

Estimating Gestational Age From Ultrasound Fetal Biometrics

Daniel W. Skupski, MD, John Owen, MD, MSPH, Sungduk Kim, PhD, Karin M. Fuchs, MD, Paul S. Albert, PhD, and Katherine L. Grantz, MD, MS, for the Eunice Kennedy Shriver National Institute of Child Health and Human Development Fetal Growth Studies

OBJECTIVE: To compare the accuracy of a new formula with one developed in 1984 (and still in common use) and to develop and compare racial and ethnic-specific and racial and ethnic-neutral formulas.

METHODS: The Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) Fetal Growth Studies-Singletons was a prospective cohort study that recruited women in four self-reported racial-ethnic groups—non-Hispanic black, Hispanic, non-Hispanic white, and Asian—with singleton gestations from 12 U.S. centers (2009–2013). Women with

a certain last menstrual period confirmed by first-trimester ultrasonogram had longitudinal fetal measurements by credentialed study ultrasonographers blinded to the gestational age at their five follow-up visits. Regression analyses were performed with linear mixed models to develop gestational age estimating formulas. Repeated cross-validation was used for validation. The estimation error was defined as the mean squared difference between the estimated and observed gestational age and was used to compare the formulas' accuracy.

RESULTS: The new formula estimated the gestational age (± 2 SD) within ± 7 days from 14 to 20 weeks of gestation, ± 10 days from 21 to 27 weeks of gestation, and ± 17 days from 28 to 40 weeks of gestation. The new formula performed significantly better than a formula developed in 1984 with an estimation error of 10.4 compared with 11.2 days from 21 to 27 weeks of gestation and 17.0 compared with 19.8 days at 28–40 weeks of gestation, respectively. Racial and ethnic-specific formulas did not outperform the racial and ethnic-neutral formula.

CONCLUSION: The NICHD gestational age estimation formula is associated with smaller errors than a well-established historical formula. Racial and ethnic-specific formulas are not superior to a racial-ethnic-neutral one.

(*Obstet Gynecol* 2017;130:433–41)

DOI: 10.1097/AOG.0000000000002137

From the Division of Maternal Fetal Medicine, Department of Obstetrics and Gynecology, New York Presbyterian Hospital, Queens, New York; the Division of Maternal-Fetal Medicine, Department of Obstetrics and Gynecology, University of Alabama at Birmingham, School of Medicine, Birmingham, Alabama; the Division of Intramural Population Health Research, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland; and the Department of Obstetrics and Gynecology, Columbia University Medical Center, New York, New York.

Supported by funding through the Eunice Kennedy Shriver National Institute of Child Health and Human Development intramural program (contracts HHSN275200800013C, HHSN275200800002I, HHSN27500006, HHSN275200800003IC, HHSN275200800014C, HHSN275200800012C, HHSN275200800028C, and HHSN275201000009C) and included funding through the American Recovery and Reinvestment Act of 2009.

Presented at the Annual Meeting of the Society for Maternal-Fetal Medicine, February 4–7, 2016, Atlanta, Georgia.

The authors thank GE Healthcare Women's Health Ultrasound for their support and training on the Voluson product over the course of this study. The authors also thank the principal investigators at the clinical sites for their support and the site sonographers for their good and careful work. For a list of principal investigators who participated in this study, see Appendix 1 online at <http://links.lww.com/AOG/A972>.

Each author has indicated that he or she has met the journal's requirements for authorship.

Corresponding author: Daniel W. Skupski, MD, Obstetrics and Gynecology, New York Presbyterian Queens, 56-45 Main Street, Room M-365, Flushing, NY 11355; email: dwskupsk@med.cornell.edu.

Financial Disclosure

The authors did not report any potential conflicts of interest.

© 2017 by The American College of Obstetricians and Gynecologists. Published by Wolters Kluwer Health, Inc. All rights reserved.

ISSN: 0029-7844/17

Accurate gestational dating is imperative for optimal maternal and neonatal outcomes. The addition of ultrasound fetal biometric measurements to a woman's reported last menstrual period (LMP) allows for improvement in gestational dating.¹ Methods for estimating gestational age from ultrasound fetal biometric measurements were reported decades ago, mostly in racially homogenous populations.² Indeed, formulas from early reports are still used today to estimate gestational age.³ Formula accuracy



was initially evaluated and then further refined using pregnancies conceived by in vitro fertilization as the referent standard.⁴ Ultrasound imaging has continued to advance since that time^{5,6} and may allow for improved gestational age estimation through better image quality and more precise caliper placement. The recent performance and completion of the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD) Fetal Growth Studies provides a unique opportunity to readdress the issue of gestational age estimation in a larger and diverse cohort, one that included a cohort of women of four different racial-ethnic groups, conditions that were optimal for fetal growth, a standardized ultrasound protocol, and extensive training and quality control of images and measurements.^{7,8} In addition, maternal health characteristics have changed in the last 30 years or more such as age and obesity rates and a reevaluation seems prudent. Because there is no consensus in the obstetric community on whether racial and ethnic-specific formulas are superior to racial and ethnic-neutral formulas, the NICHD Fetal Growth Studies provide this opportunity as well.

