Published in final edited form as:

BJOG. 2018 October ; 125(11): 1397–1404. doi:10.1111/1471-0528.15179.

Estimating gestational age at birth from fundal height and additional anthropometrics: a prospective cohort study

Sarah J. Pugh^{1,†}, Ana M. Ortega-Villa^{2,†}, William Grobman³, Roger B. Newman⁴, John Owen⁵, Deborah A. Wing^{6,7}, Paul S. Albert^{2,†}, and Katherine L. Grantz^{1,†}

¹Epidemiology Branch, Division of Intramural Population Health Research, Eunice Kennedy Shriver National Institute of Child Health and Human Development, Bethesda, MD, 20892 United States

²Biostatistics Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Medical Center Drive, Rockville, MD 20892, United States

³Department of Obstetrics and Gynecology, Feinberg School of Medicine, Northwestern University, Chicago, IL, 60208 United States

⁴Department of Obstetrics and Gynecology, Medical University of South Carolina, Charleston, SC, 29425 United States

⁵The University of Alabama at Birmingham Center for Women's Reproductive Health, Dept of Ob/Gyn, Birmingham, AL, United States

⁶Division of Maternal-Fetal Medicine, Department of Obstetrics-Gynecology, University of California, Irvine, School of Medicine, Orange, CA

⁷Miller Children's Hospital/Long Beach Memorial Medical Center, Long Beach, CA, 92617 United States

Abstract

Objective—Accurate assessment of gestational age is critical to pediatric care, yet is limited in developing countries without access to ultrasound. Our objectives were to assess the accuracy of gestational age at birth (GA) prediction and preterm birth (PTB) classification using routinely collected anthropometry measures.

Correspondence to: Katherine L. Grantz, M.D., M.S., 6710B Rockledge Drive, MSC 7004, Bethesda, MD 20817, Tel: 301-435-6935, Fax: 301-402-2084, katherine.grantz@nih.gov.

[†]Joint first and senior authors

Trial Registration Number: NCT00912132

Declaration of interests: The other authors declare no competing interests. Completed disclosure of interest forms are available to view online as supporting information.

Contribution to Authorship: All authors conceived and designed the research question. WG, RN, JO, DAW, KLG, PSA contributed to the study design and enrollment of patients. SJP, AMOV, PSA, KLG analyzed and contributed to the interpretation of the data. SJP and AMOV drafted the report, and all authors edited and revised the report. All authors are responsible for the integrity of the data and accuracy of the analysis, and all approved the final report.

Details of Ethics Approval: Human subjects' approval was obtained from all participating sites and women gave informed consent. The local institution as stated in the Materials and Methods section has approved human experimentation. Institutional Review Board Project #09-CH-N152 was approved on December 2009

Design—Prospective cohort study

Setting—United States

Population or Sample—2334 non-obese and 468 obese pregnant women

Methods—Enrollment GA was determined based on last menstrual period, confirmed by first-trimester ultrasound. Maternal anthropometry and fundal height (FH) were measured by a standardized protocol at study visits; FH alone was additionally abstracted from medical charts. Neonatal anthropometry measurements were obtained at birth. To estimate GA at delivery, we developed three predictor models using longitudinal FH alone and with maternal and neonatal anthropometry. For all predictors, we repeatedly sampled observations to construct training (60%) and test (40%) sets. Linear mixed models incorporated longitudinal maternal anthropometry and a shared parameter model incorporated neonatal anthropometry. We assessed models' accuracy under varied scenarios.

Main Outcome Measures—Estimated GA at delivery.

Results—Prediction error for various combinations of anthropometric measures ranged between 13.9 to 14.9 days. Longitudinal FH alone predicted GA within 14.9 days weeks with relatively stable prediction errors across individual race/ethnicities [whites (13.9 days), blacks (15.1), Hispanics (15.5) and Asians (13.1)], and correctly identified 75% of PTB's. The model was robust to additional scenarios.

Conclusions—In low-risk, non-obese women, longitudinal FH measures alone can provide a reasonably accurate assessment of GA when ultrasound measures are not available.

Keywords

Fundal Height; Prediction; Gestational Age; Anthropometry

Introduction

Accurate assessment of gestational age is critical to obstetric and pediatric care. Gestational age guides the timing of interventions at the time of delivery that help to balance the risk of fetal death with neonatal morbidity and mortality.¹ Infants delivering preterm (i.e., less than 37 weeks of gestation) are at risk for a host of short- and long-term morbidities.² While it is known that preterm birth is the leading cause of morbidity and mortality in developing countries³, accurate assessment of gestational age to determine an infant's risk status in these countries remains challenging.

