BJOG An International Journal of Obstetrics and Gynaecology

DOI: 10.1111/1471-0528.13575
www.bjog.org

Systematic review

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

LM Kindinger,[a,b] LC Poon,[b,c] S Cacciatore,[a] DA MacIntyre,[a] NS Fox,[d] E Schuit,[e] BW Mol,[f] S Liem,[g] AC Lim,[g] V Serra,[h] A Perales,[i] F Hermans,[g] A Darzi,[j] P Bennett,[a] KH Nicolaides,[c] TG Teoh[b]

[a] Institute of Reproductive and Developmental Biology, Imperial College London, London, UK [b] Fetal Medicine Unit, St Mary's Hospital, Imperial College Healthcare NHS Trust, London, UK [c] Harris Birthright Research Centre for Fetal Medicine, Kings College Hospital, London, UK [d] Maternal Fetal Medicine Associates, PLLC, New York, NY, USA [e] Stanford Prevention Research Center, Stanford University, Stanford, CA, USA [f] The Robinson Research Institute, School of Paediatrics and Reproductive Health, University of Adelaide, Adelaide, SA, Australia [g] Department of Obstetrics and Gynaecology, Academic Medical Centre, University of Amsterdam, Amsterdam, the Netherlands [h] Maternal-Fetal Medicine Unit, Instituto Valenciano de Infertilidad, University of Valencia, Valencia, Spain [i] Department of Paediatrics, Obstetrics and Gynaecology, La FE, University and Polytechnic Hospital, University of Valencia, Valencia, Spain [j] Department of Academic Surgery, St Marys Hospital, Imperial College Healthcare NHS Trust, London, UK
Correspondence: Dr Lindsay Kindinger, Institute of Reproductive and Developmental Biology, Imperial College London, Hammersmith Hospital Campus, Du Cane Road, W12 0NN London, UK. Email l.kindinger@imperial.ac.uk

Accepted 22 June 2015. Published online 1 September 2015.

Objective To assess the effect of gestational age (GA) and cervical length (CL) measurements at transvaginal ultrasound (TVUS) in the prediction of preterm birth in twin pregnancy.

Design Individual patient data (IPD) meta-analysis.

Setting International multicentre study.

Population Asymptomatic twin pregnancy.

Methods MEDLINE and EMBASE searches were performed and IPD obtained from authors of relevant studies. Multinomial logistic regression analysis determined probabilities for birth at $\leq 28^{+0}$, 28^{+1} to 32^{+0}, 32^{+1} to 36^{+0}, and $\geq 36^{+1}$ weeks as a function of GA at screening and CL measurements.

Main outcome measures Predicted probabilities for preterm birth at $\leq 28^{+0}$, 28^{+1} to 32^{+0}, and 32^{+1} to 36^{+0}.

Results A total of 6188 CL measurements were performed on 4409 twin pregnancies in 12 studies. Both GA at screening and CL had a significant and non-linear effect on GA at birth. The best prediction of birth at $\leq 28^{+0}$ weeks was provided by screening at $\leq 18^{+0}$ weeks ($P < 0.001$), whereas the best prediction of birth between 28^{+1} and 36^{+0} weeks was provided by screening at $\geq 24^{+0}$ weeks ($P < 0.001$). Negative prediction value of 100% for birth at $\leq 28^{+0}$ weeks is achieved at CL 65 mm and 43 mm at ultrasound GA at $\leq 18^{+0}$ weeks and at 22^{+1} to 24^{+0} weeks, respectively.

Conclusion In twin pregnancies, prediction of preterm birth depends on both CL and the GA at screening. When CL is <30 mm, screening at $\leq 18^{+0}$ weeks is most predictive for birth at $\leq 28^{+0}$ weeks. Later screening at $>22^{+0}$ weeks is most predictive of delivery at 28^{+1} to 36^{+0} weeks. In twins, we recommend CL screening in twins to commence from $\leq 18^{+0}$ weeks.

Tweetable abstract An individual patient meta-analysis assessing gestation and CL in the prediction of preterm birth in twins.

Linked article This article is commented on by JM O'Brien, p.885 in this issue. To view this article mini commentary visit http://dx.doi.org/10.1111/1471-0528.13710.

Please cite this paper as: Kindinger LM, Poon LC, Cacciatore S, MacIntyre DA, Fox NS, Schuit E, Mol BW, Liem S, Lim AC, Serra V, Perales A, Hermans F, Darzi A, Bennett P, Nicolaides KH, Teoh TG. 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 2016;123:877–884.

