BMJ Open

Prediction of gestational age with symphysis-fundal height and estimated uterine volume in a pregnancy cohort in Sylhet, Bangladesh

Anne C C Lee ,[1,2] Rachel Whelan,[1] Nazmun Nahar Bably,[3] Lauren E Schaeffer,[1] Sayedur Rahman,[4] Salahuddin Ahmed,[4] Syed Mamun Ibne Moin,[4] Nazma Begum,[4] Mohammad Abdul Quaiyum,[5] Bernard Rosner,[6,7] James A Litch,[8] Abdullah H Baqui,[9] Blair J Wylie[10]

To cite: Lee ACC, Whelan R, Bably NN, *et al*. Prediction of gestational age with symphysis-fundal height and estimated uterine volume in a pregnancy cohort in Sylhet, Bangladesh. *BMJ Open* 2020;**10**:e034942. doi:10.1136/bmjopen-2019-034942

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2019-034942).

Received 11 October 2019
Revised 09 December 2019
Accepted 13 December 2019

For numbered affiliations see end of article.

Correspondence to
Dr Anne C C Lee;
alee6@bwh.harvard.edu

ABSTRACT

Objective To improve the accuracy of the prediction of gestational age (GA) before birth with the standardised measurement of symphysis-fundal height (SFH), estimation of uterine volume, and statistical modelling including maternal anthropometrics and other factors.
Design Prospective pregnancy cohort study.
Setting Rural communities in Sylhet, Bangladesh.
Participants 1516 women with singleton pregnancies with early pregnancy ultrasound dating (<20 weeks); 1486 completed follow-up.
Methods SFH and abdominal girth were measured at subsequent antenatal care (ANC) visits by community health workers at 24 to 28, 32 to 36, and/or >37 weeks gestation. An estimated uterine volume (EUV) was calculated from these measures. Data on pregnancy characteristics and other maternal anthropometrics were also collected.
Primary outcome measure GA at subsequent ANC visits, as defined by early ultrasound dating.
Results 1486 (98%) women had at least one subsequent ANC visit, 1102 (74%) women had two subsequent ANC visits, and 748 (50%) had three visits. Using the common clinical practice of approximating the GA (in weeks) with the SFH measurement (cm), SFH systematically underestimated GA in late pregnancy (mean difference −4.4 weeks, 95% limits of agreement −12.5 to 3.7). For the classification of GA <28 weeks, SFH <26 cm had 85% sensitivity and 81% specificity; and for GA <34 weeks, SFH <29 cm had 83% sensitivity and 71% specificity. EUV had similar diagnostic accuracy. Despite rigorous statistical modelling of SFH, accounting for repeated longitudinal measurements and additional predictors, the best model without including a known last menstrual period predicted 95% of pregnancy dates within ±7.4 weeks of early ultrasound dating.
Conclusions We were unable to predict GA with a high degree of accuracy before birth using maternal anthropometric measures and other available maternal characteristics. Efforts to improve GA dating in low- and middle-income countries before birth should focus on increasing coverage and training of ultrasonography.
Trial registration number NCT01572532

Strengths and limitations of this study

- ► This study was conducted in a well-characterised prospective pregnancy cohort with early ultrasonography dating at <20 weeks gestation.
- ► Intensive training and standardisation in measurement techniques for symphysis-fundal height (SFH) were performed.
- ► The additional dimension of abdominal girth was measured in order to calculate an estimated uterine volume.
- ► Statistical modelling was performed to account for longitudinal measures and potential confounders, including maternal body mass index and parity.
- ► A study limitation was the smaller proportion of mothers with repeated (at least three) longitudinal measures of SFH.

BACKGROUND

In low-and middle-income countries (LMICs), the gestational length of a given pregnancy is commonly unknown or inaccurate. Traditionally, maternal recall of the first day of the last menstrual period (LMP) is used to date pregnancies. However, limitations to LMP include varying menstrual cycle length, misinterpretation of early bleeding and poor recall, with up to two-thirds of women in LMICs without a recorded LMP.[1–3] Early pregnancy ultrasound is considered the gold standard for pregnancy dating; however, access is variable and often late in pregnancy, when it is less reliable for dating.[4] In sub-Saharan Africa, only 7% of pregnant women are estimated to have access to ultrasonography.[5] In India, access to ultrasonography in pregnancy increased from 24% in 2005 to 61% by 2016, although coverage is higher in urban areas and among women from higher wealth quintiles.[6] In a majority of LMICs, when ultrasound or menstrual data are unavailable, measurement of symphysis-fundal height (SFH), also known as uterine fundal height, is commonly used to estimate the gestational age (GA) of pregnancy.[4]

Accurate gestational dating in pregnancy is required in order to appropriately deliver interventions for preterm labour and premature infants. Before birth, the delivery of antenatal corticosteroids is predicated on gestational age determination. The Global Network's Antenatal Corticosteroids Trial failed to show benefits among small infants (<5th percentile for birth weight) and was associated with an overall increase in neonatal mortality and stillbirth, and higher rates of suspected maternal infection in the intervention group.[7] Investigators considered inaccuracy of gestational age determination, and thus inaccurate classification of preterm labour, as a potential reason for these findings.[8] After birth, gestational age also may determine an infant's clinical management. In many settings, the gestational age determines the viability of the foetus, and infants <28 weeks are not provided supportive care. The early identification of preterm infants can aid the early receipt of respiratory support, such as continuous positive airway pressure (CPAP), and other evidence-based interventions in LMICs.[9 10]

Measuring the symphysis-fundal height (SFH), the vertical distance from the superior aspect of the symphysis pubis to the apex of the uterine fundus, is typically performed at antenatal care (ANC) visits and frequently the primary method of pregnancy dating in LMICs.[4 11] SFH is relatively simple to measure, and the standard teaching is that the number of centimetres equals the weeks of gestation.[12] However, this relationship (1 cm=1 week) is unlikely to be true in settings with high rates of pregnancy morbidity, maternal undernutrition and consequent foetal growth restriction.[12] SFH is a single dimensional measure, imprecise, variable between measurers and influenced by other factors, such as maternal parity, uterine fibroids, foetal position and station.[12] Furthermore, few studies have evaluated the accuracy of SFH and other maternal anthropometrics to predict gestational age compared with an early ultrasound gold standard.