Ultrasound is the optimal method to assess gestational age, but cost, training, and maintenance issues limit access in developing countries⁴. Therefore, other approaches to estimate gestational age at the time of birth are used, including recalled last menstrual period (LMP) and cross-sectional symphysis pubis fundal height, as well as neonatal Ballard⁵ and Dubowitz scores⁶. However, these methods all have limitations. Determining Ballard and Dubowitz scores requires technical training that may not be available.⁷ LMP relies on recall, which is often inaccurate⁸ and may be particularly difficult in countries where language or cultural barriers exist between patients and clinicians.^{9–11} Fundal height measurement

requires little training, few resources, and can be accessed in all settings, but its reliability is controversial.¹² Longitudinal measures of fundal height may improve gestational age estimation over cross-sectional measures¹³; however, it remains unknown if the accuracy of gestational age estimation may be further improved by the addition of other longitudinal measures of maternal anthropometry and/or cross-sectional measures of neonatal anthropometry at birth.

The objective of this study was to assess the accuracy of predicting gestational age at the time of birth and preterm birth classification using routinely collected, low-cost measures of longitudinal maternal and cross-sectional neonatal anthropometry. It is of great public health importance to identify a simple, reliable, and universally applicable tool to determine gestational age at birth.

Methods

The presented methods and results represent a secondary analysis of the NICHD Fetal Growth Study-Singletons, a prospective cohort of pregnant women aged 18–40 years with a pregnancy between 8 weeks 0 days and 13 weeks 6 days (mean 12.7 [SD 0.95] weeks of gestation) from 12 US sites from July 2009 through January 2013. The primary purpose of the original NICHD fetal growth study was to establish a standard for normal fetal growth (velocity) and size for gestational age in the US population. Secondary objectives included a collection of blood samples for an etiology study of gestational diabetes. To achieve the first objective, the study recruited 2334 healthy, nonobese (body mass index [BMI] between 19.0–29.9 kg/m²), low-risk women across 4 race/ethnicity groups, who conceived spontaneously and had no obvious risk factors for fetal growth restriction or overgrowth, specifically, non-Hispanic white, African American, Hispanic, and Asian/Pacific Islander women. In addition, the study recruited 468 obese women (BMI between 30–45 kg/m²) with no restriction on race/ethnicity to achieve the additional study aims. Exclusion criteria were similar between nonobese and obese women and included chronic hypertension or high blood pressure under medical supervision (obese women if requiring ≥ 2 medications), pregestational diabetes, chronic renal disease under medical supervision, autoimmune disease, psychiatric disorders, cancer, and HIV/AIDS. Additional exclusion criteria for the nonobese women included a history of preterm low birthweight (<2500 g), or macrosomic (>4000 g) neonate; history of stillbirth or neonatal death; medically assisted conception; cigarette smoking or illicit drug use in past 6 or 12 months, respectively; ≥ 1 daily alcoholic drinks; previous fetal congenital malformation; history of noncommunicable diseases (asthma requiring weekly medication, epilepsy or seizures requiring medication, hematologic disorders, hypertension, thyroid disease); or history of gravid diseases (gestational diabetes, severe preeclampsia/eclampsia, or hemolysis, elevated liver enzymes, low platelet count syndrome). Human subjects' approval was obtained from all participating sites and women gave informed consent. A core outcome set and patient or public involvement were not used in the design of this study.

After all inclusion criteria were met, a screening ultrasound was performed to confirm gestational age based on self-reported last menstrual period (LMP). The difference between ultrasound-estimated and LMP-estimated gestational age was calculated and eligible patients

were enrolled if the LMP-estimated GA was within 5 days for women between 8w0d and 10w6d, 6 days for those between 11w0d and 12w6d, and 7 days for participants between 13w0d and 13w6d. For all enrolled patients, the estimated date of delivery and the gestational ages at each study visit were based on LMP, which was also confirmed by ultrasound.

