those identified as at higher risk of preterm birth had the potential to be cost saving overall. Therefore, whilst noting that the cost of screening twin pregnancies in England and Wales might exceed £1 million, they did not anticipate a significant resource impact to the NHS overall.

Other factors the committee took into account

The committee were aware of an ongoing trial called PROSPECT (A Randomized Trial of Pessary and Progesterone for Preterm Prevention in Twin Gestation with a Short Cervix) assessing the effectiveness of an arabin pessary or vaginal progesterone compared to placebo to prevent preterm birth in women or pregnant people with twin pregnancies and with a short cervix. This trial has been in progress for 9 years and is due to be completed by February 2025, and the committee followed up with the main study contact to determine if this timescale will be achieved but were unable to obtain confirmation of this. The committee therefore agreed it was reasonable to make recommendations based on the current evidence base that they had reviewed The committee noted that the PROSPECT study included serial measurements of cervical length in women or pregnant people whose first measurement did not show a short cervix, used a cervical length cut-off of 30 mm, included the arabin pessary as an additional intervention, and that the results of this study may therefore provide evidence for further changes in practice which may require their recommendations to be updated in the future.

Based on stakeholder feedback the committee amended a recommendation in the section of the guideline on screening for preterm birth to provide information on the impact of a short cervix on the chance of a preterm birth.

Recommendations supported by this evidence review

This evidence review supports recommendations 1.5.1 to 1.5.5 in the NICE guideline.

References – included studies

Effectiveness

Conde-Agudelo 2022

Conde-Agudelo, A., Rehal, A., Da Fonseca, E., Brizot, M.L., Rode, L., Serra, V., Cetingoz, E., Syngelaki, A., Tabor, A., Perales, A., Hassan, S.S., Nicolaides, K.H., Vaginal progesterone for the prevention of preterm birth and adverse perinatal outcomes in twin gestations with a short cervix: an updated individual patient data meta-analysis, Ultrasound Obstet Gynecol, 59(2), 263–266, 2022

EPPPIC Group 2021

EPPPIC Group., Evaluating progestogens for preventing preterm birth international collaborative (EPPPIC): meta-analysis of individual participant data from randomised controlled trials, Lancet (London, England), 397(10280), 1183-1194, 2021

Rehal 2021

Rehal, A., Benkő, Z., De Paco Matallana, C., Syngelaki, A., Janga, D., Cicero, S., Akolekar, R., Singh, M., Chaveeva, P., Burgos, J., Molina, F.S., Savvidou, M., De La Calle, M., Persico, N., Rojas, M.S.Q., Sau, A., Greco, E., O'Gorman, N., Plasencia, W., Pereira, S., Jani, J.C., Valino, N., del Mar Gil, M., Maclagan, K., Wright, A., Wright, D., Nicolaides, K.H., Early vaginal progesterone versus placebo in twin pregnancies for the prevention of spontaneous preterm birth: a randomized, double-blind trial, American journal of obstetrics and gynecology, 224(1), 86.e1-86.e19, 2021

Shabaan 2018

Shabaan, O.M., Hassanin, I.M., Makhlouf, A.M., Salem, M.N., Hussein, M., Mohamed, M., Abbas, A.M., Vaginal progesterone for prevention of preterm delivery in women with twin pregnancy: a randomized controlled trial, Facts, Views & Vision in ObGyn, 10(2), 93-98, 2018

Appendices

Appendix A Review protocols

Review protocol for review question: What is the clinical and cost-effectiveness of progesterone in preventing spontaneous preterm birth in twin and triplet pregnancy?

Table 4: Review protocol

Field	Content
PROSPERO registration number	CRD42023465092
Review title	Use of progesterone for preventing spontaneous preterm birth in twins and triplets
Review question	What is the clinical and cost-effectiveness of progesterone in preventing spontaneous preterm birth in twin and triplet pregnancy?
Objective	To assess effectiveness of progesterone for preventing preterm birth in twin and triplet pregnancies.
Searches	The following databases will be searched: Clinical searches: • Cochrane Central Register of Controlled Trials (CENTRAL) • Cochrane Database of Systematic Reviews (CDSR) • Embase • MEDLINE All • Epistemonikos Economic searches: • MEDLINE ALL • Embase • International Network of Agencies for Health Technology Assessment (INAHTA) • HTA

Field	Content
	An economic evaluation filter will be applied. Searches will be restricted by: • English language • Human studies • Systematic Reviews • RCTs • Date limit: 2018 (last date searched) The full search strategies will be published in the final review.
Condition or domain being studied	Twin and triplet pregnancies
Population	Inclusion: Women at risk of preterm birth in twin and triplet pregnancy. Strata: Cervical length • Women with a shorter cervix (≤25 mm) • Women with a longer cervix (>25 mm) Previous preterm birth • Women with previous preterm birth • Women with no previous preterm birth Exclusion: • Singleton pregnancy Other exclusions: • Iatrogenic causes of preterm birth

Field	Content
	For example, interventions in response to any complications of multiple pregnancies such as twin to twin transfusion, small baby, pre-eclampsia and its sequelae. (Note: Population we are interested in this review are those at risk of spontaneous vaginal birth) • Women with a quadruplet or higher-order pregnancy as per scope • Studies that do not report results specifically for twin and/or triplet pregnancies • Women with known structural and chromosomal anomalies • Studies in which interventions are given to women in labour or women requiring imminent birth • Studies examining preterm birth in entire populations of women with a complication (for example, gestational diabetes or hypertension) Setting: any setting
Intervention	• Vaginal progesterone • Oral progesterone • Intramuscular 17-hydroxyprogesterone caproate (17-OHPC) We will include progesterone use in first, second and third trimester(part). All of them are relevant.
Comparator	With no intervention (placebo or control or standard care) or with each other
Types of study to be included	Include published full-text papers: • Systematic reviews of RCTs • RCTs Study level data from studies included in the IPD will not be included in our review. Data from studies not included in the IPD will be included in this review. If there is more than one IPD in the same population, we will include the most recent IPD in our review.

Field	Content
Other exclusion criteria	Conference abstracts will not be included
Context	During the 2019 update, the committee did not make any recommendations on vaginal progesterone for preventing preterm birth in twin pregnancies because they were awaiting evidence in this area, particularly about the use of progesterone in subgroups of women with a short cervix. As such section 1.5 Preventing preterm birth stated that NICE will carry out an exceptional update based on the new evidence when it becomes available.
Primary outcomes (critical outcomes)	• Stillbirth or neonatal death* (to report neonatal death outcome separately if reported) • Preterm birth** at 22+0 - 27+6 weeks • Preterm birth** at 28+0 - 31+6 weeks • Preterm birth** at 32+0 - 36+6 weeks • Spontaneous preterm birth <34 weeks of gestation (this will include spontaneous preterm birth <33 weeks) *Stillbirth is a baby that dies after 24 weeks of pregnancy but before they are born, and a neonatal death is death within 28 days after birth. **This includes spontaneous preterm birth and indicated preterm birth (in which a baby is delivered by early induction of labour or caesarean birth due to maternal or fetal illness).
Secondary outcomes (important outcomes)	• Composite of serious neonatal complications (for example, severe necrotising enterocolitis stages 2–3, intraventricular haemorrhage grades 3–4, retinopathy of prematurity stage 3 or worse, bronchopulmonary dysplasia, confirmed sepsis, patent ductus arteriosus, and neonatal infection) • Composite of adverse maternal outcomes (for example, gestational hypertension, pre-eclampsia, gestational diabetes, and maternal infection including chorioamnionitis)
Data extraction (selection and coding)	All references identified by the searches and from other sources will be uploaded into EPPI and de-duplicated. Titles and abstracts of the retrieved citations will be screened to identify studies that potentially meet the inclusion criteria outlined in the review protocol. Dual sifting will be performed on at least 10% of records; 90% agreement is required. Disagreements will be resolved via discussion between the two reviewers, and consultation with senior staff if necessary. Full versions of the selected studies will be obtained for assessment. Studies that fail to meet the inclusion criteria once the full version has been checked will be excluded at this

Field	Content
	stage. Each study excluded after checking the full version will be listed, along with the reason for its exclusion. A standardised form will be used to extract data from studies. The following data will be extracted: study details (reference, country where study was carried out, type and dates), participant characteristics, inclusion and exclusion criteria, details of the interventions if relevant, setting and follow-up, relevant outcome data and source of funding. One reviewer will extract relevant data into a standardised form, and this will be quality assessed by a senior reviewer.
Risk of bias (quality) assessment	Quality assessment of individual studies will be performed using the following checklists: • ROBIS tool for systematic reviews • Cochrane RoB tool v.2 for RCTs and quasi-RCTs • Wang et al checklist for assessing the methodological quality of IPD meta-analysis https://www.bmj.com/content/bmj/373/bmj.n736.full.pdf
Strategy for data synthesis	Intervention review: Depending on the availability of the evidence, the findings will be summarised narratively or quantitatively. Where possible, meta-analyses will be conducted using Cochrane Review Manager software. A fixed effect meta-analysis will be conducted, and data will be presented as risk ratios or odds ratios for dichotomous outcomes, and mean differences or standardised mean differences for continuous outcomes. Heterogeneity in the effect estimates of the individual studies will be assessed using the I2 statistic. Alongside visual inspection of the point estimates and confidence intervals, I2 values of greater than 50% and 80% will be considered as significant and very significant heterogeneity, respectively. Heterogeneity will be explored as appropriate using sensitivity analyses and pre-specified subgroup analyses. If heterogeneity cannot be explained through subgroup analysis, then a random effects model will be used for meta-analysis, or the data will not be pooled. The confidence in the findings across all available evidence will be evaluated for each outcome using an adaptation of the 'Grading of Recommendations Assessment, Development and Evaluation (GRADE) toolbox' developed by the international GRADE working group: http://www.gradeworkinggroup.org/ Importance and imprecision of findings will be assessed against minimally important differences (MIDs). The following MIDs will be used: 0.8 and 1.25 for all relative

Field	Content			
	dichotomous outcomes, for continuous outcomes any published validated MIDs, if none are available then +/- 0.5x control group SD. For neonatal mortality, statistical significance will be used for assessing clinical importance. Event rates will be used for assessing imprecision. The below cut-offs will be used for event rates: • >300 events- no imprecision • 150-300 events- serious imprecision • <150 events- very serious imprecision			
Analysis of sub-groups	No sub-groups identified			
Type and method of review	☒	Intervention		
	☐	Diagnostic		
	☐	Prognostic		
	☐	Qualitative		
	☐	Epidemiologic		
	☐	Service Delivery		
	☐	Other (please specify)		
Language	English			
Country	England			
Anticipated or actual start date	September 2023			
Anticipated completion date	January 2024			
Stage of review at time of this submission	Review stage	Started	Completed	
	Preliminary searches	☑	☑	
	Piloting of the study selection process	☑	☑	
	Formal screening of search results against eligibility criteria	☑	☑	

Field	Content		
	Data extraction	☑	☑
	Risk of bias (quality) assessment	☑	☑
	Data analysis	☑	☑
Named contact	Named contact: National Guideline Alliance development team, NICE Named contact e-mail: twintripletpregnancy@nice.org.uk Organisational affiliation of the review: National Institute for Health and Care Excellence (NICE)		
Review team members	National Guideline Alliance development team, NICE		
Funding sources/sponsor	This systematic review is being completed by NICE.		
Conflicts of interest	All guideline committee members and anyone who has direct input into NICE guidelines (including the evidence review team and expert witnesses) must declare any potential conflicts of interest in line with NICE's code of practice for declaring and dealing with conflicts of interest. Any relevant interests, or changes to interests, will also be declared publicly at the start of each guideline committee meeting. Before each meeting, any potential conflicts of interest will be considered by the guideline committee Chair and a senior member of the development team. Any decisions to exclude a person from all or part of a meeting will be documented. Any changes to a member's declaration of interests will be recorded in the minutes of the meeting. Declarations of interests will be published with the final guideline.		
Collaborators	Development of this systematic review will be overseen by an advisory committee who will use the review to inform the development of evidence-based recommendations in line with section 3 of Developing NICE guidelines: the manual. Members of the guideline committee are available on the NICE website: [NICE guideline webpage].		
Other registration details	None		
Reference/URL for published protocol	https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42023465092		

Field	Content	
Dissemination plans	NICE may use a range of different methods to raise awareness of the guideline. These include standard approaches such as: • notifying registered stakeholders of publication • publicising the guideline through NICE's newsletter and alerts • issuing a press release or briefing as appropriate, posting news articles on the NICE website, using social media channels, and publicising the guideline within NICE.	
Keywords	Progesterone, preterm birth, twins, triplets, stillbirth, neonatal death	
Details of existing review of same topic by same authors	None	
Current review status	☐	Ongoing
	☒	Completed but not published
	☐	Completed and published
	☐	Completed, published and being updated
	☐	Discontinued
Additional information	None	
Details of final publication	www.nice.org.uk	

CDSR: Cochrane Database of Systematic Reviews; CENTRAL: Cochrane Central Register of Controlled Trials; GRADE: Grading of Recommendations Assessment; INAHTA: International Network of Agencies for Health Technology Assessment; IPD: individual participant data; MEDLINE: Medical Literature Analysis and Retrieval System Online; MID: minimally important difference; NICE: National Institute for Health and Care Excellence; RCT: randomised controlled trial; RoB: risk of bias; ROBIS: Risk of Bias in Systematic Reviews; SD: standard deviation; SRDR: The Systematic Review Data Repository; 17-OHPC: 17-hydroxyprogesterone caproate

Appendix B Literature search strategies

Literature search strategies for review question: What is the clinical and cost-effectiveness of progesterone in preventing spontaneous preterm birth in twin and triplet pregnancy?

Database: Ovid MEDLINE

Date of last search: 13/09/2023

#	Searches
1	exp Pregnancy, Multiple/
2	((multiple* or triplet* or twin* or dizygotic* or monozygotic* or trizygotic*) adj3 (birth* or f?etal or f?etus* or gestation* or pregnan*)).tw,kf.
3	(chorionicity or dichorionic or monochorionic or trichorionic).tw,kf.
4	or/1-3
5	exp Progesterone/
6	Progestins/
7	Progesterone Congeners/
8	(progest* or gestagen* or acetophenide or algestone or alphasone or dihydroxyprogesterone or dihydroprogesterone or dydrogesterone or 17a hexanoate or 17a hydroxypregn or 17alphahydroxyprogest* or alphahydroxyprogest* or hydroxyprogest* or 17a-HPC or 17aHPC or 17-OH or 17OH or 17-OHP or 17OHP or OHPC or 17OHPC or 17PC).tw,kf.
9	or/5-8
10	4 and 9
11	letter/
12	editorial/
13	news/
14	exp historical article/
15	Anecdotes as topic/
16	comment/
17	case reports/
18	(letter or comment*).ti.
19	or/11-18
20	randomized controlled trial/ or random*.ti,ab.
21	19 not 20
22	animals/ not humans/
23	exp Animals, Laboratory/
24	exp Animal Experimentation/
25	exp Models, Animal/
26	exp Rodentia/
27	(rat or rats or rodent* or mouse or mice).ti.
28	or/21-27
29	10 not 28
30	meta-analysis/
31	meta-analysis as topic/
32	(meta analy* or metanaly* or metaanaly*).ti,ab.
33	((systematic* or evidence*) adj2 (review* or overview*)).ti,ab.
34	(reference list* or bibliograph* or hand search* or manual search* or relevant journals).ab.
35	(search strategy or search criteria or systematic search or study selection or data extraction).ab.
36	(search* adj4 literature).ab.
37	(medline or pubmed or cochrane or embase or psychlit or psyclit or psychinfo or psycinfo or cinahl or science citation index or bids or cancerlit).ab.
38	cochrane.jw.
39	or/30-38
40	randomized controlled trial.pt.
41	controlled clinical trial.pt.
42	pragmatic clinical trial.pt.
43	randomi#ed.ab.
44	placebo.ab.
45	drug therapy.fs.
46	randomly.ab.
47	trial.ab.
48	groups.ab.

#	Searches
49	or/40-48
50	29 and 39
51	29 and 49
52	50 or 51
53	limit 52 to english language
54	limit 53 to yr="2018 -Current"

Database: Ovid Embase

Date of last search: 13/09/2023

#	Searches
1	exp multiple pregnancy/
2	((multiple* or triplet* or twin* or dizygotic* or monozygotic* or trizygotic*) adj3 (birth* or f?etal or f?etus* or gestation* or pregnan*)).tw,kf.
3	(chorionicity or dichorionic or monochorionic or trichorionic).tw,kf.
4	or/1-3
5	hydroxyprogesterone/ or progesterone/
6	gestagen/ or progesterone derivative/
7	(progest* or gestagen* or acetophenide or algestone or alphasone or dihydroxyprogesterone or dihydroprogesterone or dydrogesterone or 17a hexanoate or 17a hydroxypregn or 17alphahydroxyprogest* or alphahydroxyprogest* or hydroxyprogest* or 17a-HPC or 17aHPC or 17-OH or 17OH or 17-OHP or 17OHP or OHPC or 17OHPC or 17PC).tw,kf.
8	or/5-7
9	4 and 8
10	letter.pt. or letter/
11	note.pt.
12	editorial.pt.
13	case report/ or case study/
14	(letter or comment*).ti.
15	or/10-14
16	randomized controlled trial/ or random*.ti,ab.
17	15 not 16
18	animal/ not human/
19	nonhuman/
20	exp Animal Experiment/
21	exp Experimental Animal/
22	animal model/
23	exp Rodent/
24	(rat or rats or rodent* or mouse or mice).ti.
25	or/17-24
26	9 not 25
27	systematic review/
28	meta-analysis/
29	(meta analy* or metanaly* or metaanaly*).ti,ab.
30	((systematic or evidence) adj2 (review* or overview*)).ti,ab.
31	(reference list* or bibliograph* or hand search* or manual search* or relevant journals).ab.
32	(search strategy or search criteria or systematic search or study selection or data extraction).ab.
33	(search* adj4 literature).ab.
34	(medline or pubmed or cochrane or embase or psychlit or psyclit or psychinfo or psycinfo or cinahl or science citation index or bids or cancerlit).ab.
35	((pool* or combined) adj2 (data or trials or studies or results)).ab.
36	cochrane.jw.
37	or/27-36
38	random*.ti,ab.
39	factorial*.ti,ab.
40	(crossover* or cross over*).ti,ab.
41	((doubl* or singl*) adj blind*).ti,ab.
42	(assign* or allocat* or volunteer* or placebo*).ti,ab.
43	crossover procedure/
44	single blind procedure/
45	randomized controlled trial/
46	double blind procedure/
47	or/38-46
48	26 and 37
49	26 and 47
50	48 or 49

#	Searches
51	limit 50 to english language
52	limit 51 to yr="2018 -Current"
53	(conference abstract* or conference review or conference paper or conference proceeding).db,pt,su.
54	52 not 53

Database: Cochrane Database of Systematic Reviews (CDSR) Issue 9 of 12, September 2023

Date of last search: 13/09/2023

ID	Search
#1	MeSH descriptor: [Pregnancy, Multiple] explode all trees
#2	((multiple* or triplet* or twin* or dizygotic* or monozygotic* or trizygotic*) near/3 (birth* or fetal or foetal or fetus or foetus* or gestation* or pregnan*)):ti,ab,kw
#3	(chorionicity or dichorionic or monochorionic or trichorionic):ti,ab,kw
#4	{or #1-#3}
#5	MeSH descriptor: [Progesterone] explode all trees
#6	MeSH descriptor: [Progestins] this term only
#7	MeSH descriptor: [Progesterone Congeners] this term only
#8	(progest* or gestagen* or acetophenide or algestone or alphasone or dihydroxyprogesterone or dihydroprogesterone or dydrogesterone or "17a hexanoate" or "17a hydroxypregn" or 17alphahydroxyprogest* or alphahydroxyprogest* or hydroxyprogest* or "17a-HPC" or 17aHPC or "17-OH" or 17OH or "17-OHP" or 17OHP or OHPC or 17OHPC or 17PC):ti,ab,kw
#9	{or #5-#8}
#10	#4 and #9 with Cochrane Library publication date Between Jan 2018 and Sep 2023, in Cochrane Reviews, Cochrane Protocols

Database: Cochrane Central Register of Controlled Trials (CENTRAL) Issue 8 of 12, August 2023

Date of last search: 13/09/2023

ID	Search
#1	MeSH descriptor: [Pregnancy, Multiple] explode all trees
#2	((multiple* or triplet* or twin* or dizygotic* or monozygotic* or trizygotic*) near/3 (birth* or fetal or foetal or fetus or foetus* or gestation* or pregnan*)):ti,ab,kw
#3	(chorionicity or dichorionic or monochorionic or trichorionic):ti,ab,kw
#4	{or #1-#3}
#5	MeSH descriptor: [Progesterone] explode all trees
#6	MeSH descriptor: [Progestins] this term only
#7	MeSH descriptor: [Progesterone Congeners] this term only
#8	(progest* or gestagen* or acetophenide or algestone or alphasone or dihydroxyprogesterone or dihydroprogesterone or dydrogesterone or "17a hexanoate" or "17a hydroxypregn" or 17alphahydroxyprogest* or alphahydroxyprogest* or hydroxyprogest* or "17a-HPC" or 17aHPC or "17-OH" or 17OH or "17-OHP" or 17OHP or OHPC or 17OHPC or 17PC):ti,ab,kw
#9	{or #5-#8}
#10	#4 and #9
#11	"conference":pt or (clinicaltrials or trialsearch):so
#12	#10 not #11 with Publication Year from 2018 to 2023, in Trials

Database: Epistemonikos

Date of last search: 13/09/2023

ID	Search
1	(((multiple* or triplet* or twin* or dizygotic* or monozygotic* or trizygotic*) and (birth* or fetal or foetal or fetus or foetus* or gestation* or pregnan*)) or chorionicity or dichorionic or monochorionic or trichorionic)
2	(progest* or gestagen* or acetophenide or algestone or alphasone or dihydroxyprogesterone or dihydroprogesterone or dydrogesterone or "17a hexanoate" or "17a hydroxypregn" or 17alphahydroxyprogest* or alphahydroxyprogest* or hydroxyprogest* or "17a-HPC" or 17aHPC or "17-OH" or 17OH or "17-OHP" or 17OHP or OHPC or 17OHPC or 17PC)
3	1 and 2
4	[Filters: min_date=20180101, max_date=20230913

Economic Search Strategies:

Database: Ovid MEDLINE

Date of last search: 13/09/2023

#	Searches
1	exp Pregnancy, Multiple/
2	((multiple* or triplet* or twin* or dizygotic* or monozygotic* or trizygotic*) adj3 (birth* or f?etal or f?etus* or gestation* or pregnan*)).tw,kf.
3	(chorionicity or dichorionic or monochorionic or trichorionic).tw,kf.
4	or/1-3
5	exp Progesterone/
6	Progestins/
7	Progesterone Congeners/
8	(progest* or gestagen* or acetophenide or algestone or alphasone or dihydroxyprogesterone or dihydroprogesterone or dydrogesterone or 17a hexanoate or 17a hydroxypregn or 17alphahydroxyprogest* or alphahydroxyprogest* or hydroxyprogest* or 17a-HPC or 17aHPC or 17-OH or 17OH or 17-OHP or 17OHP or OHPC or 17OHPC or 17PC).tw,kf.
9	or/5-8
10	4 and 9
11	letter/
12	editorial/
13	news/
14	exp historical article/
15	Anecdotes as topic/
16	comment/
17	case reports/
18	(letter or comment*).ti.
19	or/11-18
20	randomized controlled trial/ or random*.ti,ab.
21	19 not 20
22	animals/ not humans/
23	exp Animals, Laboratory/
24	exp Animal Experimentation/
25	exp Models, Animal/
26	exp Rodentia/
27	(rat or rats or rodent* or mouse or mice).ti.
28	or/21-27
29	10 not 28
30	Economics/
31	Value of life/
32	exp "Costs and Cost Analysis"/
33	exp Economics, Hospital/
34	exp Economics, Medical/
35	exp Resource Allocation/
36	Economics, Nursing/
37	Economics, Pharmaceutical/
38	exp "Fees and Charges"/
39	exp Budgets/
40	budget*.ti,ab.
41	cost*.ti,ab.
42	(economic* or pharmaco?economic*).ti,ab.
43	(price* or pricing*).ti,ab.
44	(financ* or fee or fees or expenditure* or saving*).ti,ab.
45	(value adj2 (money or monetary)).ti,ab.
46	resourc* allocat*.ti,ab.
47	(fund or funds or funding* or funded).ti,ab.
48	(ration or rations or rationing* or rationed).ti,ab.
49	ec.fs.
50	or/30-49
51	29 and 50
52	limit 51 to english language
53	limit 52 to yr="2018 -Current"

Database: Ovid Embase

Date of last search: 13/09/2023

#	Searches
1	exp multiple pregnancy/
2	((multiple* or triplet* or twin* or dizygotic* or monozygotic* or trizygotic*) adj3 (birth* or f?etal or f?etus* or gestation* or pregnan*)).tw,kf.
3	(chorionicity or dichorionic or monochorionic or trichorionic).tw,kf.
4	or/1-3
5	hydroxyprogesterone/ or progesterone/
6	gestagen/ or progesterone derivative/
7	(progest* or gestagen* or acetophenide or algestone or alphasone or dihydroxyprogesterone or dihydroprogesterone or dydrogesterone or 17a hexanoate or 17a hydroxypregn or 17alphahydroxyprogest* or alphahydroxyprogest* or hydroxyprogest* or 17a-HPC or 17aHPC or 17-OH or 17OH or 17-OHP or 17OHP or OHPC or 17OHPC or 17PC).tw,kf.
8	or/5-7
9	4 and 8
10	letter.pt. or letter/
11	note.pt.
12	editorial.pt.
13	case report/ or case study/
14	(letter or comment*).ti.
15	or/10-14
16	randomized controlled trial/ or random*.ti,ab.
17	15 not 16
18	animal/ not human/
19	nonhuman/
20	exp Animal Experiment/
21	exp Experimental Animal/
22	animal model/
23	exp Rodent/
24	(rat or rats or rodent* or mouse or mice).ti.
25	or/17-24
26	9 not 25
27	health economics/
28	exp economic evaluation/
29	exp health care cost/
30	exp fee/
31	budget/
32	funding/
33	resource allocation/
34	budget*.ti,ab.
35	cost*.ti,ab.
36	(economic* or pharmaco?economic*).ti,ab.
37	(price* or pricing*).ti,ab.
38	(financ* or fee or fees or expenditure* or saving*).ti,ab.
39	(value adj2 (money or monetary)).ti,ab.
40	resourc* allocat*.ti,ab.
41	(fund or funds or funding* or funded).ti,ab.
42	(ration or rations or rationing* or rationed).ti,ab.
43	or/27-42
44	26 and 43
45	limit 44 to english language
46	limit 45 to yr="2018 -Current"
47	(conference abstract* or conference review or conference paper or conference proceeding).db,pt,su.
48	46 not 47

Database: Health Technology Assessment (HTA)

Date of last search: 13/09/2023

#	Searches
1	MeSH DESCRIPTOR Pregnancy, Multiple EXPLODE ALL TREES

#	Searches
2	((((multiple* or triplet* or twin* or dizygotic* or monozygotic* or trizygotic*) and (birth* or fetal or foetal or fetus or foetus* or gestation* or pregnan*)) or chorionicity or dichorionic or monochorionic or trichorionic))
3	#1 OR #2
4	MeSH DESCRIPTOR Progesterone EXPLODE ALL TREES
5	MeSH DESCRIPTOR Progesterone Congeners
6	MeSH DESCRIPTOR Progestins
7	((progest* or gestagen* or acetophenide or algestone or alphasone or dihydroxyprogesterone or dihydroprogesterone or dydrogesterone or "17a hexanoate" or "17a hydroxypregn" or 17alphahydroxyprogest* or alphahydroxyprogest* or hydroxyprogest* or "17a-HPC" or 17aHPC or "17-OH" or 17OH or "17-OHP" or 17OHP or OHPC or 17OHPC or 17PC))
8	#4 OR #5 OR #6 OR #7
9	* IN HTA FROM 2018 TO 2023
10	#3 AND #8 AND #9

Database: INAHTA International HTA Database

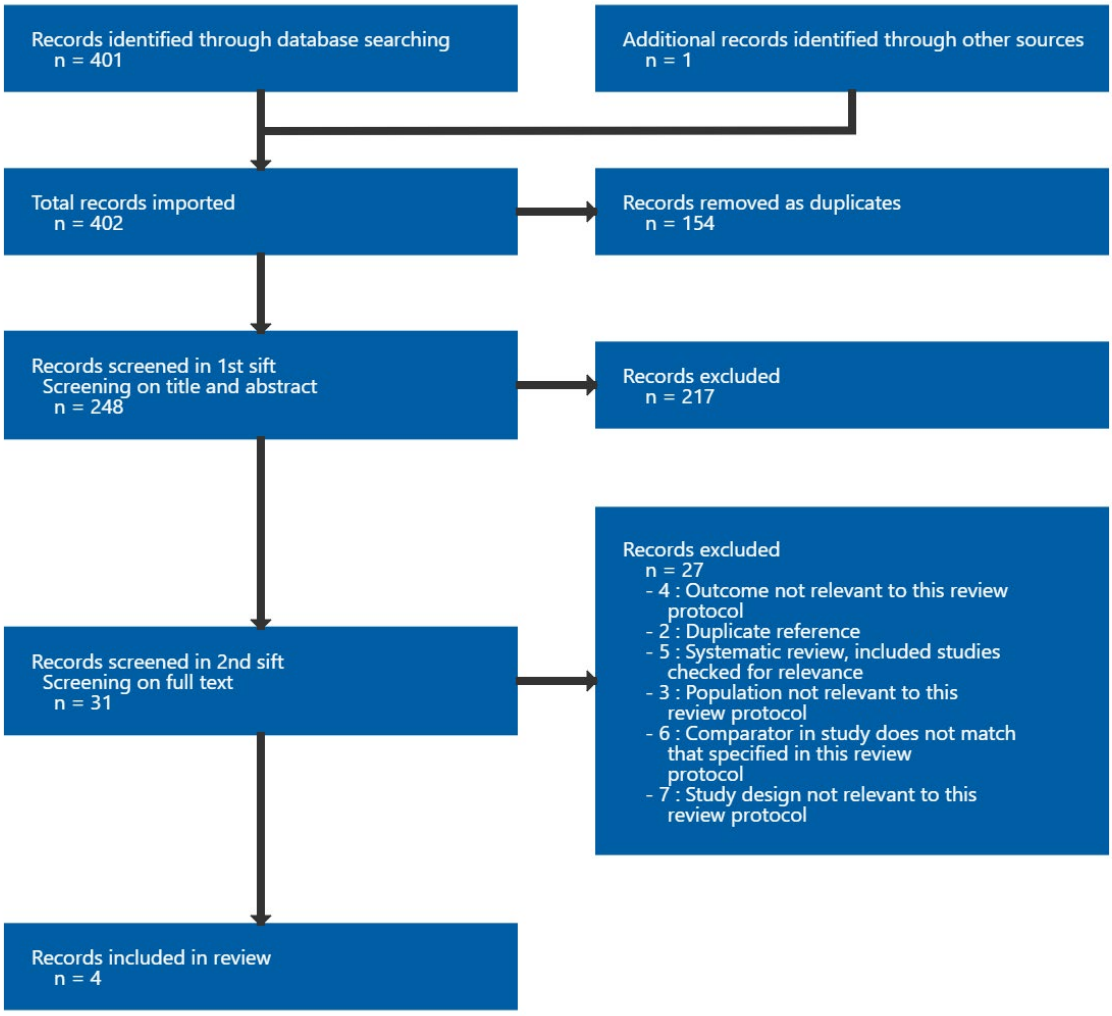
Date of last search: 13/09/2023

Line	Query
5	#3 AND #4 FROM 2018 TO 2023 AND (English)[Language]
4	(progest* or gestagen* or acetophenide or algestone or alphasone or dihydroxyprogesterone or dihydroprogesterone or dydrogesterone or "17a hexanoate" or "17a hydroxypregn" or 17alphahydroxyprogest* or alphahydroxyprogest* or hydroxyprogest* or "17a-HPC" or 17aHPC or "17-OH" or 17OH or "17-OHP" or 17OHP or OHPC or 17OHPC or 17PC)
3	#1 or #2
2	(((multiple* or triplet* or twin* or dizygotic* or monozygotic* or trizygotic*) and (birth* or fetal or foetal or fetus or foetus* or gestation* or pregnan*)) or chorionicity or dichorionic or monochorionic or trichorionic)
1	"Pregnancy, Multiple"[mh]

Appendix C Effectiveness evidence study selection

Study selection for: What is the clinical and cost-effectiveness of progesterone in preventing spontaneous preterm birth in twin and triplet pregnancy?

