BJOG An International Journal of Obstetrics and Gynaecology

Royal College of Obstetricians & Gynaecologists

DOI: 10.1111/1471-0528.16582
www.bjog.org

Systematic Review

The long-term effect of prenatal progesterone treatment on child development, behaviour and health: a systematic review

NE Simons,[a] M Leeuw,[a] J van't Hooft,[a] J Limpens,[b] TJ Roseboom,[a,c] MA Oudijk,[a] E Pajkrt,[a] MJJ Finken,[d] RC Painter[a]

[a] Department of Obstetrics, Amsterdam Reproduction and Development Research Institute, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands [b] Medical Library, Research Support, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands [c] Department of Epidemiology and Data Science, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands [d] Department of Paediatric Endocrinology, Emma Children's Hospital, Amsterdam UMC, Amsterdam, The Netherlands
Correspondence: NE Simons, Department of Obstetrics and Gynaecology, Amsterdam UMC – Location AMC, Room H4-240, Meibergdreef 9, 1105 AZ Amsterdam, The Netherlands. Email: n.e.simons@amsterdamumc.nl

Accepted 22 October 2020. Published Online 28 November 2020.

Background Progesterone is widely used in prenatal care. However, long-term effects of prenatal progesterone treatment on child development are unclear.

Objectives To evaluate long-term outcomes in children after prenatal progesterone treatment.

Search strategy MEDLINE, Embase and Cochrane Central Register of Controlled Trials from inception to 24 May 2020.

Selection criteria Randomised controlled trials (RCTs) reporting outcomes in children born to women who received progesterone treatment (compared with placebo or another intervention) during any trimester in pregnancy.

Data collection and analysis Two authors independently selected and extracted data. We used the Cochrane Risk of Bias tool for randomised trials and Quality In Prognosis Studies.

Main results Of 388 papers, we included seven articles based on five RCTs, comprising 4222 measurements of children aged 6 months to 8 years. All studies compared progesterone to placebo in second and/or third trimester for the prevention of preterm birth. Meta-analysis (two studies, n = 890 children) showed no difference in neurodevelopment as assessed by the Bayley-III Cognitive Composite score at 2 years between children exposed to progesterone versus placebo (Standardised Mean Difference −0.04, 95% Confidence Interval −0.26 to 0.19), I^2 = 22%. Heterogeneity prohibited additional meta-analyses. Other long-term outcomes showed no differences.

Conclusions Our systematic review comprising a multitude of developmental measurements with a broad age range did not find evidence of benefit or harm in offspring prenatally exposed to progesterone treatment for the prevention of preterm birth. We identified an urgent need for follow-up studies of prenatal progesterone administration in early pregnancy and effects in offspring beyond early childhood.

Keywords Development, Follow up, long-term, luteal phase support, preterm birth prevention, progesterone, systematic review.

Tweetable abstract Progesterone to prevent preterm birth: no effect on child development. Outcomes after first trimester progesterone are unclear.

Linked article This article is commented on by C Vedel and L Rode, p. 975 in this issue. To view this mini commentary visit https://doi.org/10.1111/1471-0528.16608.

Please cite this paper as: Simons NE, Leeuw M, van't Hooft J, Limpens J, Roseboom TJ, Oudijk MA, Pajkrt E, Finken MJJ, Painter RC. The long-term effect of prenatal progesterone treatment on child development, behaviour and health: a systematic review. BJOG 2021;128:964–974.

Introduction

Progesterone is a crucial hormone in the establishment and maintenance of pregnancy. In early pregnancy, progesterone is produced by the corpus luteum and suppresses the maternal immune system, enabling the embryo's survival.[1,2] Later in pregnancy, progesterone is produced by the placenta and plays a role in the relaxation of smooth muscle cells, ensuring myometrial quiescence until delivery.[1,3] Due to the potentially beneficial characteristics of progesterone (as natural progesterone or synthetic

Presentation: Presented as poster at the ISUOG Virtual World Congress on Ultrasound in Obstetrics and Gynecology 2020.

17-alpha hydroxyprogesterone caproate [17-OHPC)]) it has been broadly employed in reproductive medicine and obstetrics as luteal phase support after assisted reproductive technologies[4,5], as a strategy for (recurrent) miscarriage prevention or treatment[6,7] and for preterm birth prevention.[8–12] Despite the fact that progesterone is widely used during pregnancy, the long-term effects of fetal exposure to exogenous progesterone on child development have barely been investigated.

Endogenous and exogenous progestins are able to penetrate the blood-brain barrier and enter the fetal brain.[13,14] In several animal models, binding of progesterone, or its neurosteroid active metabolites, to progesterone receptor isoforms (which are ubiquitously expressed across the fetal brain) were found to stimulate processes such as neuronal growth, myelination and neural circuitry formation.[15] Unlike progesterone, 17-OHPC is not metabolised to the neuroprotective compound allopregnanolone. Furthermore, progesterone has the potential to downregulate the placental barrier enzyme 11β-hydroxysteroid-dehydrogenase type 2, thereby increasing fetal exposure to excess maternal cortisol. Cortisol excess in the fetus may have long-term repercussions for the limbic system.[16,17]

To date, 92 randomised controlled trials (RCTs) have been included in multiple Cochrane systematic reviews[4,10–12,18–20] addressing the efficacy of progesterone treatment in reproductive medicine and obstetrics; however, only two reviews have included long-term effects and they have reported that there is limited information on childhood outcomes.[10,11] An individual patient data meta-analysis (IPD-MA)[8] concluded that vaginal progesterone for the prevention of preterm birth had no harmful effects on child development, based upon results from only one RCT. Furthermore, it has been argued that natural progesterone administered in early pregnancy has an established safety profile in terms of embryo-fetal viability.[1] However, this was not backed up by systematic evaluation of outcomes beyond birth.[1]

The aim of this systematic review was to assess published literature on the effects of prenatal progesterone treatment, in human offspring beyond the perinatal period.

Methods

Study design

This systematic review was reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses checklist (PRISMA).[21,22] The review protocol was registered in PROSPERO (CRD42019142422).[23] The importance of long-term developmental outcomes has been shown in a core outcome set for preterm birth studies published in 2016.[24] In that core outcome set, late neurodevelopmental morbidity was one of the final 13 core outcomes. In another project organised by our team (not published), mothers of prematurely born children who participated in a focus group meeting expressed concerns regarding several developmental outcomes of their children later in life, with cognitive development as the most important outcome.