At the enrollment visit (10w0d - 13w6d), crown rump length was measured in triplicate.¹⁵ Women were randomized into one of four follow-up schedules for five targeted visits throughout gestation to capture weekly data without subjecting women to weekly follow-up visits.¹⁴ At each study visit starting at 10 weeks, per protocol fundal height was measured prior to ultrasound examination after women emptied their bladders.^{16,17} Measurements were taken using the blank side of a 76-cm (30-inch) disposable non-stretchable paper tape by research nurses blinded to gestational age. Fundal height was measured in centimeters first from the fundus to the top of the symphysis pubis and second from the symphysis pubis to the top of the fundus. Measurements were compared for consistency and then averaged for analyses to improve predictive accuracy by being inclusive of varying measurement techniques.¹⁸ Additional fundal heights were abstracted from the prenatal records. Mid-upper arm circumference was measured using a non-stretchable linen tape to the nearest 0.1cm. The triceps skinfold was measured at the point of mid-upper arm circumference and maternal subscapular skinfold was measured on a diagonal, at a 45° incline inferolaterally, at a site just inferior to inferior angle of the scapula using a Lange Skinfold Caliper (Beta Technology, Inc., Santa Cruz, CA)¹⁹. Maternal weight was measured following a standardized anthropometric protocol. Maternal weight at routine antenatal clinical visits was also abstracted from the prenatal records. Gestational weight gain (GWG) was calculated as the difference between maternal weight and self-reported prepregnancy weight.

To improve the precision of estimates by increasing the number of measurements per women, we evaluated the quality of chart abstracted fundal height and maternal weights from prenatal records to see if both the measured values as part of the study protocol and chart-abstracted values from prenatal records could be combined. We found that both fundal height and maternal weights were highly consistent between the two sources and therefore used both sources in our analyses.

Neonatal measurements were collected after delivery. All circumferences were measured using a non-stretch measuring tape. Abdominal circumference was measured at the mid-point between the xiphisternum and the umbilicus and head circumference was measured just above the eyebrows and posteriorly at the maximum protrusion of the occiput. Length was measured using a seca 416 infantometer and birthweight was abstracted from medical charts.

All anthropometric sites were measured in duplicate and if the second measure differed by a pre-specified tolerance value, specific to each anthropometric measure,^{19–21} a third measure was performed. Average readings were used for analysis.

Statistical Analysis

To estimate gestational age at delivery, we developed three predictor models using combinations of longitudinal maternal and cross-sectional neonatal anthropometry measures. For all predictor models, we used a linear mixed model framework for the longitudinal anthropometry data and for the third model, we added a shared parameter to incorporate neonatal anthropometry, which allowed us to estimate model coefficients for prediction of gestational age.

A novel statistical approach for estimating gestational age that allows for a differing number of longitudinal anthropometric predictors at different observation times across pregnancy was developed. The proposed method can be implemented in two steps. In the first step, available data with ultrasound consistent GA (i.e. study data) is used to fit a linear mixed model, where GA is used as a covariate to predict the anthropometry measure(s). In the second step, when the goal is to predict the unknown GA (i.e. when a new woman needs GA predicted), the coefficients from the fitted model in Step 1 are used to calculate the GA value. To mimic this situation, a training-test paradigm is implemented, in which the full dataset is split between a training set and a test set, whereby the model (step 1) is fit in the training set, and the predictions (step 2) and their accuracy are assessed in the test set.

The predictor models were developed in the non-obese cohort considering subjects with at least 2 fundal height measurements beginning at any gestational age, where the training set was limited to a random sample of 60% of the cohort and the test set included the remaining 40%.

For the first and primary predictor model of interest, a univariate longitudinal mixed model was used with fundal height as the anthropometric measure. The second and third predictor models were supplemental models to include additional longitudinal maternal anthropometry and cross-sectional neonatal anthropometry measures in a multivariate longitudinal mixed model and a shared random effects model, respectively. The models were fit using the women in the training data set, where GA is known and model parameters were extracted (step 1). A formula was created to predict gestational age based on the estimated model parameters (Appendix S1), and applied to all members of the test set, in which GA was treated as unknown for prediction (step 2) through the use of the conditional expected value of GA given the observed anthropometry measures.²² Prediction accuracy was evaluated using the ultrasound-consistent GA values for the subjects in the test set by computing the standard deviation of prediction, which is the squared root of the average squared difference between the ultrasound-consistent gestational age and the model-predicted gestational age for each subject. We multiplied the standard error by 1.96 to construct the prediction error, such that in 95% of cases, the difference between the model formula's predicted gestational age and the ultrasound-consistent gestational age at birth is within plus or minus this value (i.e. gestational weeks or days).