Figure 2: Study selection flow chart



Appendix D Evidence tables

Evidence tables for review question: What is the clinical and cost-effectiveness of progesterone in preventing spontaneous preterm birth in twin and triplet pregnancy?

Table 5: Evidence tables

Conde-Agudelo, 2022

Bibliographic Reference	Conde-Agudelo, A.; Rehal, Anoop.; Da Fonseca, Eduardo.; Brizot, M.L.; Rode, L.; Serra, V.; Cetingoz, E.; Syngelaki, A.; Tabor, A.; Perales, A.; Hassan, S.S.; Nicolaides, K.H.; Vaginal progesterone for the prevention of preterm birth and adverse perinatal outcomes in twin gestations with a short cervix: an updated individual patient data meta-analysis; Ultrasound Obstet Gynecol; 2022

Study details

Country/ies where study was carried out	• Brizot (2015) Brazil • Cetingoz (2011) Turkey • Fonseca (2007) International (UK, Chile, Brazil and Greece) • Rehal 2021 International (UK, Spain, Bulgaria, Italy, Belgium and France) • Rode (2011) International (Denmark and Austria) • Serra (2012) Spain
Study type	Individual participant data (IPD) review
Study dates	January 2017 - November 2021
Inclusion criteria	**IPD review - inclusion criteria** • RCTs investigating the effectiveness of vaginal progesterone for the prevention of preterm birth in women with a twin pregnancy and a mid-trimester sonographic short cervix (≤25 mm) or women with an unselected twin pregnancy and a pre-randomisation cervical length ≤25 mm **Individual study - inclusion criteria** Brizot (2015) • Naturally conceived diamniotic twin pregnancy

	• 18–21 weeks of gestation • No history of preterm delivery • No major fetal abnormalities • No allergies to progesterone or peanuts • No contraindicated health conditions, uterine malformation or prophylactic cerclage Cetingoz (2011) • Twin pregnancy • At least one spontaneous preterm birth • Malformation of uterus Fonseca (2007) • Short cervix (≤15 mm) • 20–25 weeks of gestation Rehal (2021) • Age over 18 years • Dichorionic or monochorionic diamniotic twin pregnancy • Two live fetuses at the 11-13 weeks' scan • Fluent in the local language Rode (2011) • Diamniotic twin pregnancy at 18–24 weeks • Chorionicity assessed before 16 weeks of gestation Serra (2012) • Dichorionic diamniotic twin pregnancy at 20 weeks of gestation
Exclusion criteria	**IPD review - exclusion criteria** • Study that was retracted • Study that did not collect data on cervical length before randomisation • Study that used vaginal progesterone treatment in the third trimester **Individual study - exclusion criteria** Brizot (2015) • Major fetal abnormalities • Ovular infection

	• Loss to follow up Cetingoz (2011) • Abortion • Delivery at 20 to 24 weeks of gestation prior to the study interventions • Prophylactic cervical cerclage Fonseca (2007) • Major fetal abnormalities • Painful regular contractions • History of ruptured membranes or cervical cerclage Rehal (2021) • Monoamniotic pregnancies • Monochorionic diamniotic pregnancies with early signs of twin-to-twin transfusion syndrome (>20% discordance in crown-rump length at the 11-13 weeks' scan) • Major fetal abnormality or nuchal translucency thickness >3.5 mm on 11-13 weeks' scan • Women who were unconscious or severely ill • Women with learning difficulties or serious mental illness; hypersensitivity to progesterone; regular treatment with progesterone within the previous 7 days; severe hepatic dysfunction; mammary or genital tract cancer, thrombophlebitis or thromboembolic disorders; porphyria; cerebral hemorrhage; and allergy to sunflower oil, soya lecithin, glycerol (E422), gelatin, and titanium dioxide (E171) • Women who participated in another drug trial within 28 days Rode (2011) • Age under 18 years • Known allergy to progesterone or peanuts • History of hormone associated thromboembolic disorders • Rupture of membrane • Treatment for or signs of twin-to-twin transfusion syndrome • Intentional fetal reduction • Known major structural or chromosomal fetal abnormality • Known or suspected malignancy in genitals or breasts • Known liver disease • Higher-order multiples • Women who did not understand and speak German or Danish Serra (2012)

	• Elective cervical cerclage <14 weeks of gestation • History of hepatic problems or cholestasis • Abnormal kidney function • Abnormal liver enzymes • Recurrent vaginal bleeding or infections • Fetal anomalies • Alcohol or illicit drug consumption • ≥10 cigarettes per day
Patient characteristics	IPD review - patient characteristics Not reported
Intervention(s)/control	**IPD review - intervention(s)/control** Vaginal progesterone (100-600 mg) Placebo **Individual study - intervention(s)/control** Brizot (2015) • Vaginal Progesterone 200 mg suppository nightly from 18-21 to 34 weeks of gestation • Placebo Cetingoz (2011) • Vaginal progesterone 100 mg suppository (micronized progesterone) nightly between 24 and 34 weeks of gestation • Placebo Fonseca (2007) • Vaginal progesterone 200 mg suppository nightly from 24–25 to 34 weeks of gestation • Placebo Rehal (2021) • One 300 mg vaginal progesterone capsule twice a day from 11-14 weeks to 34 weeks of gestation • One vaginal placebo capsule twice a day

	Rode (2011) • Vaginal micronised progesterone 200 mg suppository daily from 20–24 to 34 weeks of gestation • Placebo Serra (2012) • Vaginal Progesterone (1) 2 x 200 mg suppository nightly, (2) 1 x 200 mg + 1 x placebo suppository nightly 20–34 weeks of gestation • Placebo
Duration of follow-up	Not reported
Sources of funding	**IPD review - sources of funding** Not industry funded **Individual study - sources of funding** Not reported: Cetingoz (2011) Not industry funded: Brizot (2015) Fonseca (2007) Rehal (2021) Rode (2011) Serra (2012)
Sample size	**IPD review - sample size** N=95 Vaginal progesterone: n=52 (n=104 baby or fetus) Placebo: n=43 (n=86 baby or fetus) **Individual study - sample size** Brizot (2015) IPD: n=21 Cetingoz (2011)

	IPD: n=7 Fonseca (2007) IPD: n=24 Rehal (2021) IPD: n=16 Rode (2011) IPD: n=21 Serra (2012) IPD: n=6

RCT: randomised controlled trial

Outcomes

Vaginal progesterone (100-600 mg per day) versus placebo

Outcome	Vaginal progesterone vs Placebo
Stillbirth (fetal death) (baby or fetus) Vaginal progesterone=6/104 vs. Placebo=4/86 (analysis with adjustment for the non-independence between twins) Relative risk/95% CI	0.54 (0.17 to 1.77)
Neonatal death (baby or fetus) Vaginal progesterone=4/104 vs. Placebo=9/86 (analysis with adjustment for the non-independence between twins) Relative risk/95% CI	0.51 (0.2 to 1.28)
Preterm birth <28 weeks (pregnant women) Vaginal progesterone=7/52 vs. Placebo=11/43 Relative risk/95% CI	0.41 (0.19 to 0.91)

Outcome	Vaginal progesterone vs Placebo
Preterm birth <32 weeks (pregnant women) Vaginal progesterone=16/52 vs. Placebo=20/43 Relative risk/95% CI	0.56 (0.33 to 0.93)
Preterm birth <37 weeks (pregnant women) Vaginal progesterone=43/52 vs. Placebo=38/43 Relative risk/95% CI	0.91 (0.75 to 1.1)
Spontaneous preterm birth <34 weeks (pregnant women) Vaginal progesterone=20/52 vs. Placebo=28/43 Relative risk/95% CI	0.58 (0.38 to 0.89)
Composite of serious neonatal complications (composite of neonatal morbidity/mortality, including respiratory distress syndrome, intraventricular haemorrhage, necrotising enterocolitis, proven neonatal sepsis, or neonatal death) (follow-up not reported) (baby or fetus) Vaginal progesterone=24/102 vs. Placebo=31/84 (analysis with adjustment for the non-independence between twins) Relative risk/95% CI	0.59 (0.33 to 0.98)

CI: confidence interval

Critical appraisal - NGA Critical appraisal – Wang et al 2021 checklist

Methodological items	Answer
Did the research questions and inclusion criteria for the review include the components of PICO?	Low *(Study authors report all components of PICO)*
Did the report of the review contain an explicit statement that the review methods were established before conduct of the review	Low *(Study authors report study protocol or deviations from protocol)*

Methodological items	Answer
and did the report justify any significant deviations from the protocol?	
Did the review authors explain their selection of the study designs for inclusion in the review?	Low *(Study authors provide rationale for selection of included study design)*
Did the review authors use a comprehensive literature search strategy?	Low *(Study authors reported a comprehensive literature search strategy)*
Did the review authors perform study selection in duplicate?	Low *(Study selection was independently conducted by two reviewers)*
Did the review authors perform data extraction in duplicate?	Some concerns *(Unclear if study authors performed data extraction in duplicate)*
Did the review authors provide a list of excluded studies and justify the exclusions?	High *(Study authors do not provide details on this)*
Did the review authors describe the included studies in adequate detail?	High *(Study authors do not provide adequate details on this)*
Did the review authors use a satisfactory technique for assessing RoB in individual studies that were included in the review?	Low *(Study authors provide adequate details on the technique used for assessing risk of bias)*
Did the review authors report on the sources of funding for the studies included in the review?	High *(Study authors do not provide adequate details on funding acquired in included studies in the review)*
If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis?	Low *(Study authors provide adequate details on the impact of risk of bias on the overall review findings and the included RCTs have low risk of bias)*
Did the review authors account for RoB in primary studies when interpreting or discussing the results of the review?	Low *(Study authors accounted for risk of bias in primary studies in the overall interpretation of the review findings and the included RCTs have low risk of bias)*

Methodological items	Answer
Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?	Low *(There was no significant heterogeneity in the results)*
If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?	Some concerns *(Unclear whether the study assessed the impact of publication bias)*
Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?	Low *(Study authors report potential sources of conflict of interest, including any funding they received for conducting the review)*
Was the quality of time-to-event-outcome data checked?	Low *(Outcome data not relevant to review type)*
Did researchers stratify or account for clustering of participants within trials using either a one or two stage approach to meta-analysis?	Low *(Clustering of participants was accounted for by a two-stage approach to meta-analysis)*
Was the choice of one or two stage analysis specified in advance or results for both approaches provided, or both?	Low *(Methodology for two stage approach was specified in advance)*
Were IPD obtained from a large proportion of the eligible trials?	Low risk of bias *(Study authors obtained data from all eligible trials)*
Were the reasons for not obtaining IPD provided?	Low *(Study authors report reasons for why IPD data were not available)*
Were there any strategies taken to account for unavailable IPD?	Some concerns *(Study authors do not report details on strategies taken to account for unavailable IPD)*
Were the data checked for missing, invalid, out of range, or inconsistent items?	Low *(Study authors checked trial data for completeness)*
Did the author check any discrepancies with the trial report (if available)?	Some concerns *(Information unavailable)*
Were any issues queried and, if possible, resolved?	Low

Methodological items	Answer
	(Study authors contacted principal investigators to seek further information on any issues)
Were the methods of assessing whether effects of interventions vary by participant characteristics appropriate?	Low *(Study authors used appropriate methods to assessing varying effects of interventions by participant characteristics)*
Was the choice of participant level characteristics and methods of assessing participant level interactions specified in advance?	Low *(Participant level characteristics and methods of assessing participant level of interactions was specified in advance with rationale provided)*
If there was no evidence of a differential effect by trial or participant characteristic, was emphasis placed on the overall results?	Low *(Study authors took differential effects by participant characteristics into account where found)*
Were exploratory analyses highlighted as such?	*Low* *(Exploratory analyses (sensitivity) were labelled and taken into account when assessing the evidence)*
Does any report of the results adhere to the PRISMA-IPD?	Some concerns *(Information unavailable)*
Overall risk of bias and directness	Low
Overall risk of bias and directness	Directly applicable

IPD: individual participant data; PICO: population, intervention, comparison and outcome; RCT: randomised controlled trial; RoB: risk of bias

EPPPIC Group 2021

Bibliographic Reference	EPPPIC Group; Evaluating Progestogens for Preventing Preterm birth International Collaborative (EPPPIC): meta-analysis of individual participant data from randomised controlled trials.; Lancet (London, England); 2021; vol. 397 (no. 10280); 1183-1194

Study details

Country/ies where study was carried out	Vaginal progesterone trials (eight included trials; twin pregnancies) • Aboulghar (2012) Egypt • Crowther (2017) Australia • Fonseca (2007) International (UK, Chile, Brazil and Greece)

	• Norman (2009) UK • Rode (2011) International (Denmark and Austria) • Serra (2012) Spain • Wood (2012) Canada • Brizot (2015) Brazil Intramuscular 17-hydroxyprogesterone caproate (17-OHPC) trials (eight included trials, twin and triplet pregnancies) • Awaad (2015) Lebanon • Caritis (2009) US • Combs (2010) US • Combs (2011) US • Lim (2011) Netherlands • Rouse (2007) US • Senat (2013) France • Briery (2009) US
Study type	Individual participant data (IPD) review
Study dates	Primary search: April 2017 Top up search: July 30th 2019
Inclusion criteria	**IPD review - inclusion criteria** • RCTs that compared progesterone with placebo or standard care, or with other forms of progesterone in asymptomatic pregnant women with an increased risk of preterm birth **Individual study - inclusion criteria** **Vaginal progesterone trials** Aboulghar (2012) • First pregnancy with in vitro fertilisation/intracytoplasmic sperm injection • 18-24 weeks of gestation • Normal uterine, cervical and fetal anatomy Crowther (2017)

	• Spontaneous preterm birth in preceding singleton pregnancy (where labor onset spontaneous or with cervical competence or following preterm premature rupture of the membranes) • Current pregnancy between 18–24 week of gestation Fonseca (2007) • Short cervix (≤15 mm) • 20–25 weeks of gestation Norman (2009) • Twin pregnancy • Gestation and chorionicity established by scan before 20 weeks of gestation • 20-22 weeks of gestation Rode (2011) • Diamniotic twin pregnancy at 18–24 weeks • Chorionicity assessed before 16 weeks of gestation Serra (2012) • Dichorionic diamniotic twin pregnancy with 20 weeks of gestation Wood (2012) • Multifetal pregnancy with 16–21 weeks of gestation Brizot (2015) • Naturally conceived diamniotic twin pregnancy • 18–21 weeks of gestation • No history of preterm delivery • No major fetal abnormalities • No allergies to progesterone or peanuts • No contraindicated health conditions, uterine malformation or prophylactic cerclage **Intramuscular 17-hydroxyprogesterone caproate (17-OHPC) trials** Awaad (2015) • Twin pregnancy between 12–20 weeks of gestation • Age over 18 years Caritis (2009) • Triplet pregnancy between 16–20 weeks of gestation

	Combs (2010) • Trichorionic-triamniotic triplet pregnancy at 15–23 weeks of gestation • No major fetal anomalies Combs (2011) • Dichorionic-diamniotic twin pregnancy between 15–23 weeks of gestation • No major fetal anomalies Lim (2011) • Multifetal pregnancy between 15–19 weeks of gestation Rouse (2007) • Twin pregnancy between 16-20 weeks of gestation Senat (2013) • Twin pregnancy between 24–32 weeks of gestation • Cervix ≤25 mm • Age over 18 years Briery (2009) • Twin pregnancy between 20-30 weeks of gestation with intact membranes The IPD (individual participant data) also included singleton pregnancies. Data for this population was not included as it was not our protocol population.
Exclusion criteria	**IPD review - exclusion criteria** • Progestogens given to prevent early miscarriage • Progestogens given to treat symptomatic women with signs of threatened preterm labour **Individual study - exclusion criteria** **Vaginal progesterone trials** Aboulghar (2012) • Monochorionic and monoamniotic twins • Triplet pregnancies

	• Cervical cerclage Crowther (2017) • Current pregnancy with active vaginal bleeding requiring hospital admission at ≥18 weeks of gestation • Preterm premature rupture of the membranes • Active labour • Known lethal fetal anomaly or fetal demise • Progesterone treatment after 16 weeks of gestation Fonseca (2007) • Major fetal abnormalities • Painful regular contractions • History of ruptured membranes or cervical cerclage Norman (2009) • Structural or chromosomal fetal abnormality at recruitment • Planned cervical suture • Planned elective delivery <34 weeks of gestation • Planned intervention for twin-to-twin transfusion syndrome <22 weeks of gestation • Higher multiple pregnancies Rode (2011) • Age under 18 years • Known allergy to progesterone or peanuts • History of hormone associated thromboembolic disorders • Rupture of membrane • Treatment for or signs of twin-to-twin transfusion syndrome • Intentional fetal reduction • Known major structural or chromosomal fetal abnormality • Known or suspected malignancy in genitals or breasts • Known liver disease • Higher-order multiples • Women who did not understand and speak German or Danish Serra (2012) • Elective cervical cerclage <14 weeks of gestation • History of hepatic problems or cholestasis • Abnormal kidney function

- Abnormal liver enzymes
- Recurrent vaginal bleeding or infections
- Fetal anomalies
- Alcohol or illicit drug consumption
- ≥10 cigarettes per day

Wood (2012)

- Placenta praevia
- Pre-existing hypertension
- Known major fetal anomaly
- Monoamniotic monozygotic multifetals
- Other health conditions
- Known progesterone sensitivity

Brizot (2015)

- Major fetal abnormalities
- Ovular infection
- Loss to follow up

Intramuscular 17-hydroxyprogesterone caproate (17-OHPC) trials

Awaad (2015)

- Fetal anomalies
- Elective cervical cerclage prior to 14 weeks of gestation
- Contraindicated health conditions

Caritis (2009)

- Serious fetal anomalies
- Two or more fetuses in one amniotic sac
- Suspected twin-to-twin transfusion syndrome
- Ultra-sonographic growth discordance
- Cervical cerclage
- Major uterine anomaly
- Un-fractioned heparin therapy greater than 10,000 units/day
- Low molecular weight heparin therapy
- Major chronic medical diseases

Combs (2010)

	• Symptomatic uterine contractions • Rupture of membrane • Contraindication to interventions prolonging pregnancy • Contraindicated health conditions • Women <18 years of age • Cervical cerclage Combs (2011) • Age under 18 years • Progestin intake after 15 weeks of gestation • Symptomatic uterine contractions • Rupture of the fetal membranes • Contraindication to prolonging the pregnancy • Contraindicated health conditions Lim (2011) • Previous preterm birth <34 weeks of gestation • Serious congenital defects • Death of ≥1 fetus • Early signs of twin-to-twin transfusion syndrome • Primary cerclage Rouse (2007) • Serious fetal anomalies • Spontaneous death of a fetus after 12 weeks • Monoamniotic placenta • Suspected twin-to-twin transfusion syndrome • Marked ultra-sonographic growth • Cerclage • Major uterine anomaly • >10,000 units of unfractionated heparin/day • Treatment with low-molecular-weight heparin • Major chronic medical conditions • Twin gestations as result of intentional fetal reduction Senat (2013) • Cervical dilatation >3 cm

	• Preterm premature rupture of the membranes • Placenta praevia • Monochorial monoamniotic pregnancy • Signs of twin-to-twin transfusion syndrome • Severe intrauterine growth restriction • Major structural or chromosomal fetal abnormality • Death of 1 fetus • Maternal or fetal disease requiring preterm birth • Progesterone therapy before inclusion • Anticonvulsant treatment • Intentional fetal reduction Briery (2009) • Severe medical disorders • Cervical dilation ≥1cm • Intrauterine growth restriction • Growth discordancy between twins • Cerclage • Uterine abnormalities
Patient characteristics	**IPD review - patient characteristics** **Maternal age in years [mean(SD)]*** Vaginal progesterone: 31.3 (5.4) Intramuscular 17-hydroxyprogesterone caproate (17-OHPC): 31.2 (5.9) **Maternal BMI [n(%)]*** Vaginal progesterone: Underweight: 60 (2.9) Normal weight: 775 (37.5) Overweight: 305 (14.7) Obese: 185 (8.9) Unknown: 743 (35.9) Total: N=2068 Intramuscular 17-hydroxyprogesterone caproate (17-OHPC): Underweight: 73 (3.2)

	Normal weight: 1046 (46.1) Overweight: 532 (23.4) Obese: 476 (21.0) Unknown: 143 (6.3) Total: N=2270 **Cervical length [n(%)]*** Vaginal progesterone: Women with a shorter cervix (≤25 mm): 72 (3.5) Women with a longer cervix (>25 mm): 1155 (55.8) Women with unknown cervix length: 841 (40.6) Intramuscular 17-hydroxyprogesterone caproate (17-OHPC): Women with a shorter cervix (≤25 mm): 227 (10) Women with a longer cervix (>25 mm) 1121 (49.4) Women with unknown cervix length: 922 (40.6) **Previous preterm birth and parity [n(%)]*** Vaginal progesterone: Women with previous preterm birth: 61 (2.9) Women with no previous preterm birth: 611 (29.6) Nulliparous women: 1058 (51) Women with unknown parity: 338 (16.4) Intramuscular 17-hydroxyprogesterone caproate (17-OHPC): Women with previous preterm birth: 202 (8.9) Women with no previous preterm birth: 872 (38.5) Nulliparous women: 1191 (52.4) Women with unknown parity: 5 (0.2) **Pre-pregnancy diabetes [n(%)]*** Vaginal progesterone: No: 755 (36.5) Yes: 9 (0.4) Unknown: 1304 (63.1) Intramuscular 17-hydroxyprogesterone caproate (17-OHPC): No: 1418 (62.5) Yes: 11 (0.5)

	Unknown: 841 (37.0) **Chronic hypertension [n(%)]*** Vaginal progesterone: No: 435 (21.0) Yes: 39 (1.9) Unknown: 1594 (77.1) Intramuscular 17-hydroxyprogesterone caproate (17-OHPC): No: 1743 (76.8) Yes: 41 (1.8) Unknown: 486 (21.4) **Ethnicity [n(%)]*** Vaginal progesterone: Black: 38 (1.8) Asian: 14 (0.7) Hispanic: 0.0 (0.0) Middle Eastern: 95 (4.6) Other: 177 (8.6) White: 524 (25.3) Unknown: 1220 (59.0) Intramuscular 17-hydroxyprogesterone caproate (17-OHPC): Black: 223 (9.8) Asian: 72 (3.2) Hispanic: 170 (7.5) Middle Eastern: 306 (13.5) Other: 32 (1.4) White: 1278 (56.3) Unknown: 189 (8.3) *Data was not reported for control groups.
Intervention(s)/control	**IPD review - intervention(s)/control** Vaginal progesterone (90-400 mg) Intramuscular 17-hydroxyprogesterone caproate (17-OHPC) (250-500mg)

	Individual study - intervention(s)/control **Vaginal progesterone trials** Aboulghar (2012) • Vaginal progesterone 400 mg suppository daily from 18-24 to 37 weeks of gestation • Placebo Crowther (2017) • Vaginal progesterone 100 mg suppository nightly from 20–24 to 34 weeks of gestation • Placebo Fonseca (2007) • Vaginal progesterone 200 mg suppository nightly from 24–25 to 34 weeks of gestation • Placebo Norman (2009) • Vaginal progesterone 90 mg gel daily 24–34 weeks of gestation • Placebo Rode (2011) • Vaginal micronised progesterone 200 mg suppository daily from 20–24 to 34 weeks of gestation • Placebo Serra (2012) • Vaginal progesterone (1) 2 x 200 mg suppository nightly, (2) 1 x 200 mg + 1 x placebo suppository nightly 20–34 weeks of gestation • Placebo Wood (2012) • Vaginal progesterone 90 mg gel daily from 16–21 to 36 weeks of gestation • Placebo Brizot (2015) • Vaginal progesterone 200 mg suppository nightly from 18-21 to 34 weeks of gestation • Placebo **Intramuscular 17-hydroxyprogesterone caproate (17-OHPC)**

	Awaad (2015) • Intramuscular injection of 17-OHPC 250 mg weekly from 16–20 to 36 weeks of gestation • Placebo Caritis (2009) • Intramuscular injection of 17-OHPC 250 mg weekly from 16–20 to 35 weeks of gestation • Placebo Combs (2010) • Intramuscular injection of 17-OHPC 250 mg weekly from 16–24 to 34 weeks of gestation • Placebo Combs (2011) • Intramuscular injection of 17-OHPC 250 mg weekly from 16–24 to 34 weeks of gestation • Placebo Lim (2011) • Intramuscular injection of 17-OHCP 250 mg weekly from 16–20 to 36 weeks of gestation • Placebo Rouse (2007) • Intramuscular injection of 17-OHPC 250 mg weekly from 16-20 to 35 weeks of gestation • Placebo Senat (2013) • Intramuscular injection of 17-OHCP 500 mg twice weekly from 24-32 to 36 weeks of gestation • Standard care Briery (2009) • Intramuscular injection of 17-OHPC 250 mg weekly 20-34 weeks of gestation • Placebo
Duration of follow-up	Not reported in the IPD
Sources of funding	**IPD review - sources of funding** Not industry funded

	Individual study - sources of funding Part industry funded: Aboulghar (2012) Not industry funded: Awaad (2015) Briery (2009) Brizot (2015) Caritis (2009) Combs (2010) Combs (2011) Crowther (2017) Fonseca (2007) Lim (2011) Norman (2009) Rode (2011) Rouse (2007) Serra (2012) Senat (2013) Wood (2012)
Sample size	**IPD review - sample size** N=9936 (30 trials) 1 trial added in a targeted update 31 trials included in the analysis N=11644 (N=16185 baby or fetus) **Individual study - sample size** **Aboulghar (2012)** N randomised: 313 (215 singleton pregnancies, 91 twin pregnancies) IPD: n=303 N missing: 6 lost to follow up, 1 termination, 3 no information **Awaad (2015)** N randomised: 293

	IPD: n=288 N missing: 5 lost to follow up **Briery (2009)** N randomised: 30 IPD: n=30 N missing: not reported **Brizot (2015)** N randomised: 390 IPD: n=390 N missing: not reported **Caritis (2009)** N randomised: 134 IPD: n=134 N missing: not reported **Combs (2010)** N randomised: 81 IPD: n=81 N missing: not reported **Combs (2011)** N randomised: 240 IPD: n=240 N missing: 2 lost to follow up **Crowther (2017)** N randomised: 787 (705 singleton pregnancies, 12 twin pregnancies) IPD: n=787 N missing: not reported **Fonseca (2007)** N randomised: 250 (226 singleton pregnancies, 24 twin pregnancies) IPD: n=250 N missing: not reported

	Lim (2011) N randomised: 671 IPD: n=671 **Norman (2009)** N randomised: 500 IPD: n=500 N missing: not reported **Rode (2011)** N randomised: 677 IPD: n=677 N missing not reported **Rouse (2007)** N randomised: 661 IPD: n=661 **Serra (2012)** N randomised: 294 IPD: n=290 N missing: 4 lost to follow up **Senat (2013)** N randomised: 165 IPD: n=165 **Wood (2012)** N randomised: 84 IPD: n=84 N missing: not reported
Other information	In 2020, a large additional trial (PROLONG) completed outside of the meta-analysis inclusion timeframe and was included in a targeted update of initial analyses. Some of the included trials contained mixed populations of singleton and twin pregnancies

