Ultrasound Obstet Gynecol 2020; **55**: 391–400
Published online in Wiley Online Library (wileyonlinelibrary.com). **DOI:** 10.1002/uog.21913

Comparison of different methods of measuring angle of progression in prediction of labor outcome

A. FRICK[1], V. KOSTIV[1,2], D. VOJTASSAKOVA[1,2], R. AKOLEKAR[2,3#] and K. H. NICOLAIDES[1#]

[1]*Harris Birthright Research Centre for Fetal Medicine, King's College Hospital, London, UK;* [2]*Fetal Medicine Unit, Medway Maritime Hospital, Gillingham, UK;* [3]*Institute of Medical Sciences, Canterbury Christ Church University, Chatham, UK*

KEYWORDS: angle of progression; intrapartum ultrasound; labor; time to delivery; transperineal ultrasound; vaginal delivery

CONTRIBUTION

What are the novel findings of this work?
The method of measuring the angle of progression (AoP) by transperineal ultrasound during labor with the highest degree of reliability is the manual parasagittal method, when compared to both the manual sagittal and automated parasagittal approaches. This study developed models to predict time to delivery and need for Cesarean section (CS) because of failure to progress (FTP) in labor from maternal and pregnancy characteristics, intrapartum factors and ultrasound findings during the first stage of labor.

What are the clinical implications of this work?
Future research should focus on the parasagittal method of AoP measurement, as compared to the sagittal method. Over half of the variation in time to vaginal delivery can be explained by a model that combines maternal factors, pregnancy characteristics and ultrasound findings of AoP and fetal head position, but larger datasets and clinical validation studies are needed before clinical implementation of individualized labor curves. The ability of AoP to provide clinically useful prediction of CS for FTP in the first stage of labor is limited.

ABSTRACT

Objectives *First, to compare the manual sagittal and parasagittal and automated parasagittal methods of measuring the angle of progression (AoP) by transperineal ultrasound during labor, and, second, to develop models for the prediction of time to delivery and need for Cesarean section (CS) for failure to progress (FTP) in a population of patients undergoing induction of labor.*

Methods *This was a prospective observational study of transperineal ultrasound in a cohort of 512 women with a singleton pregnancy undergoing induction of labor. A random selection of 50 stored images was assessed for inter- and intraobserver reliability of AoP measurements using the manual sagittal and parasagittal and automated parasagittal methods. In cases of vaginal delivery, univariate linear, multiple linear and quantile regression analyses were performed to predict time to delivery. Univariate and multivariate binomial logistic regression analyses were performed to predict CS for FTP in the first stage of labor.*

Results *The intraclass correlation coefficient (ICC) for the manual parasagittal method for a single observer was 0.97 (95% CI, 0.95–0.98) and for two observers it was 0.96 (95% CI, 0.93–0.98), indicating good reliability. The ICC for the sagittal method for a single observer was 0.93 (95% CI, 0.88–0.96) and for two observers it was 0.74 (95% CI, 0.58–0.84), indicating moderate reliability for a single observer and poor reliability between two observers. Bland–Altman analysis demonstrated narrower limits of agreement for the manual parasagittal approach than for the sagittal approach for both a single and two observers. The automated parasagittal method failed to capture an image in 19% of cases. The mean difference in AoP measurements between the sagittal and manual parasagittal methods was 11°. In pregnancies resulting in vaginal delivery, 54% of the variation in time to delivery was explained in a model combining parity, epidural and syntocinon use during labor and the sonographic findings of fetal head position and AoP. In the prediction of CS for FTP in the first stage of labor, a model which combined maternal factors with the sonographic measurements of AoP and estimated fetal weight was superior to one utilizing maternal factors alone (area under the receiver-operating-characteristics curve, 0.80* vs *0.76).*

Correspondence to: Prof. K. H. Nicolaides, Fetal Medicine Research Institute, King's College Hospital, 16–20 Windsor Walk, Denmark Hill, London SE5 8BB, UK (e-mail: kypros@fetalmedicine.com)

#R.A. and K.H.N. are joint senior authors.

Accepted: 24 October 2019

Conclusions *First, the method of measuring AoP with the greatest reliability is the manual parasagittal technique and future research should focus on this technique. Second, over half of the variation in time to vaginal delivery can be explained by a model that combines maternal factors, pregnancy characteristics and ultrasound findings. Third, the ability of AoP to provide clinically useful prediction of CS for FTP in the first stage of labor is limited. Copyright © 2019 ISUOG. Published by John Wiley & Sons Ltd.*

INTRODUCTION

Since Friedman's seminal work in the 1950s, vaginal examination has formed the basis for assessing progress in labor, with cervical dilatation, fetal head position and fetal head descent (station) all recorded at each assessment and plotted serially on a graph over time (partogram)[1–4]. However, vaginal examination is subjective, imprecise, uncomfortable for women and associated with infection, leading to calls for research into new approaches for assessing progress in labor[5–11].

A number of new techniques have been described using transperineal ultrasound to monitor labor progress through measurements relating fetal head position to the maternal pelvis[12,13]. These techniques are non-invasive, are well tolerated by patients and have a high degree of inter- and intraobserver reliability[8,14–23]. The most widely studied measurement is that of the angle of progression (AoP), which is the angle between the leading part of the fetal skull and the maternal pubic symphysis[13]. The AoP correlates with clinical estimation of fetal station, with a higher AoP associated with a shorter time to delivery, and is useful in predicting successful instrumental delivery[17,18,24,25]. Studies assessing the utility of AoP in the first stage of labor in the prediction of vaginal delivery and time to delivery have been limited by small numbers[8,26–28]. One barrier to the uptake of the use of AoP in clinical practice has been a perception amongst obstetricians that the anatomical landmarks are not easy to identify for the non-expert[29]. This problem has been partially addressed with the development of automated software (Sono L&D; GE Healthcare, Zipf, Austria) which uses a different set of landmarks, namely the more hyperechogenic pubic rami seen in a slightly parasagittal view (Figure 1). A previous study assessing the automated technique found that it systematically overestimated AoP compared to the sagittal approach, but the study did not compare directly a manual parasagittal approach to the automated method[30]. No previous study has evaluated a manual parasagittal approach in predicting time to delivery or operative delivery.

The objectives of this study were, first, to compare the manual sagittal and parasagittal and automated parasagittal methods of measuring AoP by transperineal ultrasound during labor, and, second, to develop models for the prediction of time to delivery and need for Cesarean section (CS) for failure to progress (FTP) in a population of patients undergoing induction of labor.

METHODS

Study population

Women undergoing induction of labor at Medway Maritime Hospital, Gillingham, UK, between May 2016 and August 2017 were recruited into the study. The inclusion criteria were age ≥ 18 years and singleton pregnancy with a live fetus in cephalic presentation. We excluded multiple pregnancies and women with significant mental illness or learning difficulties.