There is an urgent need for programmatically feasible and accurate methods of gestational age dating in LMICs.[13] We hypothesised that the estimation of uterine volume by incorporating measurement of a pregnant woman's abdominal girth (AG), or circumference, may improve the estimation of foetal size, as a surrogate of gestational age. Estimated uterine volume (EUV) has previously been proposed to estimate birth weight,[14] but has not yet been studied or validated for estimating gestational age. Furthermore, we also hypothesised that accounting for other factors that may influence the relationship of SFH or EUV and gestational age, such as maternal nutritional status and parity, may improve the prediction accuracy of these maternal anthropometric measures. The main objective of this research was to evaluate maternal SFH and EUV as predictors of gestational age defined by the gold standard of early ultrasound dating. We aimed to use advanced statistical modelling to develop more accurate prediction models for gestational age using maternal anthropometrics and other variables in a well-dated pregnancy cohort in rural Bangladesh.

METHODS

Study setting and population

The study was conducted in the Projahnmo research site in Sylhet, Bangladesh.[15 16] Projahnmo is a collaboration of the Ministry of Health and Family Welfare of Bangladesh, International Centre for Diarrhoeal Disease Research-Bangladesh (icddr,b), Projahnmo Research Foundation, Shimantik (a non-governmental organisation), Brigham and Women's Hospital/Harvard Medical School and the Johns Hopkins Bloomberg School of Public Health. The study site is located in northeastern Bangladesh in two subdistricts in rural Sylhet (Kanaighat and Zakiganj: 290 km^2). The estimated population of the selected study areas was approximately 120 000 with an annual birth cohort of 2800.

Pregnancy surveillance and gold standard dating

For the current study, the Projahnmo research group was enrolling pregnancy cohorts for two ongoing studies in the site.[15 16] All women in the study area were provided menstrual calendars at study initiation, prospectively recorded the first day of their LMP, and were visited monthly by community health workers (CHWs) to inquire about menstrual dates. At the first missed period, women had a urine pregnancy test and were scheduled for an ultrasound for pregnancy dating. For this study, we consecutively enrolled women who had an early pregnancy (<20 week) ultrasound for pregnancy dating. Pregnant women were enrolled from 9 December 2014 to 12 November 2016. Baseline characteristics for the study participants were collected, including maternal/paternal age, socioeconomic status and obstetric history.

For early pregnancy gold standard ultrasound pregnancy dating, biometric parameters (crown rump length (CRL), bi-parietal diameter (BPD) or femur length) were measured three times per standard ultrasound operating procedures, and the median value for each measure was used for analysis. Quarterly external quality control reviews of ultrasound images were conducted by a maternal-foetal medicine obstetrician. The INTERGROWTH CRL formula[17] was used to determine GA for CRL measurements <95 mm, and BPD (Hadlock *et al*[18]) was used if CRL was ≥95 mm. Only singleton pregnancies were included in the analysis.

Antenatal care visits and maternal anthropometric measures

Women were visited at home by CHWs who conducted antenatal visits between 24 to 28 weeks, 32 to 36 weeks and >37 weeks gestation, based on LMP dating. During each of these home visits, CHWs performed and recorded measures of SFH, AG, weight, height and maternal mid-upper arm circumference (MUAC) three times. The median value of each measure was used for analysis. ANC visits occurred in the study cohort from 14 August 2015 to 13 April 2017.

For quality control, in a randomly selected subset of women, a study physician conducted independent, blinded repeat measurements of pregnant women of an estimated 10% of all CHWs' measurements. If discrepancies of >2 cm were noted between physician-CHW measurements, CHWs were directly observed and re-trained in measurement technique.

Measurement of symphysis-fundal height and abdominal girth

Women were instructed to empty their bladder prior to the study visit. Measurements were performed using a non-elastic measuring tape, labelled on one side with cm markings (precision 1 mm), with the pregnant women lying in the supine position. To measure the SFH, the health worker first palpated the superior rim of the pubic bone and demarcated the landmark with a ballpoint pen. She next used her hand to palpate for the uppermost point of the uterine fundus and marked the second landmark with a pen. The measuring tape was then used to measure the distance between the two pen marks with the measuring tape in contact with the skin of the abdomen and in a vertical axis crossing the umbilicus. To measure the AG, the circumference of the abdomen was measured at the level of the umbilicus, with the tape measure perpendicular to the examining table. SFH was measured first, followed by AG, and this procedure was repeated two additional times. If an error was found in demarcating the initial landmark, the initial pen mark was corrected on the subsequent measurement.

Training

CHWs were trained and standardised on the measurement of SFH and AG using methods adapted from the WHO Multicentre Growth Reference Study.[19] The initial training was conducted in May 2015. One master trainer and three medical trainers led 13 CHWs in a 1-day training and standardisation using methods adapted from the WHO's Multicentre Growth Reference Study protocol with 42 subjects in their second and third trimester. The training included a didactic and hands-on training session, during which CHWs performed SFH and AG measurements on a pregnant woman to ensure trainees could correctly identify landmarks, position the measuring tape and read the numbers on the tape. This was followed by a standardisation session during which independent measurements were recorded by the trainer and a trainee for each pregnant subject. Each trainee measured 10 subjects for the initial training-standardisation.[20 21] Precision was assessed using the intra-rater technical error of measurement, and accuracy was assessed using mean differences between trainers and trainees with 95% limits of agreement. A 1-day refresher training was performed in February 2016. The results of the precision and reliability of measurement are reported in detail elsewhere.[22]

Statistical analysis

Estimation of uterine volume

We estimated uterine volume (figure 1A) using the volume of a prolate ellipsoid (figure 1B), as per methods described by Poulos and Langstadt.[14] This method has previously been validated to estimate foetal weight.[23 24] The detailed methods for this calculation are shown in the online supplementary webappendix eMethods 1.