IPD: individual participant data; RCT: randomised controlled trial

Outcomes

Vaginal progesterone (90-400 mg per day) versus control

Outcome	Vaginal progesterone vs Control
Stillbirth (fetal death/stillbirth) (pregnant women) n=4103 pregnant women (n not reported for each group separately) Relative risk/95% CI	1.19 (0.67 to 2.1)
Neonatal death (neonatal death after live birth) (pregnant women) n=4035 pregnant women (n not reported for each group separately) Relative risk/95% CI	1.15 (0.66 to 1.98)
Preterm birth <28 weeks (pregnant women) n=2046 pregnant women (n not reported for each group separately) Relative risk/95% CI	1.22 (0.77 to 1.94)
Preterm birth <34 weeks (pregnant women) n=2046 pregnant women (n not reported for each group separately) Relative risk/95% CI	1.01 (0.84 to 1.2)
Composite of serious neonatal complications (serious neonatal complications, including severe necrotising enterocolitis stages 2–3, intraventricular haemorrhage grades 3–4, retinopathy of prematurity stage 3 or worse, bronchopulmonary dysplasia, confirmed sepsis, patent ductus arteriosus, and neonatal infection) (follow-up not reported) (pregnant women) n=3840 pregnant women (n not reported for each group separately) Relative risk/95% CI	0.94 (0.74 to 1.2)

Outcome	Vaginal progesterone vs Control
Composite of adverse maternal outcomes (maternal complications, including gestational hypertension, pre-eclampsia, gestational diabetes, and maternal infection such as chorioamnionitis) (follow-up not reported) (pregnant women) n=1938 pregnant women (n not reported for each group separately) Relative risk/95% CI	0.93 (0.73 to 1.17)

CI: confidence interval

Intramuscular 17-hydroxyprogesterone caproate (17-OHPC; 250 mg weekly or 500 mg twice weekly) versus control

Outcome	Intramuscular 17-OHPC vs Control
Stillbirth (fetal death/stillbirth) (pregnant women/infants) n=4744 pregnant women/infants (n not reported for each group separately) Relative risk/95% CI	1 (0.67 to 1.5)
Neonatal death (neonatal death after live birth) (pregnant women/infants) n=4632 pregnant women/infants (n not reported for each group separately) Relative risk/95% CI	1.13 (0.76 to 1.68)
Preterm birth <28 weeks (pregnant women) n=2253 pregnant women (n not reported for each group separately) Relative risk/95% CI	1.07 (0.78 to 1.46)
Preterm birth <34 weeks (pregnant women) n=2253 pregnant women (n not reported for each group separately) Relative risk/95% CI	1.04 (0.92 to 1.18)

Outcome	Intramuscular 17-OHPC vs Control
Composite of serious neonatal complications (serious neonatal complications, including severe necrotising enterocolitis stages 2–3, intraventricular haemorrhage grades 3–4, retinopathy of prematurity stage 3 or worse, bronchopulmonary dysplasia, confirmed sepsis, patent ductus arteriosus, and neonatal infection) (follow-up not reported) (pregnant women) n=4724 pregnant women (n not reported for each group separately) Relative risk/95% CI	1.12 (0.76 to 1.65)
Composite of adverse maternal outcomes (maternal complications, including gestational hypertension, pre-eclampsia, gestational diabetes, and maternal infection such as chorioamnionitis) (follow-up not reported) (pregnant women) n=2095 pregnant women (n not reported for each group separately) Relative risk/95% CI	1.09 (0.94 to 1.27)

CI: confidence interval

Critical appraisal - NGA Critical appraisal – Wang et al 2021 checklist

Methodological items	Answer
Did the research questions and inclusion criteria for the review include the components of PICO?	*Low* *(Study authors report all components of PICO)*
Did the report of the review contain an explicit statement that the review methods were established before conduct of the review and did the report justify any significant deviations from the protocol?	*Low* *(Study authors describe the study protocol, literature search and screening with predefined outcome metrics and analysis plan. Protocol is registered and available for view. No deviations from protocol reported)*
Did the review authors explain their selection of the study designs for inclusion in the review?	*High* *(Study authors do not provide rationale for selection of study designs for inclusion)*
Did the review authors use a comprehensive literature search strategy?	*Low* *(Study authors used a comprehensive literature search strategy with at least two databases used and keyword/MeSH terms used. Reference lists of included trials and prior systematic reviews were checked.)*

Methodological items	Answer
Did the review authors perform study selection in duplicate?	*Low* *(Study authors state that study selection was performed independently by two researchers)*
Did the review authors perform data extraction in duplicate?	*Low* *(Study authors state that data extraction was performed by one researcher and checked by a second)*
Did the review authors provide a list of excluded studies and justify the exclusions?	*Some concerns* *(Study authors provide number of excluded studies with a reason for group as a whole)*
Did the review authors describe the included studies in adequate detail?	*Some concerns* *(Study authors do not provide details of intervention and control group numbers from the included studies)*
Did the review authors use a satisfactory technique for assessing RoB in individual studies that were included in the review?	*Low* *(Study authors assess RoB satisfactorily on included studies)*
Did the review authors report on the sources of funding for the studies included in the review?	*Low* *(Study authors provide details on funding acquired in included studies in the review)*
If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis?	*Low* *(Study authors took into account the potential impact RoB of the individual studies)*
Did the review authors account for RoB in primary studies when interpreting or discussing the results of the review?	*Low* *(Study authors took into account RoB data from primary studies when discussing the results)*
Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?	*Low* *(Study authors explore heterogeneity and adequately report on it in the review findings)*
If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?	*High* *(Publication bias was not investigated)*

Methodological items	Answer
Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?	*Low* *(Study authors report potential sources of conflict of interest, including any funding they received for conducting the review)*
Was the quality of time-to-event-outcome data checked?	*Low* *(Outcome data not relevant to review type)*
Did researchers stratify or account for clustering of participants within trials using either a one or two stage approach to meta-analysis?	*Low* *(Clustering of participants was accounted for by a two-stage approach to meta-analysis)*
Was the choice of one or two stage analysis specified in advance or results for both approaches provided, or both?	*Low* *(Methodology for two stage approach was specified in advance)*
Were IPD obtained from a large proportion of the eligible trials?	*Low* *(Study authors obtained data from all eligible trials)*
Were the reasons for not obtaining IPD provided?	*Low* *(Study authors report reasons for why IPD data were not available)*
Were there any strategies taken to account for unavailable IPD?	*Unclear* *(Authors did not report on any strategies taken to account for unavailable IPD data)*
Were the data checked for missing, invalid, out of range, or inconsistent items?	*Low* *(Study authors checked trial data for completeness)*
Did the author check any discrepancies with the trial report (if available)?	*Unclear* *(The study authors do not report if any discrepancies with the trial report were checked)*
Were any issues queried and, if possible, resolved?	Low *(Study authors contacted principal investigators to seek further information on any issues)*
Were the methods of assessing whether effects of interventions vary by participant characteristics appropriate?	*Low* *(Study authors used appropriate methods to assessing varying effects of interventions by participant characteristics)*
Was the choice of participant level characteristics and methods of assessing participant level interactions specified in advance?	*Participant level characteristics and methods of assessing participant level of interactions was specified in advance with rationale provided)*

Methodological items	Answer
If there was no evidence of a differential effect by trial or participant characteristic, was emphasis placed on the overall results?	*Low* *(Study authors took differential effects by participant characteristics into account where found)*
Were exploratory analyses highlighted as such?	*Low* *(Exploratory analyses (sensitivity) were labelled and taken into account when assessing the evidence)*
Does any report of the results adhere to the PRISMA-IPD?	*Low* *(Report of results adhere to PRISMA-IPD)*
Overall risk of bias and directness	*Low*
Overall risk of bias and directness	*Directly applicable*

IPD: individual participant data; PICO: population, intervention, comparison and outcome; RoB: risk of bias

Rehal, 2021

Bibliographic Reference	Rehal, Anoop; Benko, Zsofia; De Paco Matallana, Catalina; Syngelaki, Argyro; Janga, Deepa; Cicero, Simona; Akolekar, Ranjit; Singh, Mandeep; Chaveeva, Petya; Burgos, Jorge; Molina, Francisca S; Savvidou, Makrina; De La Calle, Maria; Persico, Nicola; Quezada Rojas, Maria Soledad; Sau, Ashis; Greco, Elena; O'Gorman, Neil; Plasencia, Walter; Pereira, Susana; Jani, Jacques C; Valino, Nuria; Del Mar Gil, Maria; Maclagan, Kate; Wright, Alan; Wright, David; Nicolaides, Kypros H; Early vaginal progesterone versus placebo in twin pregnancies for the prevention of spontaneous preterm birth: a randomized, double-blind trial.; American journal of obstetrics and gynecology; 2021; vol. 224 (no. 1); 86e1-86e19

Study details

Country/ies where study was carried out	England, Spain, Bulgaria, Italy, Belgium and France
Study type	Randomised controlled trial (RCT)
Study dates	May 2017 - April 2019
Inclusion criteria	• Age over 18 years • Dichorionic or monochorionic diamniotic twin pregnancy • Two live fetuses at the 11-13 weeks' scan

	• Fluent in the local language
Exclusion criteria	• Monoamniotic pregnancies • Monochorionic diamniotic pregnancies with early signs of twin-to-twin transfusion syndrome (>20% discordance in crown-rump length at the 11-13 weeks' scan) • Major fetal abnormality or nuchal translucency thickness >3.5 mm on 11-13 weeks' scan • Women who were unconscious or severely ill • Women with learning difficulties or serious mental illness; hypersensitivity to progesterone; regular treatment with progesterone in the previous 7 days; severe hepatic dysfunction; mammary or genital tract cancer, thrombophlebitis or thromboembolic disorders; porphyria; cerebral haemorrhage; and allergy to sunflower oil, gelatin, soya lecithin, glycerol (E422), and titanium dioxide (E171) • Women who participated in another drug trial within 28 days
Patient characteristics	Median age in years (IQR): Vaginal progesterone: 34.1 (30.3-37.7) Placebo: 34.0 (30.0-37.6) Median gestational age in weeks at randomisation (IQR): Vaginal progesterone: 13.2 (12.7-13.6) Placebo: 12.2 (12.7-13.7) Dichorionic pregnancies: Vaginal progesterone: N=449 Placebo: N=453 Monochorionic pregnancies: Vaginal progesterone: N=133 Placebo: N=134 Median cervical length in mm (IQR): Vaginal progesterone: 34.4 (31.0-38.0) Placebo: 34.6 (31.5-38.0) Participants with cervical length <30 mm: Vaginal progesterone: N=85 Placebo: N=70 Chronic hypertension:

	Vaginal progesterone: N=11 Placebo: N=7 Type 1 or type 2 diabetes: Vaginal progesterone: N=8 Placebo: N=3 Natural conception: Vaginal progesterone: N=382 Placebo: N=380 Assisted conception by ovulation drugs: Vaginal progesterone: N=35 Placebo: N=44 In vitro fertilisation: Vaginal progesterone: N=165 Placebo: N=163 Median BMI in kg/m2 (IQR): Vaginal progesterone: 24.7 (21.9-28.4) Placebo: 24.3 (22.0-27.9) Nulliparous: Vaginal progesterone: N=317 Placebo: N=326 Parous with preterm birth <37 weeks: Vaginal progesterone: N=23 Placebo: N=33 Parous without preterm birth <37 weeks: Vaginal progesterone: N=242 Placebo: N=228 Ethnicity: White Vaginal progesterone: N=473 Placebo: N=492

	Ethnicity: Black Vaginal progesterone: N=69 Placebo: N=59 Ethnicity: South Asian Vaginal progesterone: N=18 Placebo: N=28 Ethnicity: East Asian Vaginal progesterone: N=8 Placebo: N=3 Ethnicity: Mixed Vaginal progesterone: N=14 Placebo: N=5
Intervention(s)/control	Vaginal progesterone (early vaginal progesterone): participants received 300 mg vaginal progesterone capsule twice a day from 11-14 weeks to 34 weeks of gestation Placebo: participants received vaginal placebo capsules, which were identical to vaginal progesterone, twice a day from 11-14 weeks to 34 weeks of gestation
Duration of follow-up	Adverse events were assessed from 20 weeks to 37 weeks of gestation in dichorionic twin pregnancies and from 16 weeks of gestation to 30 days after the last capsule was taken in monochorionic twin pregnancies. Duration of follow-up for neonatal complications was unclear.
Sources of funding	Not industry funded
Sample size	N=1194 Vaginal progesterone: n=596 Placebo: n=598 In the final analyses, n=582 pregnant women (n=1164 baby or fetus) from the vaginal progesterone group and n=587 (n=1174 baby or fetus) from the placebo group were included because n=10 withdrew consent and n=4 was lost to follow up in the vaginal progesterone group and n=11 withdrew consent in the placebo group.

IQR: interquartile range; RCT: randomised controlled trial

Outcomes

Vaginal progesterone (600 mg per day) versus placebo

Outcome	Vaginal progesterone, N=582 pregnant women (n=1164 baby or fetus)	Placebo, N=587 pregnant women (n=1174 baby or fetus)
Still birth or neonatal death (number of baby or fetus) No of events	12/1164	10/1174
Any birth between 24 weeks and <28 weeks (number of pregnant women) No of events	8/567	9/561
Any birth between 24 weeks and <32 weeks (number of pregnant women) No of events	38/567	39/561
Any birth between 24 and <37 weeks (number of pregnant women) No of events	315/567	296/561
Spontaneous birth between 24 and <34 weeks (number of pregnant women) adjusted OR: 1.35 (95% CI 0.88 to 2.05) (adjusted for the effect of participating centre, parity, chorionicity, and method of conception) No of events	56/541	44/538
Composite of serious neonatal complications (neonatal morbidity, including necrotising enterocolitis, retinopathy of prematurity, intraventricular haemorrhage, sepsis, respiratory distress syndrome and anaemia) (follow-up not reported) (number of baby or fetus) No of events	77/1125	85/1113

Outcome	Vaginal progesterone, N=582 pregnant women (n=1164 baby or fetus)	Placebo, N=587 pregnant women (n=1174 baby or fetus)
Composite of serious neonatal complications (requirement of neonatal therapy, including admission to neonatal intensive care unit and requirement of ventilation) (follow-up not reported) (number of baby or fetus) No of events	268/1125	252/1113
Composite of adverse maternal outcomes (mother with at least 1 serious adverse event, including pre-eclampsia with prolonged hospital stay [PHS] [5 days], eclampsia with PHS [10 days], pulmonary embolism with PHS [4 days], postnatal liver rupture with PHS [31 days], obstetrical cholestasis with PHS [2 days], abnormal liver function tests, postpartum haemorrhage with 3-L blood loss, gastritis with PHS [4 days], dyspnoea with PHS [2 days], restrictive cardiomyopathy with PHS [4 days], urinary tract infection with PHS [3 days], and maternal mirror syndrome in association with fetal hydrops) (follow-up up to 37 weeks of gestation or 30 days after the last capsule was taken) (number of pregnant women) No of events	8/582	7/587
Composite of adverse maternal outcomes (mother with at least 1 nonserious adverse event, including vaginal discharge, vaginal itching, vaginal pain or discomfort, vaginal bleeding, headache/dizziness, fatigue, depression, insomnia, nausea/vomiting, abdominal pain/discomfort, diarrhoea/constipation, joint pain, swelling of extremities, palpitations, itching/skin rash, urinary tract infection, and other adverse events) (follow-up up to 37 weeks of gestation or 30 days after the last capsule was taken) (number of pregnant women) No of events	200/596	186/598
Spontaneous birth between 24 and <34 weeks in women with previous preterm birth (number of pregnant women)	5/22	7/28

Outcome	Vaginal progesterone, N=582 pregnant women (n=1164 baby or fetus)	Placebo, N=587 pregnant women (n=1174 baby or fetus)
No of events		
Spontaneous birth between 24 and <34 weeks in women with previous term birth (number of pregnant women) No of events	21/228	10/213
Spontaneous birth between 24 and <34 weeks in women with long cervix ≥30 mm (number of pregnant women) No of events	48/466	32/477

CI: confidence interval; OR: odds ratio; PHS: prolonged hospital stay; RoB: risk of bias

Critical appraisal – NGA critical appraisal – Cochrane RoB2

Section	Question	Answer
Domain 1: Bias arising from the randomisation process	Risk of bias judgement for the randomisation process	Low *(Women were randomly assigned, in a 1:1 ratio by a simple permuted block provided by Besins Healthcare, Bruxelles, Belgium. In the random-sequence generation, there was stratification according to participating centre. No differences in baseline characteristics to suggest an issue with randomisation)*
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	Risk of bias for deviations from the intended interventions (effect of assignment to intervention)	Low *(Participants, investigators, pharmacists, and others involved in giving the intervention, assessing outcomes, or analysing data remained masked to treatment allocation until the end of the study. There is no reason to believe that deviations from the intended intervention arose due to trial context. Analysis was intention to treat.)*

Section	Question	Answer
Domain 3. Bias due to missing outcome data	Risk-of-bias judgement for missing outcome data	Low *(Missing data are less than 5% for all outcomes.)*
Domain 4. Bias in measurement of the outcome	Risk-of-bias judgement for measurement of the outcome	Low *(Measurement of all outcomes was appropriate. Measurement or ascertainment of the outcome unlikely to have differed between groups as the assessors were blinded to the intervention received and the outcomes measured are deemed objective outcomes.)*
Domain 5. Bias in selection of the reported result	Risk-of-bias judgement for selection of the reported result	Low *(There is clear evidence that all eligible reported results for the outcome correspond to all intended outcome measurements and analyses.)*
Overall bias and Directness	Risk of bias judgement	Low
Overall bias and Directness	Overall Directness	Directly applicable
Overall bias and Directness	Risk of bias variation across outcomes	None

RoB: risk of bias

Shabaan, 2018

Bibliographic Reference	Shabaan, O M; Hassanin, I M; Makhlouf, A M; Salem, M N; Hussein, M; Mohamed, M; Abbas, A M; Vaginal progesterone for prevention of preterm delivery in women with twin pregnancy: a randomized controlled trial.; Facts, views & vision in ObGyn; 2018; vol. 10 (no. 2); 93-98

Study details

Country/ies where study was carried out	Egypt

Study type	Randomised controlled trial (RCT)
Study dates	February 2015 - January 2017
Inclusion criteria	Pregnant women with naturally conceived dichorionic diamniotic twin gestation, uncomplicated pregnancy, 28 weeks of gestation, and no major fetal abnormalities
Exclusion criteria	Pregnant women with a contraindication to progesterone, single fetal demise or co-twin with fetal growth restriction, higher multiple pregnancies, polyhydramnios, premature rupture of membranes, threatened preterm labour, a cervical cerclage from current or previous pregnancy, chronic medical conditions and difficulty to attend follow-up visits regularly
Patient characteristics	Mean age in years (SD): Vaginal progesterone: 29.1 (4.1) Control: 28.6 (3.5) Mean gestational age in weeks by ultrasound (SD): Vaginal progesterone (first fetus): 28.9 (0.6) Vaginal progesterone (second fetus): 28.8 (0.7) Control (first fetus): 28.9 (0.3) Control (second fetus): 28.9 (0.4) Cervical length: not reported Previous preterm birth: not reported Mean parity (SD): Vaginal progesterone: 2.0 (1.7) Control: 2.1 (1.6) Participants with previous miscarriage: Vaginal progesterone: N=10 Control: N=8
Intervention(s)/control	Vaginal progesterone: participants received progesterone vaginal pessaries 400 mg per day at bedtime from 28 weeks of gestation to delivery

	Control: participants received the normal tonics during pregnancy
Duration of follow-up	Until delivery (However, duration of follow-up for neonatal complications was unclear)
Sources of funding	Not industry funded
Sample size	N=140 Vaginal progesterone: n=70 Control: n=70 n=11 pregnant women from each group were lost to follow-up, so the final analysis included n=59 (n=118 baby or fetus) from vaginal progesterone group and n=59 (n=118 baby or fetus) from control group

RCT: randomised controlled trial; SD: standard deviation

Outcomes

Vaginal progesterone (400 mg per day) versus control

Outcome	Vaginal progesterone, N=59 pregnant women (n=118 baby or fetus)	Control, N=59 pregnant women (n=118 baby or fetus)
Stillbirth or neonatal death (perinatal mortality rate) (number of baby or fetus) No of events	1/118	1/118
Preterm birth at 28-30 weeks (number of pregnant women) No of events	2/59	2/59
Preterm birth at 32-36 weeks (number of pregnant women) No of events	8/59	13/59

Critical appraisal – NGA critical appraisal – RoB2

Section	Question	Answer
Domain 1: Bias arising from the randomisation process	Risk of bias judgement for the randomisation process	Low *(Randomisation sequence generated using a computer-generated random table, and the allocation sequence was concealed. No significant differences between groups at baseline)*
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	Risk of bias for deviations from the intended interventions (effect of assignment to intervention)	Low *(Participants and the study personnel were aware of the interventions; however, there is no reason to believe that deviations from the intended intervention arose due to trial context. Appropriate analysis was used)*
Domain 3. Bias due to missing outcome data	Risk-of-bias judgement for missing outcome data	Some concerns *(About 16% of participants were lost to follow-up from each group)*
Domain 4. Bias in measurement of the outcome	Risk-of-bias judgement for measurement of the outcome	Low *(Outcomes were objective, and knowledge of the assigned intervention was not likely to influence outcome assessment)*
Domain 5. Bias in selection of the reported result	Risk-of-bias judgement for selection of the reported result	Low *(There is clear evidence that all eligible reported results for the outcome correspond to all intended outcome measurements and analyses)*
Overall bias and Directness	Risk of bias judgement	Some concerns *(The study is judged to raise some concerns in at least one domain)*
Overall bias and Directness	Overall Directness	Directly applicable
Overall bias and Directness	Risk of bias variation across outcomes	None

RoB: risk of bias

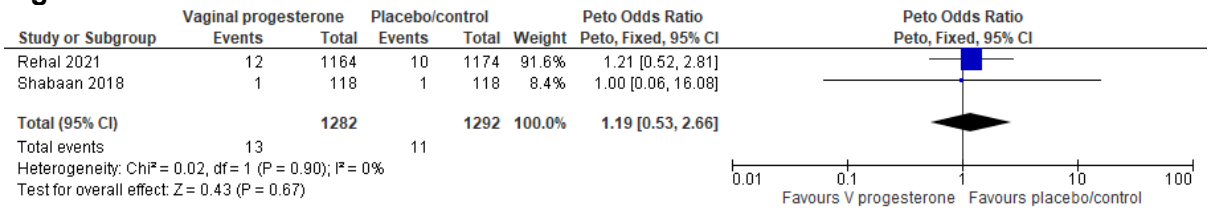
Appendix E Forest plots

Forest plots for review question: What is the clinical and cost-effectiveness of progesterone in preventing spontaneous preterm birth in twin and triplet pregnancy?

This section includes forest plots only for outcomes that are meta-analysed. Outcomes from single studies are not presented here; the quality assessment for such outcomes is provided in the GRADE profiles in appendix F.

Comparison: Vaginal progesterone (400 or 600 mg per day) versus placebo/control in all women with twin pregnancies (participants with short cervix/long cervix and with/without previous preterm birth) (study level analysis)

Figure 3: Stillbirth or neonatal death



CI: confidence interval

Appendix F GRADE tables

GRADE tables for review question: What is the clinical and cost-effectiveness of progesterone in preventing spontaneous preterm birth in twin and triplet pregnancy?

Table 6: Evidence profile for comparison: vaginal progesterone (400 or 600 mg per day) versus placebo/control in all women with twin pregnancies (participants with short cervix/long cervix and with/without previous preterm birth) (study level analysis)

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Vaginal progesterone	Placebo/control	Relative (95% CI)	Absolute		
Stillbirth or neonatal death (number of baby or fetus)												
2* (Rehal 2021, Shabaan 2018)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[1]	none	13/1282 (1%)	11/1292 (0.85%)	POR 1.19 (0.53 to 2.66)	2 more per 1000 (from 4 fewer to 14 more)	LOW	CRITICAL
Any birth between 24 weeks and <28 weeks (number of pregnant women)												
1 (Rehal 2021)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[2]	none	8/567 (1.4%)	9/561 (1.6%)	RR 0.88 (0.34 to 2.26)	2 fewer per 1000 (from 11 fewer to 20 more)	LOW	CRITICAL
Any birth between 24 weeks and <32 weeks (number of pregnant women)												
1 (Rehal 2021)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[2]	none	38/567 (6.7%)	39/561 (6.9%)	RR 0.96 (0.63 to 1.48)	3 fewer per 1000 (from 26 fewer to 33 more)	LOW	CRITICAL
Preterm birth at 28-30 weeks (number of pregnant women)												
1 (Shabaan 2018)	randomised trials	serious[3]	no serious inconsistency	no serious indirectness	very serious[2]	none	2/59 (3.4%)	2/59 (3.4%)	RR 1.00 (0.15 to 6.87)	0 fewer per 1000 (from 29 fewer to 199 more)	VERY LOW	CRITICAL
Any birth between 24 and <37 weeks (number of pregnant women)												

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Vaginal progesterone	Placebo/control	Relative (95% CI)	Absolute		
1 (Rehal 2021)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	315/567 (55.6%)	296/561 (52.8%)	RR 1.05 (0.95 to 1.17)	26 more per 1000 (from 26 fewer to 90 more)	HIGH	CRITICAL
Preterm birth at 32-36 weeks (number of pregnant women)												
1 (Shabaan 2018)	randomised trials	serious[3]	no serious inconsistency	no serious indirectness	very serious[2]	none	8/59 (13.6%)	13/59 (22%)	RR 0.62 (0.28 to 1.37)	84 fewer per 1000 (from 159 fewer to 82 more)	VERY LOW	CRITICAL
Spontaneous birth between 24 and <34 weeks (adjusted analysis) (number of pregnant women)												
1 (Rehal 2021)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[4]	none	56/541 (10.4%)	44/538 (8.2%)	OR 1.35 (0.88 to 2.05)	26 more per 1000 (from 9 fewer to 73 more)	MODERATE	CRITICAL
Composite of serious neonatal complications (neonatal morbidity, including necrotising enterocolitis, retinopathy of prematurity, intraventricular haemorrhage, sepsis, respiratory distress syndrome and anaemia) (number of baby or fetus) (follow-up not reported)												
1 (Rehal 2021)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[4]	none	77/1125 (6.8%)	85/1113 (7.6%)	RR 0.9 (0.67 to 1.21)	8 fewer per 1000 (from 25 fewer to 16 more)	MODERATE	IMPORTANT
Composite of serious neonatal complications (requirement of neonatal therapy, including admission to neonatal intensive care unit and requirement of ventilation) (number of baby or fetus) (follow-up not reported)												
1 (Rehal 2021)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	268/1125 (23.8%)	252/1113 (22.6%)	RR 1.05 (0.9 to 1.22)	11 more per 1000 (from 23 fewer to 50 more)	HIGH	IMPORTANT
Composite of adverse maternal outcomes (mother with at least 1 serious adverse event, including pre-eclampsia with prolonged hospital stay [PHS] [5 days], eclampsia with PHS [10 days], pulmonary embolism with PHS [4 days], postnatal liver rupture with PHS [31 days], obstetrical cholestasis with PHS [2 days], abnormal liver function tests, postpartum												

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Vaginal progesterone	Placebo/control	Relative (95% CI)	Absolute		
haemorrhage 3-L blood loss, gastritis with PHS [4 days], dyspnoea with PHS [2 days], restrictive cardiomyopathy with PHS [4 days], urinary tract infection with PHS [3 days], and maternal mirror syndrome in association with fetal hydrops) (number of pregnant women) (follow-up up to 37 weeks of gestation or 30 days after the last capsule was taken)												
1 (Rehal 2021)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[2]	none	8/582 (1.4%)	7/587 (1.2%)	RR 1.15 (0.42 to 3.16)	2 more per 1000 (from 7 fewer to 26 more)	LOW	IMPORTANT
Composite of adverse maternal outcomes (mother with at least 1 nonserious adverse event, including vaginal discharge, vaginal itching, vaginal pain or discomfort, vaginal bleeding, headache/dizziness, fatigue, depression, insomnia, nausea/vomiting, abdominal pain/discomfort, diarrhoea/constipation, joint pain, swelling of extremities, palpitation, itching/skin rash, urinary tract infection, and other adverse events) (number of pregnant women) (follow-up up to 37 weeks of gestation or 30 days after the last capsule was taken)												
1 (Rehal 2021)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[4]	none	200/596 (33.6%)	186/598 (31.1%)	RR 1.08 (0.92 to 1.27)	25 more per 1000 (from 25 fewer to 84 more)	MODERATE	IMPORTANT

CI: confidence interval; OR: odds ratio; PHS: prolonged hospital stay; POR: Peto odds ratio; RR: risk ratio
Information on cervical length and previous preterm birth was not reported in Shabaan 2018
**See corresponding forest plot*
[1] <150 events
[2] 95% CI crosses 2 MIDs
[3] Serious risk of bias in the evidence contributing to the outcomes as per RoB 2
[4] 95% CI crosses 1 MID

Table 7: Evidence profile for comparison: vaginal progesterone (600 mg per day) versus placebo in twin pregnancies (participants with long cervix ≥30 mm) (study level analysis)

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Vaginal progesterone	Placebo	Relative (95% CI)	Absolute		
Spontaneous birth between 24 and <34 weeks (number of pregnant women)												

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Vaginal progesterone	Placebo	Relative (95% CI)	Absolute		
1 (Rehal 2021)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	None	48/466 (10.3%)	32/477 (6.7%)	RR 1.54 (1.0001 to 2.36)	36 more per 1000 (from 0 more to 91 more)	MODERATE	CRITICAL