Information sources & search strategy

An information specialist (JL) performed a systematic search in OVID MEDLINE, Ovid Embase and the Cochrane Central Register of Controlled Trials from inception to 24 May 2020 to identify RCTs evaluating long-term effects of progesterone use for any indication during pregnancy in offspring. Search terms included controlled terms (i.e. MeSH terms in MEDLINE) and free text terms for progesterone, pregnancy/prenatal/maternal/fetus and offspring, infant, specific outcomes or long-term follow up. No language or date restrictions were applied. The complete search strategies are presented in Appendix S1. All RCTs retrieved from the search and/or included in Cochrane Reviews on progesterone in pregnancy were checked in Web of Science for follow-up studies.

Study selection

All RCTs that evaluated long-term outcomes in children (singletons and multiples) born to women who received progesterone treatment (compared with placebo or another intervention) during any trimester in pregnancy, and for any indication, were included. All studies were independently screened for eligibility on title and abstract by two reviewers (NS and ML) using Rayyan Citation.[25] Potentially eligible papers were independently assessed in full text by the same two reviewers according to the inclusion and exclusion criteria set beforehand. Any inconsistencies were discussed until consensus was reached. The reasons for exclusion are listed and described in Figure 1.

Data extraction

A piloted data extraction form was used by two review authors (NS and ML) independently. As there is no specific core outcome set for long-term follow up studies, all outcomes after hospital discharge, with no age limit at follow up, were included. If relevant data were not presented in the article, authors were contacted by email.

Multiple reporting in the same population

We included articles reporting on long-term follow-up results from the same population if they were separately published, to make sure that all outcomes assessed in these populations were included. Inclusion in the meta-analysis was assessed per outcome. In case of duplicate reporting of outcome(s) in the same population at the same age of assessment, only data of the study with the highest number of inclusions were used to avoid duplication of the same population per analysis. In case of duplicate reporting of

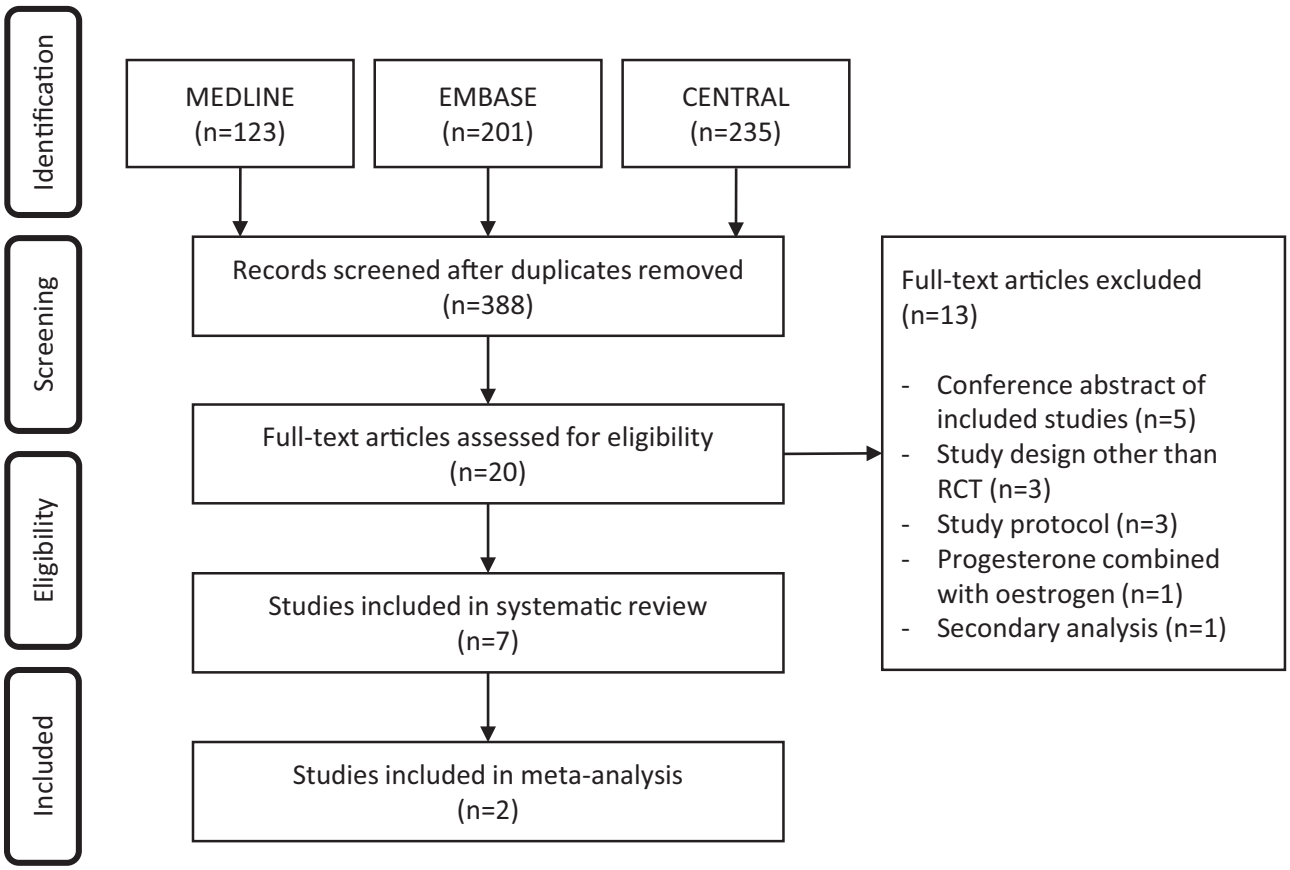

Figure 1. PRISMA 2009 flow diagram.

outcome(s) in the same population, but at different ages at assessment, analyses were categorised per age group.

