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

Alex FRICK,[1] Vira KOSTIV,[1,2] Denisa VOJTASSAKOVA,[1,2] Ranjit AKOLEKAR,[2,3 *]
Kypros H. NICOLAIDES.[1 *]

* Joint senior authors

1. Harris Birthright Research Centre of 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

Correspondence:
Professor KH Nicolaides,
Fetal Medicine Research Institute,
King's College Hospital, 16-20 Windsor Walk, Denmark Hill, London SE5 8BB
email: kypros@fetalmedicine.com

Keywords: Angle of progression, Intrapartum ultrasound, Transperineal ultrasound, Time-to-delivery, Labor, Vaginal delivery.

Running head: Angle of progression in labor

CONTRIBUTION

What does this work add to what is already known?

First, the method of measuring the angle of progression (AOP) by transperineal ultrasound during labor with highest degree of reliability is the manual para-sagittal method when compared to both the sagittal and automated approach. Second, the 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 para-sagittal method of AOP 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

Objective First, to compare the manual sagittal and para-sagittal and automated para-sagittal 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 on a cohort of 512 women with singleton pregnancies undergiong induction of labor. A random selection of 50 stored images was assessed for inter- and intra-observer reliability between methods. In the cases of vaginal delivery univariate linear, multivariate linear and quantile regression were performed to predict time-to-delivery. Univariate and multivariate binomial logistic regression were performed to predict CS for FTP in the first stage of labor.

Results The intra correlation coefficients (ICC) for the manual para-sagittal method for a single observer was 0.97 (CI 0.95-0.98) and for two observers was 0.96 (CI 0.93-0.98) indicating good reliability. The ICC for the sagittal method for a single observer was 0.93 (0.88-0.96) and for two observers was 0.74 (0.58-0.84) indicating moderate reliabilty for a single observer and poor reliability between two observers. Bland-Altman analysis demonstrated narrower limits of agreement for the manual para-saggittal approach than for the sagittal approach for both single and two observers. The automated para-sagittal method failed to capture an image in 19% of cases. The mean difference between sagittal and para-sagittal methods was 11^0. 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 labour and the sonographic findings of fetal head position and AOP. In the prediction of CS for FTP in the first stage of labour a model which combined maternal factors with the sonographic measurements of AOP and estimated fetal weight was superior to one utilising maternal factors alone (area under the curve 0.80 vs 0.76).

Conclusions First, the method of measuring AOP with greatest reliability is the manual para-sagittal 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, and third, the ability of AOP to provide clinically useful prediction CS for FTP in the first stage of labour is limited.

INTRODUCTION

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

A number of new techniques have been described using transperineal ultrasound (TPUS) to monitor labour progress through measurements relating the fetal head position to the maternal pelvis.[12,13] These techniques are non-invasive, they are well tolerated by patients and have a high degree of inter- and intra- observer reliability.[8,14-22] 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 symphysis pubis.[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.[17,18,23,24] Studies assessing the utility of AOP in the first stage of labour in the prediction of vaginal delivery and time-to-delivery have been limited by small numbers.[8,25-27] 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.[28] This problem has now been overcome with the development of automated software (Sono L&D; GE Healthcare) which uses a different set of landmarks, namely the more hyper-echogenic pubic rami seen in a slightly para-sagittal view (Figure 1). A previous study assessing the automated technique found that the automated method systematically overestimated the AOP compared to the sagittal approach but did not directly compare a manual para-sagittal approach to the automated method.[29] No previous study has evaluated a manual para-sagittal approach in predicting time-to-delivery or operative delivery.

The objectives of this study are: first, to compare the manual sagittal and para-sagittal and automated para-sagittal methods of measuring the 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, Kent between May 2016 and August 2017 were recruited into the study. The inclusion criteria were age ≥ 18 years, singleton pregnancy with alive fetus in cephalic presentation. We excluded multiple pregnancies or 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 examinations every four hours until delivery, was as recommended by national guidelines.[30] 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 ultrasound and had received training in intrapartum ultrasound. Obstetricians and midwives were not made aware of the ultrasound findings. Approval for the study approved by London-Dulwich Research Ethics Committee (REC reference 16/LO/0367).