CI: confidence interval; RR: risk ratio
[1] 95% CI crosses 1 MID

Table 8: Evidence profile for comparison: vaginal progesterone (600 mg per day) versus placebo in twin pregnancies (participants with short cervix/long cervix and with previous preterm birth) (study level analysis)

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Vaginal progesterone	Placebo	Relative (95% CI)	Absolute		
Spontaneous birth between 24 and <34 weeks in women with previous preterm birth (number of pregnant women)												
1 (Rehal 2021)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[1]	None	5/22 (22.7%)	7/28 (25%)	RR 0.91 (0.33 to 2.48)	22 fewer per 1000 (from 167 fewer to 370 more)	LOW	CRITICAL

CI: confidence interval; RR: risk ratio
[1] 95% CI crosses 2 MIDs

Table 9: Evidence profile for comparison: vaginal progesterone (600 mg per day) versus placebo in twin pregnancies (participants with short cervix/long cervix and with no previous preterm birth) (study level analysis)

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Vaginal progesterone	Placebo	Relative (95% CI)	Absolute		
Spontaneous birth between 24 and <34 weeks in women with previous term birth (number of pregnant women)												
1 (Rehal 2021)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	21/228 (9.2%)	10/213 (4.7%)	RR 1.96 (0.95 to 4.07)	45 more per 1000 (from 2 fewer to 144 more)	MODERATE	CRITICAL

CI: confidence interval; RR: risk ratio
[1] 95% CI crosses 1 MID

Table 10: Evidence profile for comparison: vaginal progesterone (90-400 mg per day) versus control in all women with twin pregnancies (participants with short cervix/long cervix and with/without previous preterm birth) (IPD analysis)

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Vaginal progesterone	Control	Relative (95% CI)	Absolute		
Stillbirth (fetal death/stillbirth) (pregnant women)												
1 (EPPPIC Group 2021)	IPD review	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	None	4103		RR 1.19 (0.67 to 2.1)	170 more per 1000 (from 400 fewer to 750 more)[2]	MODERATE	CRITICAL
Neonatal death (neonatal death after live birth) (pregnant women)												
1 (EPPPIC Group 2021)	IPD review	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	None	4035		RR 1.15 (0.66 to 1.98)	140 more per 1000 (from 420 fewer to 700 more)[2]	MODERATE	CRITICAL
Preterm birth <28 weeks (pregnant women)												

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Vaginal progesterone	Control	Relative (95% CI)	Absolute		
1 (EPPPIC Group 2021)	IPD review	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[3]	None	2046		RR 1.22 (0.77 to 1.94)	200 more per 1000 (from 260 fewer to 660 more)[2]	LOW	CRITICAL
Preterm birth <34 weeks (pregnant women)												
1 (EPPPIC Group 2021)	IPD review	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	None	2046		RR 1.01 (0.84 to 1.2)	10 more per 1000 (from 170 fewer to 190 more)[2]	HIGH	CRITICAL
Composite of serious neonatal complications (serious neonatal complications, including severe necrotising enterocolitis stage 2-3, intraventricular haemorrhage grades 3-4, retinopathy of prematurity stage 3 or worse, bronchopulmonary dysplasia, confirmed sepsis, patent ductus arteriosus, and neonatal infection) (pregnant women) (follow-up not reported)												
1 (EPPPIC Group 2021)	IPD review	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[4]	None	3840		RR 0.94 (0.74 to 1.2)	60 fewer per 1000 (from 300 fewer to 180 more)[2]	MODERATE	IMPORTANT
Composite of adverse maternal outcomes (maternal complications, including gestational hypertension, pre-eclampsia, gestational diabetes, and maternal infection such as chorioamnionitis) (pregnant women) (follow-up not reported)												
1 (EPPPIC Group 2021)	IPD review	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[4]	None	1938		RR 0.93 (0.73 to 1.17)	70 fewer per 1000 (from 310 fewer to 170 more)[2]	MODERATE	IMPORTANT

CI: confidence interval; IPD: individual participant data; OIS: optimal information size; RR: risk ratio
N=841 with unknown cervical length and N=338 with unknown history of previous preterm birth were also included in vaginal progesterone trials included in EPPPIC Group 2021
[1] Estimate may be imprecise as cannot determine if OIS criteria have been met because data on the number of events is not reported
[2] Risk difference was used to calculate absolute effect
[3] 95% CI crosses 2 MIDs
[4] 95% CI crosses 1 MID

Table 11: Evidence profile for comparison: vaginal progesterone (100-600 mg per day) versus placebo in twin pregnancies (participants with short cervix ≤25 mm) (IPD analysis)

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Vaginal progesterone	Placebo	Relative (95% CI)	Absolute		
Stillbirth (fetal death, adjusted analysis) (baby or fetus)												
1 (Conde-Agudelo 2022)	IPD review	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[1]	None	6/104 (5.8%)	4/86 (4.7%)	RR 0.54 (0.17 to 1.77)	21 fewer per 1000 (from 39 fewer to 36 more)	LOW	CRITICAL
Neonatal death (adjusted analysis) (baby or fetus)												
1 (Conde-Agudelo 2022)	IPD review	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[1]	None	4/104 (3.8%)	9/86 (10.5%)	RR 0.51 (0.2 to 1.28)	51 fewer per 1000 (from 84 fewer to 29 more)	LOW	CRITICAL
Preterm birth <28 weeks (pregnant women)												
1 (Conde-Agudelo 2022)	IPD review	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[2]	None	7/52 (13.5%)	11/43 (25.6%)	RR 0.41 (0.19 to 0.91)	151 fewer per 1000 (from 23 fewer to 207 fewer)	MODERATE	CRITICAL
Preterm birth <32 weeks (pregnant women)												
1 (Conde-Agudelo 2022)	IPD review	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[2]	None	16/52 (30.8%)	20/43 (46.5%)	RR 0.56 (0.33 to 0.93)	205 fewer per 1000 (from 33 fewer to 312 fewer)	MODERATE	CRITICAL
Preterm birth <37 weeks (pregnant women)												
1 (Conde-Agudelo 2022)	IPD review	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[2]	None	43/52 (82.7%)	38/43 (88.4%)	RR 0.91 (0.75 to 1.1)	80 fewer per 1000 (from 221 fewer to 88 more)	MODERATE	CRITICAL
Spontaneous preterm birth <34 weeks (pregnant women)												

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Vaginal progesterone	Placebo	Relative (95% CI)	Absolute		
1 (Conde-Agudelo 2022)	IPD review	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[2]	None	20/52 (38.5%)	28/43 (65.1%)	RR 0.58 (0.38 to 0.89)	273 fewer per 1000 (from 72 fewer to 404 fewer)	MODERATE	CRITICAL
Composite of serious neonatal complications (composite of neonatal morbidity/mortality, including respiratory distress syndrome, intraventricular haemorrhage, necrotising enterocolitis, proven neonatal sepsis, or neonatal death; adjusted analysis) (baby or fetus) (follow-up not reported)												
1 (Conde-Agudelo 2022)	IPD review	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[2]	None	24/102 (23.5%)	31/84 (36.9%)	RR 0.59 (0.33 to 0.98)	151 fewer per 1000 (from 7 fewer to 247 fewer)	MODERATE	IMPORTANT

CI: confidence interval; IPD: individual participant data; OIS: optimal information size; RR: risk ratio
[1] <150 events
[2] 95% CI crosses 1 MID

Table 12: Evidence profile for comparison: intramuscular 17-hydroxyprogesterone caproate (17-OHPC; 250 mg weekly or 500 mg twice weekly) versus control in all women with twin and triplet pregnancies (participants with short cervix/long cervix and with/without previous preterm birth) (IPD analysis)

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	IM 17-OHPC	Control	Relative (95% CI)	Absolute		
Stillbirth (fetal death/stillbirth) (pregnant women/infants)												
1 (EPPPIC Group 2021)	IPD review	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	4744		RR 1 (0.67 to 1.5)	0 more per 1000 (from 400 fewer to 400 more)[2]	MODERATE	CRITICAL
Neonatal death (neonatal death after live birth) (pregnant women/infants)												

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	IM 17-OHPC	Control	Relative (95% CI)	Absolute		
1 (EPPPIC Group 2021)	IPD review	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	4632		RR 1.13 (0.76 to 1.68)	12 more per 1000 (from 270 fewer to 520 more)[2]	MODERATE	CRITICAL
Preterm birth <28 weeks (pregnant women)												
1 (EPPPIC Group 2021)	IPD review	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[3]	none	2253		RR 1.07 (0.78 to 1.46)	70 more per 1000 (from 250 fewer to 380 more)[2]	LOW	CRITICAL
Preterm birth <34 weeks (pregnant women)												
1 (EPPPIC Group 2021)	IPD review	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	2253		RR 1.04 (0.92 to 1.18)	40 more per 1000 (from 80 fewer to 160 more)[2]	HIGH	CRITICAL
Composite of serious neonatal complications (serious neonatal complications, including severe necrotising enterocolitis stage 2-3, intraventricular haemorrhage grades 3-4, retinopathy of prematurity stage 3 or worse, bronchopulmonary dysplasia, confirmed sepsis, patent ductus arteriosus, and neonatal infection) (pregnant women) (follow-up not reported)												
1 (EPPPIC Group 2021)	IPD review	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[3]	none	4724		RR 1.12 (0.76 to 1.65)	110 more per 1000 (from 270 fewer to 500 more)[2]	LOW	IMPORTANT
Composite of adverse maternal outcomes (maternal complications, including gestational hypertension, pre-eclampsia, gestational diabetes, and maternal infection such as chorioamnionitis) (pregnant women) (follow-up not reported)												
1 (EPPPIC Group 2021)	IPD review	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[4]	none	2095		RR 1.09 (0.94 to 1.27)	90 more per 1000 (from 60 fewer to 230 more)[2]	MODERATE	IMPORTANT

CI: confidence interval; IPD: individual participant data; NC: not calculated; OIS: optimal information size; RR: risk ratio
N=922 with unknown cervical length and N=5 with unknown history of previous preterm birth were also included in intramuscular 17-OHPC trials included in EPPPIC Group 2021
[1] Estimate may be imprecise as cannot determine if OIS criteria have been met because data on the number of events is not reported
[2] Risk difference was used to calculate absolute effect
[3] 95% CI crosses 2 MIDs

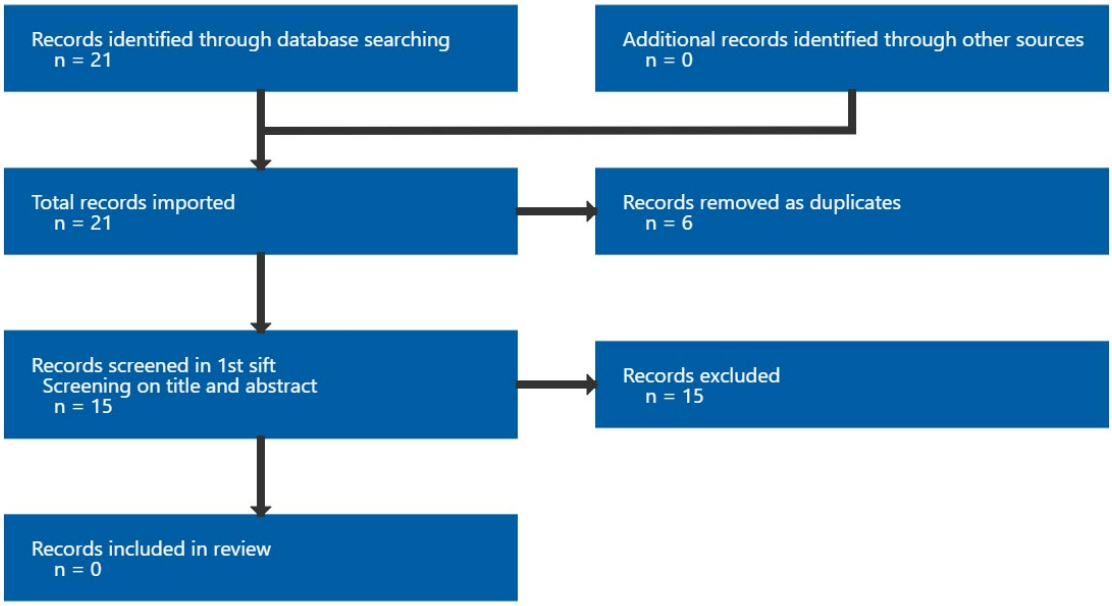
[4] 95% CI crosses 1 MID

Appendix G Economic evidence study selection

Study selection for: What is the clinical and cost-effectiveness of progesterone in preventing spontaneous preterm birth in twin and triplet pregnancy?

No economic evidence was identified which was applicable to this review question.

Figure 4: Study selection flowchart



Appendix H Economic evidence tables

Economic evidence tables for review question: What is the clinical and cost-effectiveness of progesterone in preventing spontaneous preterm birth in twin and triplet pregnancy?

No evidence was identified which was applicable to this review question.

Appendix I Economic model

Economic model for review question: What is the clinical and cost-effectiveness of progesterone in preventing spontaneous preterm birth in twin and triplet pregnancy?

Modelling cost utility of cervical length screening and vaginal progesterone treatment to prevent spontaneous preterm birth in twin pregnancies

Introduction

Twin pregnancies are associated with increased perinatal morbidity and mortality. Spontaneous preterm birth is the major contributor to these adverse outcomes. If women with twin pregnancies at higher risk of preterm birth could be identified and an effective intervention could be used to delay or prevent preterm birth, with resultant reductions in the associated adverse events, this could be cost-effective to the NHS due to the high costs of neonatal care for premature infants and by improving the survival rates and long-term health of infants from multiple pregnancies.

In the previous NICE guideline, Multiple pregnancy: antenatal care for twin and triplet pregnancies (NG137) a model was developed to assess the cost-effectiveness of cervical screening followed by treatment with vaginal progesterone in those identified with a short cervix. Although the model suggested that screening and treatment was cost-effective, the committee chose not to make a recommendation as they were aware of emerging evidence in the relevant population which could potentially have changed best estimates of treatment effectiveness.

However, as new evidence has now been published, the decision was taken to update the NICE guideline and the associated health economic model.

Methods

Setting and population

The model setting was for the NHS and the population was pregnant women with a twin pregnancy (there was insufficient evidence for triplet pregnancies to be included in the analysis). The time horizon was largely focused on the pregnancy and neonatal period, but a 2-year horizon was adopted for post-discharge NHS costs and a lifetime horizon was adopted with respect to mortality and lifelong morbidity of the babies arising from adverse health outcomes, such as cerebral palsy.

Model structure

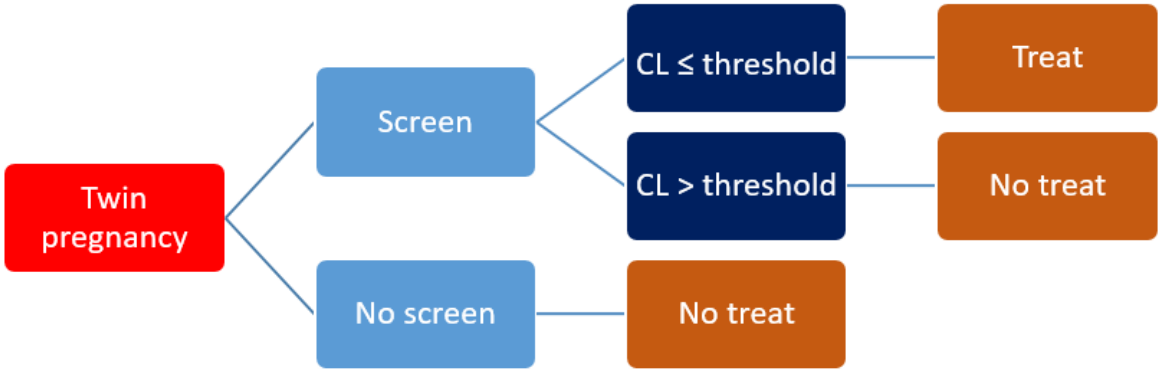
A cohort Markov decision analytic model was developed in Microsoft Excel® to evaluate the cost effectiveness of cervical length screening and vaginal progesterone treatment to prevent or delay spontaneous preterm birth in twin pregnancies, reflecting the new clinical evidence identified for this guideline on effective screening and treatment to prevent preterm birth. In the intervention women identified with a short cervix (25 mm or less) at screening would receive vaginal progesterone daily until 34 weeks or birth if earlier. Trial evidence reviewed for this guideline did not show any treatment benefit for women with a cervical length of

greater than 25 mm. Intervention was compared to a strategy of no screening and no treatment.

Screening and the starting point of progesterone therapy (if indicated) was assumed to take place by gestational age of 20 weeks to reflect that the clinical evidence suggested a treatment benefit if treatment was started by this time.

A schematic illustrating how screening was used as the basis for treatment is depicted in Figure 5.

Figure 5: Schematic chart to illustrate decision analytic approach of screening by cervical length to identify women at higher risk of preterm birth for treatment

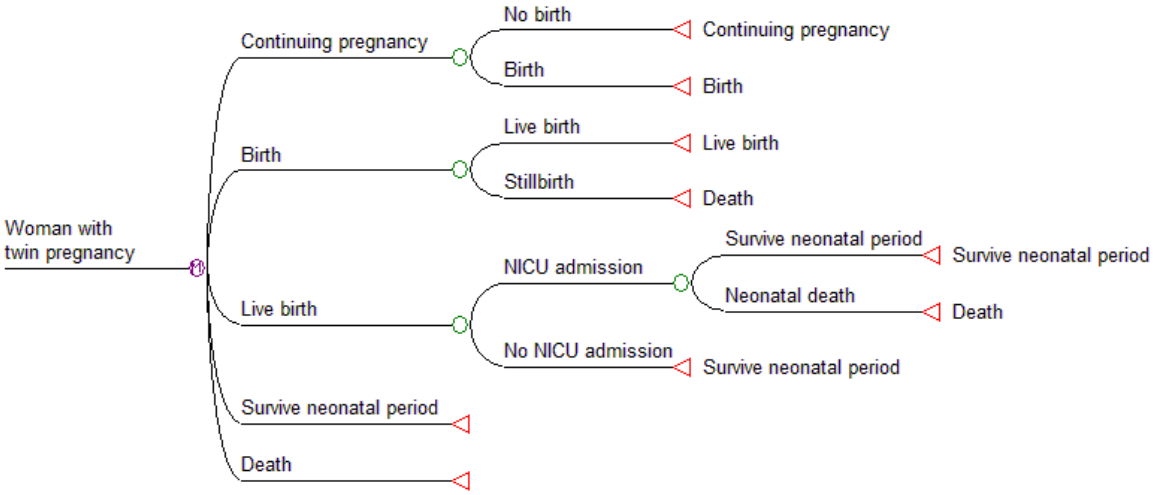


CL = cervical length

From a gestational age of 24 weeks to 37 weeks a Markov approach was used to model the impact of vaginal progesterone on the timing of birth and neonatal outcomes linked to prematurity and this is shown in Figure 6. The model assumes that all twin pregnancies will have resulted in birth by a gestational age of 37 weeks.

Pregnant women with twins enter the model in a health state of 'continuing pregnancy' but for each week of gestation, the Markov cycle duration, they can transition to the state of 'birth'. This Markov process serves as the 'birth engine' in the model with the transition probabilities dependant on gestational age, the distribution of cervical length across the model population, the probability of preterm birth at each gestational age by cervical length, the screening strategy and the effectiveness of treatment to prevent preterm birth in the women identified for treatment by screening. Figure 6 also highlights the health state transitions from 'birth' which are used to quantify the probability of various adverse neonatal outcomes, with these probabilities being tied with gestational age at birth.

Figure 6: Schematic to illustrate Markov/decision approach across pregnancy and the neonatal period



Note the outcome of stillbirth is only included as a sensitivity analysis as the committee did not think that the association between gestational age and stillbirth was causal

For babies surviving the neonatal period a basic decision analytic approach was used to assess the impact of longer-term morbidity on health-related quality of life and "downstream" costs. Figure 7, based on a published UK study (Khan, 2015) on costs associated with moderate and late preterm birth, shows the decision tree used to estimate the costs incurred by the NHS in the 2 years post-discharge for babies surviving the neonatal period. Figure 8Figure 8 illustrates the decision analytic structure used to analyse the longer-term impact of morbidity arising from adverse neonatal outcomes.

Figure 7: Schematic to illustrate decision analytic approach used to estimate costs incurred by the NHS in the 2 years from initial discharge from hospital

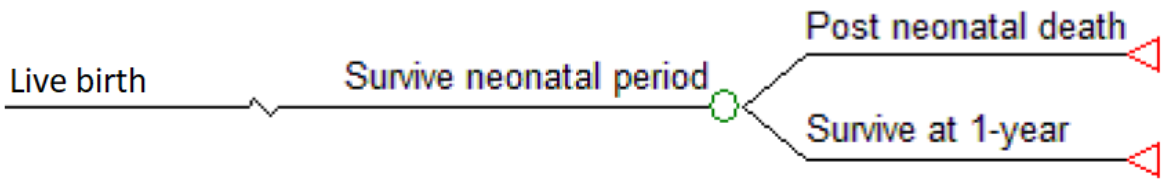
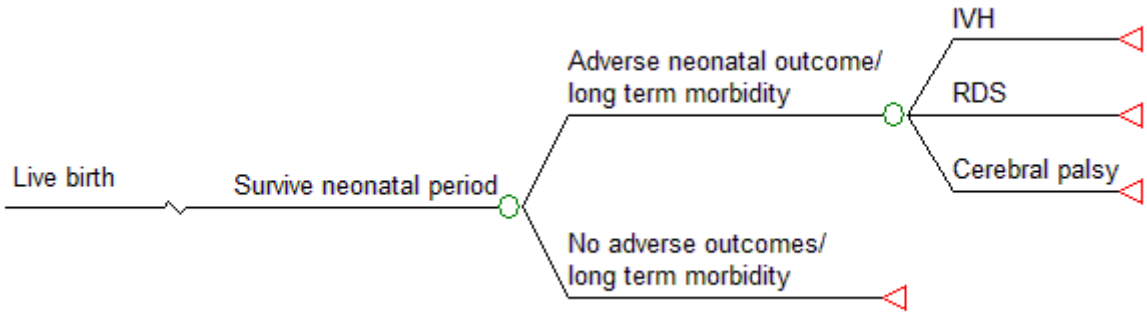


Figure 8: Schematic to illustrate decision analytic approach to long term morbidity arising from adverse neonatal outcomes



IVH = intraventricular haemorrhage; RDS = respiratory distress syndrome

Distribution of cervical length in twin pregnancies

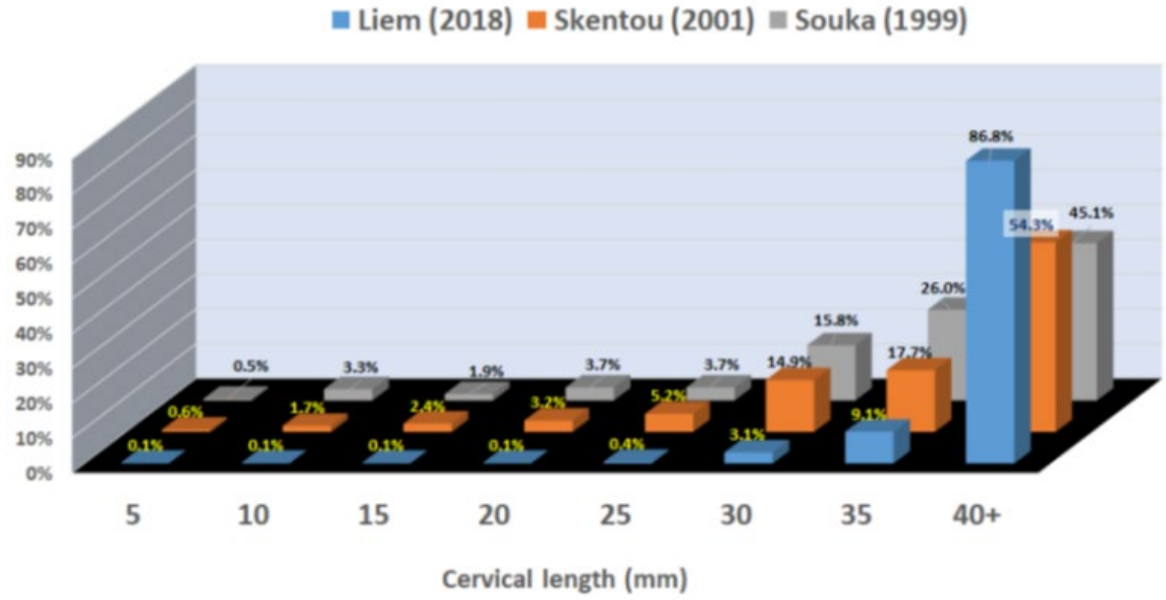
The model required that the distribution of cervical length be estimated across the population, women with a twin pregnancy, at a gestational age of 20 weeks, the time of screening. This determined the proportion of women who would receive treatment as a result of screening. It was possible to populate the model with any one of 3 distributions of cervical length, all estimated either from personal communication (e-mail correspondence between the guideline topic advisor on NG137 and Dr Sophie Liem, an obstetrician, who has published on cervical length distribution or the published literature). These distributions are summarised in Table 13 and Figure 9 below. Liem (2018) was used for the distribution of cervical length in the base case analysis as the timing of measurement of cervical length was most closely aligned with the recommended gestational age that screening would occur, which reflected the clinical evidence that all screening in the included studies occurred before 20 weeks. It was also the dataset with the largest sample size and the most recent. In addition, it was also least favourable to intervention as less women would be identified for treatment and is therefore the most conservative distribution.

Table 13: Distribution of cervical length in twin pregnancies at approximate gestational age of screening[a]

Cervical length	Liem 2018[b]	Skentou 2001[c]	Souka 1999[d]
5 mm	0.14%	0.65%	0.47%
10 mm	0.14%	1.72%	3.26%
15 mm	0.14%	2.37%	1.86%
20 mm	0.14%	3.23%	3.72%
25 mm	0.43%	5.17%	3.72%
30 mm	3.13.%	14.87%	15.81%
35 mm	9.09%	17.67%	26.05%
≥40 mm	86.79%	54.31%	45.12%

(a) All percentages had to be estimated from charts
(b) Based on a gestational age of 18-22 weeks
(c) Based on a gestational age of 23 weeks
(d) Based on a gestational age of 23 weeks

Figure 9: Chart to show distributions of cervical length in twin pregnancies at the time of screening used in the model



For probabilistic sensitivity analysis cervical length distributions were sampled using a Dirichlet distribution. A count was made for each distribution using the observed frequencies reported in Table 14 and sampled using a cumulative gamma function. The sampled cervical length proportion was calculated as its sample count ÷ sum of the sample count for all cervical length categories.

Table 14: Distribution of cervical length in twin pregnancies at approximate gestational age of screening[a]

Cervical length	Liem 2018[b]	Skentou 2001[c]	Souka 1999[d]
5 mm	1	3	1
10 mm	1	8	7
15 mm	1	11	4
20 mm	1	15	8
25 mm	3	24	8
30 mm	22	69	34
35 mm	64	82	56
≥40 mm	611	252	97
Count	**704**	**464**	**215**

Clinical outcomes

Women with twin pregnancies are at higher risk of preterm birth than women with singleton pregnancies and most of the excess morbidity and mortality of a twin pregnancy arise from this increased rate of premature birth. Delay or prevention of spontaneous preterm birth improves outcomes for babies by mitigating the adverse impact of prematurity. Therefore, the clinical outcomes assessed in the model, listed below, are important outcomes for babies related to preterm birth.

- Neonatal death

- Post neonatal death
- Neonatal care unit admission
- Cerebral palsy
- Intraventricular haemorrhage
- Respiratory distress syndrome
- *Stillbirth (sensitivity analysis only)*

These outcomes all have a potentially large impact on health-related quality of life and/or NHS costs.

Baseline

The "birth engine"

The model 'birth engine' represents the Markov process used to estimate the cumulative rate of twin births by gestational age as women in the model cohort, transition from a health state of 'continuing pregnancy' to a state of 'birth' over gestational ages 24-37 weeks. The clinical evidence considered in this review suggested that treatment is only effective in women with a short cervix (≤ 25 mm). Therefore, screening is required to identify the women to treat based on cervical length and it was important to model the twin birth rate by gestational age, according to cervical length measurement at the time of screening. A published paper (Kindinger, 2016) allowed these estimates to be made with this data summarised in Table 15.