Quality assessment

To assess the risk of bias in included studies we used the Cochrane Risk of Bias tool for randomised trials (RoB2) to appraise the original RCTs of the included studies.[26] As there is no specific risk of bias tool for follow-up methods and outcomes, the Quality In Prognosis Studies (QUIPS) was used to appraise included follow-up outcomes and/or studies.[27] QUIPS assesses the most important items in follow-up studies, i.e. participation (does the study address the representativeness of the study sample); attrition (did follow-up participants represent all participants enrolled in the original RCT); measurement (were measurements of the intervention/placebo and of the follow-up outcome(s) similar, reliable and valid for all participants); confounding (were there potential confounders) and statistical analysis and reporting of the study. As QUIPS is a tool designed for prognostic studies and not specifically for follow up of randomised controlled trials, we did not assess 'study confounding'. Considering that, by design, the included studies were all follow ups of RCTs, no confounding measures were expected, due to the randomisation process. However, selection bias may occur as a consequence of incomplete follow up and the effect of differences in background characteristics was assessed in the QUIPS appraisal. High risk of bias was no reason for exclusion but was considered in later quality assessment. Inconsistencies in quality appraisal were discussed until consensus was reached.

Data synthesis and analysis

The long-term outcomes were divided in three subcategories; general health and anthropometry; neurodevelopment; behaviour. Within these three subcategories, outcomes were divided per measurement instrument or diagnosis. Available dichotomous data were entered in a 2×2 table and relative risks (RR) or odds ratios (OR) were reported. For continuous data, means and standard deviations were extracted and reported. If dichotomous and continuous data from one outcome were reported, both were included in our review. Comparable domains of long-term outcomes of included studies were combined in a random effects meta-analysis. A random effects model was used because of expected substantial differences in clinical characteristics between studies, such as progesterone

treatment (e.g. administration methods, different doses) and type of pregnancy (singleton and multiple). Heterogeneity was measured with I^2 testing. High heterogeneity was considered if I^2 >75%, in which only a summary statistic was provided instead of pooling of the data. A standardised mean difference of >0.5 SD was considered clinically relevant. We used R Studio (version 1.2.1335: Boston, MA, USA).[28]

Results

Literature search

Of the 388 unique publications identified by the search, seven articles were included in this systematic review. For the PRISMA flowchart, see Figure 1.

Study characteristics

Characteristics of the RCTs and follow-up studies are summarised in Table 1. We included seven follow-up studies[29–35] based on five RCTs[32,34,36–38], all evaluating the use of progesterone versus placebo in second and/or third trimester for the prevention of preterm birth. No follow-up studies of progesterone administration in early pregnancy were found. Two RCTs included women with a multiple pregnancy[34,37] and the other three RCTs included women with a singleton pregnancy and a previous preterm delivery[36], a short cervical length[38] or other risk factors for preterm birth[32]. Included follow-up studies comprised 4222 unique measurements in children aged 6 months to 8 years old. Characteristics of included children are summarised in Table S1. Table S2 gives an overview of the statistical analyses used in the included follow-up studies.

Repeated or multiple reporting in the same population

The PREDICT study group reported neurodevelopmental outcomes of children at age 6 and 18 months (Rode 2011[34]) and up to 8 years of age (Vedel 2017[35]). Additionally, a subgroup analysis of children aged 6 and 18 months born to women with either a high-risk pregnancy (cervical length ≤10th percentile at randomisation or a history of spontaneous delivery before 34 weeks or miscarriage after 12 weeks) was performed (Klein 2011[30]). Results from the PREDICT studies could not be included in meta-analyses, due to differences in age and cut-off scores between studies using the same measurement instruments. The analyses of outcomes from the PREDICT studies did not overlap.

Unpublished data

The conference abstract[39] of the TripleP follow up by Cuijpers et al.[29] was retrieved from our systematic search and, after corresponding with the authors, we were able to use data from their submitted manuscript (n = 59 children). After corresponding with Rode et al.[34] we were able to use unpublished data on the number of children (total n = 1050 at 6 months and n = 991 at 18 months of age) assessed per treatment group and the number of children with an Ages and Stages Questionnaire score below a cut-off of 115 points at 18 months (total n = 37).

Risk of bias

We used the Cochrane Risk of Bias tool for the original RCTs.[26] Four[34,36–38] of five studies scored overall some concerns of bias, due to the absence of a published pre-specified analysis plan or protocol. Only one study[32] published its protocol and mentioned a pre-specified analysis plan in their manuscript; however, that separate statistical analysis plan could not be found online. (Figure S1A).

Follow-up methods and results were assessed using the Quality In Prognostic Studies tool.[27] Two studies[29,32] were of overall low bias. Study attrition was high in four studies[30,31,34,35] due to limited or no information on participants lost to follow up and one study[35] had high bias due to unblinding of participants before follow-up measurements. For details of QUIPS results see Figure S1B and Table S3.

General health and anthropometry

All studies used different methods and criteria for reporting general health. Table S4 shows the different components of disability or diagnosis. One study[32] found more impairments in the progesterone versus placebo group in renal (3 [1%] versus 1 [<1%]; OR 3.65, 95% CI 1.96–6.82), gastrointestinal (9 [2%] versus 4 [1%]; OR 2.67, 95% CI 1.37–5.20) and respiratory systems (7 [2%] versus 3 [1%]; OR 3.03, 95% CI 1.56–5.88). These, however, were of low frequency. Hearing was better in the progesterone group (1 [<1%] versus 2 [<1%]; OR 0.56, 95% CI 0.33–0.94).[32] All other studies showed inconsistent results in health impairments and anthropometry according to progesterone exposure, none of which achieved statistical significance.[29,31–33,35] Furthermore, congenital malformations, chromosomal anomalies or genital abnormalities reported in follow up did not differ between groups.[29,31,34,35] Due to severe heterogeneity in definitions regarding health/malformations/abnormalities, we were unable to aggregate results in meta-analyses. There were no outcomes reported regarding cardiovascular or mental health, sexual or gender orientation, or pubertal development.

