Trouble With the Curve

Pearls and Pitfalls in the Evaluation of Fetal Growth

Theodore J. Dubinsky, MD, Rachelle Sonneborn, MS

Objectives—Poor fetal growth is one of the most important findings on an obstetric ultrasound (US) examination. First-trimester US is the most accurate means for dating pregnancies; however, dating based on last menstrual period remains the standard for determining gestational age. Discrepancies between menstrual and US dating can lead to the overdiagnosis of in utero growth restriction (IUGR). This article will demonstrate that as long as the fetal growth falls along a curve that parallels normal growth curves, appropriate growth has occurred regardless of the gestational age and weight percentile that has been assigned to the fetus.

Methods—Following Institutional Review Board approval (HSD-00002969), 860 third-trimester fetal US examinations were retrospectively evaluated from January 1 through July 1, 2017, to determine whether they had normal growth curves. Outcome data on all cases were obtained from review of the medical records.

Results—Of 216 fetuses (25%) suspected of IUGR based on a weight below the 10th percentile, 6 developed true IUGR: 5 that led to emergent delivery of fetuses weighing less than a 1000 g and 1 in utero fetal demise. The remaining 210 fetuses all had normal outcomes.

Conclusions—As long as the fetal growth falls along a curve that parallels normal growth curves, appropriate growth has occurred regardless of the gestational age and weight percentile that has been assigned to the fetus.

Key Words—abnormal growth; fetal health; growth curves; in utero growth restriction

The determination that a fetus is not growing properly is one of the most important pieces of information that can be derived from an obstetric ultrasound (US) examination. All of the major societies involved in the performance of obstetric US examinations have developed guidelines for determining the gestational age and interval growth of fetuses.[1–3] All of the guidelines indicate that first-trimester US is the most accurate means for dating pregnancies[4–6]; however, clinical dating based on the last date of the patient's last menstrual period (LMP) remains the standard for determining gestational age. In almost every case, there is some variance between gestational age based on US and gestational age based on menstrual history.[7] As has been seen elsewhere, US dating tends to show earlier gestational ages than does menstrual dating, since menstrual dating is not indicative of when fertilization actually occurs. This discrepancy can lead to the overdiagnosis of in utero growth restriction (IUGR), and this can become clinically relevant in determining delivery timing.[8]

Guidelines have been developed outlining the variance in pregnancy dating that is a function of gestational age.[9,10] Specific recommendations for the standard deviations (in days) are presented

Received February 13, 2020, from the Departments of Radiology and Obstetrics and Gynecology, University of Washington, Seattle, Washington, USA (T.J.D.); and University of Washington, Seattle, Washington, USA (R.S.). Manuscript accepted for publication March 15, 2020.

Address correspondence to Theodore J. Dubinsky, MD, Department of Radiology, University of Washington, Box 357115, 1959 Pacific St, Seattle, WA 98195 USA.

E-mail: tdub@u.washington.edu

Abbreviations

IUGR, in utero growth restriction; LMP, last menstrual period; US, ultrasound

doi:10.1002/jum.15293

in Table 1 of these guidelines.[11] For example, up to 14 weeks' gestational age, the variance is generally accepted as ±5 days for 2 SDs, with larger variances in the second and third trimesters.[12] A fetus that is scanned for the first time in the third trimester at 30 weeks would have a variance of ±3 weeks. This reflects the greater accuracy of measuring gestational age in the first trimester based on crown-rump length[13] compared to an estimation in the third trimester averaging biparietal diameter, head circumference, abdominal circumference, and femur length.[14,15] Society guidelines also indicate that if the discrepancy between the menstrual history and US dating is greater than the 2-SD variance, then the gestational age of the fetus should be based on US dating for management purposes.[1]

However, one has to be prepared to accept that with a variance of less than 2 SDs, fetuses will frequently be placed below the 10th percentile (which is within 2 SDs: 2.3rd–97.7th percentiles) based on the size-date discrepancy alone rather than as a reflection of any real pathologic condition, and this leads to false-positive IUGR diagnoses. If being below the 10th percentile is accepted as the definition of IUGR, then a large number of fetuses that fall within the normal 2 SDs on the Gaussian distribution will end up being called growth restricted. Once defined as growth restricted, an abundance of resources is used on these patients, including counseling, rescanning, and generally delivering the fetuses earlier than would have occurred had there been no significant variance between US and menstrual dating. Early delivery has been shown repeatedly to lead to increased neonatal morbidity.