Our objectives were to develop a gestational age estimation model using ultrasound fetal parameters in women who participated in the NICHD Fetal Growth Studies, to compare this formula with one in common use³ and to assess the accuracy of racial and ethnic-specific formulas.

MATERIALS AND METHODS

The NICHD Fetal Growth Studies-Singletons was a prospective cohort study that recruited women of four self-reported racial or ethnic groups (non-Hispanic black, Hispanic, non-Hispanic white, and Asian), who had low-risk pregnancies with optimal conditions for fetal growth. To be included, women had to have a certain LMP and a first-trimester ultrasonogram that confirmed their pregnancy dating. The ultrasound estimate of gestation had to be between 8 0/7 weeks and 13 6/7 weeks and match the LMP-based gestational age within 5 days for women between 8 0/7 weeks and 10 6/7 weeks, within 6 days for those between 11 0/7 weeks and 12 6/7 weeks, and within 7 days for participants between 13 0/7 weeks and 13 6/7 weeks.¹ If these criteria were not met, the patient was excluded. Cycle length and other hormonal contraception were not assessed and were not exclusions to participation. The project gestational age was then based on the menstrual date. Women were screened at 8 0/7 weeks of gestation to 13 6/7 weeks of gestation for maternal health status associated with normal fetal growth (aged 18–40

years; body mass index [BMI, calculated as weight (kg)/[height (m)]² 19.0–29.9; healthy lifestyles and living conditions [see the exclusion criteria subsequently]; low-risk medical and obstetric history). Body mass index was calculated from self-reported prepregnancy height and weight and was confirmed by measurement at the enrollment visit. Details of the study methodology have been published elsewhere.⁷ A cohort of women with BMI 30.0–45.0 mg/kg² was recruited also, but is not detailed in the original report.⁷ A separate report comparing the obese cohort with the nonobese cohort is in development, but this comparison is not shown in our current study on gestational age assessment.

The exclusion criteria in the NICHD Fetal Growth Studies-Singletons were extensive: current pregnancy with multifetal gestation, history of preterm low birth weight (less than 2,500 g), or macrosomic (greater than 4,000 g) neonate; history of stillbirth or neonatal death; medically assisted conception; cigarette smoking or illicit drug use in the past 6 or 12 months, respectively; one or more daily alcoholic drinks; previous fetal congenital malformation; history of noncommunicable diseases (asthma requiring weekly medication, autoimmune disorders, cancer, diabetes mellitus, epilepsy or seizures requiring medication, hematologic disorders, hypertension, psychiatric disorders, renal disease, thyroid disease); or history of gravid diseases (gestational diabetes, severe preeclampsia or eclampsia, or hemolysis, elevated liver enzymes, low platelet count syndrome). Women who participated but who had complications in the current pregnancy or adverse outcomes for the neonate were excluded (post hoc) in the original report.⁷ Sample size was determined so that the 5th and 95th percentiles for each gestational week for each racial or ethnic group could be precisely estimated over gestation and assuming a 10% attrition and 30% post hoc exclusion rate for women with pregnancy or neonatal complications. We estimated needing 600 women in each group to reach our estimate of 320 women (after post hoc exclusions); our high completion and lower than expected exclusion rate resulted in larger cohorts. Women were recruited from 12 participating U.S. clinical sites from July 2009 through January 2013.

After an ultrasonogram at 8 0/7 weeks to 13 6/7 weeks of gestation, women underwent serial ultrasound assessment of fetal biometry, five targeted visits approximately once every 4 to 8 weeks, in one of four ultrasound schedules, assigned randomly, so that each gestational week from 14 to 40 was represented without exposing individual women to weekly



ultrasonography. All ultrasound examinations were performed using Voluson E8 machines. All study ultrasonographers underwent ante hoc training and credentialing, and their measurement techniques were subject to rigorous quality assurance.⁹ Ultrasonographers were blinded to the project gestational age, to the results of all prior scans, and to the gestational age of each measurement calculated in real time on the ultrasound machine to avoid measurement bias. Institutional review board approval was obtained for the NICHD and all participating clinical institutions, and the data and imaging coordinating centers. Women were enrolled from four racial-ethnic groups, 611 non-Hispanic black, 649 Hispanic, 614 non-Hispanic white, and 460 Asian or Pacific Islander women. Protocol completion rates were 92%, 93%, 93%, and 90%, respectively.