Cognition and motor development

Neurodevelopment was assessed using the Bayley Scales of Infant Development-III (Bayley-III) in two studies (n = 890 measurements),[29,32] four studies (n = 2827 measurements) used the Ages and Stages Questionnaire (ASQ)[29,30,33–35] and one study (n = 324 measurements)

Table 1. Overview of included studies

Randomised controlled trial				Long-term outcomes, follow-up study				
First author and year of publication	**Study population**	**Type of intervention (vs placebo)**	**Number of women, number of children**	**First author and year of publication**	**Age at follow up and study population**	**Number of children (follow-up rate)**	**Long-term outcome measurement instruments**	**Gestational age at delivery in weeks, mean(SD) or median (IQR) - Progesterone (Prog) - Placebo (Plac)**
Meis 2003[36]	Singleton pregnancies with a history of spontaneous preterm birth between 16 and 20 wk of gestation until 36 wk or delivery	90 mg vaginal gel Crinone	459 women 459 children	Northen 2007[33]	Between 4 and 5 y All children born to mothers enrolled in the NICH[c] and HDMFMUN[c]	278 (80%)	ASQ-T PAI[c] General Health	Delivery before 37 wk gestation[e] Prog 36.3% Plac 54.9%
Norman 2009[37] STOPPIT	Multiple pregnancies between 24 wk of gestation until 34 wk or delivery	200 mg vaginal pessary Utrogestan	494 women 988 children	McNamara 2015[31] STOPPIT	Between 4 and 5 y All children born to mothers resident in Scotland	324 (44%)	HUI[α] CDI[σ] General Health - medical records	Prog 36 (3) (n = 74) Plac 36 (2) (n = 91)
Rode 2011[34] PREDICT	Multiple pregnancies (dichorionic diamniotic) between 20 and 24 wk of gestation until 34 wk or delivery	200 mg vaginal pessary Utrogestan	677 women 1342 children	Rode 2011[34] PREDICT	6 mo 18 mo All children	1050 (79.2%) 991 (74.8%)	ASQ-T	Prog 36.0 (2.8)[d] Plac 35.9 (2.7)[d]
				Klein 2011[30] PREDICT	6 mo 18 mo Children born to women with high-risk pregnancies[b]	112 (79.2%) 102 (70.8%)	ASQ-T	Prog 36.0 (2.8)[d] Plac 35.9 (2.7)[d]
				Vedel 2016[35] PREDICT	Up to 8 y The Danish part of the PREDICT study	437 (45.8%)	ASQ-T General Health - medical records	Prog 37.3 (35.1–38.0) Plac 37.0 (35.1–38.0)

Table 1. (Continued)

Randomised controlled trial				Long-term outcomes, follow-up study				
First author and year of publication	**Study population**	**Type of intervention (vs placebo)**	**Number of women, number of children**	**First author and year of publication**	**Age at follow up and study population**	**Number of children (follow-up rate)**	**Long-term outcome measurement instruments**	**Gestational age at delivery in weeks, mean(SD) or median (IQR) - Progesterone (Prog) - Placebo (Plac)**
Van Os 2015[38] TripleP	Singleton pregnancies low risk women with short cervix (<30mm) between 16 and 20 wk of gestation until 36 wk or delivery	200 mg vaginal capsules micronised progesterone	80 women 80 children	Cuijpers[29] *2020* TripleP	2 y All children	59 (77%)	Bayley-III† ASQ⊤ CBCL[a] General Health questionnaire	Prog 38.9 (37.1–40.3) Plac 38.7 (37.9–40.1)
Norman 2016/ 2018[32,52] OPPTIMUM	Singleton pregnancies with clinical risk factors[a] for preterm birth between 22 and 24 wk of gestation until 34 wk or delivery	200mg vaginal Utrogestan	1197 women 1176 children	Norman 2018[32] OPPTIMUM	2 y All children	869 (71%)	Bayley-III† SDQ‡ General Health - questionnaire	Prog 36.7 (4.1)[d] Plac 36.9 (4.2)[d]

Face-to-face tests: † Bayley-III: Bayley Scales of Infant and Toddler Development –III.
Questionnaires: ‡ SDQ: Strength and Difficulties Questionnaire, ⊤ ASQ: Ages and Stages Questionnaire, ¤ HUI: Health Utilities Index, σ CDI: Child Development Inventory, ¥ PAI: Preschool Activities Inventory, § CBCL: Child Behavior Checklist.
[a]Any history in a previous pregnancy of preterm birth, or second trimester loss, or preterm premature fetal membrane rupture, or any history of a cervical procedure to treat abnormal smears, or cervical length of 25 mm or less.
[b]Women with a cervical length ≤10th percentile at randomisation, or a history of spontaneous delivery before 34 wk or miscarriage after 12 wk, n = 72 women.
[c]NICH and HDMFMUN: National Institute of Child Health and Human Development Maternal-Fetal Medicine Units Network.
[d]Gestational age at delivery from entire cohort (gestational age at delivery, follow up of participants not available).
[e]No report of a mean/median gestational age at delivery, only dichotomous outcome; preterm delivery before 37 wk of gestation, before 35 wk of gestation and before 32 wk of gestation.

used the Child Developmental Inventory (CDI) questionnaire[31]. The Bayley-III mean cognitive composite score in two studies[29,32] could be included in meta-analysis, comprising 438 children in progesterone group versus 452 children in placebo group. No difference was found in Bayley-III scores between children exposed to progesterone and those exposed to placebo during pregnancy (standardised mean difference [SMD] of −0.04, 95% CI −0.26 to 0.19) (Figure 2). The studies showed some, acceptable heterogeneity (I^2 = 22%). Only one study[29] assessed the Bayley-III motor composite score. Due to differences in age at assessment and definition of cut-off scores, ASQ results could not be pooled in meta-analyses.[29,30,33–35] Individual results of Bayley-III motor composite scores, mean and cut-off scores of the ASQ, and cut-off scores of the CDI are summarised in Table 2; no significant differences were found.

Composite outcome of death or moderate/severe neurodevelopmental impairment

Two studies[29,32] reported a composite of moderate/severe neurodevelopmental impairment at 2 years of age, with and without mortality rates (Table 2). The first study[32] defined neurodevelopmental impairment as 'individual component of disability (motor, cognitive, function, hearing, speech and language, vision, respiratory, gastrointestinal and renal) or hospital admission', and the second study[29] defined it as 'Bayley-III score < −1 SD, or CBCL score in the clinical range, or >1 hospital admission, or >1 surgery in the past 2 years. As imputations for missing data between studies were difficult to compare, only results of complete cases could be used for reporting of the composite outcome. The first study[32] reported moderate/severe neurodevelopmental impairment among 12.4% of children (47/379) exposed to progesterone during gestation, compared with 8.7% (35/403) exposed to placebo (OR 1.48, 95% CI 0.98–2.33); the difference did not achieve statistical significance. The composite outcome of death and moderate/severe neurodevelopmental impairment occurred in 16.8% of children (67/399) exposed to progesterone, and in 12.2% of children (51/419) exposed to placebo during gestation (OR 1.45, 95% CI 0.98–2.15), again without reaching statistical significance. The second study[29] showed no difference in abnormal developmental outcome depending on progesterone exposure during gestation, 17% (5/29) in the progesterone group versus 17% (5/30) in placebo group (OR 0.97, 95% CI 0.31–2.99) and the composite outcome of death, or in abnormal developmental outcome in 15% (6/41) in progesterone group versus 18% (7/39) in placebo group (OR 0.78, 95% CI 0.24–2.58).