The maternal weight and height were measured immediately prior to induction of labour. Patient characteristics recorded included maternal age, racial origin (White, Black, South Asian, East Asian and mixed), method of conception (spontaneous or assisted requiring the use of ovulation drugs) and parity (parous or nulliparous if no previous pregnancies at ≥ 24 weeks' gestation). We estimated fetal weight (EFW) from measurements of fetal head circumference (HC), abdominal circumference (AC) and femur length (FL) obtained by transabdominal sonography the day before induction of labor.[31,32]

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 time and date of 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, Zipf, Austria) equipped with a convex 4C-RS probe. Fetal occiput position was determined using the transabdominal technique described by Akmal *et al*.[33] Women were then placed in a modified lithotomy position with an empty bladder and the probe covered with a glove. Two dimensional TPUS was performed by placing the probe vertically between the labia to obtain a sagittal view of the fetal head in relation to the maternal symphysis pubis. The exact positioning of the probe between the labia was then adjusted to obtain clear images of the symphysis pubis, typically at a 30-40 degree 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 para-sagittal view that included a clearly identifiable length of the hyperechogenic maternal pubic rami (Figure 1). Images from both the midline and para-sagittal views were stored for later analysis.

The AOP between the symphysis and the leading edge of the fetal skull was measured, first, in a sagittal image manually,[13] second, in a parasagittal image manually and third, in a para-sagittal image by the automated technique (Sono L&D, GE Healthcare, Zipf, Austria). 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 para-sagittal 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 Kruskall-Wallis test and χ^2-square test or Fisher's exact test, respectively. Normality of distribution was assessed using probability plots and histograms. A p value of < 0.05 was considered significant.

Comparison of manual and automated measurement of AOP

A random selection of stored images was assessed for inter- and intra-observer reliability for both the sagittal and para-sagittal methods. Operator A remeasured 50 images of each method to create an intra-observer reliability dataset. Operator A remeasured 50 images originally measured by Operator B to create the inter-observer reliability dataset. Operator A was blinded to the original measurements at the time of remeasuring. The intra- and inter-observer reproducibility of measurements was examined by calculating intraclass correlation coefficients and their 95% confidence interval (CI).[34] The overlap between the 95% CI of two intraclass correlation coefficients 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; 0.70-0.90, poor; 0.90-0.95; moderate, 0.95-0.99; good, >0.99 excellent.[35] The Bland-Altman plot of the average measurement against the percentage of the differences 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.[36] The standard deviation (SD) of the differences for each method for both one and two examiners was calculated and reported. The optimal method is the one where, 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 smaller in the Bland-Altman plot. Additional Bland-Altman analysis between the sagittal, para-sagittal, and automated methods using the entire dataset was performed to assess for inter-method systematic bias.

Prediction of outcome

In women with a vaginal delivery, univariable linear regression was performed to assess the relationship between maternal age, height, weight, racial origin, parity, gestational age, EFW, epidural anesthesia, use of syntocinon, cervical dilatation, fetal head position as determined by ultrasound and all three methods of AOP with time-to-delivery in hours. Multiple linear regression 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 overfit. Due to significant heteroscedasticity seen upon visual inspection of plotted residuals, quantile regression was performed for the 5^{th}, 25^{th}, 50^{th}, 75^{th}, and 95^{th} quantiles and reported.[37]

Univariate binomial logistic regression analysis was carried out to determine which of the factors from maternal and pregnancy characteristics, clinical vaginal examination or 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. Multivariable 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 (AUROC) curves.