Introduction

Preterm birth is the leading cause of perinatal death and handicap in survivors.[1] The rate of preterm birth in twins is almost 10 times higher than in singletons.[2] Extensive studies in singleton pregnancies have established an inverse relation between mid-gestation sonographically measured cervical length (CL) and gestational age (GA) at birth,[3,4] hence CL provides effective screening for spontaneous preterm birth (sPTB). These studies in singletons have also highlighted that the relation between CL and sPTB is affected by the GA at screening; a short CL (<15 mm) at early GA at screening (<16 weeks) has a higher risk of sPTB compared with the finding of the same CL at later screening (≥24 weeks).[5]

Studies on the prediction of sPTB in twin pregnancies using CL are small in size and are not easily comparable. In these studies, there are large variations in the number of patients examined (56–1135), GA at screening (range 16–28 weeks), cut-offs for defining short CL (range 15–35 mm) and the GA thresholds for defining sPTB (range 28–37 weeks).[6–28] The largest study examined 1135 pregnant women attending for routine antenatal care at 22–24 weeks and reported that the risk of sPTB increases with decreasing CL, but even for women with a long cervix the risk is still substantially higher than in singleton pregnancies.[17] The study also showed that monochorionic and dichorionic twins have a similar incidence of early sPTB, once severe twin-to-twin transfusion syndrome has been excluded.[17]

Previous meta-analyses have confirmed the association between CL and the rate of sPTB in twins. Their methodology grouped these varied thresholds of CL and GA at birth, as defined in the original studies.[29,30] As such they were unable to evaluate the effect of GA at screening on the prediction of sPTB.

The objective of this study is to assess the effect of both GA at screening and CL in the prediction of sPTB in twin pregnancy. This meta-analysis of individual patient data (IPD) provides a novel assessment in which both CL and GA at screening are treated as continuous variables.

Methods

Literature search

Searches of MEDLINE, EMBASE, the Cochrane Central Register of Controlled Trials and Research Registers of ongoing trials were performed to identify relevant publications from inception to December 2014. The following keywords were searched: 'multiple pregnancy', 'preterm birth', 'cervical length' and their related terms. Identified studies were assessed for inclusion if they reported on CL measurements in the prediction of gestation at birth in twin pregnancy.

Among the studies, the observed thresholds for CL cut-offs varied, as did GA at screening and GA cut-offs for defining sPTB. To optimise analysis of a large sample size, individual patient data were sought from eligible studies. Corresponding authors were contacted via email and/or telephone and requests were made for original anonymised data for every individual in their study, specifically: (1) the exact GA at CL screening; (2) the CL measurement in millimetres; (3) the exact GA at birth in weeks and days. A request was also made for any subsequent and unpublished data at initial and follow-up correspondence. The additional variables maternal age, ethnicity, body mass index, smoking, chorionicity and parity were requested where available.

Eligibility criteria and study selection

The inclusion criteria were twin pregnancies with transvaginal measurement of CL at 15^{+0} to 28^{+6} weeks. Any studies unable to provide the specified original data for every individual were excluded. Further exclusion criteria were pregnancies with major fetal anomalies, iatrogenic PTB, twin-to-twin transfusion syndrome, intrauterine death, insertion of cervical cerclage or pessary and pre-pregnancy excisional cervical treatment.

Two authors (L.K. and L.P.) independently assessed eligible studies for methodological quality. In the case of eligible randomised controlled trials (RCTs) for PTB prevention (progesterone, pessary or cerclage), the study protocols were reviewed for the randomisation method and outcome reporting bias. Where the study method or participation randomisation was unclear, authors were contacted for written clarification. To rule out any treatment effect, IPD from RCTs were only included for pregnancies randomised to 'no-intervention'. All manuscripts were reviewed to ensure a standardised technique of CL attainment, ensuring the ultrasound examination was transvaginal with an empty bladder.

Statistical analysis

Clinical heterogeneity was assessed by reviewing differences in patient characteristics across studies. Statistical heterogeneity was assessed using Forest plots, the I^2 measure and Cochran's Q-test.[31] GA at measurement of CL were categorised into $\leq 18^{+0}$ weeks, 18^{+1} to 20^{+0} weeks, 20^{+1} to 22^{+0} weeks, 22^{+1} to 24^{+0} weeks, and $\geq 24^{+1}$ weeks. In the first analysis both CL and GA at birth were considered as continuous variables. In the second analysis, outcomes of GA at birth were grouped into the following categories of clinical significance: very early preterm ($\leq 28^{+0}$ weeks), early preterm (28^{+1} to 32^{+0} weeks), late preterm (32^{+1} to 36^{+0} weeks), and term ($\geq 36^{+1}$ weeks). Welch's t-test assessed statistical differences between groups, and Spearman's rank correlation assessed the relation between continuous variables. Bonferroni was used for multiple

correction and the threshold for significance was corrected to $P < 0.05$. A univariate statistical analysis was performed to assess whether the following clinical parameters were associated with GA at birth: maternal age, ethnicity, smoking, BMI, chorionicity, parity and study location.

A multinomial logistic regression model was generated where predicted probabilities of very early preterm, early preterm, late preterm, and term birth were calculated as a function of categorised GA at ultrasound and CL, considered as a continuous variable. Two independent sets (a training and validation set) were built from the entire cohort. The training set consisted of women randomly selected from each of the four GA at birth categories, and the remainder were included in the validation set. These were selected using the standard function 'sample' of the R statistical software package.[32] Pregnancies with repeat CL measurements from different screening gestations were accounted for, as participants allocated to the training and validation sets were not duplicates of the same pregnancy. We used multinomial log-linear model (via neural networks) to fit the data from the training set using the function 'multinom' of the R package 'nnet'. A model validation[33] was then performed, where predicted probabilities for categorised GA at birth were tested against observed proportions of GA at birth. Confusion matrices for predicted and true GA at birth were calculated to indicate the sensitivity and specificity of CL measurements in predicting sPTB.