Labor was induced either with a prostaglandin pessary or artificial rupture of membranes, depending on favorability of the cervix on clinical examination, and the subsequent management, which included vaginal examination every 4 h until delivery, was as recommended by national guidelines[31]. Ultrasound examination was performed immediately following the first clinical examination after the onset of regular painful contractions and admission to the delivery suite. All ultrasound examinations were conducted by doctors who had obtained The Fetal Medicine Foundation certification in obstetric

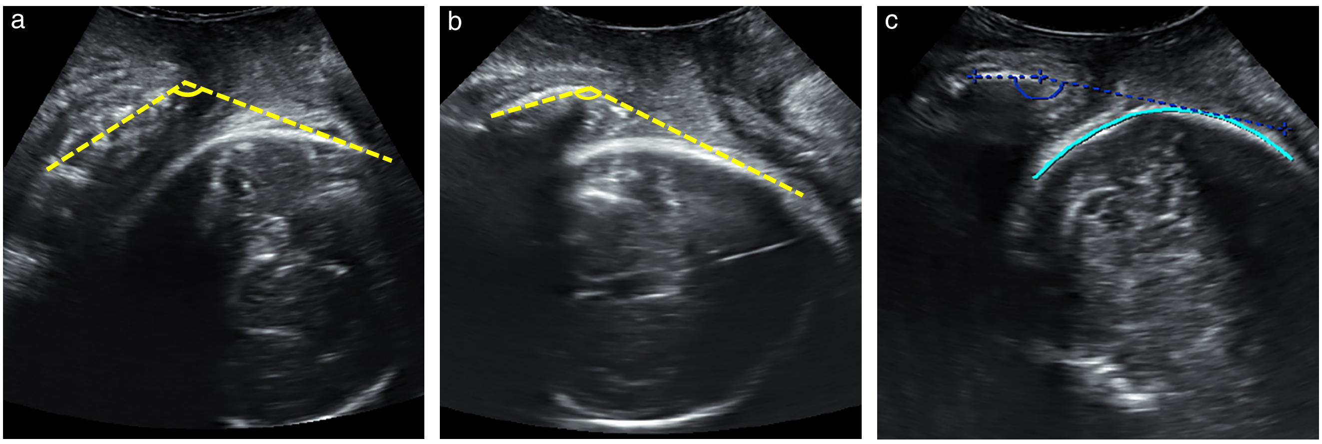

Figure 1 Ultrasound images showing representative examples of angle-of-progression measurement using manual sagittal (a), manual parasagittal (b) and automated parasagittal (c) approaches.

ultrasound and had received training in intrapartum ultrasound. Obstetricians and midwives were not made aware of the ultrasound findings. The study was approved by the London-Dulwich Research Ethics Committee (REC reference 16/LO/0367).

Maternal weight and height were measured immediately prior to induction of labor. Patient characteristics recorded included maternal age, racial origin (white, black, South Asian, East Asian or mixed), method of conception (spontaneous or assisted requiring the use of ovulation drugs) and parity (parous or nulliparous if no previous pregnancy at ≥ 24 weeks' gestation). We estimated fetal weight (EFW) from measurements of fetal head circumference, abdominal circumference and femur length obtained by transabdominal sonography the day before induction of labor[32,33].

Outcome measures

Outcome data were collected from maternal notes following delivery and stored on a secure database. Researchers collecting the data were unaware of the intrapartum ultrasound findings. Birth outcome data included gestational age at delivery, mode of delivery, indication for operative delivery and birth weight.

Intrapartum ultrasound assessment

Ultrasound measurements were taken using a portable ultrasound machine (Voluson P8; GE Healthcare) equipped with a convex 4C-RS probe. Fetal occiput position was determined using the transabdominal technique described by Akmal *et al.*[34]. Women were then placed in a modified lithotomy position with an empty bladder and the probe was covered with a glove. Two-dimensional transperineal ultrasound was performed by placing the probe vertically between the labia to obtain a sagittal view of the fetal head in relation to the maternal pubic symphysis. The exact positioning of the probe between the labia was then adjusted to obtain clear images of the pubic symphysis, typically at a 30–40° angle to the horizontal, in order to increase visualization of the maternal soft-tissue borders. The probe was then tilted the minimum amount necessary to the right or left to obtain a parasagittal view that included a clearly identifiable length of the hyperechogenic maternal pubic rami (Figure 1). Images from both the midline and parasagittal views were stored for later analysis.

The AoP between the pubic symphysis and the leading edge of the fetal skull was measured, first, on a sagittal image manually[13], second, on a parasagittal image manually and, third, on a parasagittal image by the automated technique (Sono L&D). In the manual sagittal method, a straight line was drawn through the midline of the long axis of the pubic symphysis, with the distal edge forming the vertex of the angle with the fetal head. In the manual parasagittal method, a straight line was drawn along the superior–inferior axis of the pubic bone, with the inferior end of the hyperechogenic pubic bone forming the vertex of the angle with the fetal head.

Statistical analysis

Continuous and categorical variables were compared using the Kruskal–Wallis test and χ^2 or Fisher's exact test, respectively. Normality of the distribution was assessed using probability plots and histograms. A *P*-value of < 0.05 was considered significant.

Comparison of manual and automated measurements of AoP

A random selection of stored images was assessed for inter- and intraobserver reliability for both the sagittal and parasagittal methods. Operator A remeasured 50 images of each method to create an intraobserver reliability dataset. Operator A remeasured 50 images originally measured by Operator B to create the interobserver reliability dataset. Operator A was blinded to the original measurements at the time of remeasuring. The intra- and interobserver reproducibility of measurements was examined by calculating intraclass correlation coefficients (ICC) with 95% CI[35]. Overlap between the 95% CI of two ICCs was indicative of no significant difference between them. For the interpretation of ICC values, we used published cut-off values for ultrasound measurements: ICC < 0.70, very poor reliability; 0.70–0.90, poor reliability; 0.90–0.95, moderate reliability; 0.95–0.99, good reliability; > 0.99 excellent reliability[36]. The Bland–Altman plot of the average measurement against the percentage difference between the two measurements was produced and the 95% limits of agreement (LoA) were calculated to examine the agreement and bias for a single examiner and between two examiners for each method of measurement of AoP[37]. The SD of the differences for each method for both one and two examiners was calculated and reported. The optimal method is the one in which, first, the ICCs are large and the SD of the differences between measurements is small and, second, there is no bias and the LoAs are small on the Bland–Altman plot. Additional Bland–Altman analysis between the manual sagittal and parasagittal and automated parasagittal methods using the entire dataset was performed to assess intermethod systematic bias.