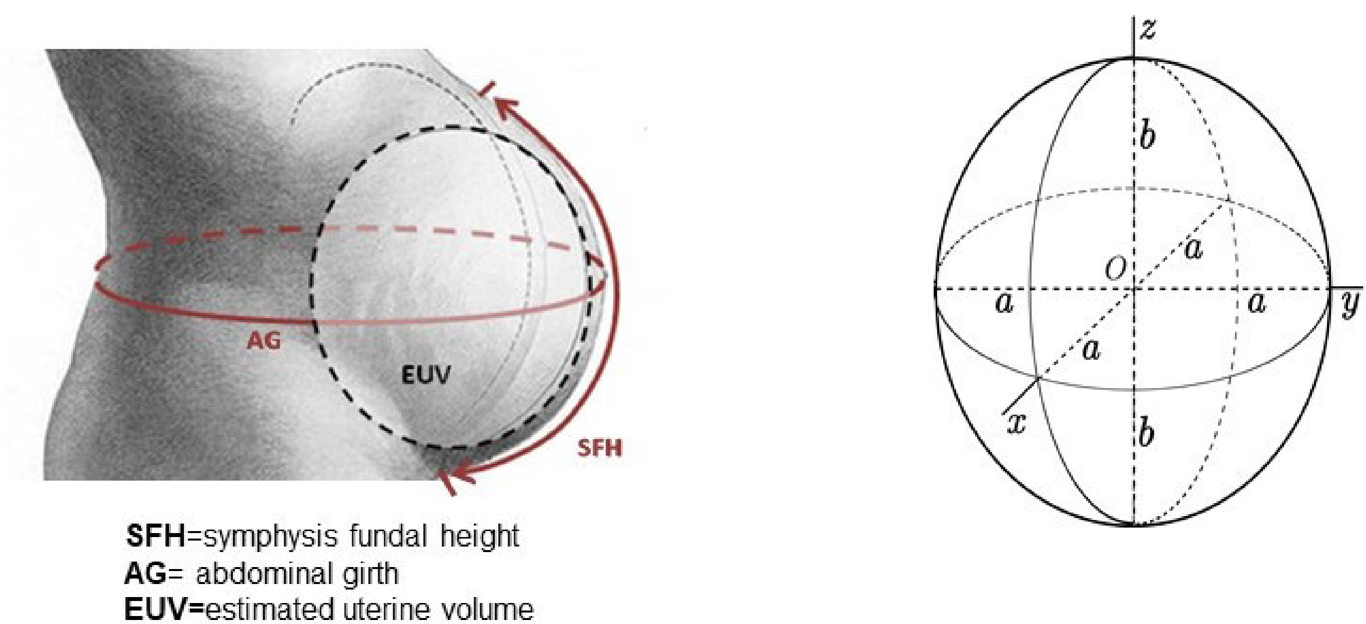

Figure 1 (A) Relationship between estimated uterine volume, symphysis-fundal height and abdominal girth. (B) Geometrical representation of a prolate ellipsoid (source: Image created by Peter Mercator, license: https://creativecommons.org/licenses/by-sa/3.0/legalcode)

Relationship of SFH, AG and EUV with GA

The percentiles (5, 10, 25, 50, 75, 90, 95) of SFH, AG and EUV for pregnant women of a particular completed gestational age (whole weeks) were calculated and plotted in line graphs. We also graphed the median SFH values for each week in gestation for the INTERGROWTH-21st study,[25] a multi-ethnic, population-based study that enrolled pregnant women with optimal pregnancy health and nutrition from eight different countries (Brazil, China, India, Italy, Kenya, Oman, UK, and USA). This was done to compare the trajectory of SFH growth in settings of optimal pregnancy nutrition to our population where the rate of maternal undernutrition is high.

Diagnostic accuracy of SFH, AG and EUV

We chose GA thresholds based on their clinical relevance for pregnancy/neonatal management: <28 weeks (limit of foetal viability in LMIC settings), <34 weeks (threshold for antenatal corticosteroids and tocolytics for imminent preterm birth) and <37 weeks (definition of preterm birth). The areas under the receiver operating curve were calculated to summarise the diagnostic accuracy of identifying these clinical GA thresholds across the range of possible cutoffs of each anthropometric measure (SFH, AG, and EUV). We also calculated the sensitivity, specificity and positive and negative predictive values to identify these thresholds for a range of cutoffs for each anthropometric measure.

Statistical modelling of gestational age

Figure 2 shows the conceptual diagram for the statistical modelling of GA including maternal anthropometric measurements as well as potential confounders, including maternal nutritional status and parity.

Online supplementary webtable 1 shows the variables included in statistical modelling. The primary outcome, or dependent variable, was the gold standard GA (continuous, in weeks) at the time of the ANC visit as determined by the enrolment in early pregnancy ultrasound (<20 weeks). We used generalised linear mixed models estimated using restricted maximum likelihood estimation with a patient-level random effect (clustering) to account for repeated measures. We examined the relationship between GA and the primary continuous predictors (SFH, EUV) with linear, logarithmic and restricted cubic spline models,[26] ultimately choosing the natural log transformation as the best fit. In sensitivity analysis, we also included a variable for time since LMP, to assess the prediction accuracy in cases when LMP is known.

Separate univariate and multivariable models were produced for each of the main predictors. To develop the multivariable models, additional variables were chosen *a priori* that might affect the relationship between maternal size and gestational age (maternal nutritional status, parity and risk factors for foetal growth restriction). Maternal body mass index (BMI), MUAC and height as continuous variables were also considered. The following variables were categorised: maternal parity (0, 1, 2, 3, 4+), maternal education (<5 years, 5 to 9 year, 10+years), substance use (tobacco, alcohol, betel nut: current, never or past use). Models were fitted with a forward stepwise approach to select additional covariates, including predictors with statistical significance