It is well known that a combination of maternal, placental, and fetal etiologies can lead to true growth restriction, and determining the etiology for growth restriction needs to be individualized in each case. Ultimately, the diagnosis of IUGR also relies on a combination of clinical and US findings such as amniotic fluid indices, umbilical cord arterial Doppler examinations, and the presence of fetal fat within the fetus[16–22] to make the diagnosis and to determine delivery timing.

The purpose of this article is to demonstrate that it is the shape of the growth curve generated on subsequent follow-up US examinations that is the best determinant of growth, regardless of how the pregnancy is dated. Numerous pitfalls regarding the diagnosis of IUGR based on clinical dating as demonstrated by fetal growth curves will be presented.

Materials and Methods

Following Institutional Review Board approval (HSD-00002969), 860 third-trimester fetal US examinations were retrospectively evaluated from January 1 through July 1, 2017, to determine whether they had normal growth curves. The necessity for written informed consent was waived by the Institutional Review Board. All examinations were performed on Philips (Amsterdam, the Netherlands) iU22 and Epic US machines using curvilinear 5/1-MHz transducers. All fetal measurements were obtained at least twice using standardized imaging planes and anatomy.[1] The criteria of Hadlock et al[5] were used for all fetal weight and weight percentile determinations.[5] All information was entered into ViewPoint, (GE Healthcare, Milwaukee, WI) for production of growth curves. Outcome data recorded included weight at delivery, Apgar scores, pH, neonatal intensive care unit admission, in utero fetal demise, and neonatal death.

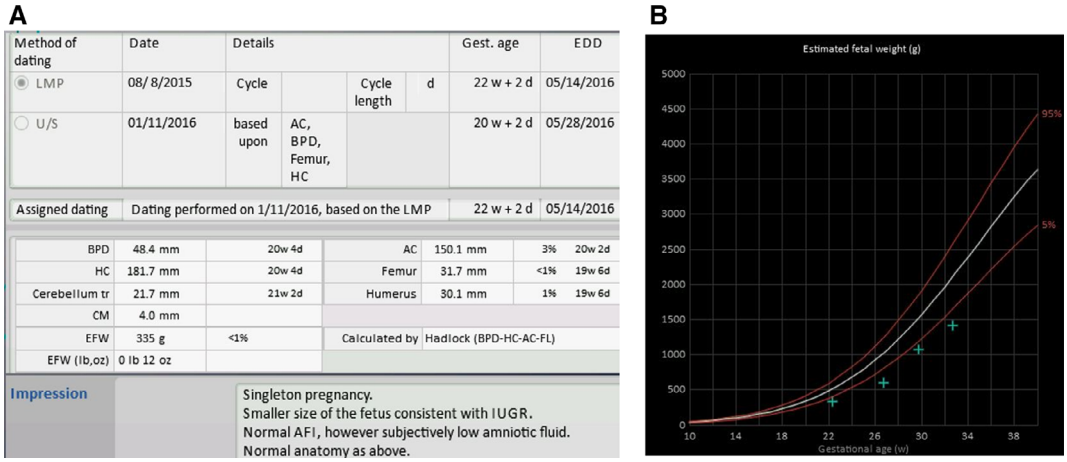
Inclusion criteria included at least having had 2 US examinations, preferably first-trimester and/or early second-trimester examinations performed within our own institution, and outcome information, including Apgar scores, neonatal pH, neonatal intensive care unit admission, and in utero or neonatal demise. Of these, 216 fetuses were found to be below the 10th percentile based on prior clinical dating (25%). Fetuses with only external assessments, congenital anomalies, and early demise before 12 weeks were excluded from the study.