We tested the accuracy of the fundal height predictor, the primary model of interest, under a number of different constraints since to ensure that one group (i.e. maternal race, term births) was not driving the accuracy of the model. We tested the robustness of the predictor, or how accurate the predictor remained, to individual races and preterm delivery < 37 weeks

by limiting the test set based on these individual characteristics and recalculated prediction errors. We also calculated the sensitivity and specificity of each predictor for classifying preterm delivery <37 weeks since identifying infants born preterm was of interest. We examined the accuracy of the fundal height model under different scenarios that may occur clinically due to the timing and frequency of measurements. We calculated the prediction error for the fundal height model when only two, three, or four measurements were available and when the measurements were taken starting at 20 weeks, consistent with most clinical guidelines,²³ and after 28 weeks gestation, consistent with late entry to prenatal care, by limiting the test set based on these constraints. In addition, since gestational age is often dated by cross-sectional fundal height in developing countries, we assessed the prediction error when fundal measurements were taken after 20 weeks as indicated by a fundal height of greater than 20 cm (i.e., when gestational age is only estimated by the initial fundal height and not by LMP).

Since the reliability of fundal height decreases with excess adiposity, we developed a separate predictor model in the training set limited to the obese only cohort (n=267) to examine the accuracy using an individualized model based on BMI. We also developed a predictor in the combined non-obese and obese cohorts (n=1601) to assess the accuracy based on a BMI-inclusive population. Analyses were conducted using R version 3.1.2.

Funding

This work was supported by the intramural program at *Eunice Kennedy Shriver* National Institute of Child Health and Human Development and included ARRA funding (Contract numbers: HHSN275200800013C, HHSN275200800002I, HHSN27500006, HHSN275200800003IC, HHSN275200800014C, HHSN275200800012C, HHSN275200800028C, HHSN275201000009C); and a Grand Challenge Exploration Grant (GCE) from the Bill and Melinda Gates Foundation. The funder did not play a role in the development of this research.

Results

We excluded a total of 133 women due to ineligibility after enrollment [10 (0.37%)], miscarriage <20 weeks [9 (0.33%)], stillbirth [6 (0.22%)], anomaly [1 (0.03%)], loss to follow up [13 (0.48%)], refusal to continue [73 (2.7%)], moved [9 (0.33%)], termination [7 (0.26%)], other [5 (0.18%)]. The final analytic sample included 2669 women (95% of the original cohort). The majority of women were 20-<39 years of age (92.6%), married (74.4%), educated beyond high school (70.1%), and held private insurance (56.5%). The racial composition was evenly distributed as per study design between non-Hispanic White, non-Hispanic Black, and Hispanic women, with a smaller proportion of Asian women (16.4%) (Table 1).

To identify the optimal predictor of gestational age as preterm (<37 weeks), we compared the sensitivity and specificity of predictors (Table 2). The specificity remained constant across all predictors at 97%. The combination of fundal height and maternal weight change resulted in the most sensitive predictor (79%), indicating that fundal height and maternal weight would correctly identify 79% of the preterm births. Various combinations of fundal

height, maternal weight change, and maternal subscapular skinfold showed similar sensitivity (77–79%). The addition of neonatal measures to fundal height lowered the sensitivity to 69–73%.

Longitudinal fundal height alone had a prediction error of 13.9 days and a sensitivity and specificity of 75% and 97%, respectively, for gestational age being preterm. Since a single anthropometry measure may be more simply and universally measured in resource-limited settings, we focused the remaining analyses to assess the robustness of fundal height as a single predictor across varying populations and clinical situations.

Overall, the prediction of gestational age from longitudinal measures of fundal height in a non-obese population was robust when applied to individual maternal races and preterm versus term deliveries with a variation of only 13.1 to 15.5 days (Table S1). Table 3 illustrates the accuracy of longitudinal measures of fundal height in predicting gestational age under varying clinical situations such as the frequency of prenatal visits or onset of prenatal care. The prediction error improved by 2.31 days as the frequency of measurements increased from two to four visits. The inclusion of fundal height measurements beginning after 20 weeks [Median (SD) number of measurements; 11 (3.3)] versus including measurements prior to 20 weeks minimally improved the accuracy of the predictor (14.1 versus 14.6 days). The error was slightly increased (15.4 days) when the gestational age of 20 weeks was estimated based on a 20cm fundal height [Median (SD) number of measurements; 11 (3.4)]. Furthermore, in potential scenarios of late prenatal care with measurements beginning after 28 weeks [Median (SD) number of measurements; 8 (2.7)], the error increased from 14.1 to 16.1 days.

The predictor developed using a non-obese cohort may have limited application to women with excess adiposity; therefore, we separately developed and assessed the accuracy of predictors in an obese-only and in a combined non-obese and obese cohort using combinations of longitudinal maternal and cross-sectional neonatal anthropometry. Predictors that were developed in and applied to an obese-only population had a higher prediction error ranging from 17.9 to 23.3 days (Table S2). When the predictor model was developed in the combined obese and non-obese cohort, the prediction error slightly improved compared with the obese-only cohort, but remained larger than the non-obese cohort with an error of 15.7 to 17.4 days (Table S3), suggesting that fundal height measures in populations with high adiposity were less accurate.