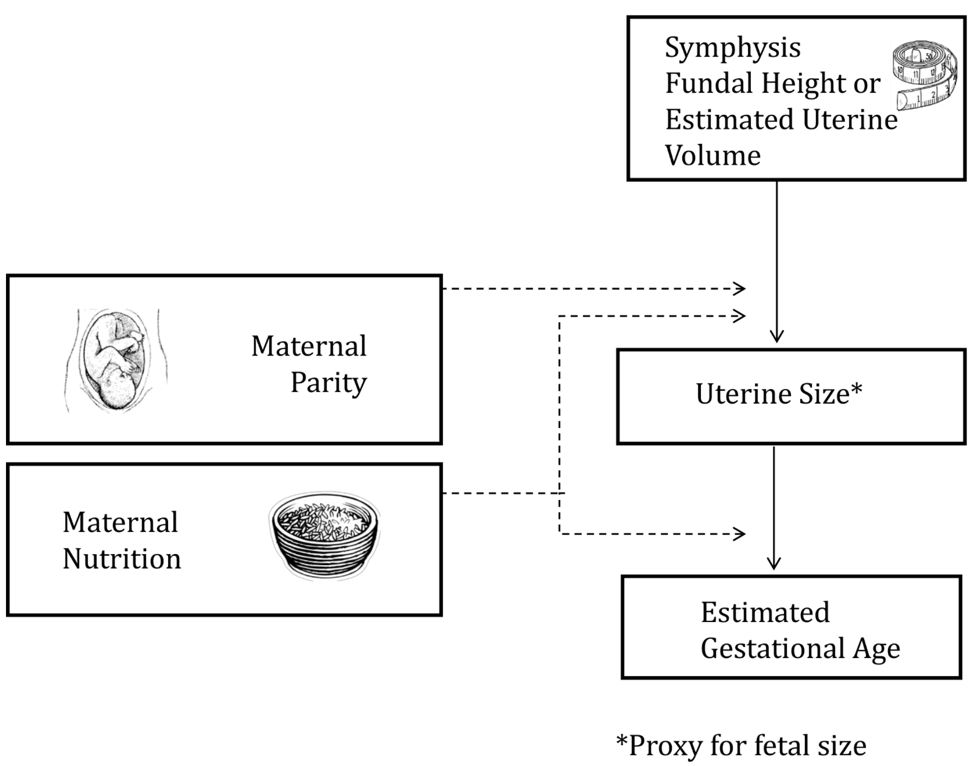

Figure 2 Diagram of relationship between gestational age, maternal anthropometrics and influencing factors.

(p<0.05). Variance inflation factors (VIFs) were calculated to assess multicollinearity among factors. R^2, adjusted R^2, Akaike information criterion (AIC), and 95% predictive intervals were calculated to compare the performance of the models.

White *et al* showed that models accounting for multiple measures of SFH had improved prediction accuracy by accounting for the rate of change in SFH in pregnancy.[27] We thus also conducted modelling to account for the multiple longitudinal measures of SFH, only including women who had three sequential SFH measurements in a sub-analysis. These methods are detailed in the online supplementary webappendix eMethods 2. Stata 14.0 (StataCorp, 2015, College Station, Texas: StataCorp LP) was used for all data analyses.

Written consent was obtained from all literate study participants. Among illiterate study participants, a woman's thumbprint was obtained as well as signature of an impartial, literate witness.

Patient and public involvement

This study included patient recruitment by CHWs. For all other parts of this study, patients, caregivers and laypeople were not involved in the development of the research question, study design or outcome measures, nor the interpretation or writing up of the results. Data from this study is available on request. Investigators may share the results with local ministries of health, patients (including original study participants) and relevant medical organisations in the communities where the study was conducted, as well as in other LMICs.

RESULTS

Two thousand four hundred and fifty-six pregnant women were enrolled in the parent studies during the study period, of which 1991 (79%) had an ultrasound at <20 weeks gestation. Among these, 1516 women with singleton pregnancies were consented and enrolled in the current study. 1486 completed follow-up with subsequent ANC anthropometric measurements and were included in final analysis. 1084 (72.9%) women attended ANC visit 1 (target 24 to 28 weeks), 1316 (88.6%) had ANC visit 2 (target 32 to 36 weeks) and 1084 (72.9%) had ANC visit 3 (≥37 weeks). 1102 (74.2%) women had at least two ANC visits, and 748 (50.3%) had three ANC visits with complete measurements; 9 (0.6%) pregnant women delivered at <30 weeks gestation, and 137 (9.2%) delivered at <37 weeks gestation.

In this population, at enrolment, the mean participant age was 23.5 (±4.5) years, height 149.7 (±5.2) cm, and weight 44.9 (±7.5) kg (table 1). Mothers had on average 6.5 (±3.0) years of education. The society is agrarian with the primary paternal occupations being farming (15.5%), daily wages (42.9%) and self-employment (20.2%) related to agriculture.

Table 1 Characteristics of mothers-infants (n=1486)

Characteristic	Summary statistic*
Maternal age (years)	23.5±4.5
Maternal education (years)	6.5±3.0
Parity	
0	7.6% (76/1003)
1	40.3% (404/1003)
2	24.6% (247/1003)
3	15.4% (154/1003)
4+	12.2% (122/1003)
Maternal height (cm)	149.7±5.2
Maternal weight at baseline visit (kg)	44.9±7.5
Maternal MUAC (cm)	22.9±2.4
Betel nut chewing status	
Never	61.7% (915/1483)
Quit pre-pregnancy	0.3% (4/1483)
Currently sniffing/chewing	38.0% (564/1483)
Gestational age at delivery (weeks)	40.0±0.1
Preterm birth (<37 weeks GA)	9.2% (137/1486)
Birth weight at newborn exam (gm)	2721±469
Small for gestational age†	23.4% (279/1194)

*Reported as mean±SD or % (n/N).
†SGA defined as <10 percentile birth weight for gestational age and sex, as classified by INTERGROWTH-21st neonatal standard.[32]
GA, gestational age; MUAC, mid-upper arm circumference; SGA, small for gestational age.

Relationship of SFH, AG and EUV and gestational age

Figure 3A shows the centiles of SFH by each completed week of gestation, with the coloured lines representing centiles of growth from our study population in Sylhet, Bangladesh. This is compared with the INTERGROWTH-21st's 50th percentile for SFH (black dotted line). The INTERGROWTH study population was from eight countries and recruited only healthy women without comorbidities and with optimal nutritional status.[25] At 34 and 37 weeks, the median centile for the INTERGROWTH population was equivalent to the upper 90th percentile SFH measures for the Sylheti population figure 3B–C.

Also in figure 3A-C, the black dashed line illustrates the standard clinical teaching that the SFH equals the GA (1 cm SFH=1 week increase in GA) after 20 weeks gestation. The rate of increase in SFH in the Sylheti population was substantially lower than this traditional obstetric teaching from high-income countries, that was more similar to the INTERGROWTH median values. A Bland-Altman plot comparing the agreement of GA estimated with the 1 cm SFH=1 week GA rule versus ultrasound based GA is shown in online supplementary webfigure 1. In our population, the SFH systematically underestimated GA with a trend of increasing underestimation in