Prediction of outcome

In women with vaginal delivery, univariate linear regression analysis was performed to assess the relationship between time to delivery in hours and maternal age, height, weight, racial origin, parity, gestational age, EFW, use of epidural anesthesia, use of syntocinon, cervical dilatation, fetal head position as determined by ultrasound and AoP measured using each of the three methods. Multiple linear regression analysis with backward elimination was then performed to develop parsimonious models to predict time to delivery. Repeat k-fold cross validation with 10 folds and three repeats was performed and the mean cross-validated R^2 reported to ensure models were not overfitted. Due to significant heteroscedasticity seen upon visual inspection of plotted residuals, quantile regression was performed and reported for the 5^{th}, 25^{th}, 50^{th}, 75^{th} and 95^{th} quantiles[38].

Univariate binomial logistic regression analysis was carried out to determine which of the factors from maternal and pregnancy characteristics, clinical vaginal examination and ultrasound measurements provided a significant contribution to the prediction of CS for FTP in the first stage of labor. Prior to analysis, continuous variables without a meaningful zero were centered around their median value. Multivariate logistic regression analysis with backward elimination was then used to determine if the three following models had significant contribution in predicting CS for FTP: first, maternal and pregnancy characteristics; second, maternal factors plus findings of the vaginal examination; and, third, maternal factors plus ultrasound findings. The performance of screening was determined by comparing the area under the receiver-operating-characteristics (ROC) curves (AUC).

The statistical software package R version 3.5.1 (R Foundation for Statistical Computing, Vienna, Austria) was used for data analyses.

RESULTS

Study population

The characteristics of the study population of 512 women are summarized in Table 1. There were 380 (74.2%) vaginal deliveries and 132 (25.8%) women had CS, including 59 (44.7%) for FTP in labor and 73 (55.3%) for presumed fetal compromise. Manual measurements of AoP were obtained in all 512 women, but the automated method successfully captured an image in only 416 (81.3%) cases.

Comparison of methods

The ICC for the manual parasagittal method for a single observer was 0.97 (95% CI, 0.95–0.98) and for two observers it was 0.96 (95% CI, 0.93–0.98), indicating good reliability. The respective values for the sagittal method were 0.93 (95% CI, 0.88–0.96) and 0.74 (95% CI, 0.58–0.84), indicating moderate reliability for a single observer and poor reliability between two observers. The SD of the difference in AoP measurements for the manual parasagittal method was 2.33° for a single observer and 3.01° for two observers; the respective values for the sagittal method were 4.32° and 8.77°.

For the manual parasagittal method, the 95% CIs of the intra- and interobserver ICCs overlapped, indicating no significant difference. The sagittal method demonstrated a lower degree of interobserver reproducibility with no overlap in the ICC 95% CIs when compared to the intraobserver ICC for the same method. The most reproducible results overall were from the manual parasagittal method as it demonstrated the highest ICC and the lowest SD for both inter- and intraobserver reliability.

Table 1 Maternal and pregnancy characteristics of study population of 512 singleton pregnancies undergoing induction of labor, according to mode of delivery

		Cesarean section for:	
Characteristic	*Vaginal delivery (n = 380)*	*Failure to progress (n = 59)*	*Presumed fetal distress (n = 73)*
Maternal age (years)*	28.0 (24.0–31.2)	29.0 (25.0–33.5)	29.0 (26.0–33.0)
Maternal weight (kg)*	86.0 (74.2–97.0)	93.0 (82.1–105.5)	89.0 (78.4–101.0)
Maternal height (cm)**	166 (161.7–170.0)	163 (158.9–167.0)	165 (160.8–168.8)
Racial origin			
White	351 (92.4)	57 (96.6)	64 (87.7)
Black	6 (1.6)	0 (0)	2 (2.7)
South Asian	15 (3.9)	1 (1.7)	5 (6.8)
East Asian	1 (0.3)	1 (1.7)	0 (0)
Mixed	7 (1.8)	0 (0)	2 (2.7)
Conception			
Spontaneous	374 (98.4)	58 (98.3)	70 (95.9)
Ovulation induction drugs	6 (1.6)	1 (1.7)	3 (4.1)
Parity***			
Nulliparous	157 (41.3)	46 (78.0)	55 (75.3)
Parous	223 (58.7)	13 (22.0)	18 (24.7)
Indication for induction***			
Medical	294 (77.4)	35 (59.3)	50 (68.5)
Postdates	86 (22.6)	24 (40.7)	23 (31.5)
Gestational age at delivery (weeks)***	39.6 (38.4–41.4)	40.9 (39.1–41.9)	40.4 (39.1–41.9)
Estimated fetal weight (g)***	3470 (3064–3807)	3753 (3486–4038)	3527 (3143–3811)
Birth weight (g)***	3365 (2999–3700)	3660 (3385–4007)	3500 (3025–3830)

Data are given as median (interquartile range) or *n* (%). Comparisons between outcome groups was by χ^2 test and Fisher's exact test for categorical variables and Kruskal–Wallis test for continuous variables: *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.

Bland–Altman plots demonstrating the degree of concordance between pairs of measurements made by the same observer and by the two different observers for the sagittal and manual parasagittal methods are illustrated in Figure 2. Bland–Altman plots demonstrating the degree of concordance between methods are illustrated in Figure S1. The results for the Bland–Altman analyses are presented in Tables S1 and S2.

Prediction of time to delivery

Multiple linear regression

On univariate regression analysis, significant predictors of time to delivery were AoP measured using any of the three methods, cervical dilation, maternal weight, parity, use of syntocinon and epidural in labor and sonographically determined occiput anterior position