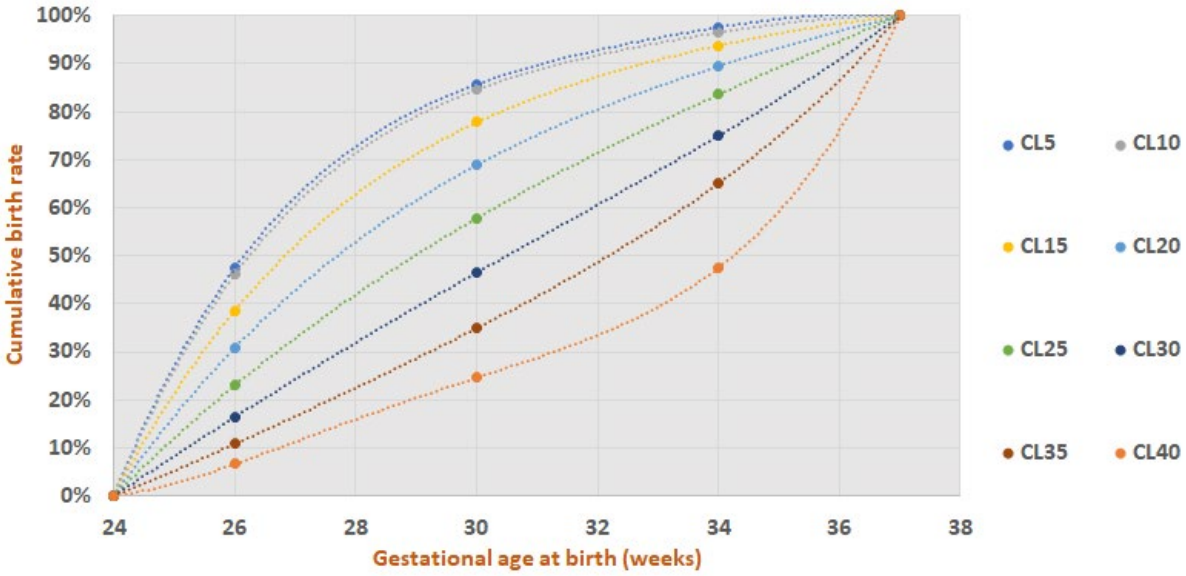
Table 15: Proportion of spontaneous birth by gestational age and cervical length[a] (cumulative in brackets)

	Gestational age			
	<28 weeks	28-32 weeks	32-36 weeks	37 weeks
5 mm	0.532 (0.532)	0.364 (0.896)	0.085 (0.981)	0.019 (1.000)
10 mm	0.461 (0.461)	0.385 (0.846)	0.119 (0.965)	0.035 (1.000)
15 mm	0.386 (0.386)	0.392 (0.778)	0.160 (0.938)	0.062 (1.000)
20 mm	0.308 (0.308)	0.381 (0.689)	0.206 (0.895)	0.105 (1.000)
25 mm	0.232 (0.232)	0.350 (0.582)	0.250 (0.832)	0.168 (1.000)

(a) Table data based on predicted probabilities reported in a published study (Kindinger, 2016) at a gestational age of 20 weeks

To estimate a weekly probability for the transition from the 'continuing pregnancy' to a state of 'birth' a best fit curve was estimated for the cumulative birth rate from gestational ages 24-37 weeks as illustrated in Figure 10 below.

Figure 10: Graph to show estimated cumulative birth rates by gestational age and cervical length at screening (e.g., 22 weeks)



CL = cervical length

It should be noted that women without a short cervix (> 25 mm) are not treated and therefore, while they are included in the analysis because of the screening costs they incur, treatment has no impact on their cumulative birth rates.

Kindinger (2016) also predicted birth probabilities by gestational age for screening results at 18 and 22 weeks as shown in Table 16 and Table 17 respectively. These were included as a sensitivity analysis.

Table 16: Proportion of spontaneous birth by gestational age and cervical length (cumulative in brackets) – screening at 18 weeks

	Gestational age			
	<28 weeks	28-32 weeks	32-36 weeks	37 weeks
5 mm	0.585 (0.585)	0.341 (0.926)	0.059 (0.985)	0.015 (1.000)
10 mm	0.518 (0.518)	0.368 (0.886)	0.085 (0.971)	0.029 (1.000)
15 mm	0.445 (0.445)	0.385 (0.830)	0.117 (0.947)	0.053 (1.000)
20 mm	0.366 (0.366)	0.386 (0.752)	0.156 (0.908)	0.092 (1.000)
25 mm	0.286 (0.286)	0.367 (0.653)	0.196 (0.849)	0.152 (1.000)

Table 17: Proportion of spontaneous birth by gestational age and cervical length[a] (cumulative in brackets) – screening at 22 weeks

	Gestational age			
	<28 weeks	28-32 weeks	32-36 weeks	37 weeks
5 mm	0.475 (0.475)	0.382 (0.858)	0.119 (0.977)	0.023 (1.000)
10 mm	0.402 (0.402)	0.394 (0.796)	0.163 (0.959)	0.042 (1.000)
15 mm	0.326 (0.326)	0.389 (0.715)	0.213 (0.928)	0.072 (1.000)
20 mm	0.252 (0.252)	0.366 (0.618)	0.265 (0.883)	0.118 (1.000)
25 mm	0.183 (0.183)	0.324 (0.507)	0.311 (0.818)	0.181 (1.000)

Neonatal deaths by gestational age

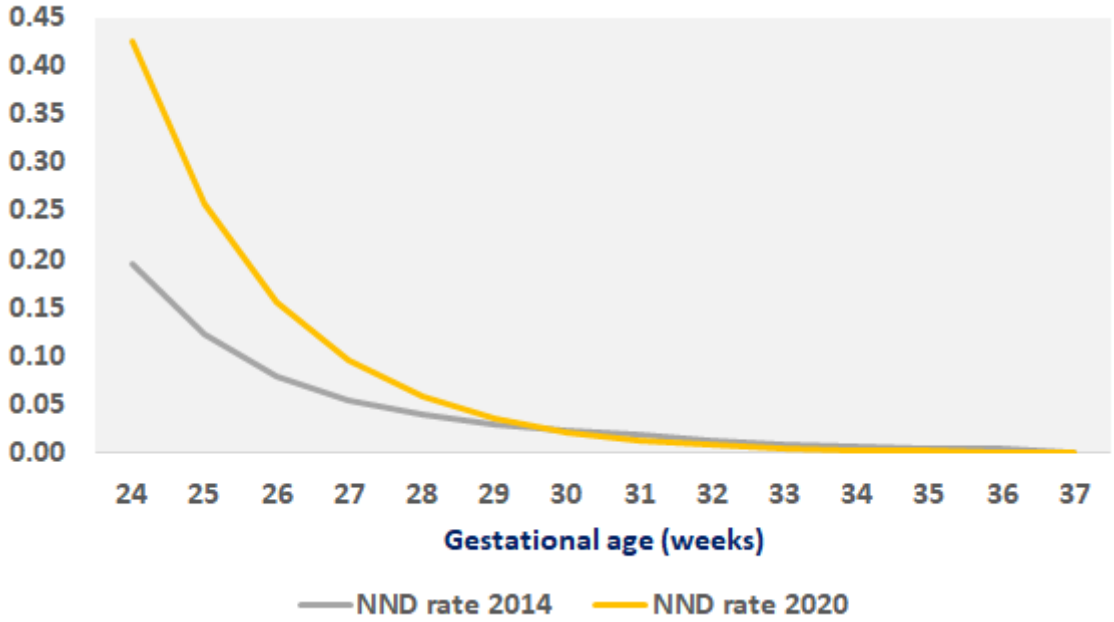
The neonatal and death rate by gestational age at birth was derived from official statistics for England and Wales in 2020 (ONS, 2023) which are summarised in Table 18 below.

Table 18: England and Wales data on twin birth and neonatal deaths by gestational age (2020)

Gestational age	Twin births	Neonatal deaths	Neonatal death rate
<24 weeks	107	81	0.757
24-27 weeks	373	52	0.139
28-31 weeks	1,180	22	0.019
32-36 weeks	8,888	20	0.0023
≥37 weeks	6,769	6	0.0009

To achieve greater granularity, a best fit curve was fitted to this data, shown in Figure 11 (NND rate 2020), to provide an estimate of neonatal mortality for each week of gestational age at birth. The chart also shows an alternative estimate for neonatal mortality by gestational age at birth derived from previous 2014 data (ONS, 2017) which was used as part of a sensitivity analysis.

Figure 11: Estimate of neonatal mortality by gestational age at birth



NND = Neonatal death
Estimate of 2020 NND: NND = 69228.92 x EXP(-0.5x) where x is gestational age in weeks

Post neonatal deaths by gestational age

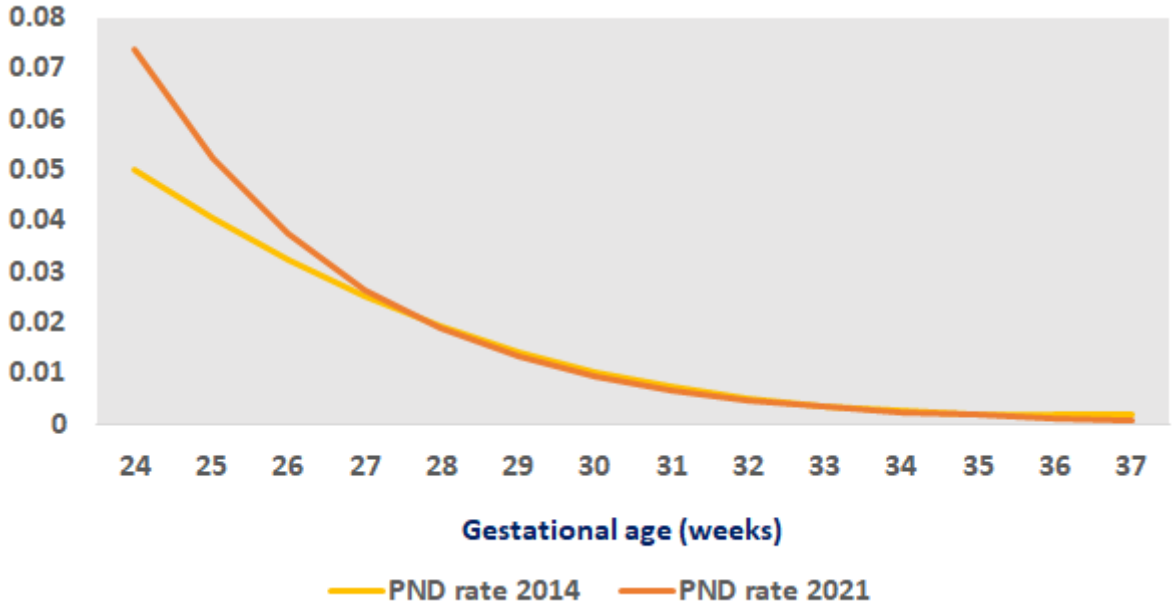
The post neonatal death rate (covering the period 28 days of life to 365 days of life) by gestational age at birth was derived from official statistics for England and Wales in 2021 (ONS, 2023) which are summarised in Table 19 below. This data is based on all pregnancies and the model assumes that twin birth does not affect post neonatal death independently of gestational age. A best fit curve was then applied to this data as depicted in Figure 12 to provide an estimate of the post neonatal death rate each week of gestational age at birth. Also, shown in best fit data for England and Wales in 2014 (ONS, 2017) which was used for sensitivity analysis.

Table 19: England and Wales data on twin birth and post neonatal deaths by gestational age (2021)[a]

Gestational age	Total births	Post neonatal deaths	Post neonatal death rate
24 weeks[b]	796	56	0.070
25 weeks	393	31	0.079
26 weeks	587	29	0.049
27 weeks	675	25	0.037
28 weeks	817	14	0.017
29 weeks	946	7	0.007
30 weeks	1,285	7	0.005
31 weeks	1,650	6	0.004
32 weeks	2,514	10	0.004
33 weeks	3,367	15	0.004
34 weeks	5,823	19	0.003
35 weeks	8,566	31	0.004
36 weeks	19,307	35	0.002
37 weeks[c]	563,377	296	0.001

(a) ONS (2023) - Child mortality (death cohort) tables in England and Wales
(b) Includes births and deaths for gestational age at birth of less than 24 weeks
(c) Includes births and deaths for gestational age at birth of greater than 36 weeks

Figure 12: Estimate of post neonatal mortality by gestational age at birth



PND = Post Neonatal death
Estimate of 2021 PND: PND = 257.72 x EXP(-0.34x) where x is gestational age in weeks

Neonatal intensive care unit admission by gestational age

Expert committee opinion was that all babies born at a gestational age of 32 weeks or less would be admitted to neonatal care.

To estimate the proportion admitted to neonatal care for births with a gestational age of greater than 32 weeks, first an estimate of the cumulative neonatal care admission rate by gestation was made using data from the Neonatal Data Analysis Unit (2017) in Table 20.

Table 20: UK neonatal care admissions in 2016

Gestational age at birth	Total neonatal care admissions	Proportion of all neonatal care admissions	Cumulative proportion of all neonatal care admissions
≤25 weeks	1,189	0.012	0.012
26 to 32 weeks	10,283	0.102	0.114
33 to 36 weeks	26,758	0.263	0.377
≥ 37 weeks	62,427	0.624	1.000

A best fit curve was applied to this data (see Figure 13) and the equation for this curve was used to estimate the cumulative neonatal care admissions for each week of gestation up to 37 weeks (see Table 21).

Figure 13: Estimated cumulative frequency of neonatal care admission by gestational age at birth

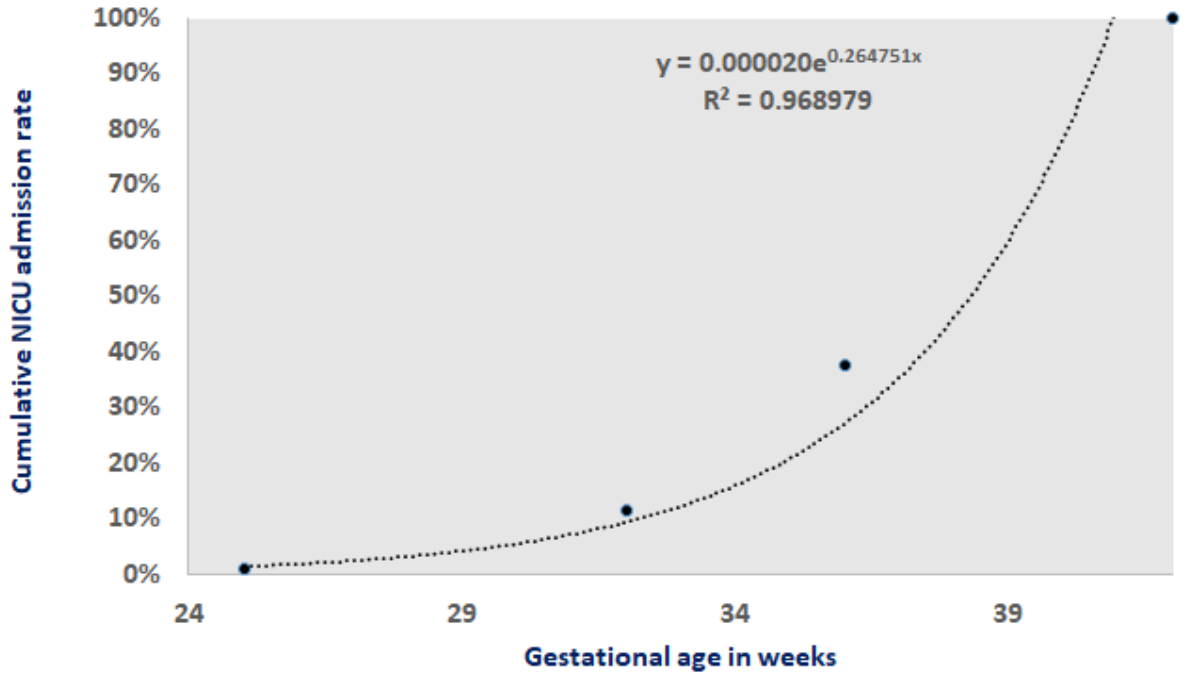


Table 21: Observed and fitted cumulative frequency distribution of neonatal care admission by gestational age at birth

Gestational age	Observed cumulative frequency	Estimated cumulative frequency
24 weeks	-	0.011
25 weeks	0.012	0.015
26 weeks	-	0.019
27 weeks	-	0.025
28 weeks	-	0.033
29 weeks	-	0.043
30 weeks	-	0.056

Gestational age	Observed cumulative frequency	Estimated cumulative frequency
31 weeks	0.114	0.072
32 weeks	-	0.094
33 weeks	-	0.123
34 weeks	-	0.160
35 weeks	-	0.209
36 weeks	0.377	0.272
37 weeks	-	0.355

The fitted cumulative frequency of neonatal care admission by gestational age reported in Table 21 were then used to estimate the total number of neonatal care admissions at each week of gestation. Then using data on the number of births at each week of gestation it was possible to derive an estimate of the proportion of birth that would be admitted to neonatal care between 33-37 weeks. Total neonatal care admission was taken from the data given in Table 20. Table 22 illustrates how the proportion of births admitted to neonatal care at each week of gestation is calculated.

Table 22: Neonatal care admission rates

Gestational age	Estimated cumulative frequency	Estimated cumulative admissions to neonatal care[a]	Estimated neonatal care admissions per week of gestation	Live births[b]	Neonatal care admission rate
33 weeks	0.123	11,585	2,695	3,755	0.72
34 weeks	0.160	15,097	3,512	6,495	0.54
35 weeks	0.209	19,673	4,576	9,555	0.48
36 weeks	0.272	25,636	5,963	21,537	0.28
37 weeks	0.355	33,406	7,770	61,019	0.13

(a) Based on an estimated total number of neonatal care admissions for 2016 in England and Wales of 94,192. This was derived from NDAU (2017) which reported a total of 100,762 neonatal admissions in England, Wales and Scotland for that year and that England and Wales accounted for 93.6% of births across the 3 jurisdictions

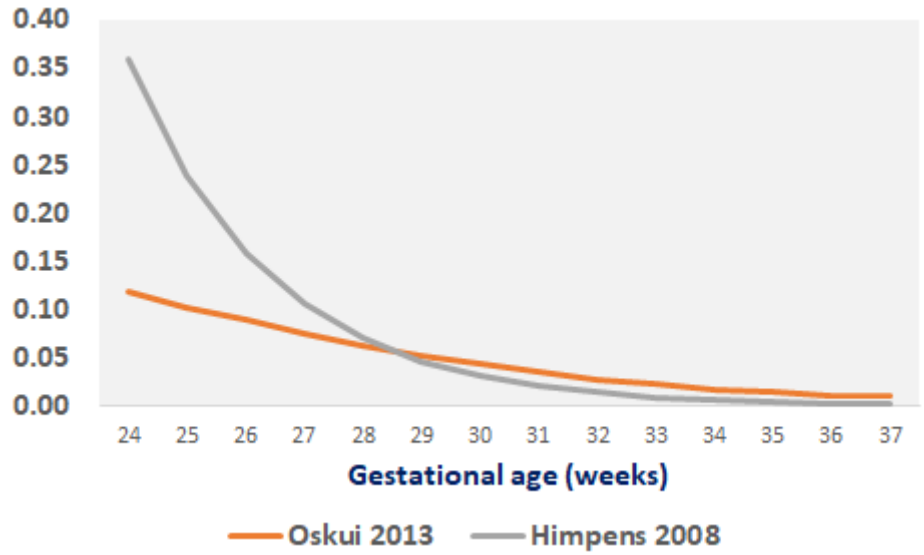
(b) Live birth rates for England and Wales in 2016 were estimated using data from 2021 for live births in England and Wales (ONS, 2023). For each gestational age the proportion of births at each gestational age was calculated and then multiplied by the total live births in England and Wales in 2016 which was 696,271 (https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/livebirths/bulletins/birthsummarytablesenglandandwales/2016)

As with other outcomes, we made the simplifying assumption for gestational ages at birth of 33-37 weeks that there was no independent effect of multiplicity on neonatal care admission over and above that of gestational age at birth.

Cerebral palsy by gestational age

A published systematic review and meta-analysis (Oskui, 2013) was used to estimate the risk of cerebral palsy by gestational age at birth in babies that survive the neonatal period. A curve was fitted to the observed data, as displayed in Figure 14. This was chosen for the base case analysis as it was more recent than an older meta-analysis (Himpens, 2008) which was used as a sensitivity analysis with the curve fitted to the observed data in that study also presented below.

Figure 14: Graph to show estimated risk of cerebral palsy by gestational age at birth



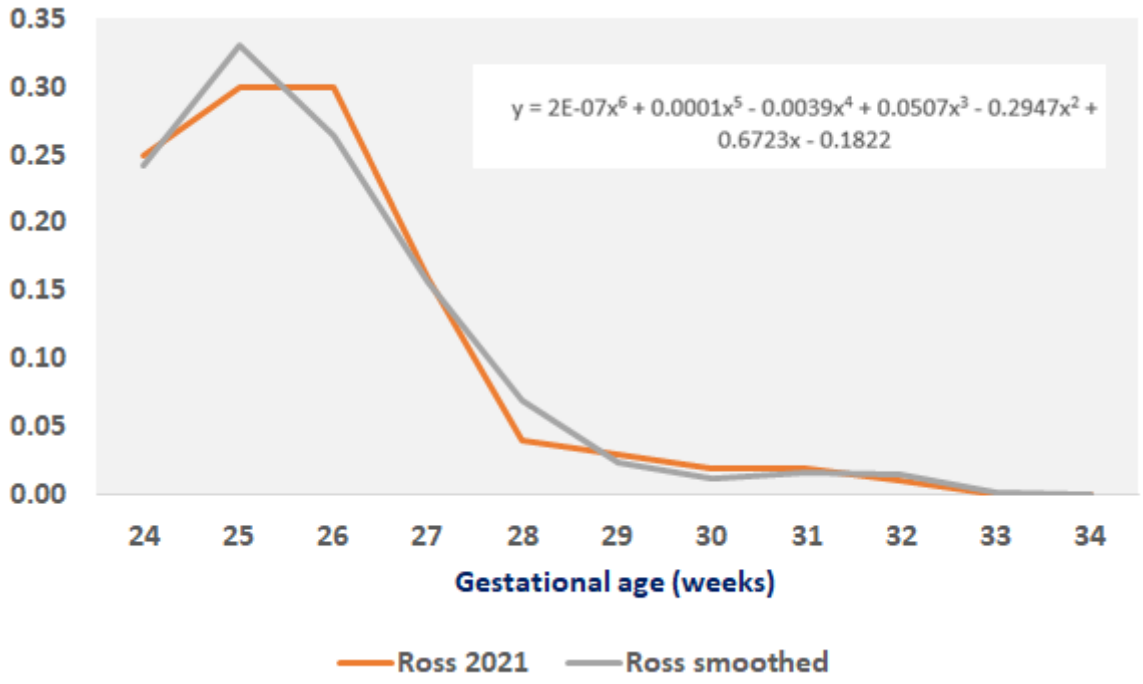
Estimated 2013 cerebral palsy rate = $0.0006x^2 - 0.0449x + 0.8501$ where x is gestational age in weeks

Intraventricular haemorrhage by gestational age

An article on preterm labour (Ross, 2021) was used to estimate the risk of IVH by gestational age at birth. A proportion of neonatal death would be accounted for by mortality due to IVH and to avoid double counting, long term costs and QALY loss associated with IVH was restricted to babies who did not die from the condition. It was assumed that the mortality rate from IVH was not related to gestational age and the risk was estimated as 30% as per the NICE guideline on preterm labour and birth (NG25).

The estimated IVH rate by gestational age at birth utilised in the model is graphed in Figure 15.

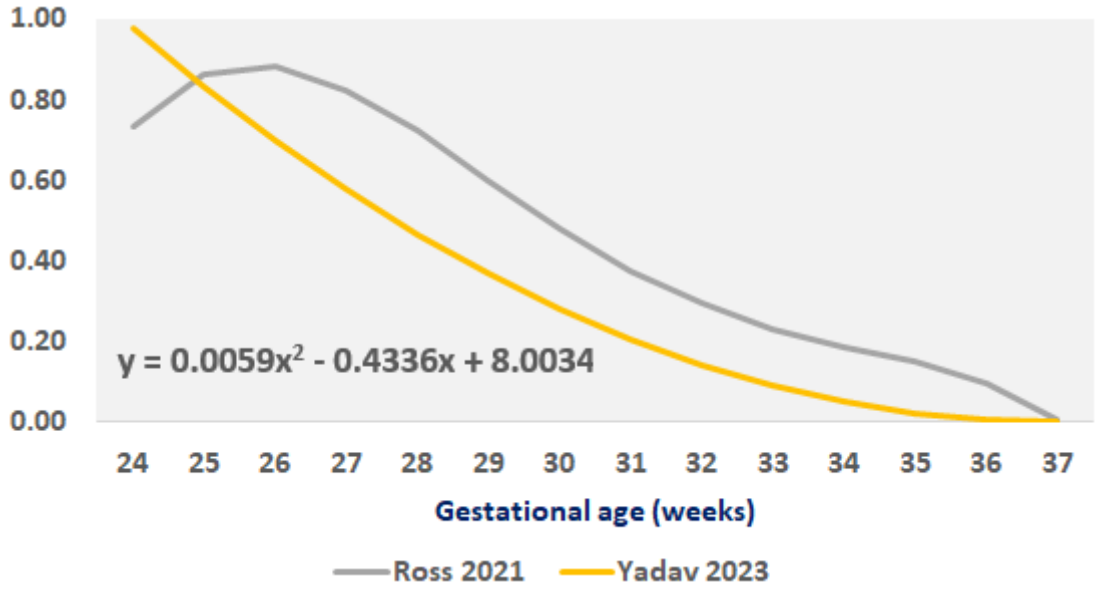
Figure 15: Graph to show estimated risk of intraventricular haemorrhage by gestational age at birth



Respiratory distress syndrome by gestational age

The RDS rate by gestational age at birth was estimated from a paper by Yadav (2023) with a best fit curve derived from the reported values at 24 weeks, 34 weeks, and 37 weeks respectively. As a sensitivity analysis we utilised the approach taken in the previous version of the guideline (NG137). The RDS rate by gestational age at birth born before 35 weeks was estimated from Ross (2021). For babies born later the RDS rate was taken from the NICE guideline on preterm labour and birth (NG25). The RDS mortality rate of 0.054 was taken from published US data (American Lung Association: Lung Disease Data 2008) and, as with IVH, it was assumed that this did not vary with gestational age. To avoid double counting, the model restricted an estimation of long-term costs and QALY losses attributable to RDS to those babies who survived the neonatal period. The alternative model estimates of RDS by gestational age at birth are shown in Figure 16.

Figure 16: Graph to show estimated risk of respiratory distress syndrome by gestational age at birth

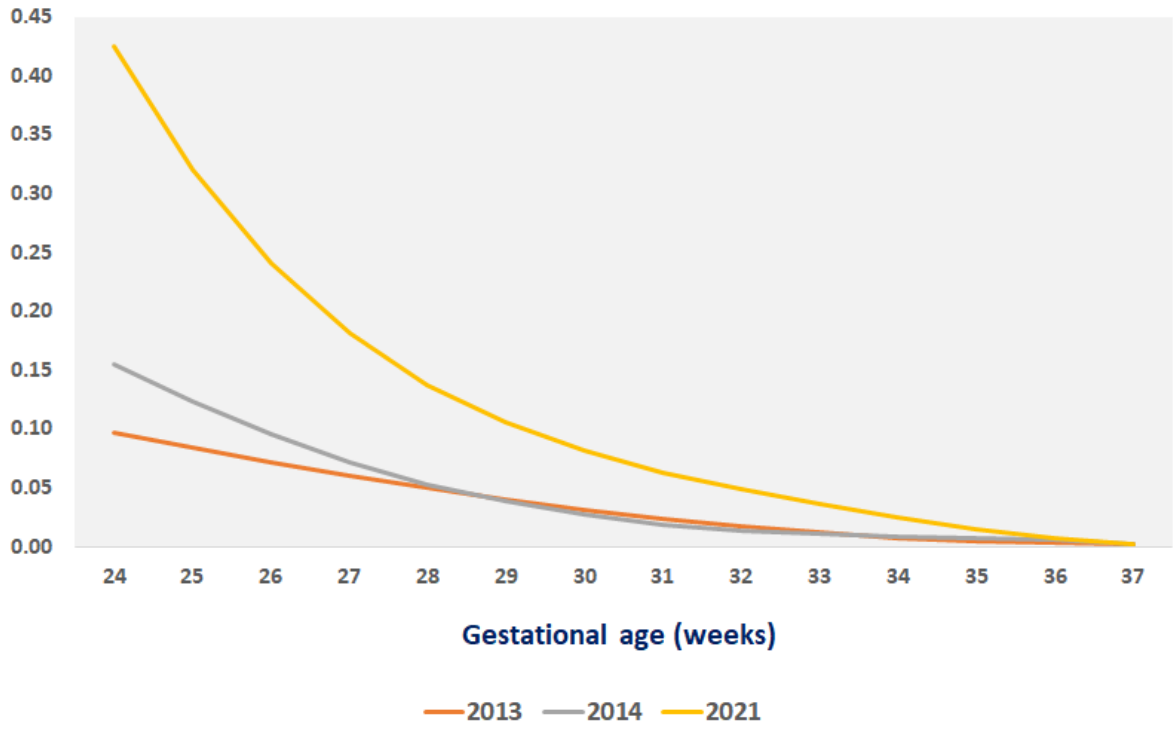


Stillbirths by gestational age

Stillbirth was not used as an outcome in the base case analysis. Whilst the committee noted the association between the stillbirth rate and gestational age, they did not believe there was a strong causal relationship and, as a result, they did not anticipate that delaying or preventing preterm birth would have much impact on the number of stillbirths.

However, as stillbirth was included as an outcome in a similar model developed for a previous NICE guideline (NG137) it was included as a sensitivity analysis, whereby a causal relationship is assumed. Using published data from the Office of National Statistics on live births and stillbirths by gestational age, best fit curves were applied to the calculated stillbirth rate for the years 2013, 2014 (ONS, 2016) and 2021 (ONS, 2023) any of which can be utilised as part of a sensitivity analysis including stillbirth as an outcome. This is illustrated in Figure 17.

Figure 17: Graph to show estimated risk of stillbirth by gestational age at birth



Treatment effectiveness

The relative treatment effectiveness of vaginal progesterone (up to 400mg daily) to prevent preterm birth, compared to no treatment, were derived from a published individual patient data meta-analysis (Conde-Agudelo, 2022). These relative treatment effects along with their 95% confidence intervals are listed in Table 23. These relative risks are applied to the baseline risks of birth for each gestational age from 24 to 36 weeks, for pregnancies identified with a cervical length of ≤25 mm and at higher risk of preterm birth by screening, to determine the weekly health state transition from on-going pregnancy to birth.