Using either a transabdominal or transvaginal approach, biparietal diameter was measured at the level of the thalami and cavum septa pellucida or the cerebral peduncles as the linear distance from the outer edge of the proximal to the inner edge of the distal skull; head circumference was measured at the same level (and often on the same images) using the ellipse function around the outer perimeter of the skull. Abdominal circumference was measured using the ellipse function circumscribing the actual or projected skin line in the transverse plane at the level of the stomach and the junction of the umbilical vein and portal sinus. Femur length was measured as the linear distance between the midpoints of each end of the calcified femoral diaphysis.

The nonobese cohort was combined with the obese cohort for the current study of gestational age assessment. A total of 468 obese women were included, 192 non-Hispanic black, 132 Hispanic, 137 non-Hispanic white, and 7 Asian or Pacific Islander. The obese cohort had the same eligibility criteria with the addition of the following exclusion criteria, which are different and not as extensive as the exclusion criteria of the singleton cohort: autoimmune diseases, cancer, chronic hypertension requiring two or more medications, chronic renal disease, diabetes while not pregnant, human immunodeficiency virus, or psychiatric disorders. These additional exclusion criteria were added because major chronic conditions such as hypertension and diabetes disproportionately affect obese women, making it hard to determine whether the elevated risks of large for gestational age or macrosomia (effects on fetal growth) are the result of those complications or the excess maternal fatness. Our study's obese cohort of women was free of these major pre-existing chronic conditions, which helps to

differentiate morbidity from obesity-related fetal effects. This will be detailed in a separate report. All participating women from the nonobese and obese cohorts were included in this analysis of gestational age estimation; there were no post hoc exclusions, which is different than the original fetal growth study report.⁷

Gestational age estimation models were developed using a backward elimination regression technique that initially contained all biometric measurements (biparietal diameter, head circumference, abdominal circumference, and femur length), including first-order, quadratic, and interaction terms, then removing the least significant terms until only terms that were significant at the .05 level remained. We used linear mixed models to account for the inpatient correlation associated with repeated measurements. The final chosen model (hereafter referred to as the NICHD model) was fit using all patients in the database and all gestational ages from 14 to 40 weeks. This model was validated using the technique of repeated cross-validation, a paradigm that includes development of the model on a random selection of 50% of the data set and validation of the model on the remaining 50%. For each formula, a single scan of each patient was used and chosen randomly on half of the population and tested on the other half of the population. We performed this 50% test–50% training procedure 1,000 times. Using the coefficients of this validated model, we derived a formula for estimating gestational age based on the statistically significant terms of biometric parameters (hereafter referred to as the NICHD formula).

The estimation error was defined as the mean squared difference between estimated and observed (project) gestational age among test set samples and was used to assess formula performance in several ways. First we evaluated the NICHD formula's SD of estimation in the second compared with third trimesters. Second, we compared the NICHD formula's SD of estimation with that of the 1984 Hadlock formula, which is in common use today.³ Finally, using the same statistical techniques described previously, we developed racial and ethnic-specific formulas for each of the four racial-ethnic groups and compared the performance of the NICHD formula's SD of estimation with racial and ethnic-specific formulas. The SDs were multiplied by 1.96 to obtain the estimation error with the interpretation that in 95% of patients, the difference between the formula's estimated gestational age and the project gestational age is within plus or minus this value (ie, gestational weeks or days).



Prior literature has focused on three smaller gestational age windows as separate entities because of the different variation in the accuracy of estimation among these windows.^{1,3,4,10–12} Therefore, we also analyzed the difference in SD of estimation error among gestational age windows (14–20, 21–27, and 28–40 weeks).

The clinical effect of more accurate gestational dating was assessed by comparing the number of women who would be outside of a prescribed range of error for gestational age prediction using either formula for each of the three gestational age windows studied (14–20, 21–27, 28–40 weeks). The designated “acceptable” range of error is based on prior literature¹ and was as follows: 14–20 weeks±7 days, 21–27 weeks±10 days, and 28–40 weeks±14 days. We used a generalized estimating equation to account for multiple measurements per individual and compare rates of estimated gestational ages outside of prescribed ranges for both the Hadlock and NICHD formulas. This allowed an assessment of any statistically significant differences between the two formulas. All analyses were implemented using SAS 9.4 or R 3.1.2.