Results

The search identified 1048 citations; following review of the articles, 23 studies met the inclusion criteria (Figure 1). The authors of 12 publications of a combined total of 3989 twin pregnancies provided IPD and were included in the study[6–17,34] (Table 1). One group, Fox et al.,[9] provided additional unpublished data on 420 twin pregnancies. Therefore, our study population included a total of 6188 transvaginal scans, performed on 4409 twin pregnancies. In all, 657 women had repeat CL measurements at various screening gestations. Authors of 11 studies, reporting on a combined total of 1846 twin pregnancies (29.5% of all eligible participants) did not respond (five studies),[24–28] could not be located, or declined to provide IPD following initial consent to contribute (six studies) (Supporting Information Table S1).[18–23] The authors of these 11 articles were contacted by phone and email up to four times over an 8-month period to maximise response rate.

Demographic and clinical characteristics of the patient cohorts are provided in Table 1. The respective mean and median at screening GA was 22^{+3} and 22^{+0} weeks (Figure S1), and GA at birth was 35^{+5} and 36^{+5} weeks. Supporting Information Figure S2 demonstrates the distribution of GA at birth, where 0.005, 2.9, 9.3, and 36.4% delivered at $\leq 22^{+0}$, $\leq 28^{+0}$, $\leq 32^{+0}$, and $\leq 36^{+0}$ weeks, respectively.

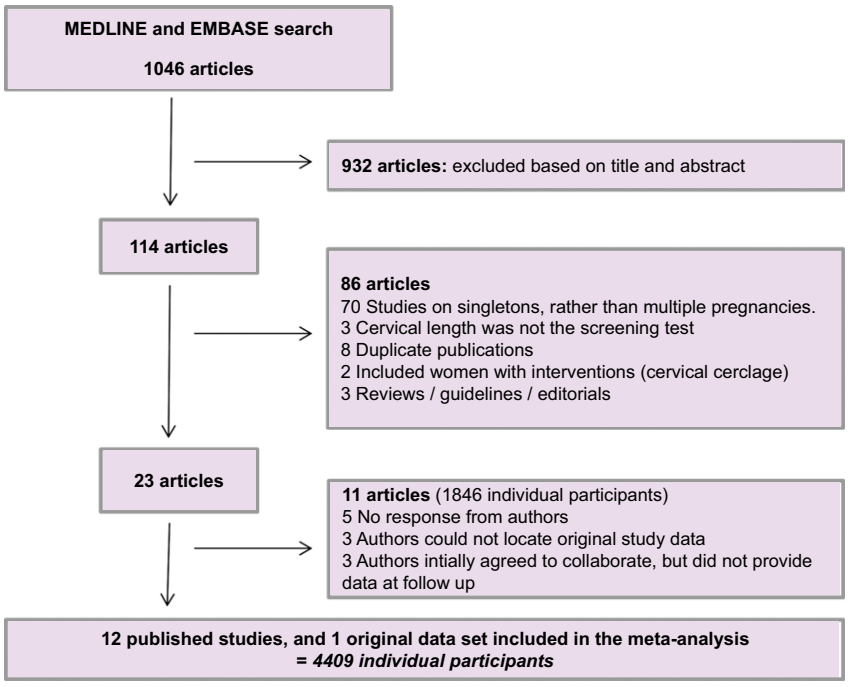

Figure 1. Search strategy flow chart.

Table 1. Study characteristics included in the IPD meta-analysis

Study	Year	Location	Study	Eligible participants, *n*	Screening GA		CL (mm)		GA at birth	
					Mean (weeks + days)	±SD (days)	Mean	±SD	Mean (weeks + days)	±SD (days)
Aboulghar et al. [6]	2008	Egypt	Observational	193	20 + 1	±2	36	±9	34 + 3	±29
Arabin et al. [7]	2006	Holland	Observational	151	22 + 4	±24	40	±9	36 + 4	±14
Brizot et al. [8]	2015	Brazil	RCT (Progesterone)	173	20 + 3	±13	38	±8	35 + 6	±19
Fox et al. [9]	2009	USA	Observational	545	22 + 2	±24	38	±9	35 + 4	±18
Hofmeister et al. [10]	2010	Brazil	Observational	385	20 + 5	±14	38	±9	35 + 5	±19
Klein et al. [11]	2008	Austria	Observational	223	23 + 0	±10	36	±8	35 + 4	±17
Liem et al. [12]	2013	Netherlands	RCT (ProTWIN – pessary)	593	18 + 5	±9	44	±9	35 + 4	±25
Lim et al. [13]	2007	Netherlands	RCT (AMPHIA – progesterone)	453	20 + 2	±16	43	±10	35 + 6	±16
Sauvanaud et al. [14]	2013	France	Observational	22	26 + 0	±12	20	±9	33 + 6	±22
Serra et al. [15]	2013	Spain	RCT (Progesterone)	151	20 + 4	±6	38	±8	36 + 3	±19
Sperling et al. [16]	2005	Denmark and Sweden	Observational	381	23 + 2	±8	38	±10	36 + 4	±18
To et al. [17]	2006	UK	Observational	1138	23 + 1	±5	35	±9	35 + 5	±20