The statistical software package *R* version 3.5.1 (2018-07-02) was used for all 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%) who had CS, including 59 (44.7%) for failure to progress 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 para-sagittal methhod for a single observer was 0.97 (CI 0.95-0.98) and for two observers was 0.96 (CI 0.93-0.98) indicating good reliability; the respective values for the sagittal method were 0.93 (CI 0.88-0.96) and 0.74 (CI 0.58-0.84), indicating moderate reliabilty for a single observer and poor reliability between two observers. The SD of the differences for the manual para-sagittal 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 para-sagittal method there was an overlap between the 95% CIs of two ICCs for both the intra-observer and inter-observer results indicating no significant difference. The sagittal method demonstrated a lower degree of inter-observer reproducibility with no overlap in the ICC CIs when compared to the intra-observer ICC for the same method. The most reproducible results overall were from the para-sagittal method as it demonstrated the highest ICC and the lowest SD for both inter- and intra-observer reliability.

Bland-Altman plots demonstrating degree of concordance between pairs of measurements made by the same observer and by the two different observers for two sagittal methods and the para-sagittal method are illustrated in Figure 2. Bland-Altman plots demonstrating the degree of concordance between methods are illustrated in supplementary Figure 1. The results for the Bland-Altman analyses are presented in supplementary Tables 1 and 2.

Prediction of time-to-delivery

Multiple linear regression

In the univariate regression analyses the only significant predictors of time-to-delivery were all three methods of AOP, parity, use of syntocinon and epidural in labor and sonographically determined occipital position (supplementary Table 3). Multiple regression was performed to predict time-to-delivery from first, parity, use of syntocinon and epidural in labour, AOP (each of the three methods of measurement) and occipital position, second, parity, use of syntocinon and epidural in labour, cervical dilatation from the clinical vaginal examination, and third, parity, use of syntocinon and epidural in labour, AOP by the manual para-sagittal 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 can be found in Table 2. Nulliparity, syntocinon use, and epidural use were associated with a longer time-to-delivery in each model. Increasing AOP, increasing vaginal 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 the each considered quantile. Representative scatterplots of the data with 5th, 25th, 50th, 75th and 95th quantile limits calculated by quantile regression are shown in Figure 3. ANOVA comparing slope coefficients of the 25th with 75th and the 5th with the 95th quantile models were significantly different ($p<0.0001$). The median time-to-delivery at a para-sagittal AOP of 125 degrees, which is approximately the mean value, was 9.7 hours for nulliparous women with epidural, 5.3 hours for nulliparous women without epidural, 3.3 hours for parous women with syntocinon use, and 1.5 hours for parous women without syntocinon use.

Prediction of cesarean section for failure to progress

Univariate regression analysis demonstrated that in prediction of CS for FTP in the first stage of labour there was a statistically significant contribution from maternal height, parity, gestational age, cervical dilatation, EFW, and sagittal, para-sagittal, and automated AOP (supplementary Table 4). The *a-priori* risk for CS for FTP is calculated from 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, Nagelkerle's R^2, and AUROC 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 model and vaginal examination model, increasing gestational age was associated with increased risk of CS for FTP, while in the vaginal model increasing vaginal dilatation lowered the risk. In the ultrasound model, increasing EFW increased the risk of CS for FTP while increasing para-sagittal AOP decreased the risk. The ROC curves demonstrating the performance of maternal, vaginal examination, and ultrasound models are shown in Figure 4.

DISCUSSION

Principle findings of the study

This prospective observational study on a cohort of women undergiong induction of labour has demonstrated that the method of AOP with highest degree of reliability is the para-sagittal method when compared to both the sagittal and automated approach. The para-sagittal approach is more reliable than the sagittal approach because first, ICC for both the same and different observers is higher, second, the mean difference on Bland-Altman analysis for the para-sagittal approach includes zero in the CI while the inter-observer mean for the sagittal approach does not, third, the LOAs for the para-sagittal approach are noticeably narrower than the sagittal approach, and fourth, the automated para-sagittal approach fails to acquire an adequate image in 19% of cases. However, the mean difference between the manual para-sagittal approach and the sagittal approach is 11^{O} 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 characeteristics and ultrasound findings, regardless of which method of AOP is used and that the exact coefficients for predictors and their signficance varies according to the quantile regressed. Significant predictors of time-to-delivery are parity, epidural and syntocinon use during labour and the sonographic findings of fetal head position and AOP. Longer labors are associated with nulliparity and use of epidural or syntcinon while shorter labors are assocaited 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 and with a larger AOP. Inclusion of ultrasonographic EFW and AOP and vaginal 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 intra- observer reliability of the manual para-sagittal method of measuring the AOP as described here. The sagittal method ICC for intra-observer 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 inter-observer reliability in our study was 0.74 which is just below the lower end of previously reported ranges of 0.77 to 0.95.[17,18,21,38] In comparing methods, we found similar mean differences between the automated para-sagittal approach and the sagittal approach (11^0 compared to 15^0) and similar failure rate of the automated methods (19% vs 15%) as in the study by Youseff *et al.*[29]