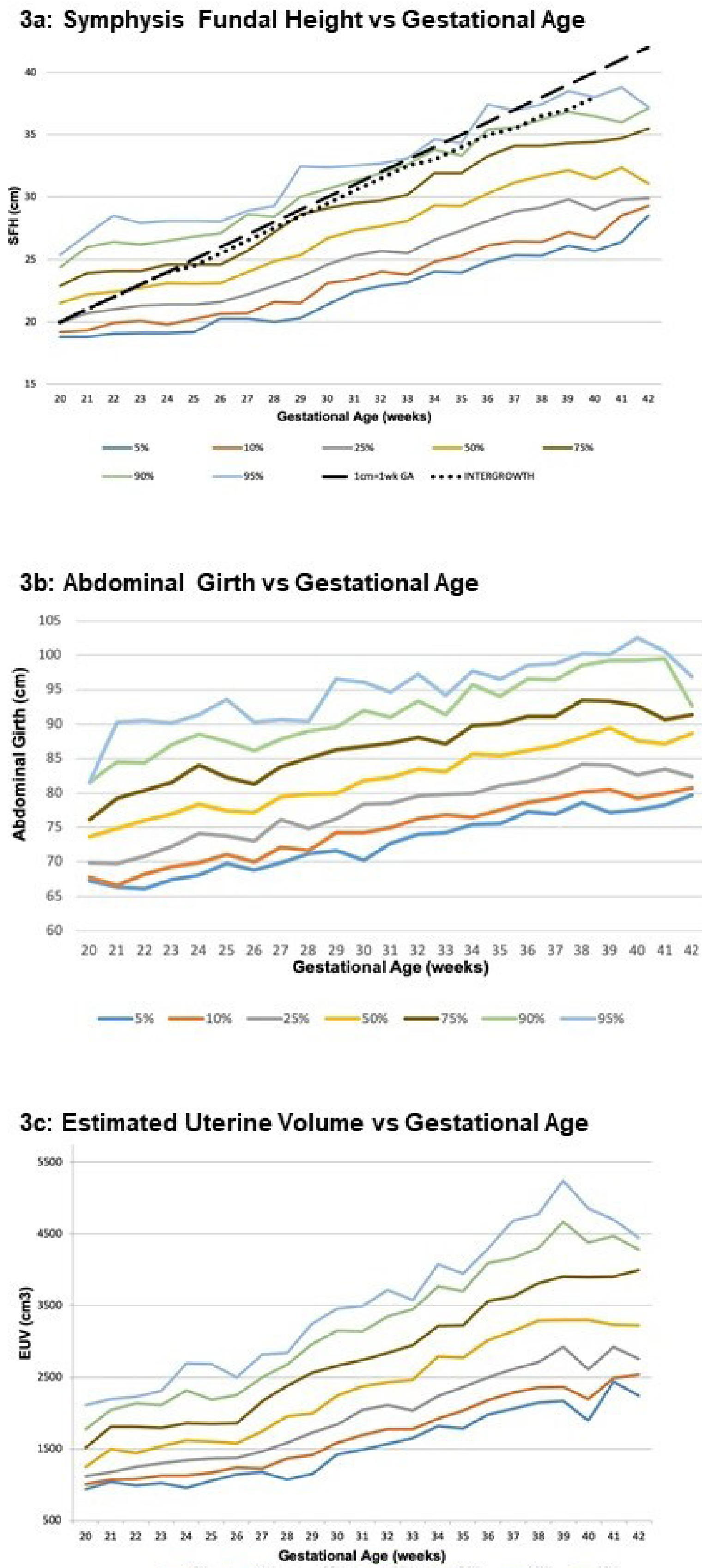

Figure 3 (A) Symphysis-fundal height by gestational age (completed weeks) in Sylhet cohort (percentiles), with INTERGROWTH median measurements (50% percentile, dotted line), and clinical assumption (1cm=1week, dashed line). (B) Abdominal girth by gestational age in Sylhet cohort (percentiles). (C) Estimated uterine volume by gestational age in Sylhet cohort (percentiles). EUV, estimated uterine volume; GA, gestational age; SFH, symphysis-fundal height.

Table 2 Validity of symphysis-fundal height for identifying different gestational age thresholds

	GA <28 weeks				GA <34 weeks				GA <37 weeks			
SFH (cm)	Sens	Spec	PPV	NPV	Sens	Spec	PPV	NPV	Sens	Spec	PPV	NPV
<23	0.497	0.959	0.828	0.827	0.280	0.992	0.982	0.480	0.225	0.997	0.995	0.296
<24	0.653	0.919	0.761	0.869	0.392	0.978	0.964	0.518	0.320	0.990	0.989	0.322
<25	0.784	0.864	0.696	0.910	0.500	0.948	0.935	0.559	0.414	0.967	0.975	0.350
<26	**0.853**	**0.806**	**0.636**	**0.932**	0.583	0.919	0.915	0.596	0.487	0.942	0.962	0.375
<27	0.902	0.713	0.556	0.948	0.676	0.858	0.877	0.639	0.579	0.896	0.945	0.410
<28	0.935	0.635	0.505	0.961	**0.745**	**0.799**	**0.847**	**0.677**	0.651	0.853	0.931	0.444
<29	0.972	0.530	0.451	0.979	**0.826**	**0.707**	**0.808**	**0.732**	**0.739**	**0.773**	**0.909**	**0.491**
<30	0.984	0.421	0.403	0.985	0.887	0.594	0.765	0.778	**0.812**	**0.667**	**0.882**	**0.537**
<31	0.989	0.333	0.371	0.987	0.929	0.495	0.733	0.823	0.866	0.569	0.860	0.580
<32	0.990	0.269	0.350	0.985	0.949	0.412	0.707	0.844	0.898	0.481	0.841	0.606

GA, gestational age; NPV, negative predictive values; PPV, positive predictive values; Sens, sensitivity; SFH, symphysis-fundal height; Spec, specificity.

later pregnancy (mean difference −4.42, 95% limits of agreement (LOA) −12.48 to 3.65 weeks).

Ffigure 3B–C shows the centiles of AG and EUV. From 20 to 38 weeks GA, mean AG increased ~1 cm/week, from 73.2 cm to 89.0 cm, respectively.

Diagnostic accuracy of SFH, AG and EUV to identify different gestational age thresholds

For SFH, the areas under the receiver operator characteristic curve (AUCs) to identify GA thresholds of 28, 34, and 37 weeks were 0.90, 0.86, and 0.83 for SFH, respectively (see online supplementary webfigure 2). For EUV, the AUCs were similar at 0.91, 0.86, and 0.84; respectively. For AG, the AUCs were lower, ranging 0.74 to 0.78.

The sensitivity and specificity of a range of measures for SFH and EUV to detect the clinical GA thresholds are shown in table 2 and online supplementary webtable 2, respectively. Cut-offs for SFH and EUV are highlighted where sensitivity and specificity are closest to 80% (a desirable minimal diagnostic accuracy for a clinicals screening tool).

Quality control measurements

Physician's conducted independent measurements of SFH and AG on 131 randomly selected pregnant women. SFH measured by the CHWs fell within 2 cm of the physician measurement in 70% of cases (mean bias 0.88 cm, 95% LOA ±3.65 cm). AG measured by CHWs were within 2 cm of physician measurements in 74% of measures (mean bias −0.28, 95% LOA ±3.26 cm).