RESULTS

The study population included 803 non-Hispanic black (28.6%), 781 Hispanic (27.9%), 751 non-Hispanic white (26.8%), and 467 Asian (16.7%), totaling 2,802 women. Table 1 includes demographic data of the population. Table 2 presents the selected fetal anthropometric parameters and their associated estimated regression coefficients obtained from model selection. The r^2 for the new NICHD formula was 0.975. The concordance among study site ultrasonographers for all measurements was in excess of $r=0.99$, demonstrating no significant difference among study sites. The equation is: gestational age (weeks) = $10.6 - 0.168 \times \text{BPD} + 0.045 \times \text{HC} + 0.03 \times \text{AC} + 0.058 \times \text{FL} + 0.002 \times \text{BPD}^2 + 0.002 \times \text{FL}^2 + 0.0005 \times (\text{BPD} \times \text{AC}) - 0.005 \times (\text{BPD} \times \text{FL}) - 0.0002 \times (\text{HC} \times \text{AC}) + 0.0008 \times (\text{HC} \times \text{FL}) + 0.0005 \times (\text{AC} \times \text{FL})$, in which BPD=biparietal diameter, HC=head circumference, AC=abdominal circumference, and FL=femur length.

Figure 1 demonstrates the observed gestational age based on the LMP (confirmed by first-trimester ultrasonogram) plotted against the gestational age estimated by the NICHD formula applied to our population for gestational weeks 14–40. The figure shows that the accuracy of estimation of the formula is higher early in gestation and decreases with advancing gestation.

Table 1. Characteristics of the Study Population (N=2,802)

Characteristic	Value
Age (y)	28.16±5.49
Self-reported height (cm)	162.60±7.08
Self-reported weight (kg)	67.34±14.98
BMI (kg/m ²)	25.41±5.13
Gestational age at ultrasound visit (wk)	
Visit 0 (enrollment)	12.7±0.95
Visit 1	19.8±2.44
Visit 2	27.1±2.08
Visit 3	32.0±1.55
Visit 4	35.9±1.36
Visit 5	38.6±0.83
Parity	
0	1,319 (47.1)
1	943 (33.7)
2 or more	540 (19.3)
Race–ethnicity	
Non-Hispanic white	751 (26.8)
Non-Hispanic black	781 (27.9)
Hispanic	803 (28.7)
Asian	467 (16.7)
Marital status	
Never married	632 (22.6)
Living as married or married	2,082 (74.4)
Divorced, separated, or widowed	84 (3.0)
Education	
No college	841 (30.0)
Some college	850 (30.3)
College graduate	645 (23.0)
Postgraduate degree	466 (16.6)
Annual family income	
Less than \$30,000	707 (29.3)
\$30,000–49,999	452 (18.7)
\$50,000–74,999	311 (12.9)
\$75,000–99,999	305 (12.6)
\$100,000 or greater	640 (26.5)
Health insurance	
Private, managed care	1,459 (56.5)
Medicaid; other	1,064 (41.2)
Self-pay	61 (2.4)
Neonatal sex	
Male	1,309 (51.2)
Female	1,249 (48.8)
Gestational age at delivery (wk)	39.17±1.91
Total no. of images	
All images together, counts	54,012
BPD, counts	13,461
HC, counts	13,470
AC, counts	13,540
FL, counts	13,541
No. of images per patient	
All images together	19.28±5.40
BPD	4.80±1.37
HC	4.81±1.37
AC	4.83±1.35
FL	4.83±1.36

BMI, body mass index; BPD, biparietal diameter; HC, head circumference; AC, abdominal circumference; FL, femur length. Data are mean±SD, n (%), or n.



Table 2. Regression Model for Estimation of Gestational Age

Variable	Estimate	SE	R ²	R ² *
Intercept	10.5992	0.1815	0.9754	0.956731
BPD	-0.1683	0.0194		
HC	0.0452	0.0058		
AC	0.0302	0.0037		
FL	0.0576	0.0168		
BPD ²	0.0025	0.0004		
FL ²	0.0017	0.0007		
BPD×AC	0.0005	0.0002		
BPD×FL	-0.0052	0.0012		
HC×AC	-0.0003	0.0001		
HC×FL	0.0008	0.0003		
AC×FL	0.0006	0.0001		

SE, standard error; BPD, biparietal diameter; HC, head circumference; AC, abdominal circumference; FL, femur length; GA, gestational age.