The index of heterogeneity among studies, I^2 (total heterogeneity/total variability), was 0% and Cochran's Q-test for CL was not significant ($P = 0.923$). A graphical representation of the distribution of the GA at screening and CL data among all studies is provided in Supporting Information Figure S3.

In the first analysis, IPD from all eligible studies were assessed for an association between CL and GA at birth. A multinomial logistic regression model with validation was performed from training ($n = 400$) and validation ($n = 4009$) sets (Supporting Information Table S2). Body mass index (BMI) was the only additional variable shown to correlate significantly with GA at birth (Supporting Information Table S3); however, when incorporated into a multinomial logistic regression model with CL and GA at ultrasound, prediction for GA at birth was not improved. As demonstrated in Supporting Information Figure S4, both CL and GA at screening have a significant and non-linear effect on predicted GA at birth ($P < 0.001$). In addition there is an overall significant interaction between GA at screening and the measurement of CL ($P < 0.001$). The implications of a short CL vary depending on GA at screening, indicated by the gradient of the curves (Figure S4). A short CL taken at $\leq 20^{+0}$ weeks indicates a probability of birth significantly earlier than if the same CL was measured at a later GA.

The second analysis provides the predicted probabilities of GA at birth within each of the clinically significant categories ($\leq 28^{+0}$, 28^{+1} to 32^{+0}, 32^{+1} to 36^{+0} and $\geq 36^{+1}$ weeks) based on CL and GA at ultrasound ($\leq 18^{+0}$, 18^{+0} to 20^{+0}, 20^{+1} to 22^{+0}, 22^{+1} to 24^{+0} and $\geq 24^{+1}$ weeks). These are shown in Table 2 and illustrated in Figure 2. The odds of birth at $\leq 28^{+0}$ weeks compared with $\geq 36^{+0}$ weeks increases by 0.77 with every 1-mm decrease in CL ($P = 0.002$). This linear relation does not persist for sPTB at 28^{+0} weeks, where there is no significant association between CL and the odds of sPTB between 28^{+1} and 36^{+0} weeks (Figure 2B,C).

Screening at $\leq 18^{+0}$ weeks is most significantly predictive for sPTB at $\leq 28^{+0}$ weeks, irrespective of CL ($P < 0.001$) (Figure 2A). In contrast, to predict sPTB at 28^{+1} to 32^{+0} weeks, screening at $\leq 18^{+0}$ weeks is most predictive if CL is ≥ 20 mm; when CL is <15 mm, a later GA at screening is superior (Figure 2B). For sPTB at $\geq 32^{+1}$ weeks (Figure 2C), the most predictive screening is at $\geq 26^{+1}$ weeks. For example, for a CL of 15 mm at 24^{+0} weeks, the probabilities of sPTB at $\leq 28^{+0}$, 28^{+1} to 32^{+0} and 32^{+1} to 36^{+0} weeks are 27, 38 and 28%, respectively (Table 2, Figure 2). If the same CL of 15 mm is obtained earlier, at 18^{+0} weeks, there is a higher probability of sPTB at $\leq 28^{+0}$ weeks (45%) than at 28^{+1} to 32^{+0} weeks (38%) or 32^{+1} to 36^{+0} weeks (12%). A 100% negative prediction for sPTB at $<28^{+0}$ weeks was achieved for the following CL thresholds: 65 mm at $\leq 18^{+0}$ weeks, 55 mm at 18^{+0} to 20^{+0} weeks, 48 mm at 20^{+1} to 22^{+0} weeks, and 43 mm at 22^{+1} to 24^{+0} weeks.