Results

A total of 860 patients with third-trimester US examinations were reviewed, and 216 had a diagnosis of being below the 10th percentile in weight for gestational age and suspicious for IUGR. Of the 216 of 860 fetuses (25%) suspected of IUGR based on a weight below the 10th percentile, a total of 192 of 216 (89%) had normal growth curves (Figure 1). All 192 of these patients had initial first-trimester US or early second-trimester US examinations placing the fetus close to or at 2 SDs below the mean for gestational age based on LMP, and all were subsequently assigned weight percentiles below 10%, raising concern for IUGR on subsequent examinations. All 192 fetuses in this group had normal outcomes.

Of 216, 18 (8.3%) fetuses had follow-up examinations showing less-than-expected growth between the initial first- and second-trimester US examinations but appropriate interval growth on all follow-up

Figure 1. A and **B** demonstrate a classic example of the size-date discrepancies occurring in the mid second trimester and the necessity of relying on the growth curve rather than gestational age and the weight percentile assigned to the fetus. Despite being within the normal 2-SD range, the discrepancy between menstrual dating and US dating of 22 weeks 2 days and 20 weeks 2 days places the fetus below the 1st percentile. All subsequent US examinations show a growth curve that is normal. Amniotic fluid and umbilical cord arterial Doppler systolic-to-diastolic ratios were normal in all subsequent examinations. Clinical notes indicate that the patient underwent extensive counseling for IUGR. Subsequently, the fetus was delivered at 35 weeks based on the menstrual history (33 weeks by US). It weighed 1894 g and had Apgar scores of 8 and 9 at 1 and 5 minutes, respectively, and was discharged without incident.



examinations. The etiology of this discrepancy was not entirely clear but was thought to be related to the change from the measurement of crown-rump length in the first trimester to measurement of the head, abdomen, and femur in the third trimester and an aggregate calculation of age and weight. This change affected the growth curves as calculated by ViewPoint. All 18 of these had normal neonatal outcomes.

Figure 2. A and **B** demonstrate declining growth curves indicative of true growth restriction in 2 different fetuses. Both fetuses were initially well within the 2-SD growth limits for gestational age, and both had growth that declined to the point where they were more than 2 SDs below the mean. Both subsequently developed oligohydramnios, and following abnormal stress testing, both were delivered on the day immediately after the last US examination. After short neonatal intensive care unit admissions, both fetuses were ultimately discharged.

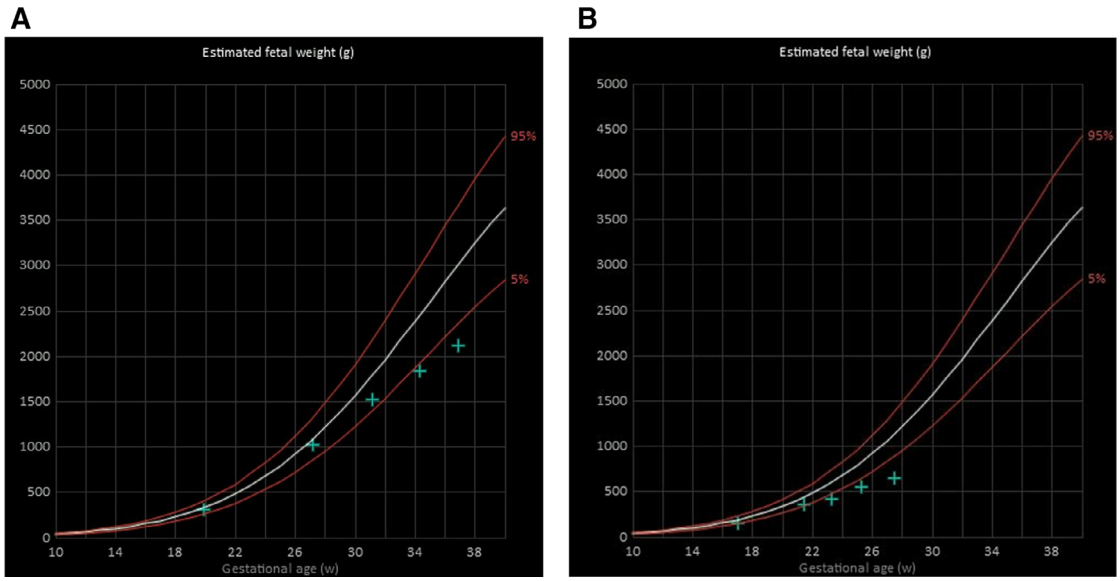
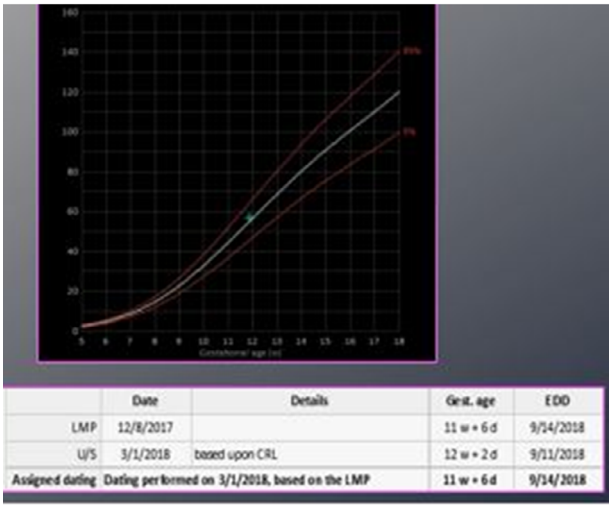