Table 23: Relative treatment effect of vaginal progesterone compared to no treatment to prevent preterm birth

Outcome	Relative risk (95% confidence interval)	Distribution	Source
Pre-term birth <28 weeks	0.41 (0.19 to 0.91)	Log-normal	Conde-Agudelo (2022)
Pre-term birth <32 weeks	0.56 (0.33 to 0.93)	Log-normal	Conde-Agudelo (2022)
Pre-term birth <36 weeks	0.89 (0.69 to 1.15)	Log-normal	Conde-Agudelo (2022)

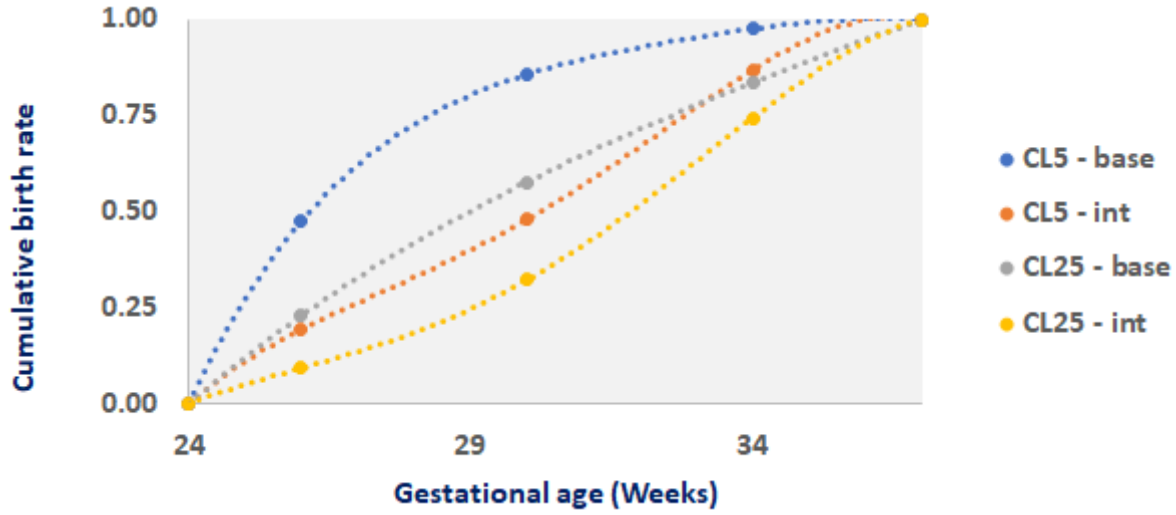
For the deterministic analysis, these relative risks are applied to the baseline risks for each of the mid-range of the 3 time periods and cumulative birth rates with treatment are estimated using a best fit curve for these points. This is illustrated for a cervical length of 5 mm and 25 mm respectively in Table 24 and Figure 18 below.

Table 24: Cumulative birth rates by gestational age with and without treatment for women with a cervical length of 5 mm and women with a cervical length of 25 mm

	Baseline		Intervention	
Gestational age	**CL = 5 mm**	**CL = 25 mm**	**CL = 5 mm**	**CL = 25 mm**
24 weeks	0.000	0.000	0.000	0.000
26 weeks	0.475	0.230	0.475 x 0.41 = 0.195	0.230 x 0.41 = 0.094
30 weeks	0.857	0.578	0.857 x 0.56 = 0.480	0.578 x 0.56 = 0.324
34 weeks	0.976	0.836	0.976 x 0.89 = 0.869	0.836 x 0.89 = 0.744
37 weeks	1.000	1.000	1.000	1.000

CL = cervical length

Figure 18: Illustrative cumulative birth rates for women with a cervix of 5 mm and with a cervix of 25 mm with and without treatment



base = baseline cumulative birth rate; int = cumulative birth rate following intervention

However, the treatment effect estimates were applied somewhat differently for probabilistic sensitivity analysis (PSA). The reason for this was to avoid non-feasible sampled cumulative birth rates and to reflect that the sampled cumulative birth rate in an earlier period cannot be independent of the sampled cumulative birth rate for earlier gestational ages.

For the PSA the cumulative birth rate was calculated as:

Cumulative births $_t$ = Probability of birth$_t$ x (1- cumulative births$_{t-1}$) + cumulative births$_{t-1}$

Table 25 illustrates how the probability of birth in time period (t) is calculated for a woman with a cervix of 5 mm without treatment. To estimate the cumulative birth rate with treatment the sampled relative risk is applied to the probability of birth$_t$ at baseline to get the treatment probability of birth in time period (t).

Table 25: Baseline probability of birth at various time periods for a women with a cervix of 5 mm

Period$_t$	Cumulative birth rate	Proportion births period$_t$	Probability of birth period$_t$
24 weeks	0.000	0.000	0

Period$_t$	Cumulative birth rate	Proportion births period$_t$	Probability of birth period$_t$
26 weeks	0.475	0.475 - 0.000 = 0.475	0.475 ÷ 1 = 47.5%
30 weeks	0.857	0.857- 0.475 = 0.382	0.382 ÷ 0.525 = 72.8%
34 weeks	0.976	0.976 – 0.857 = 0.119	0.119 ÷ 0.143 = 83.2%
37 weeks	1.000	1.000 – 0.976 = 0.024	0.024 ÷ 0.024 = 100%

Quality adjusted life years (QALYs)

In order to estimate the impact of screening and treatment on health-related quality of life, a QALY decrement was applied to the adverse health outcomes assessed within the model. These decrements are listed in Table 26. Future QALY losses were discounted at a rate of 3.5%, unless stated, in accordance with the NICE reference case.

Table 26: QALY decrement for adverse outcomes

Outcome	QALY loss	Source
Neonatal death	25.35	Kind (1999), National Life Tables for UK 2018-20; ONS 2023 [a]
Postnatal death	25.35	Kind (1999), National Life Tables for UK 2018-20; ONS 2023 [a]
Stillbirth	25.35	Kind (1999), National Life Tables for UK 2018-20; ONS 2023 [a]
Cerebral palsy	11.15	Cahill 2011[b]
Intraventricular haemorrhage	4.50	NICE 2015 [NG25] https://www.nice.org.uk/guidance/ng25
Respiratory distress syndrome	1.50	NICE 2015 [NG25] https://www.nice.org.uk/guidance/ng25

(a) A death was assumed to result in a loss of 79 years of life for males and 83 years of life for females based on current life expectancy estimates (ONS, 2023). Health state utilities, weighted by age, were taken from the literature (Kind 1999) and It was assumed the health state utilities across years of life would be as follows: age <25 = 0.94; age 25-34 = 0.93; age 35-44 = 0.91; age 45-54 = 0.85; age 55-64 = 0.81; age 65-74 = 0.78; age ≥75 = 0.71. A weighted average of male and female utilities was estimated using 355,006 male births and 337,716 female births as reported for 2021 in England and Wales (ONS, 2023) Future years were discounted at a rate of 3.5%

(b) It was assumed that each year of life with cerebral palsy would be lived with a health state utility of 0.55 and that life expectancy would be 60 years. The results in 14.2 discounted QALYs. The QALY loss from cerebral palsy was estimated by subtracting 14.2 QALYs from the discounted QALYs estimated as the QALY loss from a neonatal death

Cost and resource use

In accordance with NICE methodology a NHS and Personal Social Services (PSS) perspective was adopted for this analysis (Developing NICE guidelines: the manual). Costs were based on a 2021/22 price year reflecting the most recently available national schedule of NHS Costs at the time of writing. Any future costs were discounted at a rate of 3.5%, unless stated, in line with the NICE reference case.

Table 27 gives the unit costs related to the intervention. Screening for preterm birth was undertaken by measurement of cervical length using transvaginal ultrasound. Treatment, consisting of a 200mg daily dose of progesterone administered by a vaginal pessary, would be initiated in women screened positive and was assumed to continue until birth.

Table 27: Screening and treatment cost

Intervention	Cost	Standard error [a]	Distribution	Source
Screening	£142	£16.65	Normal	National Schedule of NHS Costs 2021-22[b]
Daily treatment	£0.60	-	Deterministic	NHS Drugs Tariff December 2023[c]

(a) Estimated using source data from the National Schedule of NHS Costs 2021-22
(b) Service code 501, Obstetrics service, Currency code MA36Z; Transvaginal ultrasound 2021-22 National Cost Collection Data
(c) https://www.nhsbsa.nhs.uk/pharmacies-gp-practices-and-appliance-contractors/drug-tariff

It was assumed that women identified as at high risk of preterm birth from cervical screening would have a counselling consultation with a consultant (see consultant obstetrician unit cost in Table 31 for cost).

Women with on-going pregnancies would continue to receive the monitoring as per the schedule recommended in this guideline, as shown for monochorionic and dichorionic twin pregnancies in Table 28 and Table 29. It is assumed that woman would receive an obstetric review for each scan. Data from the published literature was used to estimate the proportion of twins that would be either monochorionic or dichorionic as described in Table 30. The unit costs used to derive the costs from these monitoring appointments are given in Table 31.

Table 28: Monochorionic appointment schedule

	Gestational age													
Appointment	24	25	26	27	28	29	30	31	32	33	34	35	36	37
Specialist midwife follow-up	✓		✓		✓		✓		✓		✓			
Consultant obstetrician follow-up	✓								✓					
Scan for monitoring for FFTS/sIUGR/TAPS	✓		✓		✓				✓		✓			

FFTS = feto-fetal transfusion syndrome; sIUGR = selective intrauterine growth restriction; TAPS = twin anemia-polycythemia sequence

Table 29: Dichorionic appointment schedule

	Gestational age													
Appointment	24	25	26	27	28	29	30	31	32	33	34	35	36	37
Specialist midwife follow-up	✓				✓				✓		✓		✓	
Consultant obstetrician follow-up	✓								✓					
Scan for IUGR	✓				✓				✓				✓	

IUGR = intrauterine growth restriction

Table 30: Proportion of twin type

Type	Proportion	Source
Monozygotic	0.33	https://twinstrust.org/information/pregnancy-and-birth/finding-out/identical-or-non-identical.html
Monochorionic|monozygotic [a]	0.75	Shulman 2006
Dichorionic|monozygotic	0.25	Shulman 2006
Dizygotic	0.67	https://twinstrust.org/information/pregnancy-and-birth/finding-out/identical-or-non-identical.html
Monochorionic|dizygotic	0.00	Shulman 2006

Type	Proportion	Source
Dichorionic|dizygotic	1.00	Shulman 2006

(a) | denotes a conditional probability, the probability that a pregnancy is monochorionic given that it is monozygotic

Table 31: Antenatal appointment costs

Appointment	Cost	Standard error [a]	Distribution	Source
Specialist midwife follow-up	£88	£9.14	Normal	National Schedule of NHS Costs 2021-22[b]
Consultant obstetrician follow-up	£202	£11.52	Normal	National Schedule of NHS Costs 2021-22[c]
Obstetrician review	£123	13.73	Normal	National Schedule of NHS Costs 2021-22[d]
Scan	£137	£24.35	Normal	National Schedule of NHS Costs 2021-22[e]

(a) Estimated using source data from the National Schedule of NHS Costs 2021-22
(b) Consultant led; Service code 560, Midwifery service; Currency Code WF01A; Non-Admitted Face-to-Face Attendance, Follow-up 2021-22 National Cost Collection Data
(c) Consultant led; Service code 501, Obstetrics service; Currency Code WF01A; Non-Admitted Face-to-Face Attendance, Follow-up 2021-22 National Cost Collection Data
(d) Non consultant led; Service code 501, Obstetrics service; Currency Code WF01A; Non-Admitted Face-to-Face Attendance, Follow-up 2021-22 National Cost Collection Data
(e) Antenatal standard routine ultrasound scan; Service code 560, Midwifery service; Currency Code NZ21Z 2021-22 National Cost Collection Data

The model incorporates a sensitivity analysis where a causal relationship between gestational age at birth and stillbirth is assumed. This included the healthcare costs associated with stillbirth, such as postpartum care for parents and the treatment of parental anxiety and depression. In the base case analysis, it has been assumed that the costs of a neonatal and post neonatal death would be subsumed within the costs of a neonatal intensive care admission. However, the model has been devised so that additional costs related to death itself can be considered as part of a sensitivity analysis. The base case costs associated with mortality are shown in Table 32.

Table 32: Costs associated with mortality

Outcome	Cost	Distribution	Source
Stillbirths	£4,795	Deterministic	Campbell 2018
Neonatal/postnatal death	£0	Deterministic	Assumption

(a) Updated to 2021-22 prices using the NHSCII cost inflation index, with a multiplier of 1.14 applied to 2013/14 prices
(b) Assumption

Table 33 shows the unit costs associated with neonatal unit admissions. The costs of a neonatal unit admission are estimated using these costs and an estimation of length of stay by gestational and weighted by the level of care, also by gestational age. We used data from the Neonatal Data Analysis Unit (2017) as shown in Table 34, to estimate the length of stay by gestational age at birth. A best fit curve was then fitted to this data to provide an estimate for each week of gestation, as illustrated in Figure 19. We made the simplifying assumption that there was no independent effect of multiplicity on length of stay over and above that of gestational age at birth.

Table 33: Neonatal unit costs

Unit	Cost per diem	Standard error [a]	Distribution	Source
Intensive care	£1,811	£72	Normal	National Schedule of NHS Costs 2021-22[b]
High dependency	£1,226	£39	Normal	National Schedule of NHS Costs 2021-22[c]
Special care	£873	£30	Normal	National Schedule of NHS Costs 2021-22[d]

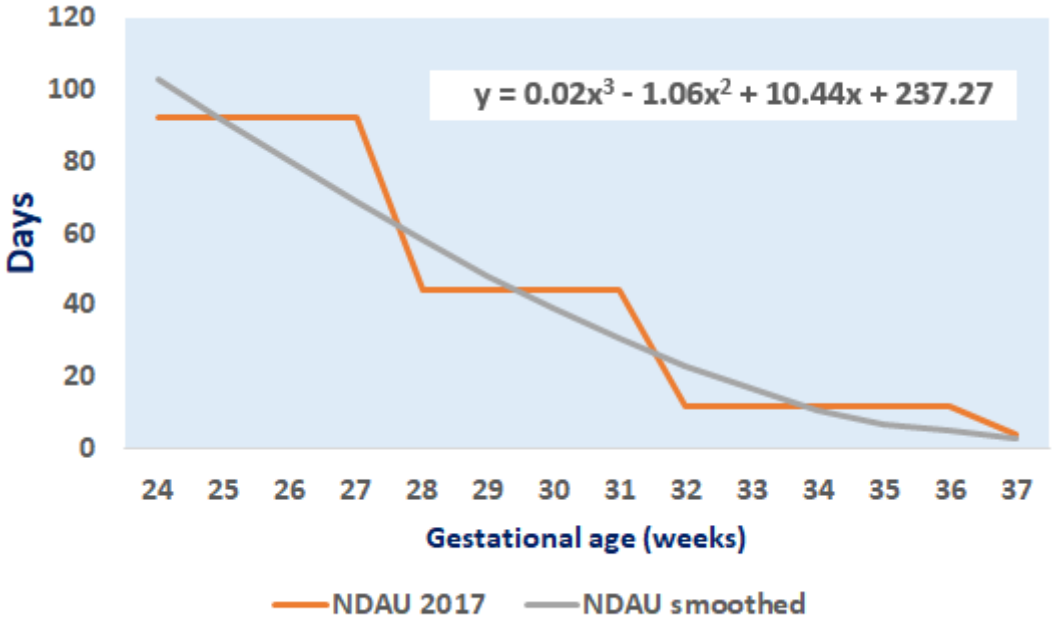
(a) Estimated using source data from the National Schedule of NHS Costs 2021-22
(b) Service code CCU13; Currency Code XA01Z; Neonatal Critical Care, Intensive Care 2021-22 National Cost Collection Data
(c) Service code CCU13; Currency Code XA02Z; Neonatal Critical Care, High Dependency 2021-22 National Cost Collection Data
(d) Service code CCU13; Currency Code XA03Z and XA04Z; Neonatal Critical Care, Special Care 2021-22 National Cost Collection Data - weighted average of with and without external carer

Table 34: Mean length of neonatal care stay by gestational age at birth

Gestational age at birth	Mean length of stay (days)
≤27 weeks	92
28-31 weeks	44
32-36 weeks	12
37 weeks	4

(a) <Insert Note here>

Figure 19: Estimate of neonatal unit length of stay by gestational age at birth



Y is the estimate of neonatal unit length of stay in days where x is gestational age in weeks

In order to weight a neonatal unit admission by the level of care we used estimates from a previous NICE guideline on multiple pregnancy (CG129) and the expert opinion of the committee, which are summarised in Table 35.

Table 35: Proportion of neonatal unit admission by level of care and gestational age at birth

Gestational age	Intensive care	High dependency	Special care
24 weeks [a]	50%	14%	36%
25 weeks [a]	46%	15%	38%
26 weeks [a]	42%	17%	42%
27 weeks [a]	36%	18%	45%
28 weeks [a]	30%	20%	50%
29 weeks	22%	22%	56%
30 weeks	22%	22%	56%
31 weeks	22%	22%	56%
32 weeks	22%	22%	56%
33 weeks	22%	11%	67%
34 weeks	22%	11%	67%
35 weeks	22%	11%	67%
36 weeks	22%	11%	67%
37 weeks	22%	11%	67%

(a) Based on expert opinion that from birth up to a gestational age 31 completed weeks, they receive Neonatal Critical Care, that they receive Neonatal Critical Care, High Dependency care for the subsequent 2 weeks (to 33 completed weeks) and until 37 completed weeks they receive Neonatal Critical Care, Special Care. Additionally, a simplifying assumption is made that discharge is at 37 weeks for these babies to calculate the proportions

Table 36 gives the costs that are assumed for model outcomes with long term morbidity while Table 37 gives the NHS costs associated with moderate and late preterm birth in the 2 years following initial discharge from hospital based on data from a UK study (Khan, 2015). The model assumed that all babies who survived the neonatal period incur the 0-6 months post discharge NHS costs, but only infants who survived to one year incur the 6-12 months and 12-24 months discharge costs. This makes the simplifying assumption that all post neonatal death occur within six months.

Table 36: Costs associated with long term morbidity

Outcome	Cost	Source
Cerebral palsy	£94,447 [a]	Kruse 2009
Intraventricular haemorrhage	£28,334 [b]	Twin Birth Costing Report (https://www.hfea.gov.uk/media/2650/nga-twin-pregnancy-costing-final.pdf)
Respiratory distress syndrome	£4,551 [c]	Marti 2016

(a) Updated to 2021-22 prices using the NHSCII cost inflation index, with a multiplier of 1.11 applied to 2016/17 prices (these costs were updated to 2016/17 prices in NICE 2019 (NG137)

(b) As per the NICE guideline on Preterm labour and birth (NG25) it was assumed that IVH would have the same cost as intracranial haemorrhage (ICH) and that Grade III and Grade IV ICH would be similar in cost to cerebral palsy. As it was estimated that 30% of ICH is of severity Grade III and Grade IV, the cost of IVH was assumed to be 30% of the cost of cerebral palsy.

(c) The costs of RDS were taken from a published UK study using the reported mean one-year post-hospital costs to the NHS in one-year survivors. Updated to 2021-22 prices using the NHSCII cost inflation index, with a multiplier of 1.17 applied to 2012 prices

Table 37: NHS costs in first 2 years after initial discharge from hospital by gestational age

Gestational age at birth	0-6 months	6-12 months	12-24 months
24-36 weeks	£1,663	£977	£512
37 weeks	£935	£824	£350

(a) Updated to 2021-22 prices using the NHSCII cost inflation index, with a multiplier of 1.11 applied to 2016/17 prices (these costs were updated to 2016/17 prices in NICE 2019 [NG137 - https://www.nice.org.uk/guidance/ng137]

Results

The results from the model are presented below. The incremental cost effectiveness ratio (ICER) is a summary measure of cost-effectiveness where a ratio is calculated by dividing the additional costs of an intervention compared to some comparator by the additional benefits of the intervention, measured by QALYs in this analysis. The ICER is then compared with some cost effectiveness threshold and if the ICER lies below this threshold the intervention is considered cost effective. If the intervention is both cheaper and more effective than the comparator then the ICER is not required, as the intervention is unambiguously cost-effective and is said to dominate the comparator.

Another summary measure of cost-effectiveness used in the presentation of results below is the net monetary benefit (NMB) which is calculated as the product of the QALYs from the intervention and the cost effectiveness threshold (which gives a monetary valuation of benefit) less the costs of the intervention. The strategy with the highest NMB is the most cost-effective strategy. The NMB statistic will always give the same conclusion with respect to cost effectiveness as the ICER, but it is also more straightforward to quantify the uncertainty around an NMB point estimate than it is for an ICER. All net monetary benefit values were calculated using a cost effectiveness threshold of £20,000 per QALY.

When an incremental net monetary benefit (iNMB) result is presented that indicates the additional NMB of screening for cervical length and daily vaginal progesterone for those identified with a short cervix compared to the comparator of no screening. Positive values indicate that intervention is cost-effective and negative values that is not.

Base Case analysis

Deterministic

The base case deterministic results using the Liem (2018) data on the distribution of cervical length in women with a twin pregnancy are shown in Table 38 and Figure 20. The costs and QALYs need to be interpreted carefully as with the exception of the screening cost, which is per twin pregnancy, they only include the proportion of the population with a short cervix as the costs and outcomes for those with a cervix >25 mm cancel out. The incremental net monetary benefit is calculated relative to the baseline strategy of no screening and no treatment and does represent that measure per twin pregnancy.

Table 38: Summary of deterministic base case analysis using Liem (2018) data on the distribution of cervical length in women with a twin pregnancy

Strategy	Cost	QALY	ICER	NMB	iNMB
No screening or treatment	£1,694	0.40	-	£6,241	-
Screen and treat CL≤25 mm	£1,175	0.45	Dominates	£7,763	£1,512

Figure 20: Cost effectiveness plane showing the incremental costs and QALYs of screening strategies when compared to no screening using Liem (2018) data on the distribution of cervical length in women with a twin pregnancy

Table 39 shows the modelled impact of the various strategies on the clinical outcomes included in the model per twin pregnancy in the proportion of women with a short cervix.

Table 39: Clinical outcomes for different screening strategies in the deterministic base case analysis using Liem (2018) data on the distribution of cervical length in women with a twin pregnancy

Strategy	Neonatal deaths	Post neonatal deaths	Cerebral palsy	IVH	RDS	Neonatal unit admissions
No screening or treatment	0.0023	0.0005	0.0013	0.0017	0.0089	0.0179
Screen and treat CL≤25 mm	0.0011	0.0003	0.0008	0.0008	0.0051	0.0158

IVH = intraventricular haemorrhage; RDS = respiratory distress syndrome

Table 40 shows the breakdown of costs by category for the intervention and comparator, noting that the costs are per twin pregnancy in that proportion with a short cervix.

Table 40: Breakdown of costs in the deterministic base case analysis

Strategy	Dx	Tx	Antenatal appts	CP	IVH	RDS	NNU	Post neonatal discharge	Total
No screening or treatment	£0	£0	£8.48	£119.07	£48.72	£40.44	£1,423	£53.23	**£1,694**
Screen and treat CL≤25 mm	£142	£2.29	£11.69	£77.91	£22.09	£23.28	£838	£57.44	**£1,175**

Appts = appointments; CP = cerebral palsy; Dx = diagnosis/screening; IVH = intraventricular haemorrhage; NNU = neonatal unit; RDS = respiratory distress syndrome; Tx = treatment

Base case probabilistic sensitivity analysis

A total of 10,000 Monte Carlo simulations were run using the Liem (2018) data on the distribution of cervical length in women with a twin pregnancy, as described in Table 13 and Figure 9, with deterministic model inputs set to their base case values. The results are summarised in Table 41. They show that screening for cervical length and treating women with a cervical length of 25 mm or less with vaginal progesterone dominates no screening and no treatment producing mean savings of nearly £500 per twin pregnancy and with higher mean QALYs. It also indicates that the cost-effectiveness of intervention is unlikely to be due to chance, with 99% of simulations being cost-effective.

Figure 21 show the plot of the 10,000 simulations on the cost-effectiveness plane and the cost-effectiveness acceptability curve (CEAC) is graphed in Figure 22, showing the probability that either strategy is cost-effective at different cost-effectiveness thresholds. The CEAC shows that the conclusion that screening and treatment are highly likely to be cost-effective is not very sensitive to different cost-effectiveness thresholds. This is because only a very small percentage of simulations occur in the northeast and southwest quadrant of the cost-effectiveness plane, where changing the threshold could make a difference.

Table 41: Probabilistic sensitivity analysis for the base case

Strategy	Mean incremental cost	Mean incremental QALY	Mean iNMB (95% CrInt)	Probability cost-effective
Screen and treat CL≤25 mm	-£485	0.05	£1,489 (£1,473 to 1,505)	99.1%

CrInt = credible intervals

Figure 21: Cost-effectiveness plane showing results of 10,000 Monte Carlo simulations of the base case analysis

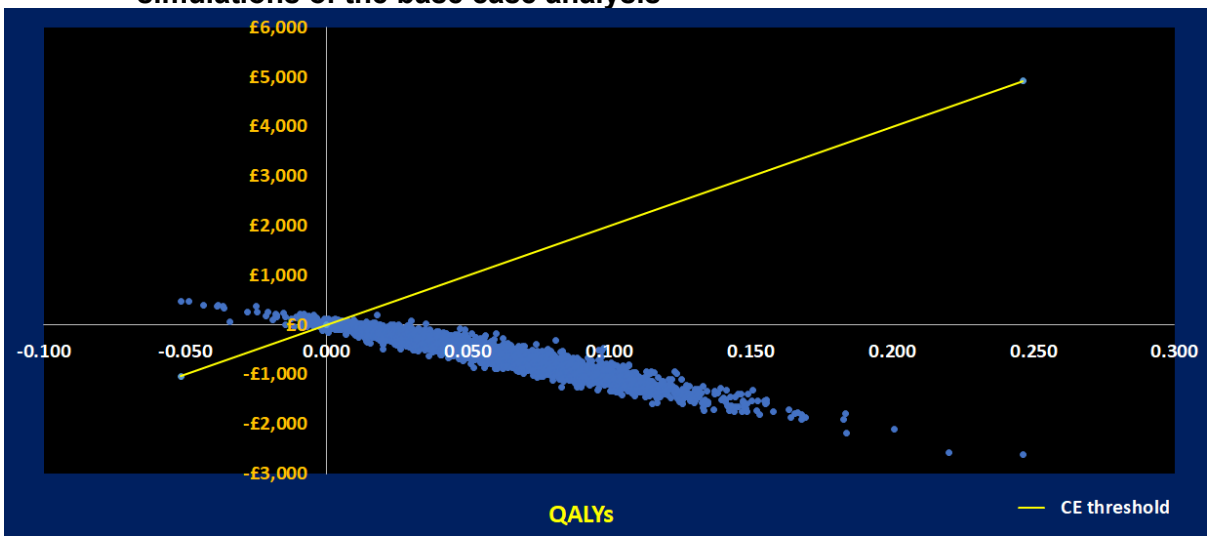
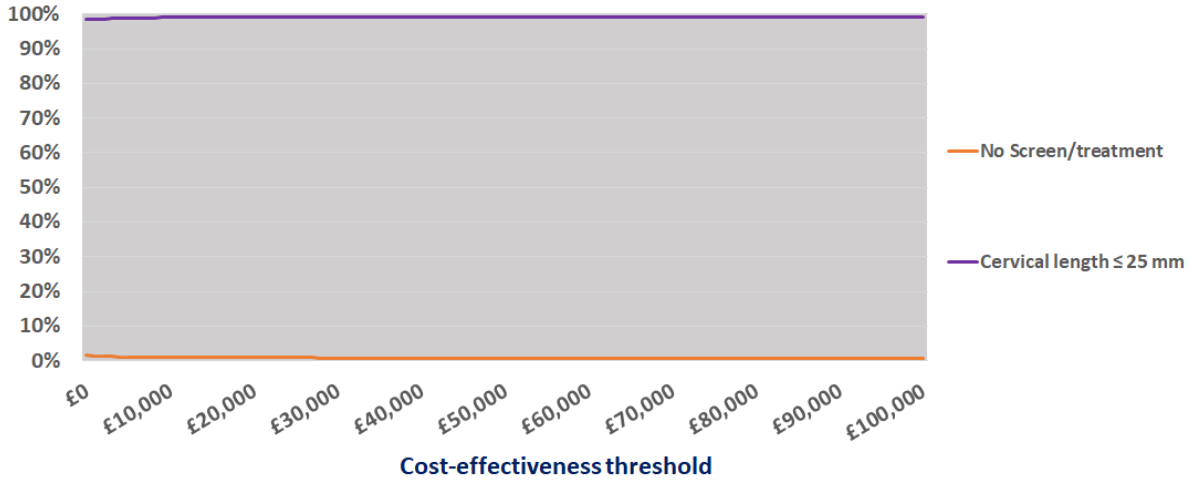


Figure 22: Cost-effectiveness acceptability curve of the base case analysis



Sensitivity analyses

Table 42 summarises the results of a number of deterministic sensitivity analyses with ICERs and iNMBs calculated relative to the baseline strategy of no screening and no treatment. A positive iNMB indicates that screening and treatment of women with a short cervix of less than or equal to 25 mm is cost-effective. In most of these analyses a single parameter or model feature is varied whilst retaining all other inputs and features at their base case value. The table notes detail exceptions. For the sensitivity analyses varying the prevalence of short cervix, the 5 cervical length categories were all assumed to have an equal frequency.