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 during the first stage of labor. Nulliparity as a predictor of longer duration of labor is well established from both older[39] and more recent work.[40] Incerti *et al.*, reported on 1,067 nulliparous women and found that a similar degree of variance (R^2=0.51) in the length of labour could be determined with gestational age, maternal ethnicity, maternal risk factors, cervical dilatation, oxytocin and epidural use as predictors in their model.[41] Nesheim analysed 5418 labors to assess the contribution of a variety of factors in predicting duration of labour and found that the most significant predictors were nulliparity, induction vs spontaneous labour, neonatal birth weight, maternal height, gestational age and occipital position, but did not report overall variance.[42] Gunnarsson *et al.,* retrospectively evaluated 1753 term multiparous women who spontaneously labored and found that maternal body mass index, neonatal birth weight, epidural and syntocinon use contributed to prediction of time to complete the first stage of labour.[43] Masturzo *et a.,* examined 270 women in spontaneous second stage of labor and reported that the highest quartile for AOP had the shortest mean time-to-delivery in the second stage of labor.[44] The differences in contributory predictor variables in our study may be due to the inclusion of the ultrasound variables, differences in sample size, or differences between study populations.

There were significant differences in maternal characteristics between the groups having a vaginal delivery, CS for fetal distress and CS for FTP as demonstrated in Table 1. The CS groups had a higer proportion of post-dates induction, higher median gestational age at induction, and higher EFW and birth weight as compared to the women with a vaginal delivery. This is consistent with the growing body of work that earlier induction of labor is associated with a decreased rate of CS,[45,46] 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,[23, 44, 47-50] but relatively few have examined predicting operative delivery from data arising in the first stage. Both Torkildsen *et al.* and Eggebo *et al.* reported higher AUROCs for AOP in a smaller sample than our study using a similar set of predictor variables.[26,27] However, their studies were based solely on nulliparous women with an already diagnosed prolonged first stage of labor and explicitly excluded CS for fetal distress which makes direct comparison difficult.

Implications for clinical practice

Labour curves based on accurate, reproducible, non-invasive measurements would be a major improvement in obstetric care and could provide potential for minimising the harms associated with VE 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 labour and future research should focus on finding additional markers or examining serial measurements for more accurate prediction. Individualised labour curves for women based on parity, epidural or syntocinon use may allow for the setting of customised thresholds for intervention but this would require larger data sets and validation prior to clinical use. As shown in Figure 3, mulitparous women laboring 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 of the study

The strengths of the 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 for contributory factors in predicting the outcomes of interest. There are a few limitations. First, inter- and intra- observer 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 the 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 greatest reliability is the manual para-sagittal 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 individualised labour curves. The ability of AOP to provide clinically useful prediction of CS for FTP in the first stage of labour is limited.

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

Conflict of interest: None

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 Gyneco*l 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, et al. 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, et al. 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; 2013.

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, et al. 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. 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.

24. 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.

25. 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.

26. 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.

27. 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.

28. 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.

29. Youssef A, Salsi G, Montaguti E, Bellussi F, Pacella G, Azzarone C, et al. Automated Measurement of the Angle of Progression in Labor: A Feasibility and Reliability Study. *Fetal Diagn Ther* 2017; **41**: 293–299.