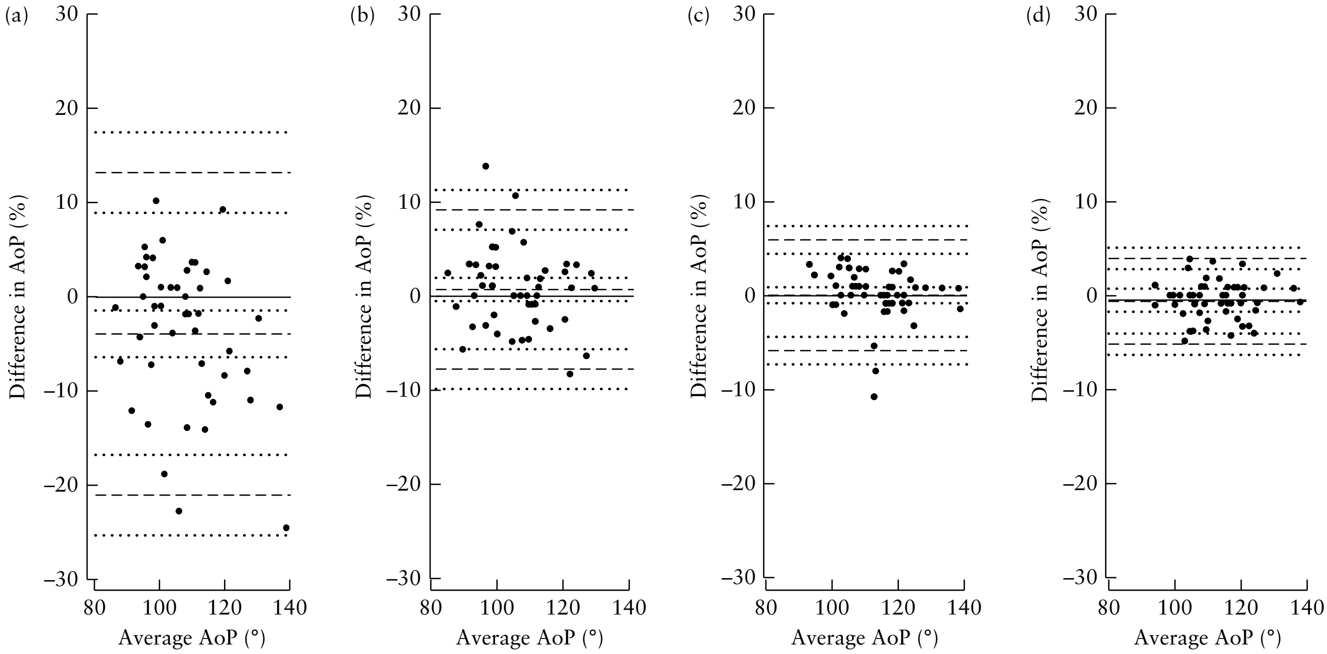

Figure 2 Bland–Altman plots demonstrating degree of concordance between pairs of measurements of angle of progression (AoP) in labor: (a) manual sagittal method, two operators; (b) manual sagittal method, single operator; (c) manual parasagittal method, two operators; (d) manual parasagittal method, single operator. Dashed lines represent mean and upper and lower limits of agreement (1.96 SD). Dotted lines represent 95% CI around means and limits of agreement.

Table 2 Multiple linear regression models for prediction of time to delivery in hours in women undergoing induction of labor

	Method of AoP measurement				
Variable	*Manual parasagittal*	*Manual sagittal*	*Automated parasagittal*	*Vaginal examination*	*Combined*†
Beta estimate					
Constant	12.07 (9.61 to 14.52)***	10.64 (8.19 to 13.08)***	12.63 (9.93 to 15.33)***	5.35 (4.57 to 6.14)***	10.03 (7.50 to 12.55)***
Nulliparous	2.40 (1.74 to 3.07)***	2.15 (1.39 to 2.90)***	2.31 (1.57 to 3.04)***	2.07 (1.42 to 2.72)***	2.34 (1.69 to 2.99)***
Syntocinon use	1.86 (1.15 to 2.59)***	2.01 (1.21 to 2.81)***	1.89 (1.09 to 2.64)***	1.91 (1.20 to 2.62)***	1.74 (1.05 to 2.44)***
Epidural use	3.36 (2.63 to 4.09)***	3.57 (2.76 to 4.39)***	3.06 (2.28 to 3.84)***	3.30 (2.57 to 4.02)***	3.27 (2.50 to 3.98)***
AoP (in °)	−0.08 (−0.10 to −0.06)***	−0.07 (−0.01 to −0.05)***	−0.08 (−0.10 to −0.06)***	—	−0.05 (−0.07 to −0.02)***
Occiput anterior	−0.99 (−1.77 to −0.21)*	−0.78 (−1.65 to 0.09)	−1.11 (−1.93 to −0.28)**	—	−0.64 (−1.41 to 0.13)
Cervical dilatation (in cm)	—	—	—	−0.62 (−0.75 to −0.50)***	−0.39 (−0.55 to −0.23)***
Observations (*n*)	380	380	316	380	380
R^2	0.54	0.54	0.52	0.55	0.56
RMSE	3.08	3.05	3.08	3.06	2.99
MAE	2.42	2.34	2.41	2.39	2.34

Values in parentheses are 95% CI. †Combined model includes manual parasagittal method for angle of progression (AoP) measurement. $*P < 0.05$. $**P < 0.01$. $***P < 0.001$. MAE, mean absolute error; RMSE, root mean squared error.

(Table S3). Multiple regression analysis was performed to predict time to delivery from, first, parity, use of syntocinon and epidural in labor, AoP (for each of the three methods of measurement) and occipital position, second, parity, use of syntocinon and epidural in labor and cervical dilatation from clinical vaginal examination, and, third, parity, use of syntocinon and epidural in labor, AoP by the manual parasagittal method, occipital position and cervical dilatation. Each model significantly predicted time to delivery, and a summary of regression coefficients, standard errors and mean R^2 is given in Table 2. Nulliparity, syntocinon use and epidural use were associated with a longer time to delivery in each model. Increasing AoP, increasing cervical dilatation and occiput anterior fetal head position were associated with a shorter time to delivery. Maternal age, height, racial origin, gestational age and EFW had no significant contribution in predicting time to delivery in any of the models.

Quantile regression

Table 3 shows the fitted regression coefficients for each of the considered quantiles. Representative scatterplots of the data with 5th, 25th, 50th, 75th and 95th quantile limits calculated by quantile regression are shown in Figure 3. ANOVA demonstrated a significant difference in slope coefficients between the 25th and 75th quantile models and the 5th and 95th quantile models ($P < 0.001$ for both). The median time to delivery at a manual parasagittal AoP of 125°, which is approximately the mean value, was 9.7 h for nulliparous women with epidural, 5.3 h for nulliparous women without epidural, 3.3 h for parous

Table 3 Beta estimates of regression models for prediction of time to delivery in hours in women undergoing induction of labor, according to quantile

	Quantile				
Variable	*5th*	*25th*	*50th*	*75th*	*95th*
Constant	6.49 (4.62 to 8.37)***	6.81 (4.68 to 8.94)***	9.55 (7.37 to 11.74)***	12.18 (8.54 to 15.82)***	18.06 (12.48 to 23.64)***
Nulliparous	0.91 (0.28 to 1.54)**	1.75 (1.11 to 2.38)***	2.30 (1.58 to 3.01)***	2.42 (1.21 to 3.62)***	2.44 (0.93 to 3.95)**
Syntocinon use	0.31 (−0.26 to 0.87)	1.07 (0.25 to 1.88)*	2.42 (1.46 to 3.37)***	2.89 (1.80 to 3.99)***	2.76 (1.17 to 4.36)***
Epidural use	1.94 (1.30 to 2.57)***	3.25 (2.48 to 4.02)***	3.25 (2.16 to 4.35)***	3.61 (2.28 to 4.95)***	5.14 (3.40 to 6.88)***
AoP (in °)†	−0.05 (−0.07 to −0.03)***	−0.05 (−0.07 to −0.03)***	−0.06 (−0.08 to −0.04)**	−0.07 (−0.10 to −0.04)***	−0.10 (−0.14 to −0.05)***
Occiput anterior	−0.75 (−1.40 to −0.11)*	−0.40 (−1.07 to 0.28)	−0.88 (−1.65 to −0.10)*	−1.07 (−2.41 to 0.27)	0.02 (−1.82 to 1.86)

Values in parentheses are 95% CI. ANOVA demonstrated significant difference in coefficients of 25th *vs* 75th and 5th *vs* 95th quantile models ($P < 0.001$ for both). †Angle of progression (AoP) measured using manual parasagittal method. *$P < 0.05$. **$P < 0.01$. ***$P < 0.001$.