Table 2. Predicted probability categories

GA (weeks)	CL (mm)	Predicted probability (%) of gestation at sPTB (weeks)			
		$\leq 28^{+0}$	28^{+1} to 32^{+0}	32^{+1} to 36^{+0}	$\geq 36^{+1}$
18	5	58.5	34.1	5.9	1.5
	10	51.8	36.8	8.5	2.9
	15	44.5	38.5	11.7	5.3
	20	36.6	38.6	15.6	9.2
	25	28.6	36.7	19.6	15.2
	30	20.9	32.6	23.1	23.4
	35	14.1	27	25.2	33.7
	40	8.9	20.7	25.6	44.8
20	5	53.2	36.4	8.5	1.9
	10	46.1	38.5	11.9	3.5
	15	38.6	39.2	16	6.2
	20	30.8	38.1	20.6	10.5
	25	23.2	35	25	16.8
	30	16.4	30.1	28.5	25.1
	35	10.8	24.1	30.2	35
	40	6.6	18	29.9	45.4
22	5	47.5	38.2	11.9	2.3
	10	40.2	39.4	16.3	4.2
	15	32.6	38.9	21.3	7.2
	20	25.2	36.6	26.5	11.8
	25	18.3	32.4	31.1	18.1
	30	12.5	27	34.3	26.2
	35	8	21.1	35.4	35.5
	40	4.8	15.4	34.4	45.3
24	5	41.6	39.2	16.4	2.8
	10	34.1	39.3	21.8	4.9
	15	26.8	37.6	27.6	8.1
	20	20	34.1	33.1	12.8
	25	14.1	29.3	37.6	19
	30	9.3	23.7	40.3	26.7
	35	5.8	18	40.7	35.4
	40	3.5	13	39	44.5
26	5	35.4	39.3	22	3.3
	10	28.2	38.1	28.3	5.5
	15	21.4	35.2	34.6	8.8
	20	15.4	30.9	40.3	13.4
	25	10.5	25.7	44.3	19.4
	30	6.8	20.3	46.3	26.6
	35	4.2	15.2	45.9	34.7
	40	2.5	10.9	43.5	43.1

In a third analysis, the predictive accuracy of the meta-regression model for sPTB was assessed. The model demonstrates improved accuracy of CL and GA at screening in predicting term compared with preterm birth; 68.2% of those predicted to deliver at $\geq 36^{+1}$ weeks were correctly classified (true negative rate) compared with 26.2, 13.3 and 36.2% correctly predicted to deliver at $\leq 28^{+0}$, 28^{+1} to 32^{+0} and 32^{+1} to 36^{+0} weeks, respectively (true positive rate) (Table S4).

Discussion

Main findings

This IPD meta-analysis has shown for the first time the importance of considering GA at screening in the prediction of sPTB in twins based on measurement of CL. Using this model, probabilities may be projected for GA at birth given any CL measurement and GA at screening. These two variables provide accurate prediction of probabilities for sPTB at $\leq 28^{+0}$ weeks, at 28^{+1} to 32^{+0} weeks and at 32^{+1} to 36^{+0} weeks.

There is an inverse and linear relation between CL and sPTB at $\leq 28^{+0}$ weeks. Uniquely, we demonstrate this relation changes to a non-linear association when predicting later sPTB at $\geq 28^{+0}$ weeks.

Strengths and limitations

The main strength of our study resides in the large number of cases examined and the methodology of the IPD meta-analysis, in which both CL and GA at screening are treated as continuous variables. A further strength is the differentiation of predicted GA at birth into very early ($\leq 28^{+0}$), early (28^{+1} to 32^{+0}) and late (32^{+1} to 36^{+0}) preterm birth. This is clinically preferable to single thresholds such as sPTB at <34 weeks as described in previous studies. Provided with the risks of very early, early and late preterm birth, a personalised and cost-effective antenatal management plan may be implemented, including optimal timing of corticosteroid administration and mobilisation to appropriate neonatal units.

The limitations of the study are the lack of data on other variables beyond CL and GA that are known to be associated with an increased risk for sPTB. Our model found that none of the other variables (maternal age, smoking, ethnicity, BMI, chorionicity and parity) predicted sPTB. In addition, although some authors commented on the presence of clinical symptoms at the time of screening, these data were not available for all individuals and therefore could not be tested.

Of the 23 eligible studies, we were unable to obtain IPD from 11 studies, equivalent to 29.5% of the potential participants (n = 1849/6258). The studies included and excluded were largely dependent on availability of raw data; several authors of older studies[21–23] did not have their data stored electronically and were unable either to locate or to transfer hard copies of patient data (accounting for 20% of unavailable IPD). Data collection may therefore be biased towards more recently published studies. Indeed, a quarter of included IPD were from studies published in the last 2 years.[8,12,14,15] A further source of potential bias is the inclusion of unpublished IPD. We had originally requested