Statistical models

Table 3 shows the results of model performance of several statistical models including SFH, EUV, and LMP as the main predictors of GA.

In Model A with SFH alone, the average model prediction error across pregnancy was 7.69 weeks (ie, the difference between observed-predicted GA was within ±7.7 weeks in 95% of women). Online supplementary web figure 3 shows the model prediction error by month of gestation. The model tended to overestimate GA in the earlier months of pregnancy, and underestimate GA in later pregnancy. The magnitude of prediction error was similar across months 5 to 9 of pregnancy.

In the multivariable models, in the initial stepwise model, both maternal BMI and MUAC were significant with p values of <0.01; however, there was evidence of moderate correlation between these factors, with VIFs~2 for these factors. BMI was removed from the final multivariable model and in the final model B, there was no evidence of multicollinearity (VIFs <1.1). In model B (including MUAC, maternal parity, and betel nut use), the prediction accuracy was similar to Model A (±7.62weeks). Models C and D that used EUV as a primary predictor did not demonstrate better performance than the SFH models. Finally, model E that accounted for repeat sequential measures resulted in only marginal improvement in prediction accuracy.

In sensitivity analysis, the inclusion of LMP in the model substantially reduced the prediction error of the models. In Model F, LMP alone predicted the early ultrasound date within ±4.65 weeks in 95% of visits. The addition of SFH or other variables to the LMP model (Model G) did not significantly improve the prediction accuracy beyond LMP alone (95% prediction interval ±4.56 weeks).

DISCUSSION

In many LMIC settings, SFH is commonly used to estimate the gestational length of pregnancy and is the primary measurement used for clinical decision-making, such as provision of antenatal corticosteroids, determining viability, and clinical care for the newborn. The reliability and accuracy of SFH to date pregnancies has traditionally been characterised as poor in the published literature. In our population, the traditional clinical rule that SFH equals the GA after 20 weeks gestation, systematically

Table 3 Comparison of statistical models

Model	Predictors	Equation	R^2	Adjusted R^2	AIC	95% prediction interval (±weeks)
A	lnSFH	GA=23.34201(lnSFH)−45.0998	0.5047	0.5045	5.5716	7.69
B	lnSFH, parity, MUAC, betel nut use	GA=23.86946(lnSFH)−0.2278491(parity)−0.1294864(MUAC) +0.114807−43.73981	0.5131	0.5122	5.5554	7.61
C	lnEUV	GA=10.25312(lnEUV)−47.42831	0.5072	0.5070	5.5647	7.66
D	lnEUV, parity, MUAC, betel nut use	GA=10.84176(lnEUV)−0.2485541(parity)−0.3396929(MUAC) +0.1303466(betel)−44.08014	0.5161	0.5151	5.5176	7.47
E	lnSFH for three sequential visits	See online supplementary webappendix	NA*	NA*	NA*	7.39
F	LMP GA	GA=0.9364955(LMP GA)+2.678059	0.8216	0.8216	4.5658	4.65
G	lnSFH, LMP, parity	GA=3.854183(lnSFH) +0.8436906(LMP GA)−0.1075177−6.980482	0.8252	0.8250	4.5307	4.56

*For Model E, refer to online supplementary webappendix eMethods 2 for statistical methods of longitudinal modelling. This involved derivation of the GA at the first SFH measurement using two separate regression models, thus a single R^2, adjusted R^2 and AIC are not reported.

AIC, Akaike information criterion; GA, gestational age; LMP, last menstrual period gestational age (weeks from first day of LMP); lnEUV, natural log transformation of estimated uterine volume; lnSFH, natural log transformation of symphysis fundal height; MUAC, maternal mid-upper arm circumference (cm); parity, previous number of deliveries of a live child.

underestimated GA on average by 4 weeks in the latter half of pregnancy, with greater bias in late pregnancy due to the influence of foetal growth restriction. We evaluated whether statistical methods considering three-dimensional estimated uterine volume, controlling for potential confounders (such as parity or nutritional status), and accounting for the longitudinal nature of the measurements could improve the accuracy of the gestational age prediction. In any model without LMP, we could only achieve an average prediction accuracy of approximately ±7 weeks compared with early ultrasound-defined GA.

Rates of foetal growth restriction are high in the population in Sylhet, Bangladesh, and this affects the rate of uterine growth during pregnancy. In the cohort included in this study, the rate of small for gestational age (SGA) defined by the INTERGROWTH-21st standard (<10th percentile birth weight for GA and sex) was 23%. Accordingly, the rate (or slope) of increase of SFH in the latter half of pregnancy (>20 weeks) was much slower compared with that described in the USA, or the recently published INTERGROWTH SFH curves.[25] This is likely due to the poor nutritional status of women, and subsequently smaller foetuses. Foetal growth restriction makes the prediction of GA, based on infant size, much more challenging. We had anticipated that controlling for maternal nutritional status as well as other risk factors for SGA, such as tobacco/betel nut use, would improve prediction accuracy. However, including those covariates in our statistical models only slightly improved prediction accuracy.

Despite multiple statistical approaches to model GA using our data, we did not achieve a statistical model that had adequate precision to accurately estimate GA. The best performing model with maternal SFH alone, yielded relatively good fit (Model A: adjusted R^2=0.5045), but the prediction accuracy was inadequate, dating 95% of pregnancies within ±7.69 weeks of the gold standard ultrasound GA. Models including estimated uterine volume and/or other potential confounders (such as parity and maternal MUAC) only modestly improved model fit and prediction accuracy. Modelling that accounted for three sequential SFH measures over time also did not improve GA prediction accuracy. A study showed in Myanmar that relatively good prediction accuracy of models could be achieved if women had six repeated measures.[27] In this study the use of six sequential SFH measures had a prediction accuracy of ±15 days, while the use of three measures had a prediction accuracy of ±33 days. Sequential SFH measures are often impractical in low-income countries, with poor ANC attendance and a limited number of women seeking sequential visits. In sensitivity analysis, LMP collected in our study predicted GA with improved accuracy, with LMP dates predicting early ultrasound GA within ±4.65 weeks in 95% of women. However, a caveat is that in our study LMP was rarely missing and was rigorously captured as part of an intervention trial with prospective LMP calendars and monthly CHW home visits. LMP collected as part of routine ANC in health systems is commonly missing, and most likely would not have this degree of accuracy. Furthermore, the combination of LMP and SFH did not improve prediction accuracy in our models compared with LMP alone.