Behaviour

Two studies assessed behavioural development.[29,32] One study (n = 597 measurements) used the Strengths and Difficulties Questionnaire (SDQ)[32] and the other study (n = 54 measurements) the Child Behavior Checklist (CBCL)[29]. Meta-analysis was not possible due to the use of different questionnaires. No significant differences depending on progesterone exposure in utero were found in individual outcomes (Table S5).

Discussion

Main findings

In this systematic review and meta-analysis, we found no evidence that progesterone treatment for preterm birth prevention in pregnancy caused alterations in child outcomes at age 6 months to 8 years.

Strengths and limitations

The studies included in our systematic review were all of moderate to good quality, compared progesterone with placebo, and maintained double-blinding during trial. Five studies[29–32,34] maintained complete blinding during follow-up measurements and four studies[29,31–33] had low/moderate bias in four or more (of five) domains of the QUIPS. Furthermore, the follow-up rate was over 70% in five of seven studies[29,30,32–34], which is high in comparison with other follow-up studies in the field of obstetrics.[40]

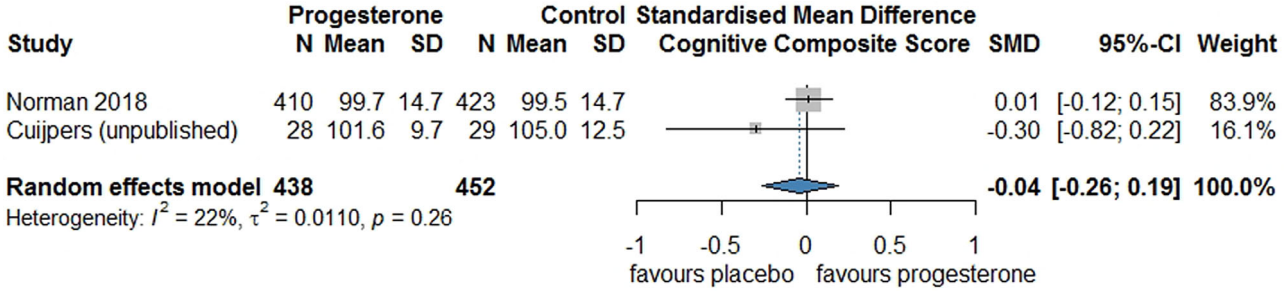

Figure 2. Forest plot of meta-analysis of unadjusted Bayley-III mean cognitive composite score.

Table 2. Summary of offspring outcomes in neurodevelopment

Neurodevelopment						
	Age at follow-up	**Score**	**Number of children progesterone/ placebo**	**Progesterone mean (SD) or n (%)**	**Placebo mean (SD) or n (%)**	**Mean difference (MD) (95% CI) or OR (95% CI) or *P*-value**
Bayley-III Cognitive Composite Score						
Norman 2018	2 y	Mean	410/423	99.7 (14.7)	99.5 (14.7)	n/a
		Mean*	430/439	97.3 (17.9)	97.7 (17.5)	MD −0.48 (−2.77 to 1.81)
Cuijpers 2020	2 y	Mean	28/29	101.6 (9.7)	105.0 (12.5)	MD −3.4 (−9.3 to 2.6)
	2 y	Cut-off ≤1SD	29/30	1 (3.6)	1 (3.4)	OR 1.04 (0.06–17.43)
Bayley-III Motor Composite Score						
Cuijpers 2020	2 y	Mean	27/29	102.4 (10.9)	107.3(12.6)	MD −4.9 (−11.2 to 1.4)
	2 y	Cut-off ≤ 1SD	27/29	0	0	n/a
Ages and Stages Questionnaire						
Rode 2011	6 mo	Mean	514/536	215 (37.5)	218 (36.7)	*P*-value 0.45
Rode 2011	18 mo	Mean	501/490	193 (42.6)	194 (40.6)	*P*-value 0.89
	18 mo	Cut-off <15 points	501/490	19 (3.8)	18 (3.7)	
Vedel 2016	48 or 60 mo	Mean	225/212	269.0 (28.2)	261.7 (31.4)	*P*-value 0.03
	48 or 60 mo	Cut-off <10th percentile	225/212	14 (6.2)	26 (12.3)	OR 0.47 (0.21–1.06)
Northen 2007	4 y	Cut-off 2 SD	193/82	53 (27.5)	23 (28)	*P*-value 0.92
Cuijpers 2020	2 y	Mean	27/27	250.7 (34.7)	256.7 (30.6)	MD 5.98 (−11.89 to 23.86)
	2 y	Cut-off 1 SD and −2 SD**	27/27	5 (18.5)	5 (18.5)	OR 1.00 (0.33–3.06)
Child Developmental Inventory score						
McNamara	5 y	≥1.5SD	140/184	60 (43)	104 (57)	OR 0.67 (0.35–1.28)
2015	5 y	≥2SD	140/184	42 (30)	65 (35)	OR 0.87 (0.46–1.63)
Composite outcome of moderate-to-severe neurodevelopmental impairment at 2 y***						
Norman 2018	2 y		379/403	47 (12.4)	35 (8.7)	OR 1.48 (0.98–2.33)
Cuijpers 2020	2 y		29/30	5 (17)	5 (17)	OR 0.97 (0.31–2.99)
Composite outcome of death or moderate-to-severe neurodevelopmental impairment at 2 y**						
Norman 2018	2 y		399/419	67 (17)	51 (12)	OR 1.45 (0.98–2.15)
Cuijpers 2020	2 y		41/39	6 (15)	7 (18)	OR 0.78 (0.24–2.58)

For used statistical analyses per follow-up study, see Table S2.
*Scores imputed for deaths.
**A score of 1 SD below the normative mean in ≥2 domains or a score of 2 SD below the normative mean on at least 1 domain were considered abnormal.
***Norman defined neurodevelopmental impairment as 'individual component of disability (motor, cognitive, function, hearing, speech and language, vision, respiratory, gastrointestinal and renal), or hospital admission', and Cuijpers as 'Bayley-III score < −1 SD, or CBCL score in clinical range, or >1 hospital admission, or >1 surgery in the past 2 y'.
****As Norman et al. did not impute abnormal cut-off scores for missing data (solely mean cognitive composite score), only cut-off results without imputation are shown.