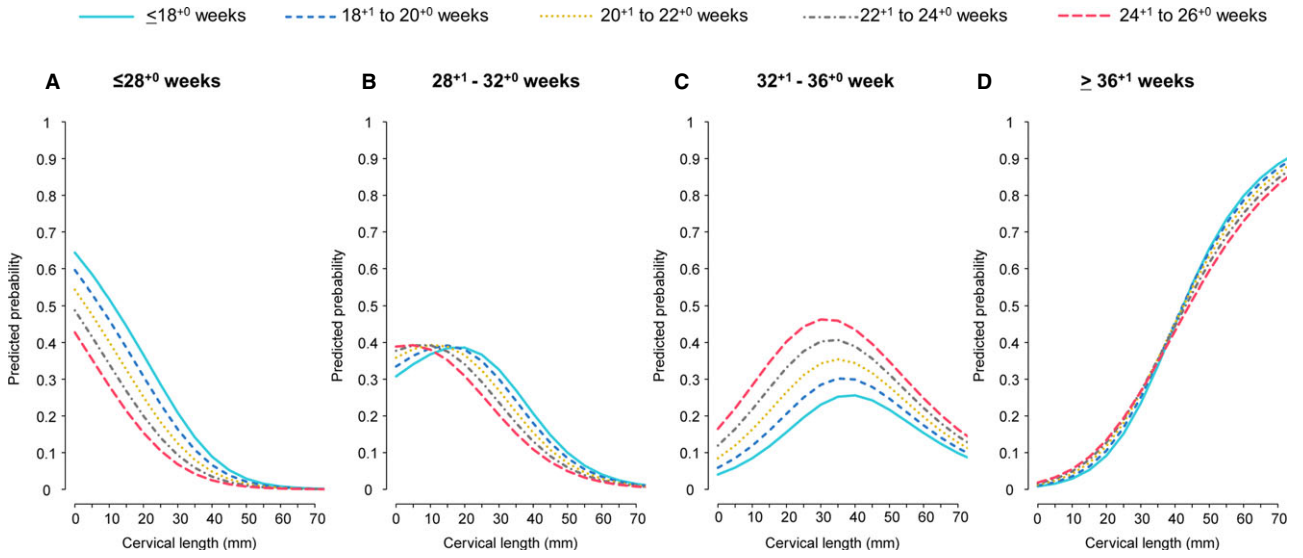

Figure 2. Predicted probability of birth at (A) ≤28^{+0} weeks, (B) 28^{+1} to 32^{+0} weeks, (C) 32^{+1} to 36^{+0} weeks and (D) ≥36^{+1} weeks based on cervical length measurements (*x*-axis) and gestational age at ultrasound screening.

additional unpublished data from all authors; however, only one author[9] offered a contribution of 450 additional participants. At heterogeneity assessment, these data were comparable to the published IPD. As our study is a collation of observational data in a multicentre international setting, it is unlikely that the missing data or the addition of unpublished data would have had a significant impact on the study findings overall.

A further limitation of our study was the inability to evaluate earlier screening at <16 weeks or the predictive value of rate of change in cervical length, as we were restricted by availability of previously collected data. Likewise the miscarriage rate at <22 weeks (0.005%) is likely to be underestimated; as the median GA at ultrasound was 22 weeks, many delivering at <22 weeks were not included in the original observational studies because they had not met the gestational criteria to receive a transvaginal ultrasound. Consequently, this analysis cannot comment on the predictive value of CL screening for late miscarriage at <22 weeks.

Future studies may consider evaluating the predictive value of CL measurement at earlier (<16 weeks) screening gestation, as well as the predictive value of rate of CL change in sequential measurements in twin pregnancies.

Interpretation in light of other evidence

Previous meta-analyses in twins confirmed the association between CL and sPTB, using thresholds of CL and GA set by the original study authors.[29,30,35] This is the first study evaluating individual patient data in the prediction of preterm in twins where CL and GA at birth are investigated as continuous variables. Our findings conclude that risk of sPTB is dependent on the GA at which the CL is obtained. Furthermore, given the GA at screening and CL, we can differentially predict the probability of early versus late preterm birth; a novel and valuable clinical tool.

Berghella et al.[5] emphasised the importance of considering GA at screening in the prediction of sPTB in singletons. A comparison between the twins in this IPD meta-analysis study and Berghella's singleton population, indicates a higher risk of sPTB in twins when comparative GA at screening and CL measurements are taken.

Currently, effective interventions for sPTB in twin pregnancies are limited. The focus of prematurity surveillance in twins remains in the antenatal preparation of targeted pregnancies considered most at risk. Timing, in particular for corticosteroid administration, is key to optimising neonatal outcome.[36] Therefore this study provides evidence justifying serial CL screening in twin pregnancy as, depending on GA at measurement, early and late sPTB may be differentially predicted. We recommend commencement of initial CL screening at ≤18^{+0} weeks with repeat screening at >22^{+0} weeks; this best identifies pregnancies at risk of rarer prematurity-associated mortality (most common with early sPTB, ≤28^{+0} weeks), as well as the more prevalent prematurity-associated morbidity from later sPTB (28^{+0} to 36^{+0} weeks).

Conclusion

This IPD meta-analysis of 12 international twin cohorts has shown that both CL and GA at screening contribute to the

prediction of GA at sPTB. To optimise prediction of preterm birth at $\leq 28^{+0}$ weeks, CL screening should commence before 18^{+0} weeks. At this stage, any CL <30 mm has a higher risk of sPTB at $\leq 28^{+0}$ weeks in twins than in singletons. Prediction of later sPTB between 28^{+0} and 36^{+0} weeks, improves with later GA, with measurement at $\geq 22^{+0}$ weeks. We therefore recommend CL screening in twins to commence from $\leq 18^{+0}$ weeks.

Disclosure of interests

None declared. Completed disclosure of interests form available to view online as supporting information.

Contribution to authorship

LK, LP, PB, KN, AD, TGT contributed to the concept of the study. LK and LP planned the study protocol and collection of data. SC, LK and LP analysed the data. NF, ES, BM, Al, AP, VS, LS, FM, FH provided published data. LK, LP, SC, KN, PB, DM and TGT wrote the manuscript. All authors reviewed and approved the final version of the manuscript.