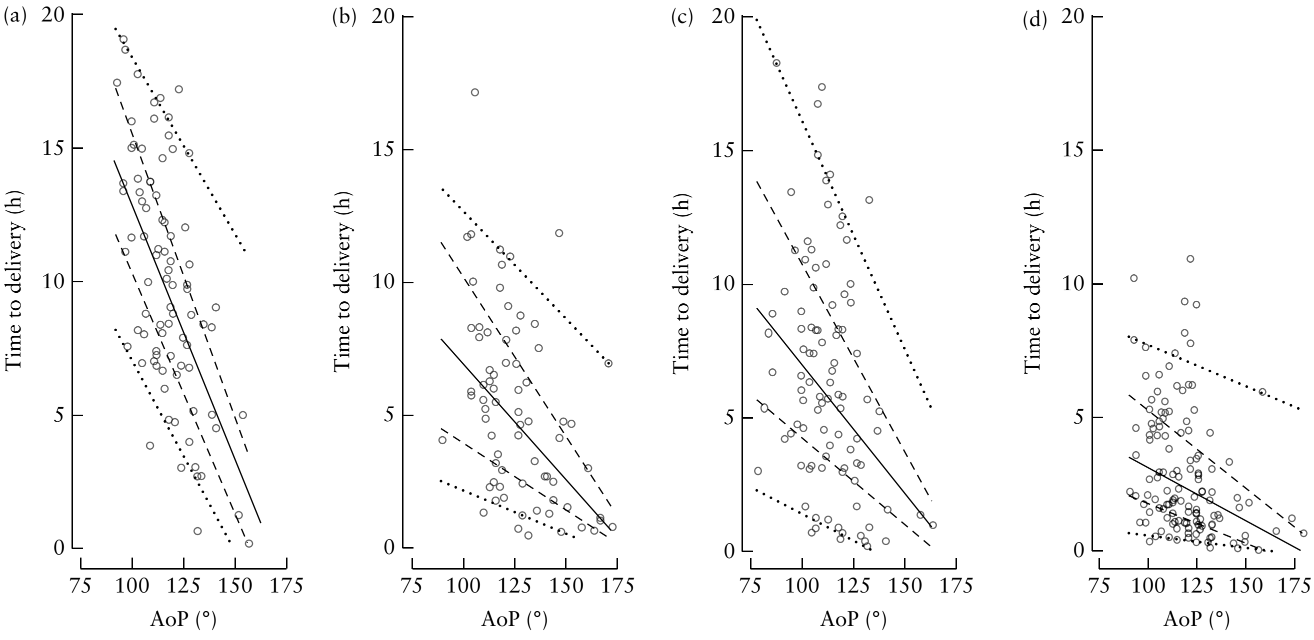

Figure 3 Prediction of time to delivery by angle of progression (AoP) measured using manual parasagittal method, with 5th, 25th, 50th, 75th and 95th quantile limits calculated by quantile regression, in: (a) nulliparous women with epidural, (b) nulliparous women without epidural, (c) parous women with syntocinon use and (d) parous women without syntocinon use.

Table 4 Multivariate logistic regression analysis for prediction of Cesarean section for failure to progress in first stage of labor

	Maternal-factors model		*Vaginal-examination model*		*Ultrasound model*	
Characteristic	*Value*	P	*Value*	P	*Value*	P
Odds ratio						
Maternal factors						
Height −165 (in cm)	0.92 (0.86–0.97)	0.006	0.92 (0.86–0.98)	0.008	0.92 (0.86–0.98)	0.014
Parous	0.27 (0.12–0.57)	< 0.001	0.28 (0.12–0.58)	0.001	0.22 (0.09–0.48)	< 0.001
Pregnancy factors						
Gestational age −40 (in weeks)	1.33 (1.08–1.66)	0.010	1.36 (1.10–1.71)	0.006	—	—
Cervical dilatation (in cm)	—	—	0.78 (0.61–0.96)	0.024	—	—
Ultrasound findings						
Estimated fetal weight −3.5 (in kg)	—	—	—	—	3.69 (1.65–8.59)	0.002
AoP −113 (in °)*	—	—	—	—	0.96 (0.94–0.99)	0.005
Model summary						
Nagelkerke R^2	0.16		0.18		0.23	
AIC	279		274		256	
AUC	0.76		0.78		0.80	

Values in parentheses are 95% CI. ANOVA demonstrated significant difference between vaginal-examination model and maternal-factors model ($P < 0.01$) and ultrasound model and maternal-factors model ($P < 0.001$). *Angle of progression (AoP) measured using manual parasagittal method. AIC, Akaike's information criterion; AUC, area under the curve.

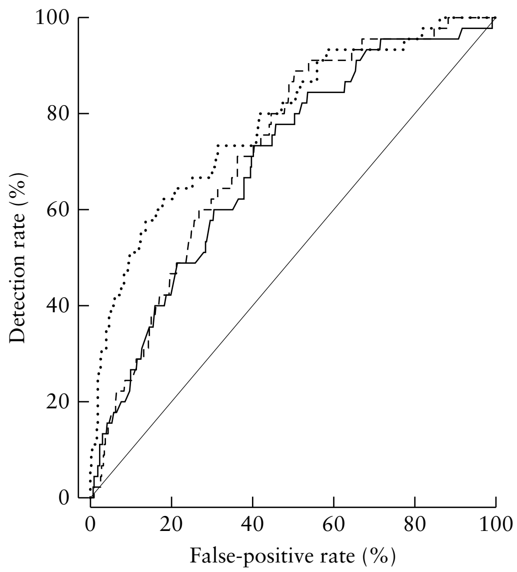

Figure 4 Receiver-operating-characteristics curves for prediction of Cesarean section for failure to progress in first stage of labor by maternal factors (——), maternal factors and findings of vaginal examination (- - -), and maternal factors and ultrasound findings (·····).

women with syntocinon use, and 1.5 h for parous women without syntocinon use.

Prediction of CS for FTP

Univariate regression analysis demonstrated that, in prediction of CS for FTP in the first stage of labor, there was a statistically significant contribution from maternal height, parity, gestational age, cervical dilatation, EFW and sagittal, manual parasagittal and automated parasagittal AoP (Table S4). The *a-priori* risk for CS for FTP is calculated using the following formula: odds/(1 + odds), where odds = e^Y and Y is derived from multivariate logistic regression analysis. Adjusted odds ratios and their 95% CI, Nagelkerke's R^2, and AUC for each of the three multiple regression models are shown in Table 4. In each model, the risk of CS for FTP decreased with a previous vaginal delivery and increasing maternal height. In the maternal-factors model and vaginal-examination model, increasing gestational age was associated with an increased risk of CS for FTP, while, in the vaginal-examination model, increasing cervical dilatation lowered the risk. In the ultrasound model, increasing EFW increased the risk of CS for FTP while increasing manual parasagittal AoP decreased the risk. The ROC curves demonstrating the performance of the maternal-factors, vaginal-examination and ultrasound models are shown in Figure 4.