The major limitation of our review was the heterogeneity of the included studies. All studies evaluated different outcomes in children at different ages, using different measurement instruments. Data pooling was therefore possible for only one outcome, which resulted in our conclusions being based on individual trials and small numbers, increasing the likelihood of a type 1 error.

Furthermore, the majority of outcomes used by included studies were subjective. They used parent-reported questionnaires, which have the advantage that they are relatively inexpensive and highly feasible. However, these developmental screening tools are less useful in detecting mild problems and are susceptible to the parental opinion of their child. Only two studies used the Bayley-III test for a

face-to-face assessment to evaluate development in children.

Neurodevelopmental impairments are common after preterm birth.[41,42] Even though there is still considerable heterogeneity between progesterone trials, a large recent IPD-MA showed a reduction of preterm birth after progesterone in singleton pregnancies at high risk for preterm birth.[43] Therefore, we anticipated improvements in neurodevelopmental outcomes among prenatally progesterone exposed children. However, in four[32,34,37,38] of five studies, preterm birth rate was comparable between the progesterone and the placebo group. Subsequently, our review included predominantly outcomes of children in studies with null-findings, which in turn presents an important limitation in the external validity of our conclusions and generalisability of our results to children who benefitted, in terms of preterm birth reduction, from prenatal progesterone. On the other hand, the similarities in gestational age between exposed and unexposed groups across studies allowed us to study the impact of progesterone, without taking the effect of prematurity into account.

Lastly, our search only retrieved RCTs evaluating progesterone use in the second and/or third trimester for the prevention of preterm birth. We found no evidence of long-term effects after progesterone treatment in the first trimester.

Interpretation

It is important to consider how many women use progesterone nowadays. International guidelines advise progesterone as luteal phase support for assisted reproductive technologies (ART) and for preterm birth prevention. The exact percentage of pregnancies in which progesterone is employed is estimated at 5–12% of all pregnancies (around 2–3% of babies are born through ART in western countries[5] and the incidence of preterm birth is 5–12%).[44,45]

The expanding list of indications for progesterone in pregnancy have led to growing numbers of children being exposed to progesterone at various stages of pregnancy. Two Cochrane systematic reviews on progesterone[10,11] included long-term outcomes of five studies,[30,31,33–35] but had to conclude that information relating to longer-term childhood outcomes are still limited. No meta-analyses were performed on long-term outcomes in these systematic reviews.

Our **systematic review found seven articles** evaluating effects of progesterone in second or third trimester for preterm birth prevention, and however, no RCTs evaluated effects of progesterone in first trimester for luteal phase support. Progesterone is frequently employed for this indication, with over 90% of ART cycles worldwide reporting the use of progesterone as luteal phase support in 2013.[5] More importantly, trials evaluating progesterone as luteal phase support in ART generally consider ongoing pregnancy or live birth rate as primary outcome(s), with limited neonatal outcomes reported.[4,5,46] Recent observational studies found an increased rate of large for gestational age or macrosomia in singletons after frozen embryo transfer with a programmed cycle (oestrogen and progesterone) and luteal phase support, as compared to natural or stimulated cycles. Furthermore, more post-term births and maternal hypertensive disorders were seen.[47–49] However, these results are mostly from retrospective data registries and systematic evaluation is minimal. Furthermore, it may be difficult to determine whether effects are progesterone related or ART related, although ample RCTs on this topic have been performed, which, if followed-up, could fulfil this knowledge gap.[4] Other indications for progesterone in first trimester are prevention of recurrent miscarriage or bleeding in early pregnancy. Two recent trials investigating these indications did not find beneficial effect of progesterone.[6,7] Without systematic evaluation of short-term and long-term effects of progesterone in first trimester (preferable after RCTs), safety of progesterone as a frequently used treatment cannot be guaranteed.[1]

Furthermore, it is important to consider that our results are based on a limited number of heterogeneous studies, precluding aggregation of evidence. This review again stresses the importance of structured long-term follow-up after perinatal intervention studies, preferably using a core outcome set as has been established for preterm birth prevention research.[24] Late neurodevelopmental morbidity is defined as one of the final 13 core outcomes in the core outcome set of preventive interventions for preterm birth.[24] However, there is no consensus on the measurements and outcomes that define long-term neurodevelopmental morbidity. Furthermore, before the core outcome set was published, only 16% of all obstetric trials followed children after discharge from the hospital.[40] Although the number of studies reporting long-term outcomes is rapidly increasing, methods and outcomes still vary remarkably. Our study illustrates the urgent need for guidance in the recommended measures to be used to assess the core outcome of late neurodevelopmental morbidity.[50] Despite the fact that some of the included studies measured the same outcome, the fact that different measures were used, still resulted in our inability to perform meta-analyses, which could be considered research waste.[51]

Conclusion

In this systematic review evaluating the long-term effect of prenatal progesterone treatment in children, we found no evidence of long-term beneficial or harmful effects after administration of progesterone in the second and/or third trimester for preterm birth prevention. Our findings are highly relevant due to the increasing and widespread implementation of progesterone treatment in preterm birth prevention guidelines. Although our findings raise no concerns

about any possible harmful effects of progesterone in mid- to late pregnancy for preterm birth prevention, we have identified an urgent need for assessment of effects in offspring after progesterone administration in early pregnancy.

Disclosure of interests
None declared. Completed disclosure of interests forms are available to view online as supporting information.

Contribution to authorship
NS and ML contributed to protocol development, data collection, data analysis, interpretation and writing. JL developed and performed the literature search. RP conceived the study. JvtH, TR, MO, EP, MF and RP contributed to protocol development, data analysis, data interpretation and writing. All authors were sent the paper as prepared for submission and given the opportunity to comment on the draft manuscript. All authors saw and approved the final version.