The equation is: $GA \text{ (weeks)} = 10.6 - 0.168 \times BPD + 0.045 \times HC + 0.03 \times AC + 0.058 \times FL + 0.002 \times BPD^2 + 0.002 \times FL^2 + 0.0005 \times (BPD \times AC) - 0.005 \times (BPD \times FL) - 0.0002 \times (HC \times AC) + 0.0008 \times (HC \times FL) + 0.0005 \times (AC \times FL)$.

The regression equation was obtained using linear mixed modeling to account for the use of multiple anthropometric measurements per fetus.

R² is based on randomly selected one sample for each participant.

* Calculated from whole dataset from fitted model based on one sample for each participant.

Figure 2 shows the gestational ages estimated by the NICHD formula compared with the Hadlock formula. The figure suggests that early in gestation, the two approaches give similar estimates and demonstrate increased differences in estimates with advancing gestation.

Table 3 presents the error of estimation results for the NICHD formula for different gestational age windows. The estimation of gestational age in weeks is more precise over the interval of 14–20 weeks compared with a later interval of 21–27 weeks. Furthermore, the accuracy is lower from 28 to 40 weeks of gestation as compared with the other two windows. This indicates more accurate estimation earlier in gestation. The NICHD formula's performance was similar to Hadlock at 14–20 weeks of gestation, whereas it was significantly more accurate after 20 weeks of gestation. The accuracy of the NICHD formula is ± 7 days from 14 to 20 weeks of gestation, ± 10 days from 21 to 27 weeks of gestation, and ± 17 days from 28 to 40 weeks of gestation. The new formula performed better than a formula developed in 1984 with an estimation error of 10.4 compared with 11.2 days from 21 to 27 weeks of gestation and 17.0 compared with 19.8 days at 28–40 weeks of gestation, respectively. Table 3 also shows the comparison of racial and ethnic-specific formulas and the

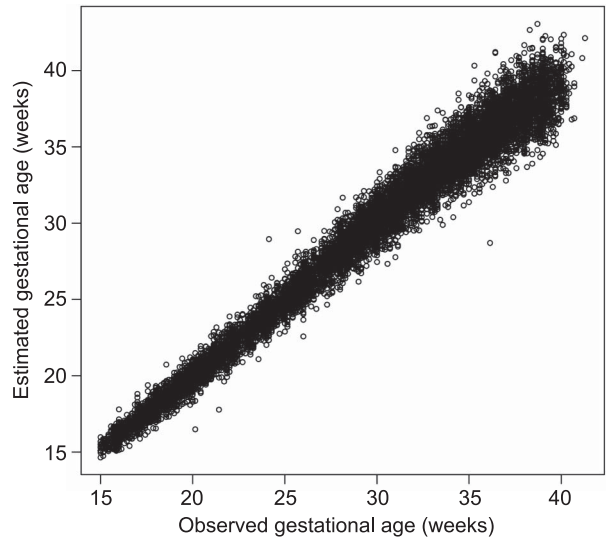


Fig. 1. Estimated gestational age compared with observed gestational age from 15 to 40 weeks. Use of the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD) model produced the estimated gestational age. Observed gestational age is that obtained from the last menstrual period.

Skupski. *NICHD Gestational Age Estimation. Obstet Gynecol* 2017.

NICHD formula applied to each racial-ethnic group. Results were similar overall (gestational weeks 14–40) and for each smaller gestational age window (14–20 weeks, 21–27 weeks, 28–40 weeks).

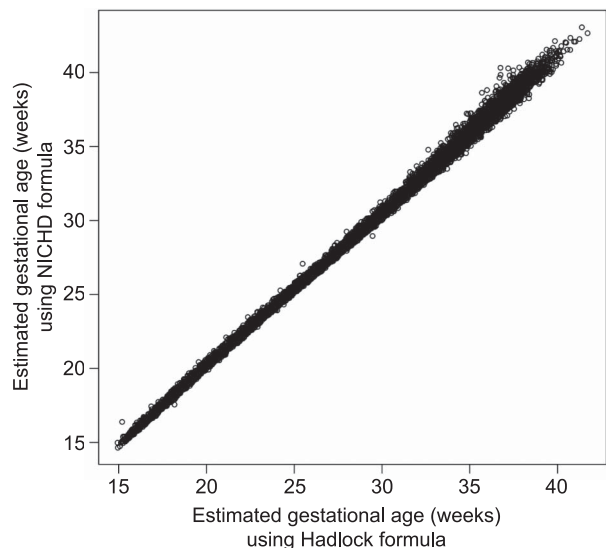


Fig. 2. Estimated gestational age. *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD) model compared with the Hadlock model.