30. National Institute for Health and Care Excellence. Intrapartum care for healthy women and babies. Clinical guideline [CG190]. NICE; 2014 https://www.nice.org.uk/guidance/cg190

31. Snijders RJM, Nicolaides KH. Fetal biometry at 14–40 weeks’ gestation. *Ultrasound Obstet Gynecol* 1994; **4**: 34–48.

32. 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.

33. 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.

34. Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. *Psychol Bull* 1979; **86**: 420–428.

35. Martins WP, Nastri CO. Interpreting reproducibility results for ultrasound measurements. *Ultrasound Obstet Gynecol* 2014; **43**: 479–480.

36. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; **327**: 307–310

37. Koenker R. Quantile Regression. Cambridge: Cambridge University Press; 2005

38. 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.

39. Friedman EA, Sachtleben MR. Dysfunctional Labor I. Prolonged latent phase in the nullipara. *Obstet Gynecol* 1961; **17**: 135–148.

40. 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.

41. Incerti M, Locatelli A, Ghidini A, Ciriello E, Malberti S, Consonni S, et al. Prediction of duration of active labor in nulliparous women at term. *Am J Perinatol* 2008; **25**: 85–89.

42. Nesheim B. Duration of labor. *Acta Obstet Gynecol Scand* 1988; **67**: 121–124.

43. 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.

44. 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.

45. 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;3 **379**: 513–23.

46. 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.

47. 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.

48. 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.

49. 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.

50. 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.

FIGURE LEGENDS

Figure 1 Representative examples of angle of progression measurements for the manual sagittal (left), manual para-sagittal (middle), and automated para-sagittal (right) approach.

Figure 2 Bland-Altman plots demonstrating the degree of concordance between pairs of measurements of angle of progression in labor: sagittal method, two operators (a), sagittal method, single operator (b), para-sagittal method, two operators (a), para-sagittal method, single operator (d). Dashed lines represent mean and upper and lower limits of agreement (1.96 SD). Dotted lines represent 95% confidence intervals around means and limits of agreement.

Figure 3 Prediction of time-to-delivery with 5^{th}, 25^{th}, 50^{th}, 75^{th} and 95^{th} quantile limits calculated by quantile regression for (a) nulliparous women with epidural, (b) nulliparous women without epidural (c) parous women with syntocinon use, and (d) parous women without syntocinon use.

Figure 4 Receiver operator characteristics curves on the prediction of cesarean section for failure to progress in labour by maternal factors (black curve), maternal factors and findings of vaginal examination (blue curve), and maternal factors and ultrasound models (red curve).

sFigure 1 Bland-Altman plots demonstrating the degree of concordance between methods of measurements of angle of progression: sagittal and para-sagittal method (left), para-sagittal and automated method (middle), and sagittal and automated method (right). Dashed lines represent mean and upper and lower limits of agreement (1.95 SD). Dotted lines represent 95% confidence intervals around means and limits of agreement.

Table 1. Maternal and pregnancy characteristics of the study populations.

Maternal and pregnancy characteristic	Vaginal delivery (n=380)	Cesarean section	
		Failure to progress (n=59)	Presumed fetal distress (n=79)
Maternal age in years*	28.0 (24.0–31.2)	29.0 (25.0–33.5)	29.0 (26.0–33.0)
Maternal weight in Kg*	86.0 (74.2–97.0)	93.0 (82.1–105.5)	89.0 (78.4–101.0)
Maternal height in 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 (0.0)	2 (2.7)
South Asian	15 (3.9)	1 (1.7)	5 (6.8)
East Asian	1 (0.3)	1 (1.7)	0 (0.0)
Mixed	7 (1.8)	0 (0.0)	2 (2.7)
Conception			
Spontaneous	374 (98.4)	58 (98.3)	70 (95.9)
Ovulation 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 (77.6)	35 (9.2)	50 (13.2)
Post-dates	86 (64.7)	24 (18.0)	23 (17.3)
Gestation at delivery in weeks***	39.6 (38.4–41.4)	40.9 (39.1–41.9)	40.4 (39.1–41.9)
Estimated fetal weight in grams***	3470 (3064–3807)	3753 (3486–4038)	3527 (3143–3811)
Birth weight in grams***	3365 (2999–3700)	3660 (3385–4007)	3500 (3025–3830)

Continuous variables are presented in median (interquartile range) and categorical variables are presented in number (%). Comparisons between outcome groups was by the χ^2-test and Fisher's exact test for categorical variables and Kruskall-Wallis test for continuous variables. * P<.05, **P <.01,***P<.001

Table 2: Statistical results comparing multiple linear regression models predicting time-to-delivery in hours.