DISCUSSION

Principal findings

This prospective observational study in a cohort of women undergoing induction of labor has demonstrated that the method of measurement of AoP with the highest degree of reliability is the manual parasagittal method when compared to both the sagittal and automated parasagittal approaches. The manual parasagittal approach is more reliable than the sagittal and automated parasagittal approaches because, first, ICC for both the same and different observers is higher, second, on Bland–Altman analysis, the 95% CI of the intra- and interobserver mean differences for the parasagittal approach include zero while the 95% CI of the interobserver mean difference for the sagittal approach does not, third, the LoAs for the manual parasagittal approach are noticeably narrower than those for the sagittal approach, and, fourth, the automated parasagittal approach fails to acquire an adequate image in 19% of cases. The mean difference between the manual parasagittal approach and the sagittal approach is 11°, which should be accounted for in comparing results from studies using different AoP methods.

We found that, in women having a vaginal delivery, over half of the total variation in time to delivery can be explained by a combination of maternal and pregnancy characteristics and ultrasound findings, regardless of which method of AoP measurement is used, and that the exact coefficients for predictors and their significance varies according to the quantile regressed. Significant predictors of time to delivery are parity, epidural and syntocinon use during labor and the sonographic findings of fetal head position and AoP. Longer labor is associated with nulliparity and use of epidural or syntocinon while shorter labor is associated with increasing cervical dilatation, occiput-anterior position and increasing AoP.

In the best performing model for prediction of mode of delivery, the risk for having a CS for FTP is higher in nulliparous women with increasing EFW and lower for taller women with a larger AoP. Inclusion of ultrasonographic EFW and AoP and cervical dilatation significantly improved models over those relying on maternal factors alone; however, the predictive performance of these models is only modest.

Comparison with findings from previous studies

No previous study has reported on the inter- and intraobserver reliability of the manual parasagittal method of measuring AoP as described here. The sagittal method ICC for intraobserver reliability in our study was 0.93, which is within the range reported in previous studies of 0.90 to 0.98, while the ICC for interobserver reliability in our study was 0.74, which is just below the lower end of previously reported ranges of 0.77 to 0.95[17,18,21,39]. On comparing methods, we found similar mean differences between the automated parasagittal approach and the sagittal approach (12° *vs* 15°) and a similar failure rate of the automated method (19% *vs* 15%) as in the study of Youssef *et al.*[30].

This is the first study to develop models for prediction of time to delivery in a population of women undergoing induction of labor using a combination of maternal and pregnancy characteristics, intrapartum factors and intrapartum ultrasound findings during the first stage of labor. Nulliparity as a predictor of longer duration of labor is well established from both older[40] and more recent work[41]. Incerti *et al.* reported on 1067 nulliparous women and found that a similar degree of variance ($R^2 = 0.51$) in the length of labor could be determined with gestational age, maternal ethnicity, maternal risk factors, cervical dilatation, oxytocin use and epidural use as predictors in their model[42]. Nesheim analyzed 5418 women in labor to assess the contribution of a variety of factors in predicting duration of labor, and found that the most significant predictors were nulliparity, induction *vs* spontaneous labor, neonatal birth weight, maternal height, gestational age and occipital position, but did not report overall variance[43]. Gunnarsson *et al.* evaluated retrospectively 1753 term parous women who had spontaneous labor and found that maternal body mass index, neonatal birth weight and epidural and syntocinon use contributed to prediction of time to complete the first stage of labor[44]. Masturzo *et al.* examined 270 women in spontaneous second stage of labor and reported that the highest quartile for AoP had the shortest mean time to delivery[45]. The differences in contributory predictor variables in our study may be due to the inclusion of ultrasound variables, differences in sample size or differences between study populations.

There were significant differences in maternal characteristics between the group having a vaginal delivery, those having CS for fetal distress and those having CS for FTP, as demonstrated in Table 1. The CS groups had a higher proportion of postdates inductions and higher median gestational age at induction, EFW and birth weight, as compared to women with a vaginal delivery. This is consistent with the growing body of work demonstrating that earlier induction of labor is associated with a decreased rate of CS[46,47], although direct comparison is difficult due to different study designs.

Previous studies have examined the ability of AoP to predict operative delivery in the second stage of labor[24,45,48–51], but relatively few have examined the prediction of operative delivery from data arising in the first stage. Both Eggebø *et al.* and Torkildsen *et al.* reported higher AUCs for AoP in a smaller sample than in our study using a similar set of predictor variables[27,28]. However, their studies were based solely on nulliparous women already diagnosed with prolonged first stage of labor and explicitly excluded CS for fetal distress, which makes direct comparison difficult.

Implications for clinical practice

Labor curves based on accurate, reproducible, non-invasive measurements would be a major improvement in obstetric care and could provide potential for minimizing the harms associated with vaginal examination and lowering unnecessary obstetric intervention. Based on our findings, nearly half of the variation in time to delivery remains unexplained from a single measurement in the first stage of labor, and future research should focus on finding additional markers or examining serial measurements for more accurate prediction. Individualized labor curves for women based on parity and epidural or syntocinon use may allow for the setting of customized thresholds for intervention, but this would require larger datasets and validation prior to clinical use. As shown in Figure 3, parous women in labor without syntocinon use had both low AoP and short time to delivery, which may indicate a population with limited benefit from the use of intrapartum ultrasound. While not yet robust enough for routine clinical practice, models predicting mode of delivery incorporating intrapartum ultrasound outperformed those based on vaginal examination alone.

Strengths and limitations

The strengths of this study are the prospective nature of the data collection, the relatively large sample size for intrapartum ultrasound, and the development of models using appropriate statistical techniques to weight contributory factors in predicting the outcomes of interest. There are a few limitations. First, inter- and intraobserver variability analyses were performed by remeasuring previously stored images. While this gives insight into reliability of measurement, it does not give insight into variation arising from differences in ultrasound technique during image capture. Second, the data arose from an ethnically homogeneous population undergoing induction of labor at a single institution and this may limit generalization to other settings or populations. In particular, the fact that median maternal weight in our population was high may have biased against maternal weight in multivariate analysis. Finally, there was significant incomplete recording of fetal head station, position and cervical consistency by clinical staff, necessitating the use of cervical dilatation as the sole component of vaginal examination on multivariate analysis. This may have biased against more robust models including a fuller set of vaginal-examination findings.