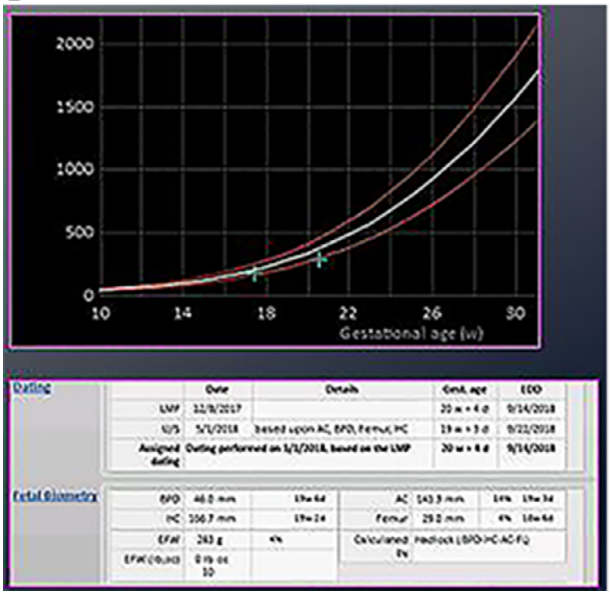
Figure 3. A and **B** demonstrate the importance of always reviewing the initial prior US examination to assess interval growth and not relying on size-date concordance alone. In this case, the initial examination at 12 weeks was concordant with the menstrual history, with the US being 3 days later. Growth had already fallen off by 17 weeks, and on a follow-up examination at 20 weeks, the US showed that the fetus had gone from being 3 days ahead to 7 days behind: a change of 10 days. Technically these last 2 examinations are still concordant with dates, but the interval growth has fallen off significantly. This is a pitfall that gets missed frequently, and declining growth between subsequent US examinations needs to be followed to determine whether growth continues to fall off, or whether the growth stabilizes. In this case, the fetus died in utero shortly after the second examination.

A



	Date	Details	Gest. age	EDD
LMP	12/8/2017		11 w + 6 d	9/14/2018
U/S	3/1/2018	based upon CRL	12 w + 2 d	9/11/2018
Assigned dating	Dating performed on 3/1/2018, based on the LMP		11 w + 6 d	9/14/2018