Discussion

Main Findings

The presented results are a secondary analysis of the NICHD Fetal Growth Study-Singletons. In non-obese populations, longitudinal measures of fundal height predicted gestational age within a margin of error similar to that of a formula used to estimate gestational age based on fetal biometrics from a third trimester ultrasound between 28–40 weeks (17 days)²³ and classified gestational ages as being preterm with 75% sensitivity and 97% specificity. Predictive accuracy was best when four compared with two or three fundal height measurements were taken in non-obese populations, indicating an improvement in

accuracy as the number of fundal height measurements obtained increases. Additionally, predictive accuracy was improved when measurements were taken beginning after 20 weeks, confirming the low correlation prior to 20 weeks between fundal height measures and gestational age.²⁴ The addition of longitudinal maternal and cross-sectional neonatal anthropometry measures only minimally improved gestational age prediction, highlighting the value of repeated measurements of fundal height alone. Longitudinal measures of fundal height alone predicted gestational age at birth with similar accuracy across each of the maternal race/ethnicities, though performance was best in non-Hispanic White women. The prediction model was not as accurate in obese women. The proposed methodology can be implemented by using the R function provided in Appendix S2, which uses parameters obtained from analyzing the NICHD Fetal Growth Study-Singletons in step 1. The user inputs longitudinal fundal height measurements, the corresponding measurement dates (MM/DD/YY), and the date of delivery and the program output suggests a gestational age at delivery (in days).

Strengths and Limitations

Despite the accuracy of longitudinal fundal height as a predictor, the number of measurements required for accuracy may not always be feasible in developing countries. Still, with only 3 repeated measures, the prediction error remained at plus or minus 2.5 weeks. The predictor was developed using a cohort of healthy women from a developed country, which may not be generalizable to populations with substantial demographic differences. However, despite selecting a cohort of healthy women, a number of women still developed complications during pregnancy and were incorporated into the prediction method. It is also unclear how accurately this method would perform in a population with a high prevalence of intrauterine growth restriction or large for gestational age fetuses, which may influence the fundal height trajectory. While the diagnostic accuracy of the models was estimated using a repeated test-training set paradigm that eliminates the potential for overfitting by evaluating the prediction on the same subjects we developed the predictor, future research is needed to assess the validity of our model to cohorts in developing countries. However, the major strengths of this study include the measurements being taken in a blinded fashion, the racially and demographically diverse population used to develop and assess the predictors, and the novel statistical methods. We developed and tested predictors in mutually exclusive populations to prevent bias.

Interpretation (in light of other evidence)

Routine clinical fundal measurements are recommended by the WHO to identify at-risk fetuses.²⁶ Therefore, much of the current literature focuses on the development of standard fundal height curves to classify under or over growth and estimate gestational age.^{27–30} However, in practice, standard curves poorly predict gestational age.¹⁷ This high variability is often attributed to differences in how gestational age was determined in the development of the curves and demographic variation, which has prompted work to develop population-specific fundal height growth curves.³¹ Yet, a more substantial limitation may be that a one-time cross-sectional measure does not perform as well as incorporating all longitudinal measures, as our findings indicate. Our study builds on the current literature by incorporating the longitudinal nature of fundal height changes and using a single model, in a

diverse cohort, to predict gestational age with reasonable accuracy. Our findings are applicable to clinical practice in developing countries to identify, at the time of birth, the gestational age of the newborn to determine the appropriate treatment for optimal neonatal health. For example, in a developing country, a woman is admitted to the clinic in labor and delivers her child, yet the physician is not aware if the patient was <37 or ≥ 37 weeks. In this situation, the physician could use the woman's chart, inclusive of several fundal height measurements to predict GA, at the time of delivery, with reasonable accuracy.

One previous study also assessed the utility of longitudinal measurements of fundal height.¹³ In a cohort of 2437 Thai-Burmese women, White et al. (2012) used serial fundal height measurements (at least 3 measurements) beginning at 10 weeks gestation to predict gestational age at birth based on the combination of individual linear models. The prediction accuracy improved with an increasing number of measurements, consistent with our study. When 10 fundal height measurements were obtained, the model predicted gestational age within 4 weeks and correctly identified 46% of preterm births. Our study expands upon this evidence to include additional longitudinal maternal anthropometry and neonatal measures in a racially/ethnically diverse population. Furthermore, our study used a novel statistical method that incorporated the longitudinal trajectory of fundal height along with random effects which tailored the estimation for each individual. We additionally calculated individualized estimates of specific racial/ethnic and BMI groups, which refined the prediction error to within 2 weeks and correctly identified 75% of preterm births.