Conclusions

The method of measuring AoP with the greatest reliability is the manual parasagittal technique, and future research on intrapartum sonography should focus on this technique. Over half the variation in time to vaginal delivery can be explained by a combination of maternal factors, pregnancy characteristics and ultrasound findings, but larger datasets will be required to create accurate ultrasound-based individualized labor curves. The ability of AoP to provide clinically useful prediction of CS for FTP in the first stage of labor is limited.

ACKNOWLEDGMENT

This study was supported by a grant from The Fetal Medicine Foundation (Charity No: 1037116).

REFERENCES

1. Friedman EA. The graphic analysis of labor. *Am J Obstet Gynecol* 1954; **68**: 1568–1575.
2. Friedman EA. Primigravid labor: A graphicostatistical analysis. *Obstet Gynecol* 1955; **6**: 567–589.
3. Friedman EA. Labor in Multiparas: A graphicostatistical analysis. *Obstet Gynecol* 1956; **8**: 686–703.
4. Cohen WR, Friedman EA. The assessment of labor: A brief history. *J Perinat Med* 2018; **46**: 1–8.
5. Buchmann EJ, Libhaber E. Accuracy of cervical assessment in the active phase of labour. *BJOG* 2007; **114**: 833–837.
6. Dupuis O, Silveira R, Zentner A, Dittmar A, Gaucherand P, Cucherat M, Redarce T, Rudigoz RC. Birth simulator: Reliability of transvaginal assessment of fetal head station as defined by the American College of Obstetricians and Gynecologists classification. *Am J Obstet Gynecol* 2005; **192**: 868–874.
7. Hassan SJ, Sundby J, Husseini A, Bjertness E. The paradox of vaginal examination practice during normal childbirth: Palestinian women's feelings, opinions, knowledge and experiences. *Reprod Health* 2012; **9**: 1–9.
8. Chan YTV, Ng VKS, Yung WK, Lo TK, Leung WC, Lau WL. Relationship between intrapartum transperineal ultrasound measurement of angle of progression and head-perineum distance with correlation to conventional clinical parameters of labor progress and time to delivery. *J Matern Neonatal Med* 2015; **28**: 1476–1481.
9. Seaward PG, Hannah ME, Myhr TL, Farine D, Ohlsson A, Wang EE, Haque K, Weston JA, Hewson SA, Ohel G, Hodnett ED. International multicentre term prelabor rupture of membranes study: Evaluation of predictors of clinical chorioamnionitis and postpartum fever in patients with prelabor rupture of membranes at term. *Am J Obstet Gynecol* 1997; **177**: 1024–1029.
10. Westover T, Knuppel RA, Westover T. Modern Management of Clinical Chorioamnionitis. *Infect Dis Obstet Gynecol* 1995; **3**: 123–132.
11. Downe S, Gyte GML, Dahlen HG, Singata M. Routine vaginal examinations for assessing progress of labour to improve outcomes for women and babies at term. *Cochrane Database Syst Rev* 2013; **7**: CD010088.
12. Eggebø TM, Gjessing LK, Heien C, Smedvig E, Økland I, Romundstad P, et al. Prediction of labor and delivery by transperineal ultrasound in pregnancies with prelabor rupture of membranes at term. *Ultrasound Obstet Gynecol* 2006; **27**: 387–391.
13. Barbera AF, Pombar X, Peruginoj G, Lezotte DC, Hobbins JC. A new method to assess fetal head descent in labor with transperineal ultrasound. *Ultrasound Obstet Gynecol* 2009; **33**: 313–319.
14. Alvarez-Colomo C, Gobernado-Tejedor JA. The validity of ultrasonography in predicting the outcomes of labour induction. *Arch Gynecol Obstet* 2016; **293**: 311–316.
15. Seval MM, Yuce T, Kalafat E, Duman B, Aker SS, Kumbasar H, Koc A. Comparison of effects of digital vaginal examination with transperineal ultrasound during labor on pain and anxiety levels: a randomized controlled trial. *Ultrasound Obstet Gynecol* 2016; **48**: 695–700.
16. Usman S, Barton H, Wilhelm-Benartzi C, Lees CC. Ultrasound is better tolerated than vaginal examination in and before labour. *Aust New Zeal J Obstet Gynaecol* 2019; **59**: 362–366.
17. Molina FS, Terra R, Carrillo MP, Puertas A, Nicolaides KH. What is the most reliable ultrasound parameter for assessment of fetal head descent? *Ultrasound Obstet Gynecol* 2010; **36**: 493–499.
18. Tutschek B, Braun T, Chantraine F, Henrich W. A study of progress of labour using intrapartum translabial ultrasound, assessing head station, direction, and angle of descent. *BJOG* 2011; **118**: 62–69.
19. Ghi T, Contro E, Farina A, Nobile M, Pilu G. Three-dimensional ultrasound in monitoring progression of labor: a reproducibility study. *Ultrasound Obstet Gynecol* 2010; **36**: 500–506.
20. Youssef A, Maroni E, Ragusa A, De Musso F, Salsi G, Iammarino MT, Paccapelo A, Rizzo N, Pilu G, Ghi T. Fetal head-symphysis distance: A simple and reliable ultrasound index of fetal head station in labor. *Ultrasound Obstet Gynecol* 2013; **41**: 419–424.
21. Sainz JA, Fernández-Palacín A, Borrero C, Aquise A, Ramos Z, García-Mejido JA. Intra and interobserver variability of intrapartum transperineal ultrasound measurements with contraction and pushing. *J Obstet Gynaecol* 2018; **38**: 333–338.
22. Montaguti E, Rizzo N, Pilu G, Youssef A. Automated 3D ultrasound measurement of the angle of progression in labor. *J Matern Neonatal Med* 2018; **31**: 141–149.
23. Ghi T, Eggebø T, Lees C, Kalache K, Rozenberg P, Youssef A, Salomon LJ, Tutschek B. ISUOG Practice Guidelines: intrapartum ultrasound. *Ultrasound Obstet Gynecol* 2018; **52**: 128–139.
24. Kalache KD, Dückelmann AM, Michaelis SAM, Lange G Cichon J, Dudenhausen JW. Transperineal ultrasound imaging in prolonged second stage of labor with occipitoanterior presenting fetuses: how well does the 'angle of progression' predict the mode of delivery? *Ultrasound Obstet Gynecol* 2009; **33**: 326–330.
25. Dückelmann AM, Bamberg C, Michaelis SAM, Lange J, Nonnenmacher A, Dudenhausen JW, Kalache KD. Measurement of fetal head descent using the 'angle of progression' on transperineal ultrasound imaging is reliable regardless of fetal head station or ultrasound expertise. *Ultrasound Obstet Gynecol* 2010; **35**: 216–222.
26. Chor CM, Poon LCY, Leung TY. Prediction of labor outcome using serial transperineal ultrasound in the first stage of labor. *J Matern Neonatal Med* 2019; **32**: 31–37.
27. Eggebø TM, Hassan WA, Salvesen KA, Lindtjørn E, Lees CC. Sonographic prediction of vaginal delivery in prolonged labor: a two-center study. *Ultrasound Obstet Gynecol* 2014; **43**: 195–201.
28. Torkildsen EA, Salvesen KÅ, Eggebø TM. Prediction of delivery mode with transperineal ultrasound in women with prolonged first stage of labor. *Ultrasound Obstet Gynecol* 2011; **37**: 702–708.
29. Youssef A, Ghi T, Awad EE, Maroni E, Montaguti E, Rizzo N, Pilu G. Ultrasound in labor: a caregiver's perspective. *Ultrasound Obstet Gynecol* 2013; **41**: 469–470.
30. Youssef A, Salsi G, Montaguti E, Bellussi F, Pacella G, Azzarone C, Farina A, Rizzo N, Pilu G. Automated Measurement of the Angle of Progression in Labor: A Feasibility and Reliability Study. *Fetal Diagn Ther* 2017; **41**: 293–299.
31. National Institute for Health and Care Excellence. *Intrapartum care for healthy women and babies*. RCOG Press: London, 2014.
32. Snijders RJM, Nicolaides KH. Fetal biometry at 14–40 weeks' gestation. *Ultrasound Obstet Gynecol* 1994; **4**: 34–48.
33. Hadlock FP, Harrist RBB, Sharman RS, Deter RL, Park SK. Estimation of fetal weight with the use of head, body, and femur measurements – A prospective study. *Am J Obstet Gynecol* 1985; **151**: 333–337.
34. Akmal S, Tsoi E, Kametas N, Howard R, Nicolaides KH. Intrapartum sonography to determine fetal head position. *J Matern Neonatal Med* 2002; **12**: 172–177.
35. Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. *Psychol Bull* 1979; **86**: 420–428.
36. Martins WP, Nastri CO. Interpreting reproducibility results for ultrasound measurements. *Ultrasound Obstet Gynecol* 2014; **43**: 479–480.
37. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; **327**: 307–310
38. Koenker R. *Quantile Regression*. Cambridge University Press: Cambridge, 2005.
39. Tutschek B, Torkildsen EA, Eggebø TM. Comparison between ultrasound parameters and clinical examination to assess fetal head station in labor. *Ultrasound Obstet Gynecol* 2013; **41**: 425–429.
40. Friedman EA, Sachtleben MR. Dysfunctional Labor I. Prolonged latent phase in the nullipara. *Obstet Gynecol* 1961; **17**: 135–148.
41. Vahratian A, Hoffman MK, Troendle JF, Zhang J. The impact of parity on course of labor in a contemporary population. *Birth* 2006; **33**: 12–17.