B

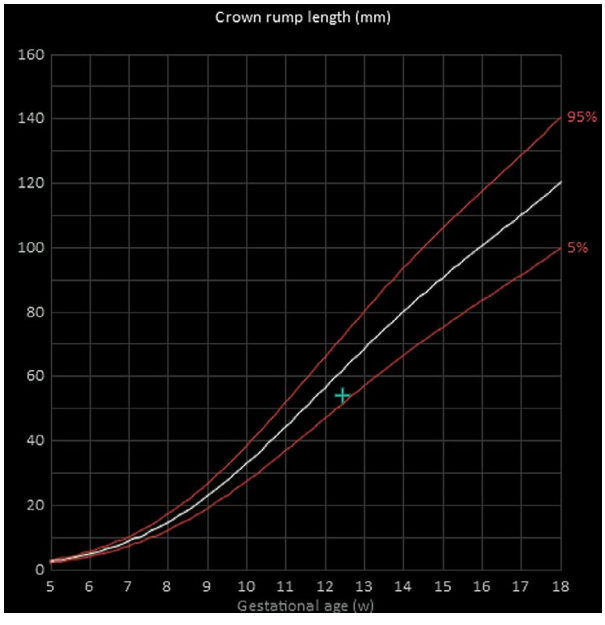


A total of 6 fetuses with abnormal growth curves developed fetal distress suggestive of true IUGR: 5 that led to emergent delivery of fetuses weighing less than 1000 g and 1 in utero fetal demise (Figure 2). Three of these 6 (50%) had gestational ages based on US initially greater than dating based on the menstrual history but had declining growth on subsequent examinations based on US but remained concordant with clinical dating (Figure 3).

On average, each patient included in this series averaged 4 complete US examinations, 6 additional limited amniotic fluid and cord Doppler examinations, and 8 clinical appointments.

We subsequently discovered that the discrepancies in dating between the LMP and first-trimester US were producing standard deviations in ViewPoint that were different from societal guidelines. For example, we use ±5 days for 2 SDs for the range of normal in the first trimester, yet we discovered that ViewPoint was using 4 days for 2 SDs and that a difference as small as 3 days would place the fetus below the 10th percentile (Figure 4).

Figure 4. This figure demonstrates that automated reporting systems may not be using the same standard deviations as outlined in societal guidelines., In this case, a discrepancy as little as 3 days in the first trimester placed the fetus at the 10th percentile. A discrepancy of 5 days, which most guidelines recommend, places the fetus greater than 2 SDs below the mean for the assigned gestational age. Since all subsequent examinations are based on the initial dating, these fetuses have an assigned gestational age that places them at or below the 10th percentile, and subsequently, all are followed for suspicion of growth restriction.



Discussion

Our data indicate that the assignment of weight percentiles based on clinical estimates of gestational age can be erroneous, and this finding is consistent with that from prior publications indicating that IUGR is frequently overdiagnosed when LMP dating is used. Our data show that 25% of all third-trimester fetuses were found to be below the 10th percentile early in gestation. Although some referral bias was due to the referral of more high-risk patients to a university than would be seen in the community, we believe that this was caused mainly because discrepancies between US and the menstrual history were placing the fetuses below the 10th percentile, which is still within 2 SDs of the mean. Once placed at the 10th percentile, as long as the same dates are used, this discrepancy is propagated forward, and appropriate interval growth will still place the fetus at the 10th percentile throughout the rest of the pregnancy.

Compounding the problem, we discovered that we were actually using standard deviations that were smaller than those recommended by societal guidelines. This places more fetuses below the 10th percentile than would be expected. For example, if a fetus is dated in the third trimester for the first time with a 2-week discrepancy between clinical and US dates, one would expect this fetus to somewhere around the 25th percentile, as one would expect a variance of $\pm$3 weeks at this gestational age. However, 2 weeks is being used by the reporting software, and subsequently, all of these fetuses are placed below the 10th percentile, raising concern for IUGR.

Accurate assignment of gestational age reduces postdated pregnancies[23] and allows the optimal timing of interventions as well as the prevention of unnecessary procedures. Even if menstrual dates are considered "certain," US is still more accurate for calculating the expected date of delivery.[24] Numerous studies have shown that dating by US biometry in the first half of pregnancy is more accurate than using menstrual data alone.[1,11,25] Savitz et al[26] found that the LMP assigns gestation 2.8 days longer on average than US dating, yields substantially more post-term births (12.1% versus 3.4%), and predicts delivery among term births less accurately.