Variable	Measurement of angle of progression			Vaginal Examination	Combined^
	Para-Sagittal	Sagittal	Automated		
Constant	12.07*** (9.61,14.52)	10.64*** (8.19, 13.08)	12.63*** (9.93, 15.33)	5.35*** (4.57, 6.14)	10.03*** (7.50, 12.55)
Nulliparity	2.40*** (1.74, 3.07)	2.15*** (1.39, 2.90)	2.31*** (1.57, 3.04)	2.07*** (1.42, 2.72)	2.34*** (1.69, 2.99)
Syntocinon	1.86*** (1.15, 2.59)	2.01*** (1.21, 2.81)	1.89*** (1.09, 2.64)	1.91*** (1.20, 2.62)	1.74*** (1.05, 2.44)
Epidural	3.36*** (2.63, 4.09)	3.57*** (2.76, 4.39)	3.06*** (2.28, 3.84)	3.30*** (2.57, 4.02)	3.27*** (2.50, 3.98)
Angle of progression	-0.08*** (-0.10, -0.06)	-0.07*** (-0.01, -0.05)	-0.08*** (-0.10, -0.06)		-0.05*** (-0.07, -0.02)
Occiput anterior	-0.99* (-1.77, -0.21)	-0.78 (-1.65, 0.09)	-1.11** (-1.93, -0.28)		-0.64 (-1.41, 0.13)
Cervical dilatation				-0.62*** (-0.75, -0.50)	-0.39*** (-0.55, -0.23)
Observations	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

Data are presented as estimates (β) and 95% confidence intervals (in parentheses) of models. RMSE = root mean squared error. MAE = mean absolute error. ^ Combined model includes manual para-sagittal method for angle of progression measurements. ***P < 0.001. **P < 0.01. *P < 0.05.

Table 3: Comparison of regression models by quantile predicting time-to-delivery in hours.