Table 42: Summary of deterministic sensitivity analyses

Analysis	ICER	iNMB
Liem CL distribution – base case	Screening/treatment dominates	£1,522
Skentou (2001) CL distribution	Screening/treatment dominates	£20,957
Souka (1999) CL distribution	Screening/treatment dominates	£22,0146
Stillbirths included (ONS 2021)	Screening/treatment dominates	£1,707
Kindinger 18 weeks	Screening/treatment dominates	£1,748
Kindinger 22 weeks	Screening/treatment dominates	£1,296
Short cervix prevalence 0.1%	£11,961 per QALY	£46
Short cervix prevalence 0.2%	Screening/treatment dominates	£235
Short cervix prevalence 0.3%	Screening/treatment dominates	£423
Short cervix prevalence 0.4%	Screening/treatment dominates	£610
Short cervix prevalence 0.5%	Screening/treatment dominates	£799
Short cervix prevalence 0.6%	Screening/treatment dominates	£987
Short cervix prevalence 0.7%	Screening/treatment dominates	£1,176
Short cervix prevalence 0.8%	Screening/treatment dominates	£1,363
Short cervix prevalence 0.9%	Screening/treatment dominates	£1,552
Neonatal death rate (ONS 2017)	Screening/treatment dominates	£1,187
Post neonatal death rate (ONS 2014)	Screening/treatment dominates	£1,501
Cerebral palsy rate (Himpens 2018)	Screening/treatment dominates	£1,739

Analysis	ICER	iNMB
Respiratory distress syndrome (Ross 2021)	Screening/treatment dominates	£1,512
"Worst case" scenario – relative risk only[a]	£20,306 per QALY	-£1
"Worst case" scenario – relative risk and outcomes[b]	£33,656 per QALY	-£37

(a) All relative risks were set to the upper limit of their 95% confidence interval (see Table 23)
(b) All relative risks were set to the upper limit of their 95% confidence interval (see Table 23). In addition neonatal death rates were estimated using ONS 2017, post neonatal death rates were estimated using ONS 2014 and respiratory distress syndrome rates were estimated using Ross 2021

Only in the very "worst case" scenario sensitivity analyses was intervention found not to be cost-effective and in most of the analyses, intervention dominated no screening and no treatment.

In addition to the deterministic sensitivity analyses undertaken, further probabilistic sensitivity analyses were undertaken for estimates of cervical length using Skentou (2001) and Souka (1999). Also, PSA was performed for short cervix prevalence of 0.1% and 0.5% and for Kindinger (2016) predictions of premature birth based on cervical length at 18 weeks and 22 weeks. The results of these analyses are reported in Table 43. Cost effectiveness planes and cost-effectiveness acceptability curves for these analyses are also provided.

Table 43: Summary of PSA using alternative features and assumptions to the base case analysis

Analysis	Incremental			Probability cost-effective	CE plane	CEAC
	Costs	QALYs	NMB (95% CrInt)			
Base case	-£485	0.050	£1,489 (£1,473 to 1,505)	99.1%	Figure 21	Figure 22
Skentou (2001)	-£7,869	0.629	£20,438 (£20,311 to £20,564)	99.4%	Figure 23	Figure 24
Souka (1999)	-£8,081	0.656	£21,199 (£21,053 to £21,345)	99.3%	Figure 25	Figure 26
Short cervix prevalence 0.1%	£74	0.006	£38 (£37 to £40)	67.3%	Figure 27	Figure 28
Short cervix Prevalence 0.5%	-£200	0.028	£765 (£759 to £770)	98.9%	Figure 29	Figure 30
Kindinger 18 weeks	-£628	0.063	£1,896 (£1,879 to £1,914)	99.7%	Figure 31	Figure 32
Kindinger 22 weeks	-£334	0.037	£1,071 (£1,057 to £1,085)	97.1%	Figure 33	Figure 34

CE = Cost-effectiveness; CEAC = Cost-effectiveness acceptability curve

All these PSA found that screening for a short cervix and treatment with vaginal progesterone in those women identified was cost-effective and with a very high probability. Although the cost-effectiveness conclusion is less clear cut when a short cervix prevalence of 0.1% is assumed, there is still a probability of 67% that intervention is cost-effective.

Figure 23: Cost-effectiveness plane showing results of 10,000 Monte Carlo simulations using Skentou (2001) estimated cervical length frequency distribution

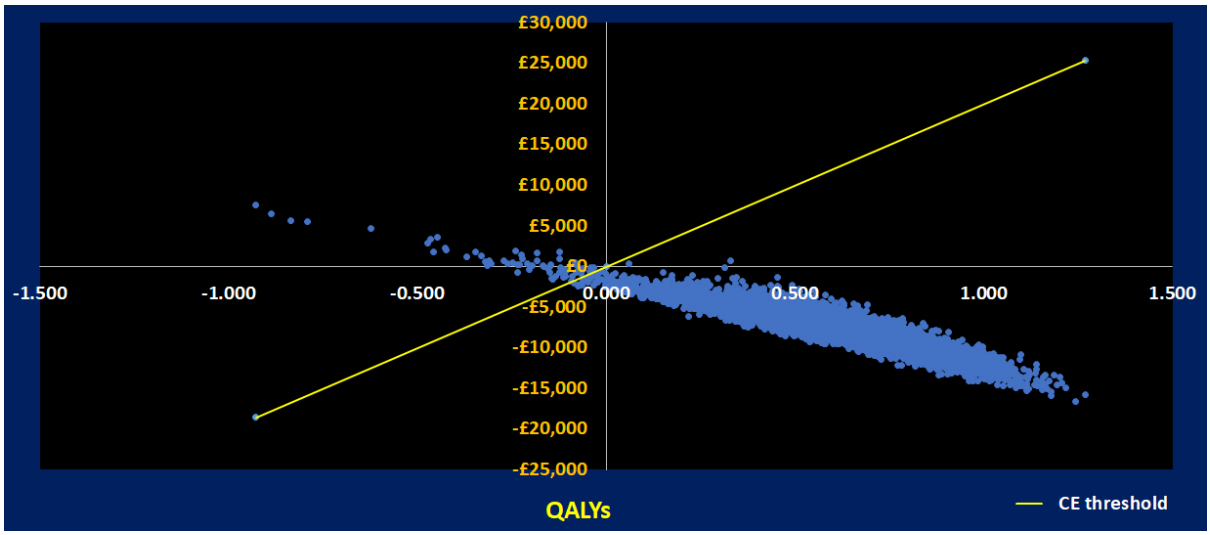


Figure 24: Cost-effectiveness acceptability curve using Skentou (2001) estimated cervical length frequency distribution

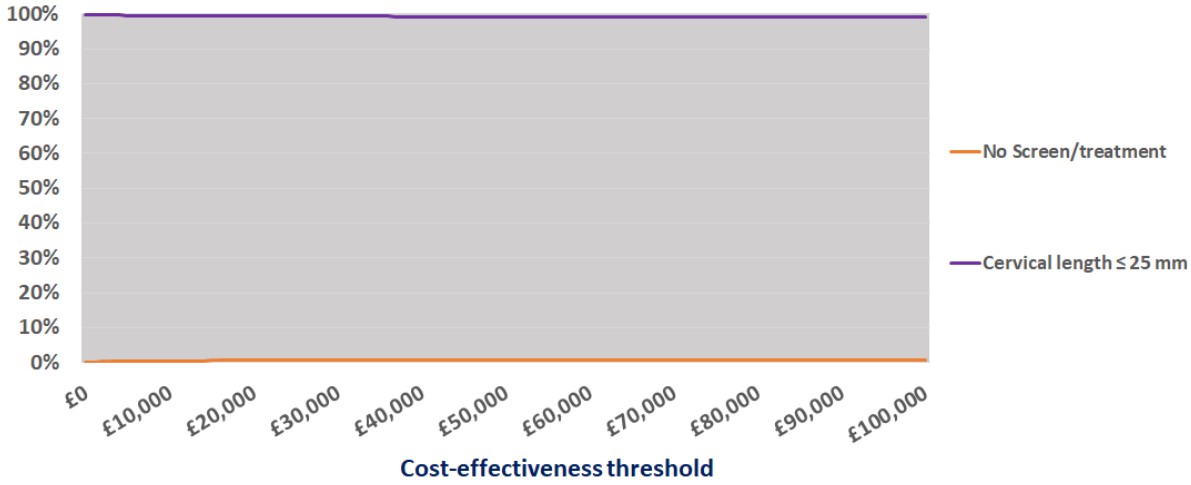


Figure 25: Cost-effectiveness plane showing results of 10,000 Monte Carlo simulations using Souka (1999) estimated cervical length frequency distribution

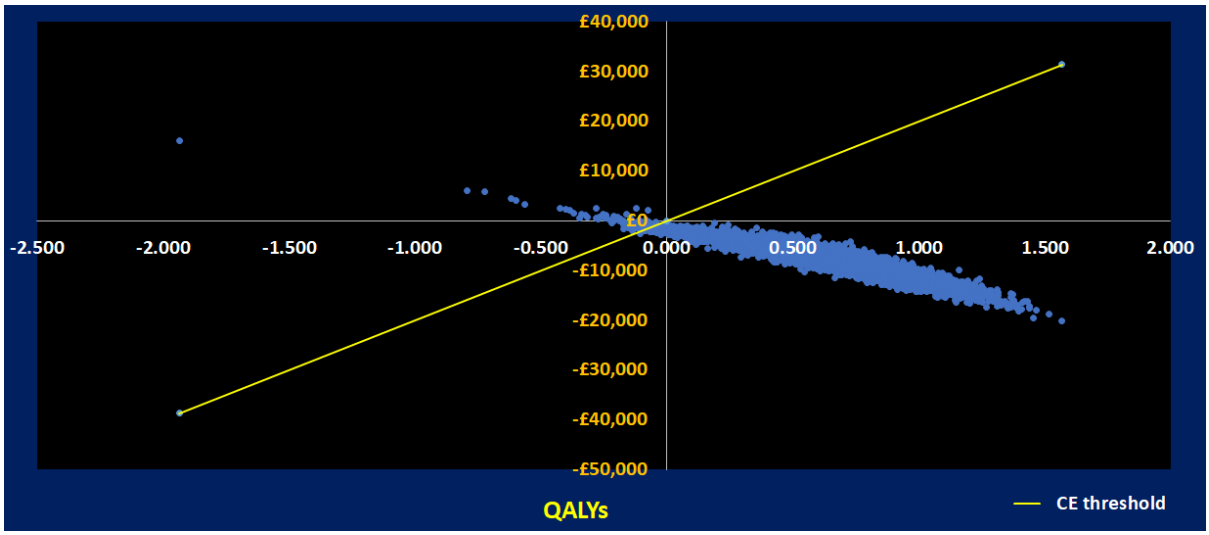


Figure 26: Cost-effectiveness acceptability curve using Souka (1999) estimated cervical length frequency distribution

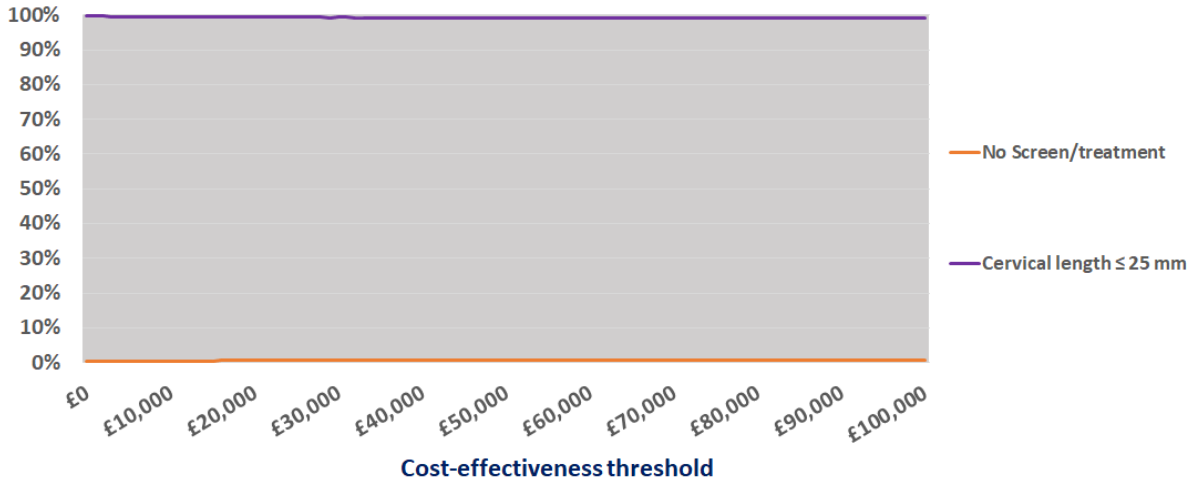


Figure 27: Cost-effectiveness plane showing results of 10,000 Monte Carlo simulations with a 0.1% short cervix (≤25 mm) prevalence

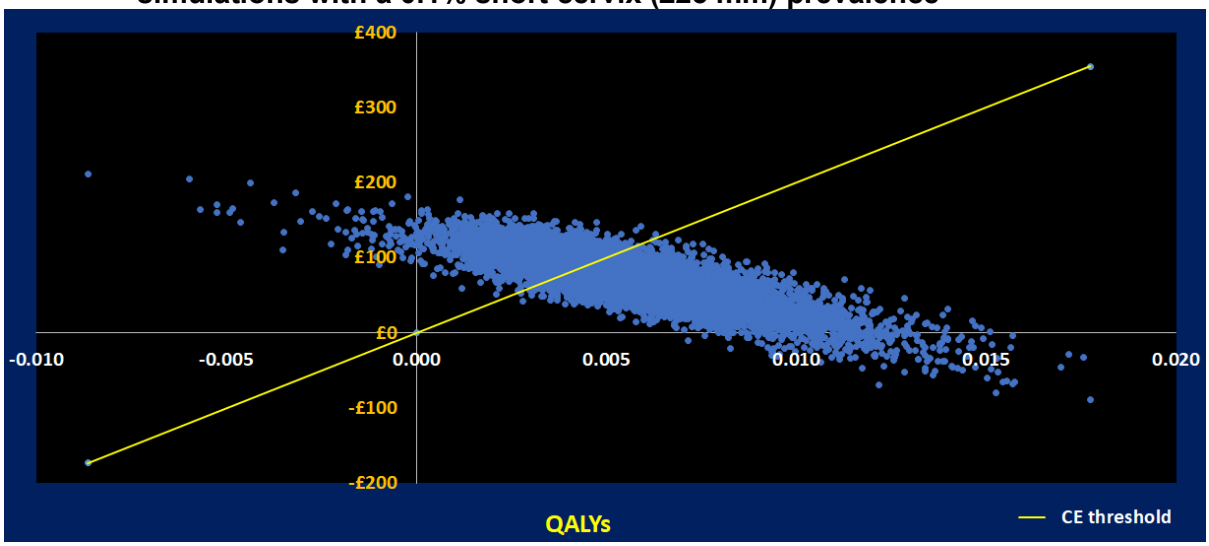


Figure 28: Cost-effectiveness acceptability curve with a 0.1% short cervix (≤25 mm) prevalence

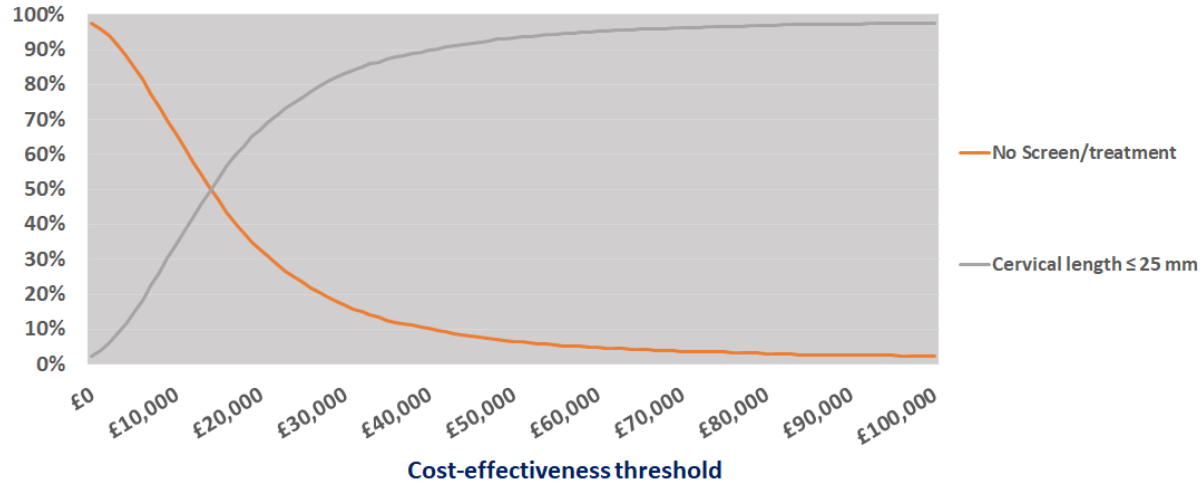


Figure 29: Cost-effectiveness plane showing results of 10,000 Monte Carlo simulations with a 0.5% short cervix (≤25 mm) prevalence

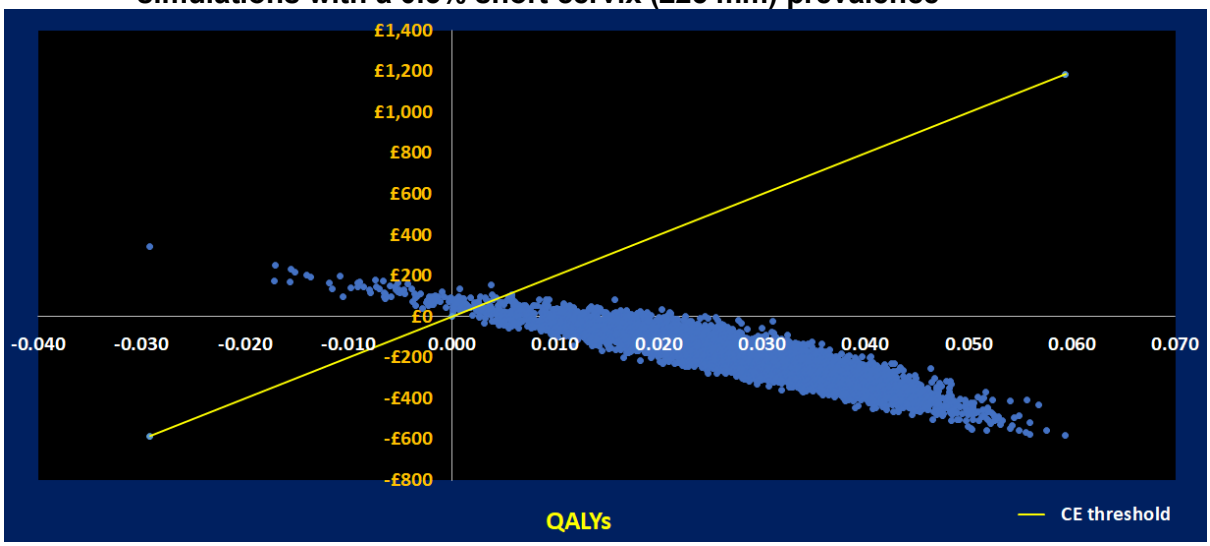


Figure 30: Cost-effectiveness acceptability curve with a 0.5% short cervix (≤25 mm) prevalence

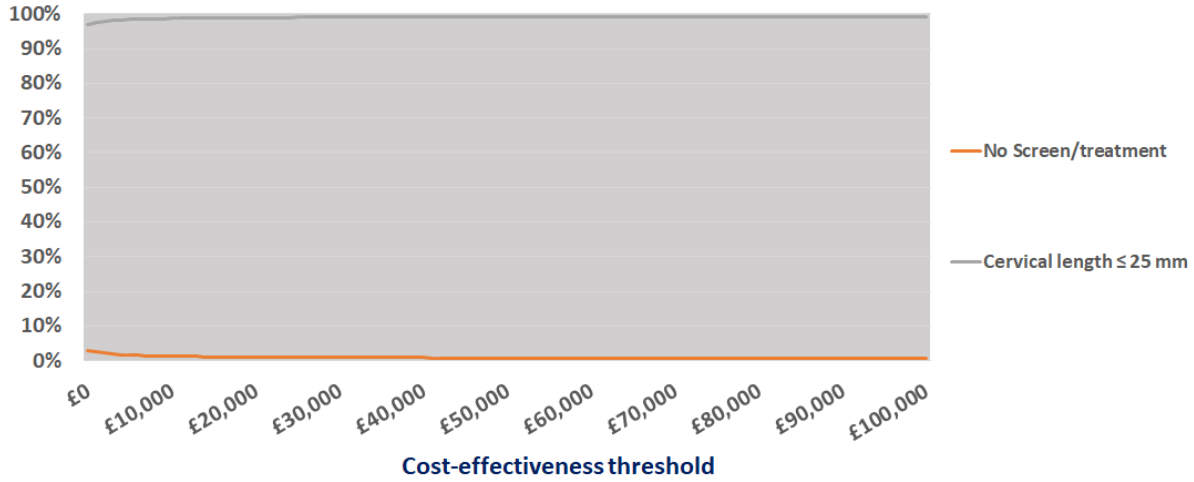


Figure 31: Cost-effectiveness plane showing results of 10,000 Monte Carlo simulations using Kindinger (2016) predictions of premature birth based on cervical length at 18 weeks

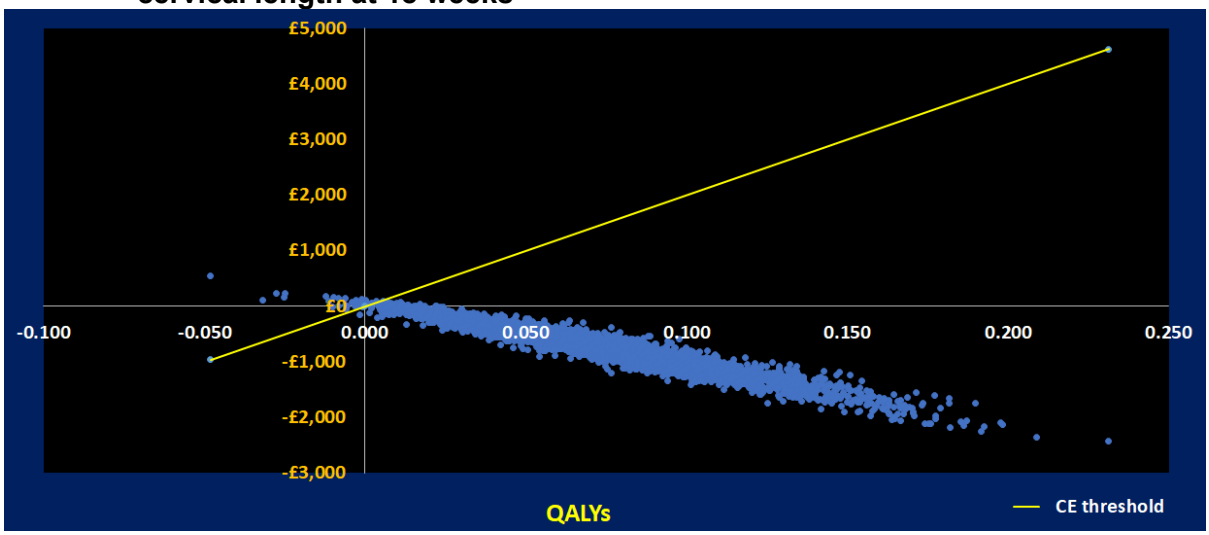


Figure 32: Cost-effectiveness acceptability curve using Kindinger (2016) predictions of premature birth based on cervical length at 18 weeks

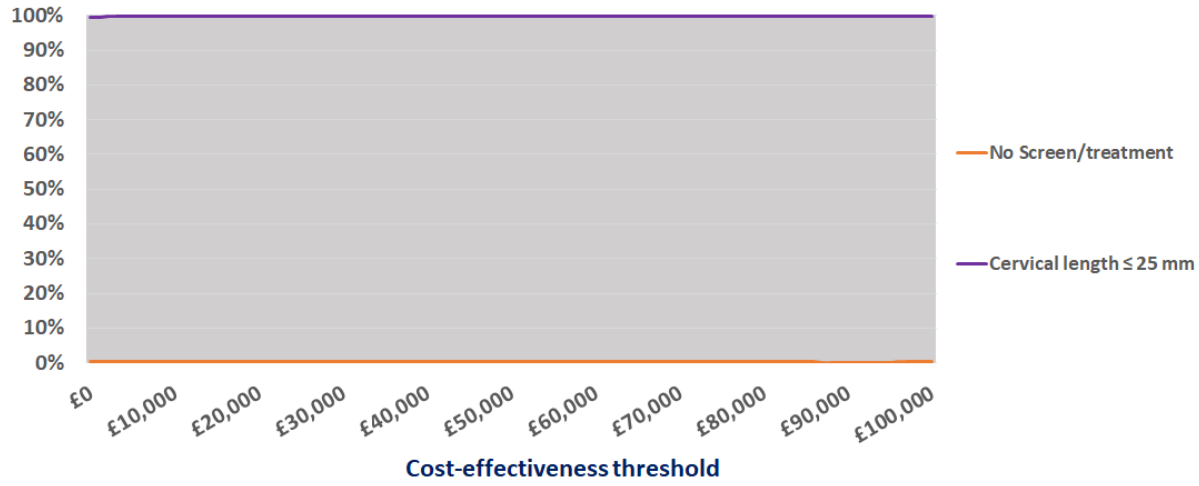


Figure 33: Cost-effectiveness plane showing results of 10,000 Monte Carlo simulations using Kindinger (2016) predictions of premature birth based on cervical length at 22 weeks

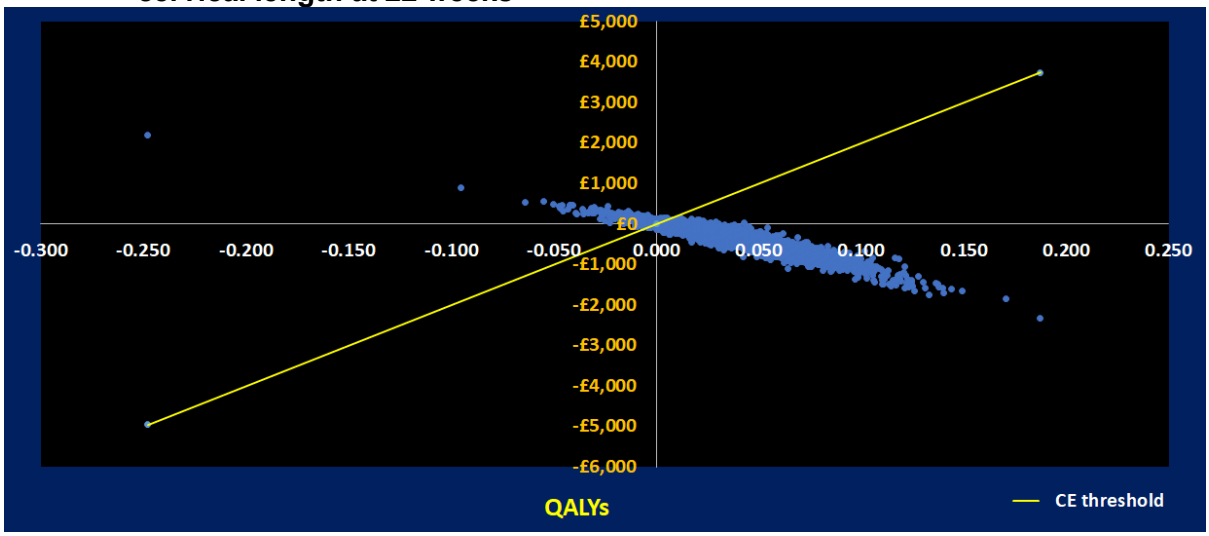
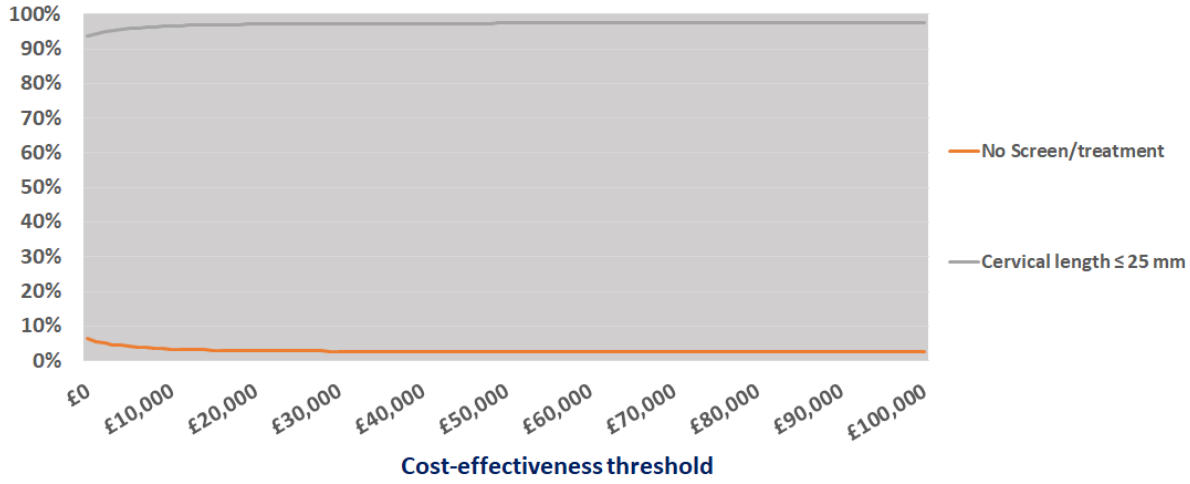


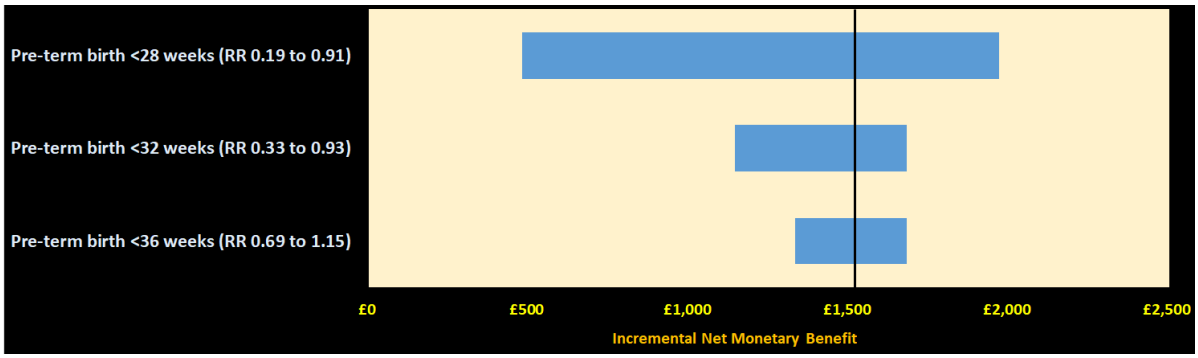
Figure 34: Cost-effectiveness acceptability curve using Kindinger (2016) predictions of premature birth based on cervical length at 22 weeks



Tornado analysis of relative risks

The "worst case" scenario indicated, not surprisingly, that the cost-effectiveness of intervention was highly dependent on the relative treatment effect size. In this analysis the relative risk for each time point was varied between the upper and lower limit of the 95% confidence intervals whilst keeping the other relative risks estimates along with other model parameters and features at their base case setting. These results are shown in the Tornado diagram in Figure 35. The solid black vertical line indicates the base case incremental net monetary benefit.