Conclusions

Gestational age assessment presents challenges in developing countries, but is particularly important due to the disproportionately high rate of preterm births³. Our novel statistical method using longitudinal fundal height measures alone provides a simple, low-cost tool for clinicians to identify gestational ages being preterm and predict gestational age at birth within a window of approximately 2 weeks.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

None

Funding: This work was supported by the intramural program at *Eunice Kennedy Shriver* National Institute of Child Health and Human Development and included ARRA funding (Contract numbers: HHSN275200800013C, HHSN275200800002I, HHSN27500006, HHSN275200800003IC, HHSN275200800014C, HHSN275200800012C, HHSN275200800028C, HHSN275201000009C); and a Grand Challenge Exploration Grant (GCE) from the Bill and Melinda Gates Foundation.

References

1. Method for Estimating Due Date. Committee Opinion No. 611. American College of Obstetricians and Gynecologists. *Obstet Gynecol.* 2014; (124):863–6. [PubMed: 25244460]
2. Stark AR. American Academy of Pediatrics Committee on F, Newborn. Levels of neonatal care. *Pediatrics.* 2004; 114(5):1341–7. [PubMed: 15520119]

3. Blencowe H, Cousens S, Oestergaard MZ, Chou D, Moller AB, Narwal R, et al. National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: a systematic analysis and implications. *Lancet*. 2012; 379(9832):2162–72. [PubMed: 22682464]
4. Shah S, Bellows BA, Adedipe AA, Totten JE, Backlund BH, Sajed D. Perceived barriers in the use of ultrasound in developing countries. *Crit Ultrasound J*. 2015; 7(1):28. [PubMed: 26123609]
5. Ballard JL, Khoury JC, Wedig K, Wang L, Eilers-Walsman BL, Lipp R. New Ballard Score, expanded to include extremely premature infants. *J Pediatr*. 1991; 119(3):417–23. [PubMed: 1880657]
6. Dubowitz LM, Dubowitz V, Goldberg C. Clinical assessment of gestational age in the newborn infant. *J Pediatr*. 1970; 77(1):1–10. [PubMed: 5430794]
7. Rosenberg RE, Ahmed AS, Ahmed S, Saha SK, Chowdhury MA, Black RE, et al. Determining gestational age in a low-resource setting: validity of last menstrual period. *J Health Popul Nutr*. 2009; 27(3):332–8. [PubMed: 19507748]
8. Waller DK, Spears WD, Gu Y, Cunningham GC. Assessing number-specific error in the recall of onset of last menstrual period. *Paediatr Perinat Epidemiol*. 2000; 14(3):263–7. [PubMed: 10949219]
9. Alexander GR, Tompkins ME, Petersen DJ, Hulsey TC, Mor J. Discordance between LMP-based and clinically estimated gestational age: implications for research, programs, and policy. *Public Health Rep*. 1995; 110(4):395–402. [PubMed: 7638326]
10. Mongelli M, Wilcox M, Gardosi J. Estimating the date of confinement: ultrasonographic biometry versus certain menstrual dates. *Am J Obstet Gynecol*. 1996; 174(1 Pt 1):278–81. [PubMed: 8572021]
11. Kramer MS, McLean FH, Boyd ME, Usher RH. The validity of gestational age estimation by menstrual dating in term, preterm, and postterm gestations. *JAMA*. 1988; 260(22):3306–8. [PubMed: 3054193]
12. Engstrom JL. Measurement of fundal height. *Journal of obstetric, gynecologic, and neonatal nursing : JOGNN / NAACOG*. 1988; 17(3):172–8.
13. White LJ, Lee SJ, Stepniewska K, Simpson JA, Dwell SL, Arunjerda R, et al. Estimation of gestational age from fundal height: a solution for resource-poor settings. *J R Soc Interface*. 2012; 9(68):503–10. [PubMed: 21849388]
14. Buck Louis GM, Grewal J, Albert PS, Sciscione A, Wing DA, Grobman WA, et al. Racial/ethnic standards for fetal growth: the NICHD Fetal Growth Studies. *Am J Obstet Gynecol*. 2015; 213(4):449–e41. e1–e41. [PubMed: 26410205]
15. Hadlock FP, Harrist RB, Martinez-Poyer J. In utero analysis of fetal growth: a sonographic weight standard. *Radiology*. 1991; 181(1):129–33. [PubMed: 1887021]
16. Jelks A, Cifuentes R, Ross MG. Clinician bias in fundal height measurement. *Obstet Gynecol*. 2007; 110(4):892–9. [PubMed: 17906025]
17. Jimenez JM, Tyson JE, Reisch JS. Clinical measures of gestational age in normal pregnancies. *Obstet Gynecol*. 1983; 61(4):438–43. [PubMed: 6828273]
18. Griffiths A, Pinto A, Margarit L. A survey of methods used to measure symphysis fundal height. *Journal of obstetrics and gynaecology : the journal of the Institute of Obstetrics and Gynaecology*. 2008; 28(7):692–4. [PubMed: 19065362]
19. Lohman TGRA, Martorell R. *Anthropometric Standardization Reference Manual*. Champaign, IL: Human Kinetics Books; 1988.
20. Ulijaszek SJ, Kerr DA. Anthropometric measurement error and the assessment of nutritional status. *Br J Nutr*. 1999; 82(3):165–77. [PubMed: 10655963]
21. Johnson TS, Engstrom JL, Gelhar DK. Intra- and interexaminer reliability of anthropometric measurements of term infants. *J Pediatr Gastroenterol Nutr*. 1997; 24(5):497–505. [PubMed: 9161941]
22. Ortega-Villa AM, Grantz KL, Albert PS. Estimating onset time from longitudinal and cross-sectional data with an application to estimating gestational age from longitudinal maternal anthropometry during pregnancy and neonatal anthropometry at birth. *Journal of the Royal Statistical Society*. 2017 Under Review.