42. Incerti M, Locatelli A, Ghidini A, Ciriello E, Malberti S, Consonni S, Pezzullo JC. Prediction of duration of active labor in nulliparous women at term. *Am J Perinatol* 2008; **25**: 85–89.
43. Nesheim B. Duration of labor. *Acta Obstet Gynecol Scand* 1988; **67**: 121–124.
44. Gunnarsson B, Skogvoll E, Jónsdóttir IH, Røislien J, Smárason AK. On predicting time to completion for the first stage of spontaneous labor at term in multiparous women. *BMC Pregnancy Childbirth* 2017; **17**: 1–8.
45. Masturzo B, Piazzese A, Paracchini S, Quezada MS, Todros T, Farina A. Time remaining in labor and probability of vaginal delivery as a function of the angle of progression in a low risk population with a normal first stage of labor. In-house observational study and comparison with the data in the literature. *Minerva Ginecol* 2018; **70**: 35–43.
46. Grobman WA, Rice MM, Reddy UM, Tita ATN, Silver RM, Mallett G, Hill K, Thom EA, El-Sayed YY, Perez-Delboy A, Rouse DJ, Saade GR, Boggess KA, Chauhan SP, Iams JD, Chien EK, Casey BM, Gibbs RS, Srinivas SK, Swamy GK, Simhan HN, Macones GA; Eunice Kennedy Shriver National Institute of Child Health and Human Development Maternal–Fetal Medicine Units Network. Labor induction versus expectant management in low-risk nulliparous women. *N Engl J Med* 2018; **379**: 513–523.
47. Sotiriadis A, Petousis S, Thilaganathan B, Figueras F, Martins WP, Odibo AO, Dinas K, Hyett J. Maternal and perinatal outcomes after elective induction of labor at 39 weeks in uncomplicated singleton pregnancy: a meta-analysis. *Ultrasound Obstet Gynecol* 2019; **53**: 26–35.
48. Bultez T, Quibel T, Bouhanna P, Popowski T, Resche-Rigon M, Rozenberg P. Angle of fetal head progression measured using transperineal ultrasound as a predictive factor of vacuum extraction failure. *Ultrasound Obstet Gynecol* 2016; **48**: 86–91.
49. Sainz JA, García-Mejido JA, Aquise A, Bonomi MJ, Borrero C, De La Fuente P, Fernández-Palacín A. Intrapartum transperineal ultrasound used to predict cases of complicated operative (vacuum and forceps) deliveries in nulliparous women. *Acta Obstet Gynecol Scand* 2017; **96**: 1490–1497.
50. Bibbo C, Rouse CE, Cantonwine DE, Little SE, McElrath TF, Robinson JN. Angle of Progression on Ultrasound in the Second Stage of Labor and Spontaneous Vaginal Delivery. *Am J Perinatol* 2018; **35**: 413–420.
51. Ghi T, Youssef A, Maroni E, Arcangeli T, De Musso F, Bellussi F, Nanni M, Giorgetta F, Morselli-Labate AM, Iammarino MT, Paccapelo A, Cariello L, Rizzo N, Pilu G. Intrapartum transperineal ultrasound assessment of fetal head progression in active second stage of labor and mode of delivery. *Ultrasound Obstet Gynecol* 2013; **41**: 430–435.

SUPPORTING INFORMATION ON THE INTERNET

The following supporting information may be found in the online version of this article:

Figure S1 Bland–Altman plots demonstrating degree of concordance between methods of measurements of angle of progression: (a) sagittal and manual parasagittal methods; (b) manual and automated parasagittal methods; and (c) sagittal and automated parasagittal methods. Dashed lines represent mean and upper and lower limits of agreement (1.96 SD). Dotted lines represent 95% CI around means and limits of agreement.

Table S1 Bland–Altman analysis for intra- and interobserver repeatability for paired measurements of angle of progression by sagittal and manual parasagittal methods

Table S2 Bland–Altman analysis for comparisons of angle of progression by sagittal, manual parasagittal and automated methods

Table S3 Univariate linear regression analysis for prediction of time to delivery in hours

Table S4 Univariate logistic regression analysis for prediction of Cesarean section for failure to progress in first stage of labor