Details of ethics approval

This study was an analysis of previously collected, anonymised data and did not require ethics approval.

Funding

The study was supported by the Imperial Healthcare NHS Trust Biomedical Research Centre (Grant Ref P45272).

Acknowledgements

We are grateful for the provision of patient level data from the following authors: M. Aboulghar, B. Arabin, M. Brizot, C. Hofmeister, K. Klein, C. Sauvanaud and L. Sperling.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1. Eligible studies, not included in the IPD meta-analysis.

Table S2. Model validation characteristics.

Table S3. Variables predictive of GA at birth.

Table S4. Cross-tabulation of predictive accuracy: percentage of deliveries within predicted gestational age groups.

Figure S1. Box and whisker plot of GA at screening (weeks) and CL measurements (mm).

Figure S2. Gestational age distribution at birth.

Figure S3. Distribution of CL measurements and GA at ultrasound by study.

Figure S4. Association between GA at birth and CL measurements according to GA at screening. ■

References

1 Health and Social Care Centre. NHS maternity statistics – 2012–13. Hospital episode statistics. 2013. [www.hscic.gov.uk/catalogue/PUB12744]. Accessed 17 March 2015.

2 ONS. Gestation-specific infant mortality in England and Wales. 2011. In: Statistics OfN, editor. [www.ons.gov.uk/ons/child-health/gestation-specific-infant-mortality-in-englang-and-wales/20112013]. Accessed 17 March 2015.

3 Iams JD, Goldenberg RL, Meis PJ, Mercer BM, Moawad A, Das A, et al. The length of the cervix and the risk of spontaneous premature delivery. National Institute of Child Health and Human Development Maternal Fetal Medicine Unit Network. *N Engl J Med* 1996;334:567–72.

4 Heath VC, Southall TR, Souka AP, Elisseou A, Nicolaides KH. Cervical length at 23 weeks of gestation: prediction of spontaneous preterm delivery. *Ultrasound Obstet Gynecol* 1998;12:312–7.

5 Berghella V, Roman A, Daskalakis C, Ness A, Baxter JK. Gestational age at cervical length measurement and incidence of preterm birth. *Obstet Gynecol* 2007;110(2 Pt 1):311–7.

6 Aboulghar MM, Aboulghar MA, Mourad L, Serour GI, Mansour RT. Ultrasound cervical measurement and prediction of spontaneous preterm birth in ICSI pregnancies: a prospective controlled study. *Reprod Biomed Online* 2009;18:296–300.

7 Arabin B, Roos C, Kollen B, van Eyck J. Comparison of transvaginal sonography in recumbent and standing maternal positions to predict spontaneous preterm birth in singleton and twin pregnancies. *Ultrasound Obstet Gynecol* 2006;27:377–86.

8 Brizot ML, Hernandez W, Liao AW, Bittar RE, Francisco RP, Krebs VL, et al. Vaginal progesterone for the prevention of preterm birth in twin gestations: a randomized placebo-controlled double-blind study. *Am J Obstet Gynecol* 2015;213:82.e1–9.

9 Fox NS, Rebarber A, Klauser CK, Peress D, Gutierrez CV, Saltzman DH. Prediction of spontaneous preterm birth in asymptomatic twin pregnancies using the change in cervical length over time. *Am J Obstet Gynecol* 2010;202:155.e1-4.

10 Hofmeister C, Brizot M, Liao A, Francisco RP, Zugaib M. Two-stage transvaginal cervical length screening for preterm birth in twin pregnancies. *J Perinat Med* 2010;38:479–84.

11 Klein K, Gregor H, Hirtenlehner-Ferber K, Stammler-Safar M, Witt A, Hanslik A, et al. Prediction of spontaneous preterm delivery in twin pregnancies by cervical length at mid-gestation. *Twin Res Hum Genet* 2008;11:552–7.

12 Liem S, Schuit E, Hegeman M, Bais J, de Boer K, Bloemenkamp K, et al. Cervical pessaries for prevention of preterm birth in women with a multiple pregnancy (ProTWIN): a multicentre, open-label randomised controlled trial. *Lancet* 2013;382:1341–9.

13 Lim AC, Schuit E, Papatsonis D, van Eyck J, Porath MM, van Oirschot CM, et al. Effect of 17-alpha hydroxyprogesterone caproate on cervical length in twin pregnancies. *Ultrasound Obstet Gynecol* 2012;40:426–30.

14 Sauvanaud C, Equy V, Faure C, Boussat B, Hoffmann P, Sergent F. Transvaginal sonographic cervical length and prediction of preterm delivery in twin pregnancies with preterm labor. *J Gynecol Obstet Biol Reprod (Paris)* 2013;42:488–92.

15 Serra V, Perales A, Meseguer J, Parrilla JJ, Lara C, Bellver J, et al. Increased doses of vaginal progesterone for the prevention of preterm birth in twin pregnancies: a randomised controlled double-blind multicentre trial. *BJOG* 2013;120:50–7.