23. Skupski DW, Owen J, Kim S, Fuchs KM, Albert PS, Grantz KL. Estimating Gestational Age From Ultrasound Fetal Biometrics. *Obstet Gynecol.* 2017; 130(2):433–41. [PubMed: 28697101]
24. Corton MM, Leveno K, Bloom S, Hauth J, Rouse D, Spong C. *Williams Obstetrics*. 23. McGraw-Hill Education; 2009.
25. Rossavik IK, Fishburne JI. Conceptional age, menstrual age, and ultrasound age: a second-trimester comparison of pregnancies of known conception date with pregnancies dated from the last menstrual period. *Obstet Gynecol.* 1989; 73(2):243–9. [PubMed: 2643067]
26. Buchmann E. Routine symphysis-fundal height measurement during pregnancy: RHL commentary. Geneva: World Health Organization; 2003.
27. Papageorgiou AT, Ohuma EO, Gravett MG, Hirst J, da Silveira MF, Lambert A, 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:i5662. [PubMed: 27821614]
28. Belizan JM, Villar J, Nardin JC, Malamud J, De Vicuna LS. Diagnosis of intrauterine growth retardation by a simple clinical method: measurement of uterine height. *Am J Obstet Gynecol.* 1978; 131(6):643–6. [PubMed: 686050]
29. Medhat WM, Fahmy SI, Mortada MM, Sallam HN, Nofal LM, Nosseir SA, et al. Construction of a local standard symphysis fundal height curves for monitoring intrauterine fetal growth. *J Egypt Public Health Assoc.* 1991; 66(3–4):305–31. [PubMed: 1791406]
30. Deeluea J, Sirichotiyakul S, Weerakiet S, Arora R, Patumanond J. Fundal height growth curve for underweight and overweight and obese pregnant women in Thai population. *ISRN Obstet Gynecol.* 2013; 2013:657692. [PubMed: 24455289]
31. Deeluea J, Sirichotiyakul S, Weerakiet S, Buntha R, Tawichasri C, Patumanond J. Fundal height growth curve for thai women. *ISRN Obstet Gynecol.* 2013; 2013:463598. [PubMed: 23691342]
32. Laird NM, Ware JH. Random-effects models for longitudinal data. *Biometrics.* 1982; 38(4):963–74. [PubMed: 7168798]
33. Pinheiro J, Bates D, DebRoy S, Sarkar D. R Core Team (2014) nlme: linear and nonlinear mixed effects models. R package version 3.1-117. 2014. See <http://CRAN.R-project.org/package=nlme>
34. Fieuws S, Verbeke G. Pairwise fitting of mixed models for the joint modeling of multivariate longitudinal profiles. *Biometrics.* 2006; 62(2):424–31. [PubMed: 16918906]