Skupski. *NICHD Gestational Age Estimation. Obstet Gynecol* 2017.



Table 3. Error of Estimation in Days for Different Gestational Age Windows, Comparing the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development Model, the Hadlock Model⁴ Normal and Obese Cohorts, and Four Racial- and Ethnic-Specific Models

GA (wk)	NICHD Model (Entire Cohort)	Hadlock Model (Entire Cohort)	NICHD Model		Hadlock Model		Non-Hispanic Black Model	Hispanic Model	Non-Hispanic White Model	Asian Model
			Normal	Obese	Normal	Obese				
14–40	14.66	16.82	14.54	15.22	14.54	15.22	14.51	15.06	14.32	13.80
14–20	7.15	7.08	6.94	8.06	6.94	8.06	6.81	7.28	7.05	7.55
21–27	10.36	11.22	10.29	10.74	10.29	10.74	10.11	11.07	9.79	9.75
28–40	17.01	19.75	16.88	17.61	16.88	17.61	16.96	17.34	16.68	15.92

GA, gestational age; NICHD, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development. The numbers were generated by multiplying the SD times 1.96. In 95% of cases, the error of the predictor is within plus or minus this value.

Figure 3 and Table 4 present the error of estimation for each gestational week when applying the NICHD and Hadlock formulas to our population. This can be used to establish gestational age when LMP differs from the ultrasound biometry outside of the estimated error range. The results are better with the NICHD compared with the Hadlock formula after 35 weeks of gestation. Before 35 weeks of gestation, the Hadlock and NICHD formulas are very similar.

Table 5 presents an example of the effect of more accurate gestational dating using the NICHD formula compared with the Hadlock formula. There is a significantly smaller proportion of patients outside the prescribed range of acceptable error using the

NICHD formula at 21–27 weeks and 28–40 weeks of gestation.

DISCUSSION

We have developed a model for gestational age estimation and compared this new formula with a commonly used formula (Hadlock 1984).³ We found the new formula to have improved estimation in the second and third trimesters and used a robust method for evaluating racial and ethnic-specific formulas and found that they are not superior to the new formula.

Similar to prior literature, gestational age estimation is better with the NICHD formula from 14 to 20 weeks of gestation compared with later windows, which indicates less biological variation earlier in gestation.^{2–4} The accuracy of the NICHD formula is ± 7 days from 14 to 20 weeks of gestation, ± 10 days from 21 to 27 weeks of gestation, and ± 17 days from 28 to 40 weeks of gestation, although we also provide individual week errors, which can be used to establish gestational age when LMP differs from the ultrasound biometry outside of the estimated error range (Table 4). The large population also allows us to support a recommendation to use ultrasonography to establish gestational age when fetal biometry at 14–15 weeks of gestation is 8 days or more different than the LMP. This type of recommendation has previously been lacking in the literature.¹³

We believe several factors are responsible for the observed improvements. The rigorous credentialing process for ultrasonographers, including masking gestational age at each examination, provided an unbiased method for obtaining each measurement and allowed minimization of random and systematic error.^{4,14} Also, the ongoing quality improvement program provided sustained use of a consistent measurement method.^{9,15} Modern ultrasound equipment with better image quality also possibly improved the

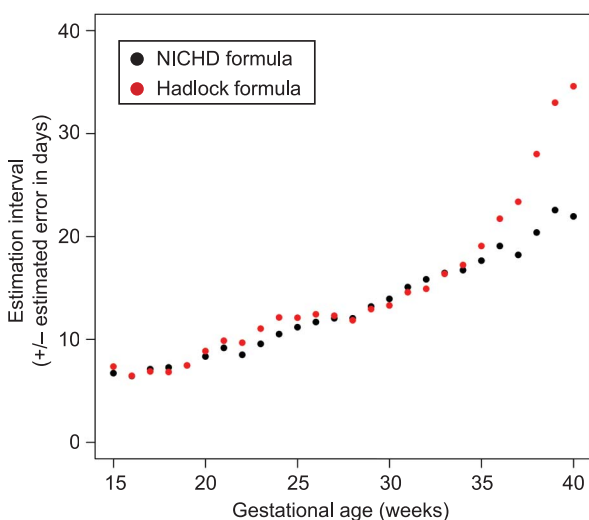


Fig. 3. Plot comparing the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD) model with the Hadlock model for estimation interval (\pm estimated error in days) for each week of gestation.