16 Sperling L, Kiil C, Larsen LU, Qvist I, Bach D, Wojdemann K, et al. How to identify twins at low risk of spontaneous preterm delivery. *Ultrasound Obstet Gynecol* 2005;26:138–44.

17 To MS, Fonseca EB, Molina FS, Cacho AM, Nicolaides KH. Maternal characteristics and cervical length in the prediction of spontaneous early preterm delivery in twins. *Am J Obstet Gynecol* 2006;194:1360–5.

18 Goya M, Pratcorona L, Merced C, Rodó C, Valle L, Romero A, et al. Cervical pessary in pregnant women with a short cervix (PECEP): an open-label randomised controlled trial. *Lancet* 2012;379:1800–6.

19 Wennerholm UB, Holm B, Mattsby-Baltzer I, Nielsen T, Platz-Christensen J, Sundell G, et al. Fetal fibronectin, endotoxin, bacterial vaginosis and cervical length as predictors of preterm birth and neonatal morbidity in twin pregnancies. *Br J Obstet Gynaecol* 1997;104:1398–404.

20 Ehsanipoor RM, Haydon ML, Lyons Gaffaney C, Jolley JA, Petersen R, Lagrew DC, et al. Gestational age at cervical length measurement and preterm birth in twins. *Ultrasound Obstet Gynecol* 2012;40:81–6.

21 Goldenberg RL, Iams JD, Miodovnik M, Van Dorsten JP, Thurnau G, Bottoms S, et al. The preterm prediction study: risk factors in twin gestations. National Institute of Child Health and Human Development Maternal-Fetal Medicine Units Network. *Am J Obstet Gynecol* 1996;1:1047–53.

22 Guzman ER, Walters C, O'reilly-Green C, Kinzler WL, Waldron R, Nigam J, et al. Use of cervical ultrasonography in prediction of spontaneous preterm birth in twin gestations. *Am J Obstet Gynecol* 2000;183:1103–7.

23 Gibson JL, Macara LM, Owen P, Young D, Macauley J, Mackenzie F. Prediction of preterm delivery in twin pregnancy: a prospective, observational study of cervical length and fetal fibronectin testing. *Ultrasound Obstet Gynecol* 2004;23:561–6.

24 Asnafi N, Basirat Z, Hajian-Tilaki K, Dadvar S. Assessment of cervical length by transvaginal ultrasonography to predict preterm delivery in twin pregnancy. *J Matern Fetal Neonatal Med* 2013;26:1435–8.

25 Boyer A, Cameron L, Munoz-Maldonado Y, Bronsteen R, Comstock CH, Lee W, et al. Clinical significance of amniotic fluid sludge in twin pregnancies with a short cervical length. *Am J Obstet Gynecol* 2014;211:506.e1-9.

26 Senat MV, Porcher R, Winer N, Vayssière C, Deruelle P, Capelle M, et al. Prevention of preterm delivery by 17 alpha-hydroxyprogesterone caproate in asymptomatic twin pregnancies with a short cervix: a randomized controlled trial. *Am J Obstet Gynecol* 2013;208:194.e1-8.

27 Imseis HM, Albert TA, Iams JD. Identifying twin gestations at low risk for preterm birth with a transvaginal ultrasonographic cervical measurement at 24 to 26 weeks' gestation. *Am J Obstet Gynecol* 1997;177:1149–55.

28 Vayssière C, Favre R, Audibert F, Chauvet MP, Gaucherand P, Tardif D, et al. Cervical length and funneling at 22 and 27 weeks to predict spontaneous birth before 32 weeks in twin pregnancies: a French prospective multicenter study. *Am J Obstet Gynecol* 2002;187:1596–604.

29 Lim AC, Hegeman MA, Huis In TVMA, Opmeer BC, Bruinse HW, Mol BW. Cervical length measurement for the prediction of preterm birth in multiple pregnancies: a systematic review and bivariate meta-analysis. *Ultrasound Obstet Gynecol* 2011;38:10–7.

30 Conde-Agudelo A, Romero R, Hassan SS, Yeo L. Transvaginal sonographic cervical length for the prediction of spontaneous preterm birth in twin pregnancies: a systematic review and metaanalysis. *Am J Obstet Gynecol* 2010;203:128 e1–12.

31 Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;327:557–60.

32 Ripley BD. *Stochastic Simulation*. New York: Wiley-Interscience, 1987.

33 Venables W, Ripley B. *Modern Applied Statistics with S*. 4th edn. New York: Springer, 2002.

34 Lim AC, Bloemenkamp KW, Boer K, Duvekot JJ, Erwich JJ, Hasaart TH, et al. Progesterone for the prevention of preterm birth in women with multiple pregnancies: the AMPHIA trial. *BMC Pregnancy Childbirth* 2007;7:7.

35 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.

36 Roberts D, Dalziel S. Antenatal corticosteroids for accelerating fetal lung maturation for women at risk of preterm birth. *Cochrane Database Syst Rev* 2006;(3):CD004454.