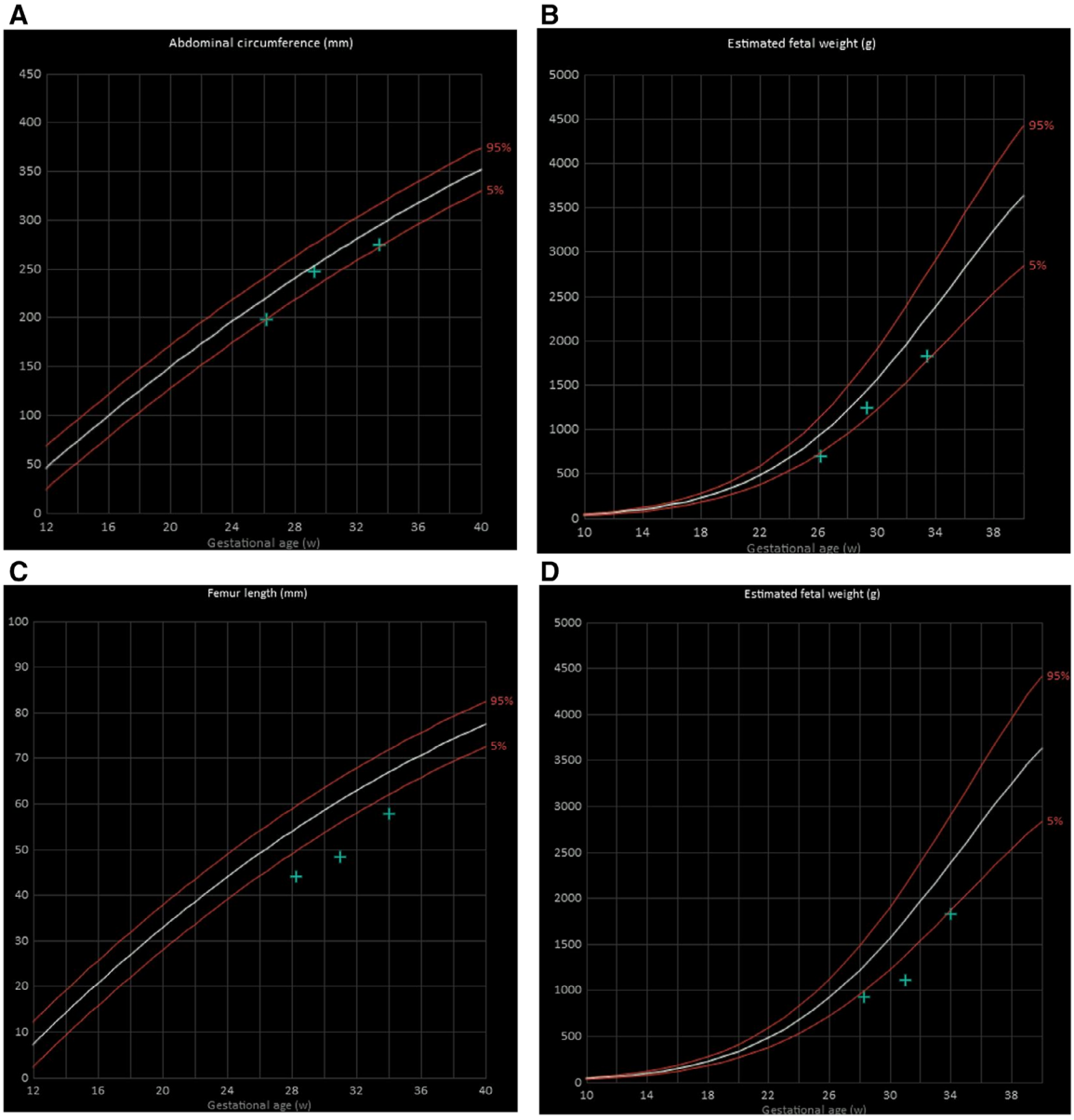
Misclassification of births as post-term was more common in women who reported the LMP by preferred dates of the month. Savitz et al[26] concluded that LMP estimates of gestational age are subject to greater random errors and in general tend to overstate the duration of gestation because of delayed ovulation. Butt and Lim[4] concluded that when performed with quality and precision, US alone is more accurate than a "certain" menstrual date for determining gestational age in the first and second trimesters (≤23 weeks) in spontaneous conceptions.