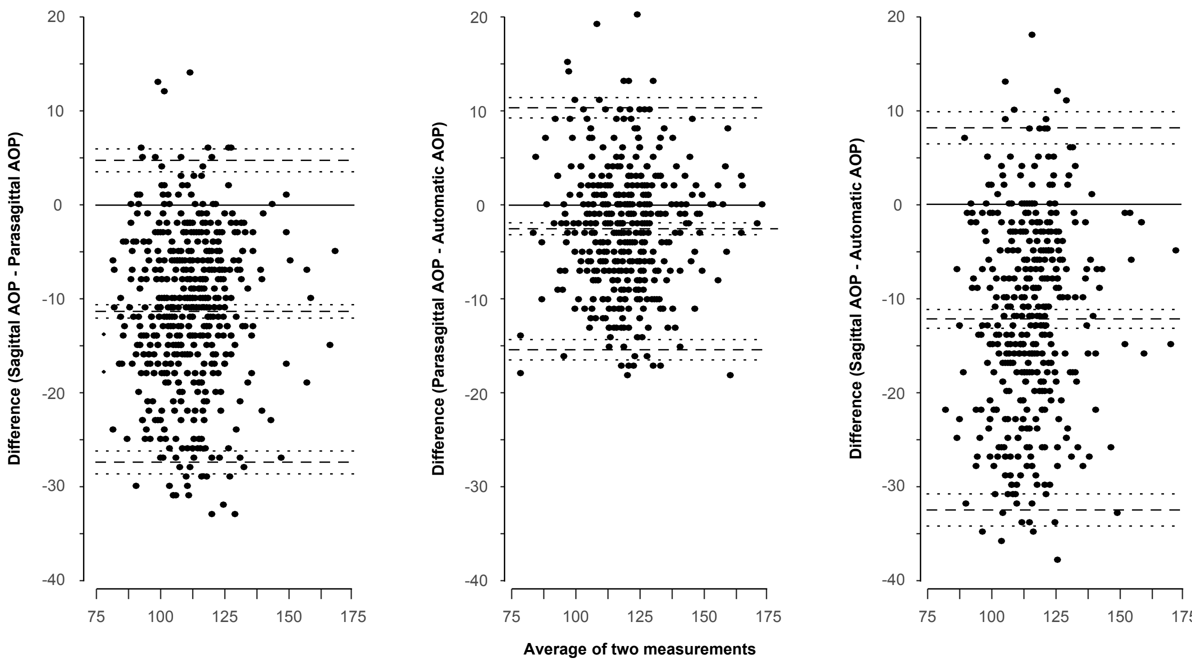
Variable	**5th**	**25th**	**50th**	**75th ***1**	**95th*** 2**
Angle of progression	-0.05*** (-0.07, -0.03)	-0.05*** (-0.07, -0.03)	-0.06** (-0.08, -0.04)	-0.07*** (-0.10, -0.04)	-0.10*** (-0.14, -0.05)
Nulliparity	0.91** (0.28, 1.54)	1.75*** (1.11, 2.38)	2.30*** (1.58, 3.01)	2.42*** (1.21, 3.62)	2.44** (0.93, 3.95)
Occiput anterior	-0.75* (-1.40, -0.11)	-0.40 (-1.07, 0.28)	-0.88* (-1.65, -0.10)	-1.07 (-2.41, 0.27)	0.02 (-1.82, 1.86)
Syntocinon	0.31 (-0.26, 0.87)	1.07* (0.25, 1.88)	2.42*** (1.46, 3.37)	2.89*** (1.80, 3.99)	2.76*** (1.17, 4.36)
Epidural	1.94*** (1.30, 2.57)	3.25*** (2.48, 4.02)	3.25*** (2.16, 4.35)	3.61*** (2.28, 4.95)	5.14*** (3.40, 6.88)
Constant	6.49*** (4.62, 8.37)	6.81*** (4.68, 8.94)	9.55*** (7.37, 11.74)	12.18*** (8.54, 15.82)	18.06*** (12.48, 23.64)

Data are presented as estimates (β) and 95% confidence intervals (in parentheses) of models.
[1] ANOVA comparing coefficients of 25th to 75th quantiles and [2] 5th to the 95th quantiles.
***P < 0.001. **P < 0.01. *P < 0.05.

Table 4 Multivariate logistic regression results for the prediction of CS for failure to progress in the first stage of labor

Maternal, pregnancy and ultrasound characteristics	Maternal model		Vaginal examination Model[1]		Ultrasound Model[2]	
	OR (95%CI)	P value	OR (95%CI)	P value	OR (95%CI)	P value
Maternal factors						
Maternal height in cm (-165)	0.92 (0.86-0.97)	0.006	0.92 (0.86-0.98)	0.008	0.92 (0.86-0.98)	0.014
Parity						
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						
Gestation in weeks (-40)	1.33 (1.08-1.66)	0.010	1.36 (1.10-1.71)	0.006		
Cervical dilation in cm			0.78 (0.61-0.96)	0.024		
Ultrasound findings						
Estimated fetal weight in kg (-3.5)					3.69 (1.65-8.59)	0.002
Para-sagittal AOP in degrees (-113)					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	

OR, odds ratio. CI, confidence interval. AIC, Akaike's Information Criteria. AUC, area under curve. [1] ANOVA comparing vaginal examination model to maternal model, $p < .01$. [2]ANOVA comparing ultrasound model to maternal model, $P < .001$.


Difference (Sagittal AOP - Parasagittal AOP)
Difference (Parasagittal AOP - Automatic AOP)
Difference (Sagittal AOP - Automatic AOP)
20
10
0
-10
-20
-30
-40
75
100
125
150
175
Average of two measurements

sFigure1