Details of ethics approval
Not applicable.

Funding
None.

Acknowledgements
Not applicable.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Appendix S1. Search strategies.

Figure S1. Risk of bias of included studies.

Table S1. Characteristics of children included in studies comparing progesterone with placebo.

Table S2. Overview of reported statistical analyses of included follow-up studies.

Table S3. Quality In Prognostic Studies appraisal (QUIPS).

Table S4. Summary of offspring general health outcomes.

Table S5. Summary of offspring behavioural outcomes. ■

References

1 Di Renzo GC, Giardina I, Clerici G, Mattei A, Alajmi AH, Gerli S. The role of progesterone in maternal and fetal medicine. *Gynecol Endocrinol* 2012;28:925–32.
2 Druckmann R, Druckmann MA. Progesterone and the immunology of pregnancy. *J Steroid Biochem Mol Biol* 2005;97:389–96.
3 Christian MS, Brent RL, Calda P. Embryo-fetal toxicity signals for 17alpha-hydroxyprogesterone caproate in high-risk pregnancies: a review of the non-clinical literature for embryo-fetal toxicity with progestins. *J Matern Fetal Neonatal Med* 2007;20:89–112.
4 van der Linden M, Buckingham K, Farquhar C, Kremer JA, Metwally M. Luteal phase support for assisted reproduction cycles. *Cochrane Database Syst Rev* 2015;(10):CD009154. https://doi.org/10.1002/14651858.CD009154.pub3
5 Child T, Leonard SA, Evans JS, Lass A. Systematic review of the clinical efficacy of vaginal progesterone for luteal phase support in assisted reproductive technology cycles. *Reprod Biomed Online* 2018;36:630–45.
6 Coomarasamy A, Williams H, Truchanowicz E, Seed PT, Small R, Quenby S, et al. A randomized trial of progesterone in women with recurrent miscarriages. *N Engl J Med* 2015;373:2141–8.
7 Coomarasamy A, Devall AJ, Cheed V, Harb H, Middleton LJ, Gallos ID, et al. A randomized trial of progesterone in women with bleeding in early pregnancy. *N Engl J Med* 2019;380:1815–24.
8 Romero R, Conde-Agudelo A, Da Fonseca E, O'Brien JM, Cetingoz E, Creasy GW, et al. Vaginal progesterone for preventing preterm birth and adverse perinatal outcomes in singleton gestations with a short cervix: a meta-analysis of individual patient data. *Am J Obstet Gynecol* 2018;218:161–80.
9 Schuit E, Stock S, Rode L, Rouse DJ, Lim AC, Norman JE, et al. Effectiveness of progestogens to improve perinatal outcome in twin pregnancies: an individual participant data meta-analysis. *BJOG* 2015;122:27–37.
10 Dodd JM, Grivell RM, OBrien CM, Dowswell T, Deussen AR. Prenatal administration of progestogens for preventing spontaneous preterm birth in women with a multiple pregnancy. *Cochrane Database Syst Rev* 2017;(10):CD012024.
11 Dodd JM, Jones L, Flenady V, Cincotta R, Crowther CA. Prenatal administration of progesterone for preventing preterm birth in women considered to be at risk of preterm birth. *Cochrane Database Syst Rev* 2013;(7):CD004947.
12 Su LL, Samuel M, Chong YS. Progestational agents for treating threatened or established preterm labour. *Cochrane Database Syst Rev* 2014;(1):CD006770.
13 Paris JJ, Brunton PJ, Russell JA, Frye CA. Immune stress in late pregnant rats decreases length of gestation and fecundity, and alters later cognitive and affective behaviour of surviving pre-adolescent offspring. *Stress* 2011;14:652–64.
14 Hill M, Paskova A, Kanceva R, Velikova M, Kubatova J, Kancheva L, et al. Steroid profiling in pregnancy: a focus on the human fetus. *J Steroid Biochem Mol Biol* 2014;139:201–22.
15 Gonzalez-Orozco JC, Camacho-Arroyo I. Progesterone Actions During Central Nervous System Development. *Front Neurosci.* 2019;13:503.
16 Murphy VE, Smith R, Giles WB, Clifton VL. Endocrine regulation of human fetal growth: the role of the mother, placenta, and fetus. *Endocr Rev* 2006;27:141–69.
17 Pasqualini JR, Chetrite GS. The formation and transformation of hormones in maternal, placental and fetal compartments: biological implications. *Horm Mol Biol Clin Investig* 2016;27:11–28.
18 Haas DM, Hathaway TJ, Ramsey PS. Progestogen for preventing miscarriage in women with recurrent miscarriage of unclear etiology. *Cochrane Database Syst Rev* 2018;(10):CD003511.
19 Meher S, Duley L. Progesterone for preventing pre-eclampsia and its complications. *Cochrane Database Syst Rev* 2006;(4):CD006175.

20 Wahabi HA, Fayed AA, Esmaeil SA, Bahkali KH. Progestogen for treating threatened miscarriage. *Cochrane Database Syst Rev* 2018;(8):CD005943.

21 Higgins JPTTJ, Chandler J, Cumpston M, Li T, Page MJ.Welch VA (editors). Cochrane handbook for systematic reviews of interventions version 6.0 (updated July 2019); 2019. [www.training.cochrane.org/handbook:Cochrane].

22 Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *J Clin Epidemiol* 2009;62:1006–12.

23 Simons N, Leeuw M, Oudijk M, Limpens J, Roseboom T, Pajkrt E, et al. The effect of prenatal progesterone exposure on offspring growth, development, behavior and health. PROSPERO 2019. Available from: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42019142422

24 van't Hooft J, Duffy JM, Daly M, Williamson PR, Meher S, Thom E, et al. A core outcome set for evaluation of interventions to prevent preterm birth. *Obstet Gynecol* 2016;127:49–58.

25 Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan– a web and mobile app for systematic reviews. *Syst Rev* 2016;5:210.

26 Sterne JAC, Savovic J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019;366:l4898.

27 Hayden JA, van der Windt DA, Cartwright JL, Cote P, Bombardier C. Assessing bias in studies of prognostic factors. *Ann Intern Med* 2013;158:280–6.