Given the poor prediction accuracy of SFH to estimate GA, future efforts to improve gestational age dating before birth should focus on increasing coverage of first and second trimester ultrasonography in LMICs (before 24 weeks), consistent with current WHO guidelines.[28] While ultrasonography in third trimester pregnancy has traditionally been considered less accurate due to the influence of foetal growth restriction, recent studies from the INTERGROWTH-21st and WHO AMANHI[29] research groups have validated new measures and developed new equations with improved prediction accuracy, dating pregnancies to within approximately 2 weeks of first trimester ultrasound dating.[25] Challenges must be considered when scaling up of ultrasonography in LMICs, including the training and standardisation of sonographers, cost of ultrasound equipment, the potential implications of sex selection and adequate health system capacity to manage complications identified. This process requires substantial strengthening of the health systems and engagement of local stakeholders and ministries of health.

In similar rural South Asian populations, where there are no other options to accurately estimate GA (such as reliable LMP or ultrasound), we identified measurements that could be considered potential thresholds for referral from primary to secondary/tertiary level facilities for evaluation and clinical obstetric or neonatal management. In LMICs, 28 weeks is often considered the threshold for viability and provision of neonatal resuscitation and supportive care. SFH <26 cm classified GA <28 weeks with 85% sensitivity and 81% specificity. For 34 weeks, the threshold below which antenatal corticosteroid administration is indicated, SFH <29 cm classified GA <34 weeks with 83% sensitivity and 71% specificity. However, given the relative inaccuracy of SFH, we do not recommend that SFH to be used to determine the provision of clinical interventions, but rather to determine that referral to a higher-level facility is indicated.

In the prior literature, abdominal girth and estimated uterine volume have been used to predict estimated foetal weight.[30 31] Taken alone, abdominal girth was not predictive of GA and did not have high diagnostic accuracy for identifying GA thresholds. EUV had slightly improved diagnostic accuracy compared with SFH. However, the minimal improvement in prediction accuracy does not justify the programmatic burden of training and standardising this additional measure in antenatal care.

There are limitations to this study and analysis. The health workers conducting the measurements in this study were CHWs and not physicians or trained medical personnel, and it is plausible that the accuracy of maternal anthropometric measures could be higher when performed by medical personnel. We conducted quality control with a study physician independently re-measuring anthropometrics in a random 10%. The rate of small for gestational age, a proxy for foetal growth restriction, is high in this population, and it is possible that SFH may be more accurate for the prediction of GA in settings with lower rates of SGA. Finally, it is possible that more accurate dating with anthropometric measurements could be established with six or more subsequent ANC visits as shown by White *et al*,[27] and participants in this study only had up to three longitudinal measurements.

CONCLUSIONS

In primary care facilities in low- and middle-income settings, gestational dating is frequently unknown and assessment of the uterine size by external clinical measurements may be the only method available to estimate gestational age before birth. In this study, despite intensive training of health workers to reliably measure SFH, estimating uterine volume and using longitudinal measures and advanced statistical modelling, we were unable to predict gestational age to high levels of accuracy from maternal pregnancy anthropometric measures. Efforts to improve gestational age dating before birth should focus on increasing coverage and training of ultrasonography. This will require concerted efforts by stakeholders, ministries of health and funders to increase access to and coverage of these services, particularly in low-income and hard to reach communities.

Author affiliations
[1]Pediatric Newborn Medicine, Brigham and Women's Hospital, Boston, Massachusetts, USA
[2]Harvard Medical School, Boston, Massachusetts, USA
[3]Health Section, United Nations Childrens Fund, Rangpur, Bangladesh
[4]Projahnmo Research Foundation, Dhaka, Bangladesh
[5]Centre for Reproductive Health, International Centre for Diarrhoeal Disease Research, Bangladesh (icddr,b), Dhaka, Bangladesh
[6]Biostatistics, Harvard Medical School, Boston, Massachusetts, USA
[7]Channing Department of Network Medicine, Brigham and Women's Hospital, Boston, Massachusetts, USA
[8]Global Alliance to Prevent Prematurity and Stillbirth (GAPPS), Lynnwood, Washington, USA
[9]Department of International Health, Johns Hopkins University Bloomberg School of Public Health, Baltimore, Maryland, USA
[10]Department of Obstetrics and Gynecology, Beth Israel Deaconess Medical Center, Boston, Massachusetts, USA

Acknowledgements We would like to acknowledge the entire Projahnmo team who made this study possible, including the management, medical officers, sonographers, project officers, field staff, and community health workers. We would like to thank Ingrid Olson and Lian Folger with their assistance in preparing the article for submission. Finally, we would like to thank the women and families who participated in this study.

Contributors ACCL, BJW, and RW designed the study, obtained funding, designed data collection tools, and monitored data collection. ACCL developed the statistical analysis plan, analysed data, and drafted the paper and is the guarantor. RW and NNB conducted training and monitored and oversaw data collection for the study; NNB oversaw quality control measures and re-training of community health workers during the study. LES conducted data analysis and helped draft and revise the paper. SR and SA were responsible for the overall study implementation in

the field site, and monitored data collection and coordination of ongoing study activities. SMIM and NB helped develop data collection tools, and oversaw and coordinated data collection. BR helped develop the statistical analysis plan. AHB, MAQ, and JAL provided inputs in the study design, data analysis, and interpretation of results. All authors provided critical inputs into the intellectual content, revisions of the paper, and approved the final version for publication.

Funding This study was supported by the United States Agency for International Development (USAID), under the terms of the Cooperative Agreement AID-OAA-A-14-00049, and the Every Preemie–SCALE programme consortium of the Global Alliance of Prematurity and Stillbirth (GAPPS), Project Concern International (PCI) and the American College of Nurse Midwives (ACNM), that provides practical, catalytic, and scalable approaches to expand the uptake of preterm birth and low birth weight interventions in 25 USAID maternal and child health priority countries in Africa and Asia. James A. Litch was principal investigator of the Every Preemie–SCALE Implementation Research Program, and provided inputs on the study design and interpretation of the data/results, as well as feedback on drafts of the manuscript.

Competing interests JAL was the principal investigator of Every Preemie–SCALE Implementation Research programme.

Patient consent for publication Not required.

Ethics approval Ethical approvals for the study were obtained from Partners HealthCare (Boston, Massachusetts, USA), Johns Hopkins Bloomberg School of Public Health (Baltimore, Maryland, USA), and icddr,b (Dhaka, Bangladesh).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request.