Skupski. NICHD Gestational Age Estimation. *Obstet Gynecol* 2017.



Table 4. Error of Estimation for Gestational Age by Each Gestational Week, Comparing the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development and Hadlock Models

Gestational Age (wk)	NICHD		Hadlock	
	Error of Estimation (wk)*	Error of Estimation (d) [†]	Error of Estimation (wk)*	Error of Estimation (d) [†]
15 to less than 16	0.96	7	1.05	7
16 to less than 17	0.92	6	0.92	6
17 to less than 18	1.02	7	0.99	7
18 to less than 19	1.04	7	0.98	7
19 to less than 20	1.07	7	1.07	7
20 to less than 21	1.19	8	1.27	9
21 to less than 22	1.31	9	1.41	10
22 to less than 23	1.22	9	1.38	10
23 to less than 24	1.37	10	1.58	11
24 to less than 25	1.50	11	1.73	12
25 to less than 26	1.60	11	1.73	12
26 to less than 27	1.67	12	1.78	12
27 to less than 28	1.72	12	1.76	12
28 to less than 29	1.72	12	1.70	12
29 to less than 30	1.89	13	1.85	13
30 to less than 31	2.00	14	1.90	13
31 to less than 32	2.16	15	2.09	15
32 to less than 33	2.26	16	2.13	15
33 to less than 34	2.35	16	2.34	16
34 to less than 35	2.39	17	2.46	17
35 to less than 36	2.52	18	2.73	19
36 to less than 37	2.73	19	3.10	22
37 to less than 38	2.60	18	3.34	23
38 to less than 39	2.92	20	4.00	28
39 to less than 40	3.23	23	4.72	33
40 to less than 41	3.14	22	4.94	35

NICHD, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development.

* Estimated SD×1.96.

† Estimated SD×1.96×7.

accuracy. Improved image quality in all clinical situations (eg, previous abdominal scar, increased BMI) has been a focus of ultrasound vendors for many years.^{16–18}

The results of Table 5 demonstrate the clinical importance of more accurate estimation. Because the SDs of estimated errors between the two groups cannot be compared using inferential statistics, and thus *P* values cannot be displayed, we created Table 5 to demonstrate that the clinical effect of improved estimation with the NICHD formula is significant.¹⁹ These findings are particularly important for women who present late to prenatal care, when ultrasonography has a wider range of error. Currently, 5–15% of women fall outside the prescribed range of acceptable error. Using the NICHD formula, 2–5% of women would have more accurate dating, thus preventing interventions for preterm labor or postterm pregnancy (Table 5). Up to 25% of the entire population in the United States undergoes induction of labor,²⁰ providing a significant potential to prevent morbidity

from unnecessary interventions. The error range for the third trimester (+17 days) argues that caution in gestational dating is necessary late in gestation.

We believe the finding that the racial and ethnic-specific formulas were similar to a racial and ethnic-neutral formula is the result of the adequately represented racial–ethnic diversity in the population from which our formula was developed.^{21–23} In addition, this fact simplifies the workflow in ultrasound units because a single formula for gestational age estimation is accurate for all races and ethnicities.

The NICHD Fetal Growth Studies showed previously that there is a significant difference in many fetal measurements when comparing different racial–ethnic groups.⁷ The results of the current study found no difference among different racial–ethnic groups. This discrepancy is explained by the fact that growth trajectories, as assessed in the parent trial, and the relationship between cross-sectional measures and gestational age, as assessed in the current trial, reflect different biological processes.



Table 5. Comparison of the Error in Gestational Age Estimation of the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development and Hadlock Formulas

Formula	In Range*	Out of Range*	P
14–20 wk			
Hadlock	1,352 (94.5)	79 (5.5)	.77
NICHD	1,350 (94.3)	81 (5.7)	
21–27 wk			
Hadlock	2,312 (92.2)	196 (7.8)	<.001
NICHD	2,376 (94.7)	132 (5.3)	
28–40 wk			
Hadlock	5,712 (84.4)	1,058 (15.6)	<.001
NICHD	6,076 (89.7)	694 (10.3)	

NICHD, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development.

Data are n (%) unless otherwise specified.

* In range, 14–20 weeks \pm 7 days, 21–27 weeks \pm 10 days, 28–40 weeks \pm 14 days. Out of range, when measurements are outside these ranges.