28 RT. *RStudio: Integrated Development for R*. Boston: RStudio, Inc.; 2018.

29 Cuijpers CJJ, Van't Hooft J, Schneeberger C, Van Der Lee JH, Simons NE, Os MA, et al. Progesterone for preterm birth prevention in women with short cervical length: outcomes in children at 2 years. *Ultrasound Obstet Gynecol* 2020. https://doi.org/10.1002/uog.23126 [Epub ahead of print].

30 Klein K, Rode L, Nicolaides KH, Krampl-Bettelheim E, Tabor A, PREDICT Group. Vaginal micronized progesterone and risk of preterm delivery in high-risk twin pregnancies: secondary analysis of a placebo-controlled randomized trial and meta-analysis. *Ultrasound Obstet Gynecol* 2011;38:281–7.

31 McNamara HC, Wood R, Chalmers J, Marlow N, Norrie J, MacLennan G, et al. STOPPIT Baby Follow-up Study: the effect of prophylactic progesterone in twin pregnancy on childhood outcome. *PLoS One* 2015;10:e0122341.

32 Norman JE, Marlow N, Messow C-M, Shennan A, Bennett PR, Thornton S, et al. Does progesterone prophylaxis to prevent preterm labour improve outcome? A randomised double-blind placebo-controlled trial (OPPTIMUM). *Health Technol Assess* 2018;22:1–304.

33 Northen AT, Norman GS, Anderson K, Moseley L, Divito M, Cotroneo M, et al. Follow-up of children exposed in utero to 17 alpha-hydroxyprogesterone caproate compared with placebo. *Obstet Gynecol* 2007;110:865–72.

34 Rode L, Klein K, Nicolaides KH, Krampl-Bettelheim E, Tabor A. Prevention of preterm delivery in twin gestations (PREDICT): a multicenter, randomized, placebo-controlled trial on the effect of vaginal micronized progesterone. *Ultrasound Obstet Gynecol* 2011;38:272–80.

35 Vedel C, Larsen H, Holmskov A, Andreasen KR, Uldbjerg N, Ramb J, et al. Long-term effects of prenatal progesterone exposure: neurophysiological development and hospital admissions in twins up to 8 years of age. *Ultrasound Obstet Gynecol* 2016;48:382–9.

36 Meis PJ, Klebanoff M, Thom E, Dombrowski MP, Sibai B, Moawad AH, et al. Prevention of recurrent preterm delivery by 17 alpha-hydroxyprogesterone caproate. *N Engl J Med* 2003;348:2379–85.

37 Norman JE, Mackenzie F, Owen P, Mactier H, Hanretty K, Cooper S, et al. Progesterone for the prevention of preterm birth in twin pregnancy (STOPPIT): a randomised, double-blind, placebo-controlled study and meta-analysis. *Lancet* 2009;373:2034–40.

38 Van Os MA, Van Der Ven AJ, Kleinrouweler CE, Schuit E, Kazemier BM, Verhoeven CJ, et al. Preventing preterm birth with progesterone in women with a short cervical length from a low-risk population: a multicenter double-blind placebo-controlled randomized trial. *Am J Perinatol* 2015;32:993–1000.

39 Jvt H, Cuijpers C, Schneeberger C, van der Lee JH, Opmeer BC, Steenis L, et al. 861: Preventing preterm birth with progesterone in women with short cervical length, outcomes in children at 24 months of age. *Am J Obstet Gynecol* 2017;216:S492.

40 Teune MJ, van Wassenaer AG, Malin GL, Asztalos E, Alfirevic Z, Mol BW, et al. Long-term child follow-up after large obstetric randomised controlled trials for the evaluation of perinatal interventions: a systematic review of the literature. *BJOG* 2013;120:15–22.

41 MacKay DF, Smith GC, Dobbie R, Pell JP. Gestational age at delivery and special educational need: retrospective cohort study of 407,503 schoolchildren. *PLoS Med* 2010;7:e1000289.

42 Allotey J, Zamora J, Cheong-See F, Kalidindi M, Arroyo-Manzano D, Asztalos E, et al. Cognitive, motor, behavioural and academic performances of children born preterm: a meta-analysis and systematic review involving 64 061 children. *BJOG* 2018;125:16–25.

43 Stewart LA, Simmonds M, Duley L, Dietz KC, Harden M, Hodkinson A, et al. Evaluating progestogen for prevention of preterm birth international collaborative (EPPPIC): individual participant data meta-analysis. https://www.pcori.org/research-results/2017/evaluating-hormone-treatments-women-increased-risk-preterm-birth-%E2%80%93-epppic

44 Martin JA, Hamilton BE, Osterman MJK, Driscoll AK, Drake P. Births: final data for 2017. *Natl Vital Stat Rep* 2018;67:1–50.

45 Blencowe H, Cousens S, Oestergaard MZ, Chou D, Moller AB, Narwal R, et al. National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: a systematic analysis and implications. *Lancet* 2012;379:2162–72.

46 Land JA, Evers JL. Risks and complications in assisted reproduction techniques: Report of an ESHRE consensus meeting. *Hum Reprod* 2003;18:455–7.

47 Ginstrom Ernstad E, Wennerholm UB, Khatibi A, Petzold M, Bergh C. Neonatal and maternal outcome after frozen embryo transfer: Increased risks in programmed cycles. *Am J Obstet Gynecol* 2019;221:126e1–18.

48 Saito K, Kuwahara A, Ishikawa T, Morisaki N, Miyado M, Miyado K, et al. Endometrial preparation methods for frozen-thawed embryo transfer are associated with altered risks of hypertensive disorders of pregnancy, placenta accreta, and gestational diabetes mellitus. *Hum Reprod* 2019;34:1567–75.

49 Orvieto R, Kirshenbaum M, Gleicher N. Is embryo cryopreservation causing macrosomia– and what else? *Front Endocrinol (Lausanne)* 2020;11:19.

50 Doyle LW, Anderson PJ, Battin M, Bowen JR, Brown N, Callanan C, et al. Long term follow up of high risk children: who, why and how? *BMC Pediatr* 2014;14:279.

51 Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. *Lancet* 2009;374:86–9.

52 Norman JE, Marlow N, Messow CM, Shennan A, Bennett PR, Thornton S, et al. Vaginal progesterone prophylaxis for preterm birth (the OPPTIMUM study): a multicentre, randomised, double-blind trial. *Lancet* 2016;387:2106–16.