Table 1

Demographic characteristics

	Total (Overall) (n = 2669)	Non-Obese (n = 2224)	Obese (n = 445)
Maternal Age (years), no.(%)			
<20	154 (6.3)	128 (6.3)	26 (6.4)
20 – <29	1205 (49.6)	983 (48.6)	222 (54.4)
30 – <39	1044 (43.0)	891 (44.1)	153 (37.5)
40 – <44	27 (1.1)	20 (1.0)	7 (1.7)
Gestational Age at enrollment, mean (SD)	12.69 (0.96)	12.69 (0.96)	12.65 (0.97)
Prepregnancy BMI (kg/m ²), mean (SD)	25.47 (5.22)	23.64 (3.09)	34.56 (4.02)
Prepregnancy weight (kg), mean (SD)	67.40 (15.00)	62.51 (9.61)	91.83 (13.09)
Maternal height (cm), mean (SD)	162.54 (6.86)	162.48 (6.91)	162.84 (6.58)
Race/Ethnicity, no.(%)			
Non-Hispanic White	721 (27.0)	590 (26.5)	131 (29.4)
Non-Hispanic Black	745 (27.9)	581 (26.1)	164 (36.9)
Hispanic	766 (28.7)	623 (28.0)	143 (32.1)
Asian & Pacific Islander	437 (16.4)	430 (19.3)	7 (1.6)
Parity, no.(%)			
0	1253 (46.9)	1091 (49.1)	162 (36.4)
1	1416 (53.1)	1133 (50.9)	283 (63.6)
Marital status, no.(%)			
Never Married	604 (22.6)	478 (21.5)	126 (28.3)
Married/Living as Married	1984 (74.4)	1685 (75.8)	299 (67.2)
Divorced/Separated/Widowed	79 (3.0)	59 (2.7)	20 (4.5)
Education, no.(%)			
Less than high school	307 (11.5)	238 (10.7)	69 (15.5)
High school diploma or equivalent	488 (18.3)	384 (17.3)	104 (23.4)
Some college or Associate degree	810 (30.3)	651 (29.3)	159 (35.7)
Bachelors degree	625 (23.4)	548 (24.6)	77 (17.3)
Masters or Advanced degree	439 (16.4)	403 (18.1)	36 (8.1)
Family income (\$), no.(%)			
<30,000	671 (29.0)	531 (27.8)	140 (34.5)
30,000–<50,000	432 (18.7)	325 (17.0)	107 (26.4)
50,000–<75,000	295 (12.8)	231 (12.1)	64 (15.8)
75,000– <100,000	295 (12.8)	257 (13.5)	38 (9.4)
100,000	620 (26.8)	563 (29.5)	57 (14.0)
Health insurance, no.(%)			
Private/managed care	1457 (56.5)	1237 (57.6)	220 (51.0)
Medicaid/Self pay	1062 (41.2)	863 (40.2)	199 (46.2)
Unknown	61 (2.4)	49 (2.3)	12 (2.8)

Abbreviations: BMI, body mass index; SD, standard deviation

Table 2

The sensitivity and specificity for detecting gestational age preterm (<37 weeks) in the non-obese group (n=2224) by combinations of longitudinal maternal and neonatal anthropometry^a

Model	Sensitivity (%)	Specificity (%)
Fundal Height	75	97
Fundal Height + Maternal Weight Change	79	97
Fundal Height + Subscapular Skin Fold	77	97
Maternal Weight Change + Subscapular Skin Fold	32	93
Fundal Height + Maternal Weight Change + Subscapular Skin Fold	78	97
Fundal Height +Abdominal Circumference	70	97
Fundal Height + Head Circumference	70	97
Fundal Height + Length	73	97
Fundal Height + Birthweight	71	97
Fundal Height All neonatal anthropometry measures	69	97
Fundal Height +Maternal Weight Change +Head Circumference	73	97

^aTraining set included non-obese women from all four racial/ethnic groups with 2 fundal height measures beginning at any time in gestation

Table 3

The prediction error for gestational age using fundal height alone by the timing and frequency of measurements in the non-obese group (n=2224)^a

Fundal Height Model Measurement Constraints^b	Prediction Error (weeks)	Prediction Error (days)
Frequency of Measurements ^c		
Only 2 measurements	2.69	18.83
Only 3 measurements	2.51	17.57
Only 4 measurements	2.36	16.52
Timing of Measurements		
>20 weeks	2.02	14.14
>20 cm	2.20	15.40
>28 weeks	2.30	16.10

^a Training set included non-obese women from all four racial/ethnic groups with 2 fundal height measures beginning at any time in gestation

^b Limited in the test set only

^c Randomly selected