Figure 35: Tornado diagram showing the impact of varying relative risks for vaginal progesterone treatment on the incremental net monetary benefit of treatment



This analysis showed that intervention remains cost-effective when varying the relative risks one at a time between their lower and upper 95% confidence intervals. It showed that the cost-effectiveness was more sensitive to changes in the relative risks for extreme prematurity (babies born at a gestational age of less than 28 weeks), but that vaginal progesterone remains cost-effective even when a much lower treatment benefit than in the base case was assumed.

Threshold analyses

In these one-way sensitivity analyses, a single input value was varied from its base case value up to the point or threshold where screening at a cervical length of 25 mm or less would no longer be cost effective at a cost effectiveness threshold of £20,000 per QALY when compared to a baseline strategy of no screening. These threshold analyses are summarised in Table 44, showing the default base case value and the threshold value at which point an ICER of £20,000 per QALY would be exceeded.

Table 44: Base case and threshold values for cost effectiveness for model input parameters varied one at a time

Variable	Base case value	Threshold value
Relative risk preterm birth < 28 weeks	0.41	1.12
Cost of screening	£142	£1,664
Per diem treatment cost	£0.60	£1,595
Specialist midwife follow up appointment cost	£88	£187,000
Consultant obstetrician follow up appointment cost	£202	£71,000
Obstetrician Review	£123	£206,000
Scan	£86	£206,000

Discussion

The results from the base case analysis demonstrate that screening for preterm birth using ultrasound determined cervical length measurement (at 20 weeks) and treatment with vaginal progesterone is cost effective. The probabilistic analysis showed that a cervical length screening threshold of 25 mm and treatment of those identified, was both cheaper than no screening and no treatment and generated more QALYs. As shown in the plots of the Monte Carlo simulations (e.g., Figure 21) it can be seen that there is a strong inverse correlation between incremental QALY gains and incremental costs which is why dominance

is such a common outcome in most of the PSA. The larger the sampled reduction in preterm birth, the greater the reduction in adverse outcomes. This in turn results in larger QALY gains and larger “downstream” cost savings.

Considerable uncertainty exists as to the prevalence of short cervix in a population of women with twin pregnancies screened at 20 weeks. As would be expected, the sensitivity analyses show that a higher prevalence results in increasing cost-effective as a higher proportion of pregnancies benefit from treatment and the reduction in adverse outcomes. The base case analysis utilised the lowest prevalence of the 3 available cervical length frequency distributions and whilst generating a much smaller iNMB than using alternative estimates of cervical length distribution (Souka 1999; Skentou 2001), PSA suggested there was still a 99% probability that screening for preterm birth using a cervical length threshold of 25 mm or less and treatment in those identified at higher risk of preterm birth was cost-effective. Sensitivity analyses using even lower values for short cervix prevalence continued to show that intervention was cost-effective. This is explained by the fact that any absolute reduction or delay in preterm birth, no matter how small, produces positive health and QALY gains, and the averted costs of adverse outcomes associated with prematurity are substantial and will offset the low costs of screening and treatment. Therefore, even with very small differences in adverse outcomes between intervention and its comparator, the additional costs of intervention can represent good value of the money for the NHS. Clearly there exists a theoretically low short cervix prevalence at which screening and treatment of those identified would cease to be cost-effective, but the model suggests that this would have to be more than an order of magnitude lower than observed in our 3 real world estimates.

Clearly the cost-effectiveness of screening for a short cervix and treatment of those identified by screening as being at higher risk of preterm birth is predicated on treatment being successful at preventing preterm birth. The Tornado diagram presented in Figure 35 illustrates the extent to which the incremental net monetary benefit varies with different assumptions about treatment effectiveness at different gestational ages. It shows that, when varied one at a time, conclusions about the cost-effectiveness of screening for preterm birth and treatment of those identified as being at higher risk of preterm birth remain robust with respect to uncertainty in the treatment effect size.

To further stress test the cost-effectiveness conclusions of the model, 2 sensitivity analyses were undertaken using “worst case” scenarios for screening for preterm birth and treatment of those identified as being at higher risk of preterm birth. The first of these involved setting all 3 relative risks for the different gestational time periods to their least favourable value as gauged by the upper 95% confidence limit. Even in this “worst case” the cost-effectiveness decision was borderline at a cost-effectiveness threshold of £20,000 per QALY with the screening and treatment strategy having an ICER of £20,306 per QALY and an iNMB of -£1. Of course, it is highly unlikely that the true treatment effect size for all 3 variables varied would be given by the value of the upper 95% confidence limit from their sampled distribution and the uncertainty in relative treatment effectiveness is best assessed by the PSA which showed very high probabilities that intervention would be cost-effective given the best available evidence on clinical effectiveness.

An even more “worst case” scenario was assessed where, as well as using the above assumption about a lower relative treatment effect, less favourable assumptions for intervention were made about the relationship between neonatal death, post neonatal death, cerebral palsy, and respiratory distress syndrome and only in this analysis did the model find that the benefits of intervention did not warrant the additional costs with an ICER of £33,656. However, whilst these “what-if” analyses were useful for model validation, checking that unfavourable changes to parameters for intervention have the expected impact on the direction of results, they are less useful for gauging cost-effectiveness. For example, clearly there is a theoretical threshold cost of screening at which the intervention would cease to be cost-effective but there is good data from national NHS sources to indicate that the cost of

screening is much lower than this theoretical threshold value. The extent to which these threshold values can depart from plausible valuation was shown in the threshold analyses reported in Table 44. This, further to the PSA and other sensitivity analyses, gives further confidence in the robustness of the model's conclusions whilst recognising that the PSA provides a better overall assessment of cost-effectiveness in the context of parameter uncertainty.

Table 39 and Table 40 give important insights into what drives the cost-effectiveness conclusions of the model as it shows the impact that screening for preterm and treatment of those identified as being at higher risk has on important clinical outcomes having a large bearing on health-related quality of life and "downstream" costs. The absolute reduction in these outcomes is relatively small although that needs to be considered in the context of the model that suggests, based on the available data on the distribution of cervical length, that only 1% to 13% of women would be identified for treatment based on a cervical screening length threshold of 25 mm or less. The reduction in the number of adverse outcomes from such a small, treated population is derived from the best available clinical evidence (Conde-Agudelo, 2022) which indicates that treatment can substantially reduce the risk of preterm birth in women with a twin pregnancy and short cervix.

Whether a cervical length screening threshold of 25 mm or less will be cheaper and hence dominant when compared to no screening depends on whether the savings from reduced prematurity more than offset the costs of screening and treatment as well as the additional monitoring costs incurred as a result delaying or preventing preterm birth. Most of the analyses undertaken demonstrated such dominance but not all.

We recognise that this model has a number of limitations. Perhaps most importantly is the uncertainty with respect to the actual distribution of cervical length in women who will be screened as a result of this guideline's recommendation. The distributions of cervical length used in this model were derived from the published literature and personal communication but in all cases, percentages had to be estimated from a histogram bar chart. Nevertheless, it is important to recognise that this model demonstrated the cost-effectiveness of screening for preterm birth even when only 0.1% of twin pregnancies were identified as being at higher risk of preterm birth.

Pivotal to the analysis, are the modelled relationship between gestational age at birth and adverse outcomes. The model makes a simplifying assumption that the risks of adverse clinical outcomes relate solely to prematurity and that there are no independent risks from the twin pregnancy. Whilst recognising that there may be other fetal risks associated with twin pregnancy, such as fetal weight and co-existing pathologies, the purpose of the intervention is to prevent spontaneous preterm birth and the committee considered that it was reasonable to assume that preterm birth is the major concern at the point of care and that spontaneous preterm birth is the major risk to perinatal mortality and morbidity if it occurs. Therefore, much of the data used to inform the relationship between gestational age and adverse outcomes is derived from preterm singleton pregnancies in the absence of equivalent data for twin pregnancies. Whilst there was very good data for stillbirths and neonatal deaths by gestational age, somewhat crude estimates were necessary for some outcomes such as neonatal unit admissions and cerebral palsy.

Although these estimates of the risk of adverse outcomes were sourced from the literature, some additional simplifying assumptions were also required. For example, although the model includes a number of important baby outcomes known to be related to prematurity, it does not, because of complexity and the lack of available data, model all outcomes, such as all the neurodevelopmental problems that may result. Even with the outcomes that are included, the real-world relationship with gestational age is more complicated than could be modelled. For example, the severity of cerebral palsy will also be related to gestational age at birth but in the model cerebral palsy is treated as a single entity. Nevertheless, these

simplifying assumptions made with respect to gestational age and outcomes are unlikely to invalidate the findings of the model. Importantly, the fact that there is a relationship between the adverse outcomes in the model and gestational age at birth is not disputed.

Indeed, the fact that not all outcomes linked to preterm birth are included in the analysis may mean that the cost effectiveness is potentially underestimated. It is also true that there are societal benefits beyond the health care sector from reducing preterm birth that are not factored into this analysis. Although an approach not common in other NICE obstetric guidance, other NICE guidance, such as NG195, have incorporated the impact on parents' quality of life arising from NICU admission. This was not done in this guideline as it was considered to represent a caring externality which are not usually taken into account in economic evaluation. Were such benefits to be taken into account it would reinforce the cost-effectiveness conclusions of the analysis.

So, whilst the model has a number of limitations and simplifying assumptions, we think it unlikely that they result in a misleading interpretation of the cost effectiveness of screening for preterm birth and treatment in those identified as being at higher risk of prematurity. Cost effectiveness in the model is driven by evidence-based estimates on treatment efficacy and the undisputed fact that preterm birth is associated with adverse outcomes leading to reductions in health-related quality of life and "downstream" costs to the NHS.

Conclusions

The available distributions of cervical length in women with a twin pregnancy identified for the health economic model suggested that only a relatively small proportion of women pregnant with twins would be identified as at a higher risk of preterm birth by screening. Despite this, the analysis demonstrated that, with the best available clinical evidence, screening for preterm birth using a cervical length threshold of 25 mm or less followed by daily treatment with vaginal progesterone in those women identified as at higher risk of preterm birth was highly cost effective. Sensitivity analysis which subjected this finding to rigorous challenge suggested that the cost-effectiveness conclusion was robust with respect to parameter uncertainty within the model. Therefore, the result of this economic evaluation provides good cost effectiveness evidence to support the recommendations made by the committee.

In addition to providing evidence on cost effectiveness the analyses also suggested that a screening strategy using a cervical length threshold of 25 mm would be cost saving to the NHS.

References (appendix I)

Cahill 2011

Cahill A.G., Odibo A.O., Stout M.J., Grobman W.A., Macones G.A., Caughey A.B. Magnesium sulfate therapy for the prevention of cerebral palsy in preterm infants: a decision-analytic and economic analysis. Am J Obstet Gynecol. 2011 Dec;205(6):542.e1-7

Campbell 2018

Campbell, H. E., Kurinczuk, J. J., Heazell, A., Leal, J. & Rivero-Arias, O., Healthcare and wider societal implications of stillbirth: a population-based cost-of-illness study. BJOG, 125,108-117, 2018

Conde-Agudelo 2022

Conde-Agudelo, A., Rehal, A., Da Fonseca, E., Brizot, M.L., Rode, L., Serra, V., Cetingoz, E., Syngelaki, A., Tabor, A., Perales, A., Hassan, S.S., Nicolaides, K.H., Vaginal progesterone for the prevention of preterm birth and adverse perinatal outcomes in twin gestations with a short cervix: an updated individual patient data meta-analysis, Ultrasound Obstet Gynecol, 59(2), 263–266, 2022

Himpens 2008

Himpens, E., Van den Broeck, C., Oostra, A., Calders, P., Vanhaesebrouck, P. Prevalence, type, distribution, and severity of cerebral palsy in relation to gestational age: a meta-analytic review. Dev Med Child Neurol. 50(5):334-40, 2008.

Khan 2015

Khan, K.A., Petrou, S., Dritsaki, M., Johnson, S.J., Manktelow, B., Draper, E.S., Smith, L.K., Seaton, S.E., Marlow, N., Dorling, J., Field, D.J., Boyle, E.M. Economic costs associated with moderate and late preterm birth: a prospective population-based study. BJOG; 122:1495–1505; 2015

Kind 1999

Kind, P., Hardman, G. & Macran, S., Centre for Health Economics Discussion Paper 172: UK population norms for EQ-5D, Centre for Health Economics, University of York, UK, 1999

Kindinger 2016

Kindinger, L.M., Poon, L.C., Cacciatore, S., MacIntyre, D.A., Fox, N.S., Schuit, E., Mol, B.W.,Liem, S., Lim, A.C., Serra, V., Perales, A., Hermans, F., Darzi, A., Bennett, P., Nicolaides, K.H., Teoh, T.G. The effect of gestational age and cervical length measurements in the prediction of spontaneous preterm birth in twin pregnancies: an individual patient level meta-analysis. BJOG: An International Journal of Obstetrics & Gynaecology. 123(6):877-84, 2016.

Kruse 2009

Kruse, M., Michelsen, S. I., Flachs, E. M., Bronnum-Hansen, H., Madsen, M., & Uldall, P. Lifetime costs of cerebral palsy. Dev Med Child Neurol, 51(8), 622-628, 2009

Marti 2016

Marti J., Hall P., Hamilton P., Lamb S., McCabe C., Lall R., Darbyshire J., Young D., Hulme C. One-year resource utilisation, costs and quality of life in patients with acute respiratory distress syndrome (ARDS): secondary analysis of a randomised controlled trial. J Intensive Care. 2016 Aug 11;4:56.

Oskui 2013

Oskoui M., Coutinho F., Dykeman J., Jetté N., Pringsheim T. An update on the prevalence of cerebral palsy: a systematic review and meta-analysis. Dev Med Child Neurol. 2013 Jun;55(6):509-19.

Shulman 2006

Shulman, Lee S.; Vugt, John M. G. van (2006). Prenatal medicine. Washington, DC: Taylor & Francis. pp. Page 447. ISBN 0-8247-2844-0.CS1 maint:

Skentou 2001

Skentou C, Souka AP, To MS, Liao AW, Nicolaides KH. Prediction of preterm delivery in twins by cervical assessment at 23 weeks. Ultrasound Obstet Gynecol, 17(1):7-10, 2001

Souka 1999

Souka, A.P, Sevastopoulou, I., Nicolaides K.H. Cervical length at 23 weeks in twins in predicting spontaneous preterm delivery.

Yadav 2023

Yadav S., Lee B., Kamity R. Neonatal Respiratory Distress Syndrome. [Updated 2023 Jul 25]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2023 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK560779/

Appendix J Excluded studies

Excluded studies for review question: What is the clinical and cost-effectiveness of progesterone in preventing spontaneous preterm birth in twin and triplet pregnancy?

Excluded effectiveness studies

Table 45: Excluded studies and reasons for their exclusion

Study	Code [Reason]
Abdel Wahab, Ahmed S, Abdelmonaem, Mostafa I, Mahmoud, Walaa M et al. (2022) A randomized controlled trial of two-doses of vaginal progesterone 400 vs. 200 mg for prevention of preterm labor in twin gestations. Journal of perinatal medicine 50(3): 294-299	- Comparator in study does not match that specified in this review protocol *The study compared different doses of vaginal progesterone (400 mg vs. 200 mg)*
Agra, Isabela K R, Carvalho, Mario H B, Hernandez, Wagner R et al. (2019) The effect of prenatal vaginal progesterone on cervical length in nonselected twin pregnancies. The journal of maternal-fetal & neonatal medicine : the official journal of the European Association of Perinatal Medicine, the Federation of Asia and Oceania Perinatal Societies, the International Society of Perinatal Obstetricians 32(8): 1245-1249	- Outcome not relevant to this review *The effect of prenatal vaginal progesterone on cervical length was reported. The study was a secondary publication of a study included in previous guideline*
Boiko, Volodymyr I, Nikitina, Irina M, Babar, Tetyana V et al. (2018) The problem of miscarriage in multiple pregnancy. Wiadomosci lekarskie (Warsaw, Poland : 1960) 71(7): 1195-1199	- Study design not relevant to this review *Not a RCT*
Chan, Diana Man Ka, Cheung, Ka Wang, Ko, Jennifer Ka Yee et al. (2021) Use of oral progestogen in women with threatened miscarriage in the first trimester: a randomized double-blind controlled trial. Human reproduction (Oxford, England) 36(3): 587-595	- Population not relevant to this review *Only 0.2% (n=1) of participants had twin pregnancy, and there were no participants with triplet pregnancy*
Conde-Agudelo, A (2019) Pessary Compared With Vaginal Progesterone for the Prevention of Preterm Birth in Women With Twin Pregnancies and Cervical Length Less Than 38 mm: a Randomized Controlled Trial. Obstetrics and gynecology 134(2): 421-422	- Study design not relevant to this review *A comment on a RCT that compared vaginal pessary with vaginal progesterone*
Conde-Agudelo, A. (2019) Pessary Compared With Vaginal Progesterone for the Prevention of Preterm Birth in Women With Twin Pregnancies and Cervical Length Less Than 38 mm: A Randomized Controlled Trial. Obstetrics and Gynecology 134(2): 421-422	- Study design not relevant to this review *A comment on a RCT that compared vaginal pessary with vaginal progesterone*
Conde-Agudelo, Agustin, Romero, Roberto, Rehal, Anoop et al. (2023) Vaginal progesterone for preventing preterm birth and adverse perinatal	- Systematic review, included studies checked for relevance *Included studies published before 2018*

Study	Code [Reason]
outcomes in twin gestations: a systematic review and meta-analysis. American journal of obstetrics and gynecology	
Coomarasamy, A, Devall, AJ, Cheed, V et al. (2019) A Randomized Trial of Progesterone in Women with Bleeding in Early Pregnancy. The New England journal of medicine 380(19): 1815-1824	- Population not relevant to this review *Unclear whether women with twin and triplet pregnancies were included*
D'Antonio, Francesco, Berghella, Vincenzo, Di Mascio, Daniele et al. (2021) Role of progesterone, cerclage and pessary in preventing preterm birth in twin pregnancies: A systematic review and network meta-analysis. European journal of obstetrics, gynecology, and reproductive biology 261: 166-177	- Systematic review, included studies checked for relevance *Included studies published before 2018 and studies assessed cerclage and vaginal pessary*
Dang, Vinh Q, Nguyen, Linh K, Pham, Toan D et al. (2019) Pessary Compared With Vaginal Progesterone for the Prevention of Preterm Birth in Women With Twin Pregnancies and Cervical Length Less Than 38 mm: A Randomized Controlled Trial. Obstetrics and gynecology 133(3): 459-467	- Comparator in study does not match that specified in this review protocol *The study compared Arabin pessary with vaginal progesterone (400 mg once daily)*
Dodd, Jodie M, Grivell, Rosalie M, OBrien, Cecelia M et al. (2019) Prenatal administration of progestogens for preventing spontaneous preterm birth in women with a multiple pregnancy. The Cochrane database of systematic reviews 2019(11)	- Systematic review, included studies checked for relevance *Included studies published before 2018*
Easter, Sarah Rae, Little, Sarah E, Robinson, Julian N et al. (2018) Obstetric History and Likelihood of Preterm Birth of Twins. American journal of perinatology 35(11): 1023-1030	- Outcome not relevant to this review *The study investigates the association between preterm birth in a prior pregnancy and preterm birth in a twin pregnancy. This is a secondary publication of a study included in previous guideline*
Fayyaz, S., Sadaf, J., Hafeez, S. et al. (2022) Comparison of Efficacy of Cervical Cerclage and Vaginal Progesterone in the Prevention of Preterm Labour. Medical Forum Monthly 33(1): 44-47	- Comparator in study does not match that specified in this review protocol *The study compared cervical cerclage with vaginal progesterone and included participants with singleton pregnancy*
Frey, Heather A, Stout, Molly J, Abdelwahab, Mahmoud et al. (2022) Vaginal progesterone for preterm birth prevention in women with arrested preterm labor. The journal of maternal-fetal & neonatal medicine : the official journal of the European Association of Perinatal Medicine, the Federation of Asia and Oceania Perinatal Societies, the International Society of Perinatal Obstetricians 35(25): 8160-8168	- Population not relevant to this review *Participants also received other interventions (tocolytic therapy), that are given in labour and could delay preterm birth*
Johnsson, Vilma L, Pedersen, Nina G, Worda, Katharina et al. (2019) Plasma progesterone, estradiol, and unconjugated estriol concentrations in twin pregnancies: Relation with cervical length	- Outcome not relevant to this review *The association between plasma hormone concentrations, cervical length, and preterm delivery in twin pregnancies was reported*

Study	Code [Reason]
and preterm delivery. Acta obstetricia et gynecologica Scandinavica 98(1): 86-94	
Le, K D, Nguyen, L K, Nguyen, L T M et al. (2020) Cervical pessary vs vaginal progesterone for prevention of preterm birth in women with twin pregnancy and short cervix: economic analysis following randomized controlled trial. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology 55(3): 339-347	- Comparator in study does not match that specified in this review protocol *The study compared cervical pessary with vaginal progesterone*
Lin, X and Nie, Y (2022) Pregnant populations which benefit from vaginal progesterone for preventing preterm birth < 34 weeks and neonatal morbidities: a systematic review and meta-analysis. American journal of perinatology	- Systematic review, included studies checked for relevance *Included studies published before 2018 and participants with singleton pregnancy*
Lin, Xiaobin and Nie, Yu (2022) Pregnant Populations which Benefit from Vaginal Progesterone for Preventing Preterm Birth at <34 Weeks and Neonatal Morbidities: A Systematic Review and Meta-analysis. American journal of perinatology	- Systematic review, included studies checked for relevance *Included studies published before 2018 and participants with singleton pregnancy*
Megli, Christina, Combs, C Andrew, Venkataramanan, Raman et al. (2022) Recurrent Preterm Birth Reduction by 17-Hydroxyprogesterone Caproate in Dichorionic/Diamniotic Twin Gestation. American journal of perinatology 39(11): 1183-1188	- Study design not relevant to this review *The study is a post-hoc secondary analysis of individual patient-level data from two RCTs but not an individual patient data meta-analysis and did not use the systematic approach*
Mol, B.W., Wood, S., Rode, L. et al. (2021) Evaluating Progestogens for Preventing Preterm Birth International Collaborative (EPPPIC): Meta-analysis of Individual Participant Data from Randomised Controlled Trials. Obstetrical and Gynecological Survey 76(8): 464-466	- Study design not relevant to this review *Editorial comment on an included study (EPPPIC Group 2021)*
Mourad, Mirella, Too, Gloria, Gyamfi-Bannerman, Cynthia et al. (2021) Hypertensive disorders of pregnancy in twin gestations complicated by gestational diabetes. The journal of maternal-fetal & neonatal medicine : the official journal of the European Association of Perinatal Medicine, the Federation of Asia and Oceania Perinatal Societies, the International Society of Perinatal Obstetricians 34(5): 720-724	- Outcome not relevant to this review *The study investigates the association between hypertensive disorders and gestational diabetes*
Pacagnella, Rodolfo C, Silva, Thais V, Cecatti, Jose G et al. (2022) Pessary Plus Progesterone to Prevent Preterm Birth in Women With Short Cervixes: A Randomized Controlled Trial. Obstetrics and gynecology 139(1): 41-51	- Comparator in study does not match that specified in this review protocol *The study compared pessary plus progesterone with vaginal progesterone alone, and data not reported separately for those with twin pregnancy*
Romanenko, Tamara G and Sulimenko, Olha M (2020) Prevention of preeclampsia in women with	- Study design not relevant to this review

Study	Code [Reason]
multiple pregnancy after assisted reproduction. Wiadomosci lekarskie (Warsaw, Poland : 1960) 73(3): 494-497	*Not a RCT. The study used correlation analysis and assessed the combination of vaginal progesterone and other interventions (that is, angioprotector diosmin and acetylsalicylic acid).*
Romero, R., Conde-Agudelo, A., Rehal, A. et al. (2022) Vaginal progesterone for the prevention of preterm birth and adverse perinatal outcomes in twin gestations with a short cervix: an updated individual patient data meta-analysis. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology 59(2): 263-266	- Duplicate reference *A duplicate of an included study (Conde-Agudelo 2022)*
Romero, R, Conde-Agudelo, A, Rehal, A et al. (2021) Vaginal progesterone for prevention of preterm birth and adverse perinatal outcomes in twin gestation with a short cervix: updated individual patient data meta-analysis. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology	- Duplicate reference *A duplicate of an included study (Conde-Agudelo 2022)*
Romero, R, Conde-Agudelo, A, Rode, L et al. (2021) Vaginal progesterone in twin gestation and a short cervix: revisiting an individual patient data systematic review and meta-analysis. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology	- Study design not relevant to this review *Letter to the Editor*
Tran, Van T T, Nguyen, Nghia A, Nguyen, Nam T et al. (2023) Development of children born to women with twin pregnancies treated with cervical pessary or vaginal progesterone: Follow-up of a randomized controlled trial. Acta obstetricia et gynecologica Scandinavica 102(5): 626-634	- Comparator in study does not match that specified in this review protocol *The study compared Arabin pessary with vaginal progesterone*

RCT: randomised controlled trial

Excluded economic studies

No economic evidence was identified for this review.

Appendix K Research recommendations – full details

Research recommendation for review question: What is the clinical and cost-effectiveness of progesterone in preventing spontaneous preterm birth in women with twin and triplet pregnancies with a history of previous preterm birth?

K.1.1 Research recommendation

What is the clinical and cost-effectiveness of progesterone in preventing spontaneous preterm birth in women or pregnant people with twin and triplet pregnancies and a history of previous preterm birth?

K.1.2 Why this is important

Spontaneous preterm birth occurs more frequently in twin and triplet pregnancies than in singleton pregnancies, and women or pregnant people with twin and triplet pregnancies who also have a history of preterm birth are thought to be at an increased risk. Current NICE guidance recommends progesterone to prevent preterm birth in women or pregnant people with singleton pregnancies and a history of preterm birth and/or a short cervix, but the evidence for the use of progesterone in twin or triplet pregnancies has only confirmed benefit in those women or pregnant people with a short cervix. It is therefore important to assess the benefits and harms associated with progesterone use in twin and triplet pregnancies with a history of preterm birth.

K.1.3 Rationale for research recommendation

Table 46: Research recommendation rationale

Importance to 'patients' or the population	There is very little information on the effectiveness of progesterone in preventing spontaneous preterm birth in women or pregnant people with twin and triplet pregnancies with a history of previous preterm birth. New research may provide evidence to change current care by offering progesterone to reduce the risk of spontaneous preterm birth in this population.
Relevance to NICE guidance	This question would potentially change guidance in terms of if progesterone should be given to those with twin and triplet pregnancies with a history of previous preterm birth.
Relevance to the NHS	Recommendations in this area may reduce the likelihood of spontaneous preterm birth which is associated with an increased risk of morbidity and mortality, and this will have implications on NHS resources.
National priorities	High
Current evidence base	Evidence from 1 RCT (Rehal 2021) reported data for a small number of women with twin pregnancies with previous preterm birth (N=50). Evidence showed that there was no important difference between vaginal progesterone (600 mg per day) and placebo for spontaneous birth between 24 and <34 weeks. The quality of

	evidence was low. Given the lack of robust evidence, the committee did not make a recommendation for this group and decided to make a research recommendation to help inform future guidelines.
Equality considerations	None known

RCT: randomised controlled trial

K.1.4 Modified PICO table

Table 47: Research recommendation modified PICO table

Population	Inclusion: Women or pregnant people at risk of preterm birth in twin and triplet pregnancy with a history of previous preterm birth Exclusion: • Singleton pregnancies with a history of previous preterm birth • Women or pregnant people with a quadruplet or higher-order pregnancy
Intervention	• Vaginal progesterone (suggested dose: 200 mg once daily) • Oral progesterone (dose to be decided)
Comparator	With control or with each other
Outcome	Primary outcomes • Stillbirth or neonatal death • Preterm birth at 22+0 - 27+6 weeks • Preterm birth at 28+0 - 31+6 weeks • Preterm birth at 32+0 - 36+6 weeks • Spontaneous preterm birth <34 weeks of gestation Secondary outcomes • Serious neonatal complications (for example, severe necrotising enterocolitis stages 2–3, intraventricular haemorrhage grades 3–4, retinopathy of prematurity stage 3 or worse, bronchopulmonary dysplasia, confirmed sepsis, patent ductus arteriosus, and neonatal infection) • Adverse maternal outcomes (for example, gestational hypertension, pre-eclampsia, gestational diabetes, and maternal infection including chorioamnionitis)
Study design	• RCTs • Cohort studies (prospective and retrospective) • Routinely collected NHS or registry data. Study should be adequately powered and use a large sample size.
Timeframe	Follow up to 2 years after birth

Additional information	Sub-group analysis: • Women or pregnant people with a short cervix • Women or pregnant people with a long cervix

RCT: randomised controlled trial