Our study has strengths. We included a large number of women in multiple racial–ethnic groups in many centers. We controlled for bias inherent in using repeated measurements with the use of linear mixed regression models. We used statistical methodology to control for the problem of developing the predictive model in the same population in which we then tested the model—cross-validation with 1,000 iterations of a 50% test–50% training algorithm. The richness of our data allows us to suggest that our results will be more generally applicable to the U.S. obstetric population than any other formula.

Our study has limitations. Our study population did not have a known date of conception such as with in vitro fertilization that allows us to pinpoint the “true” gestational age. Even with a known date of conception, as is seen with in vitro fertilization pregnancies, ultrasonography is associated with an unavoidable range of error, which is similar to the range of error seen in our study.⁴ In addition, both the NICHD and Hadlock formulas were subject to the same estimation errors because their estimates were both compared with the same project gestational age. So, compared with a commonly used formula, the NICHD formula is associated with significantly smaller error ranges when gestational age is estimated beyond 20 weeks. We cannot comment on gestational age assessment in pregnancies conceived by assisted reproductive technologies because they were not included in our cohort.

Using the NICHD formula has the potential to decrease unnecessary obstetric interventions with

associated decreases in morbidity and cost savings. Given the large number of annual births in the United States and the significant rate of interventions, these improvements could have important implications at a population level.

REFERENCES

1. Methods for estimating the due date. Committee Opinion No. 700. American College of Obstetricians and Gynecologists. *Obstet Gynecol* 2017;129:e150–4.
2. Harrison RF, Roberts AP, Campbell S. A critical evaluation of tests used to assess gestational age. *Br J Obstet Gynaecol* 1977; 84:98–107.
3. Hadlock FP, Deter RL, Harrist RB, Park SK. Estimating fetal age: computer-assisted analysis of multiple fetal growth parameters. *Radiology* 1984;152:497–501.
4. Chervenak FA, Skupski DW, Romero R, Myers MK, Smith-Levitin M, Rosenwaks Z, et al. How accurate is fetal biometry in the assessment of fetal age? *Am J Obstet Gynecol* 1998;178:678–87.
5. Morales-Roselló J, Hervás-Marin D, Stirrup O, Perales-Marín A, Khalil A. International standards for fetal growth: relevance of advances in ultrasound technology. *Ultrasound Obstet Gynecol* 2015;46:631–2.
6. Papageorgiou AT, Ohuma EO, Altman DG, Todros T, Cheikh Ismail L, Lambert A, et al. International standards for fetal growth based on serial ultrasound measurements: the Fetal Growth Longitudinal Study of the INTERGROWTH-21st Project. *Lancet* 2014;384:869–79.
7. Buck-Louis G, 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; 213:449.e1–41.
8. Grantz KL, Grewal J, Albert PS, Wapner R, D’Alton ME, Sciscione A, et al. Dichorionic twin trajectories: the NICHD Fetal Growth Studies. *Am J Obstet Gynecol* 2016;215:221.e1–16.
9. Hediger ML, Fuchs KM, Grantz KL, Grewal J, Kim S, Gore-Langton RE, et al. Ultrasound quality assurance for singletons in the National Institute for Child Health and Human Development Fetal Growth Studies. *J Ultrasound Med* 2016;35:1725–33.
10. Sabbagha RE, Tamura RK, Socol ML. The use of ultrasound in obstetrics. *Clin Obstet Gynecol* 1982;25:735–52.
11. Romero R, Quintero R, Brekus C. Assessment of gestational age. In: Divon M, editor. *Abnormal fetal growth*. New York (NY): Elsevier Science; 1991.
12. Kalish RB, Thaler HT, Chasen ST, Gupta M, Berman SJ, Rosenwaks Z, et al. First- and second-trimester ultrasound assessment of gestational age. *Am J Obstet Gynecol* 2004;191:975–8.
13. Spong CY. Redefining ‘term’ pregnancy: recommendations from the defining ‘term’ pregnancy workgroup. *JAMA* 2013; 309:2445–6.
14. Perni SC, Chervenak FA, Kalish RB, Magherini-Rothe S, Predanic M, Streltsoff J, et al. Intraobserver and interobserver reproducibility of fetal biometry. *Ultrasound Obstet Gynecol* 2004;24:654–8.
15. Boamah EA, Asante K, Ae-Ngibise K, Kinney PL, Jack DW, Manu G, et al. Gestational age assessment in the Ghana Randomized Air Pollution and Health Study (GRAPHS): ultrasound capacity building, fetal biometry protocol development, and ongoing quality control. *JMIR Res Protoc* 2014;3:e77.