ORCID iD

Anne C C Lee http://orcid.org/0000-0003-2654-9862

REFERENCES

1 Rijken MJ, Lee SJ, Boel ME, *et al*. Obstetric ultrasound scanning by local health workers in a refugee camp on the Thai-Burmese border. *Ultrasound Obstet Gynecol* 2009;34:395–403.
2 Hall MH, Carr-Hill RA, Fraser C, *et al*. The extent and antecedents of uncertain gestation. *Br J Obstet Gynaecol* 1985;92:445–51.
3 Savitz DA, Terry JW, Dole N, *et al*. Comparison of pregnancy dating by last menstrual period, ultrasound scanning, and their combination. *Am J Obstet Gynecol* 2002;187:1660–6.
4 Blencowe H, Cousens S, Chou D, *et al*. Born too soon: the global epidemiology of 15 million preterm births. *Reprod Health* 2013;10:S2.
5 Aliyu LD, Kurjak A, Wataganara T, *et al*. Ultrasound in Africa: what can really be done? *J Perinat Med* 2016;44:119–23.
6 *Nathional family health survey (NFHS-4) 2015-16: India*. Mumbai: International Institute for Population Sciences (IPPS) and ICF, 2017.
7 Althabe F, Belizán JM, McClure EM, *et al*. A population-based, multifaceted strategy to implement antenatal corticosteroid treatment versus standard care for the reduction of neonatal mortality due to preterm birth in low-income and middle-income countries: the ACT cluster-randomised trial. *Lancet* 2015;385:629–39.
8 McClure EM, Goldenberg RL, Jobe AH, *et al*. Reducing neonatal mortality associated with preterm birth: gaps in knowledge of the impact of antenatal corticosteroids on preterm birth outcomes in low-middle income countries. *Reprod Health* 2016;13:61.
9 Bhutta ZA, Das JK, Bahl R, *et al*. Can available interventions end preventable deaths in mothers, newborn babies, and stillbirths, and at what cost? *Lancet* 2014;384:347–70.
10 *WHO recommendations on interventions to improve preterm birth outcomes*. Geneva, Switzerland: World Health Organizaiton (WHO), 2015.
11 Neilson JP. Symphysis-fundal height measurement in pregnancy. *Cochrane Database Syst Rev* 2000;2:CD000944.
12 Fetal. *Growth restriction: current evidence and clinical practice*. Springer, 2019.
13 Greensides D, Robb-McCord J, Noriega A, *et al*. Antenatal corticosteroids for women at risk of imminent preterm birth in 7 sub-Saharan African countries: a policy and implementation landscape analysis. *Glob Health Sci Pract* 2018;6:644–56.
14 Poulos PP, Langstadt JR. The volume of the uterus during labor and its correlation with birth weight. I. A method for the prediction of birth weight. *Am J Obstet Gynecol* 1953;65:233–44.
15 Lee AC, Mullany LC, Quaiyum M, *et al*. Effect of population-based antenatal screening and treatment of genitourinary tract infections on birth outcomes in Sylhet, Bangladesh (mist): a cluster-randomised clinical trial. *Lancet Glob Health* 2019;7:e148–59.
16 AMANHI study group. Maternal, neonatal deaths and stillbirths mortality cohort study. Burden, timing and causes of maternal and neonatal deaths and stillbirths in sub-Saharan Africa and South Asia: protocol for a prospective cohort study. *J Glob Health* 2016;6:020602.
17 Papageorghiou AT, Kennedy SH, Salomon LJ, *et al*. International standards for early fetal size and pregnancy dating based on ultrasound measurement of crown-rump length in the first trimester of pregnancy. *Ultrasound Obstet Gynecol* 2014;44:641–8.
18 Hadlock FP, Shah YP, Kanon DJ, *et al*. Fetal crown-rump length: reevaluation of relation to menstrual age (5-18 weeks) with high-resolution real-time US. *Radiology* 1992;182:501–5.
19 WHO. The who multicentre growth reference study (MGRS) 2003, 2017. Available: http://www.who.int/childgrowth/mgrs/en/ [Accessed 9 Nov 2017].
20 Ulijaszek SJ, Kerr DA. Anthropometric measurement error and the assessment of nutritional status. *Br J Nutr* 1999;82:165–77.
21 Gibson RS. *Principles of nutritional assessment*. 2nd edn. New York: Oxford University Press, 2005.
22 *Intra- and inter-rater reliability of maternal anthropometric measurements in pregnancy by health workers in rural Bangladesh*. New Haven, CT, USA: Unite for Sight Conference, 2016.
23 Devi PK, Mokadam N. Estimation of foetal weight. *J Obstet Gynaecol India* 1967;539.
24 Buchmann E, Tlale K. A simple clinical formula for predicting fetal weight in labour at term--derivation and validation. *S Afr Med J* 2009;99:457–60.
25 Papageorghiou AT, Ohuma EO, Gravett MG, *et al*. International standards for symphysis-fundal height based on serial measurements from the fetal growth longitudinal study of the INTERGROWTH-21st project: prospective cohort study in eight countries. *BMJ* 2016;355.
26 Stata. mkspline — linear and restricted cubic spline construction. In: LLC S, ed. *stata.com: Stata Manuals*.
27 White LJ, Lee SJ, Stepniewska K, *et al*. Estimation of gestational age from fundal height: a solution for resource-poor settings. *J R Soc Interface* 2012;9:503–10.
28 WHO. WHO recommendations on antental care for a positive pregnancy experience 2016.
29 Wylie BJ. 84: performance of third trimester biometry for gestational age assessment in low resource settings: results of the multi-country, population-based who AMANHI late pregnancy dating sStudy. *Am J Obstet Gynecol* 2018;218:S62.
30 Woo JS, Ngan HY, Au KK, *et al*. Estimation of fetal weight in utero from symphysis-fundal height and abdominal girth measurements. *Aust N Z J Obstet Gynaecol* 1985;25:268–71.
31 Dare FO, Ademowore AS, Ifaturoti OO, *et al*. The value of symphysio-fundal height/abdominal girth measurements in predicting fetal weight. *Int J Gynaecol Obstet* 1990;31:243–8.
32 Villar J, Ismail LC, Victora CG, *et al*. International standards for newborn weight, length, and head circumference by gestational age and sex: the newborn cross-sectional study of the INTERGROWTH-21st project. *Lancet* 2014;384:857–68.