In evaluating growth and weight, both maternal characteristics and fetal size need to be considered. In a heterogeneous population, the accurate assessment of an individual fetus's weight and weight percentile needs to be customized[27] to reduce the false-positive rate for the diagnosis of growth restriction, especially with regard to the fetal abdominal circumference and femur length (Figure 5). In one of the largest studies of its kind, Wilcox et al[8] found that there were significant differences between the birth weights of 37,336 neonates of mothers of English-European origin (3357 g), 1008 neonates of Afro-Caribbean origin (3173 g), and 1547 neonates of Indian Subcontinent origin (3096 g). There were also significant differences between ethnic groups for gestational length, parity, maternal height, weight, and smoking.

Furthermore, antenatal US estimations of fetal weight need to be projected forward to the expected time of birth.[27] In other words, the delay from the time of the US to the time of delivery needs to be considered. Wilcox et al[27] concluded that a standardized comparison between the predicted birth weight at US and the actual birth weight is a more logical reflection of the normality of intrauterine growth and therefore more logical as an outcome measure for pregnancy than crude birth weight alone.

Our data indicate that with regard to the diagnosis of IUGR, the exact method used to date the pregnancy initially is less important than the shape of the growth curves subsequently. Based on discrepancies between first-trimester US dating, the menstrual history, or external assessments of gestational age, if the fetus lies within 2 SDs (5th–95th percentiles) of the mean, this should be accepted as normal initially. Subsequent follow-up US examinations need to show normal growth curves whether the fetus started at the 90th percentile or 10th percentile, and if further declining growth is shown, then concern for potential IUGR would indicate the need for continued surveillance and testing.

Figure 5. A–D demonstrate that care must be taken to ensure the quality of all fetal measurements used in the calculation of fetal weight. Undermeasurement of the head, abdomen, or femur can lead to underestimation of fetal weight and the weight percentile. The images illustrate that small abdominal circumferences and short femurs can both subsequently lead to a diagnosis of IUGR. However, both of these fetuses had normal amniotic fluid, cord Doppler findings, and subcutaneous fetal fat. On a subsequent examination, when the abdominal circumference was remeasured, the fetal growth curve returned to normal. In the other patient, the extremities continued to show diminishing growth, and musculoskeletal dysplasia was diagnosed, although the birth weight was normal. Whenever a fetus is deemed to be small, it is important to look at each individual parameter to determine whether the size is being brought down by one particular measurement.



There are pitfalls associated with the assignment of weight percentiles to fetuses based on clinical dating. If using an automated reporting system, one needs to ensure that the standard deviations being reported by the system are consistent with the societal standards that one thinks one is using. The way to

ensure this is to find cases that are exactly 2 SDs discrepant and check where on the growth curve the fetus lies. If a 30-week fetus (for whom the SD should be ±3 weeks) is put at the 1st percentile when there is a 2-week discrepancy between the LMP and US dates, then the standard deviations being reported do not match those of the societal guidelines.

Another pitfall in assessing growth is to assume that size-date concordance is the same thing as appropriate interval growth. Special care must be taken to look at the initial prior US examinations to determine what due date was actually given based on US, since follow-up examinations of fetuses initially greater than dates can show significant decreases in growth yet still be consistent with dates. Given the limited amount of resources available and diminished funding for the care of these patients, it is essential to be as efficient as possible and recognize that the shape of the growth curve is more indicative of fetal health or growth restriction than the dating and, especially, the weight percentile assigned to the fetus when a size-date discrepancy exists.

In conclusion, as long as the fetal growth falls along a curve that parallels normal growth curves, appropriate growth has occurred regardless of the gestational age and weight percentile that have been assigned to the fetus, and concerns regarding growth restriction should be alleviated by this result. When a fetus is determined to have a declining growth curve irrespective of gestational age, and without reliance on arbitrarily determined weight percentiles, then secondary signs of fetal distress, including oligohydramnios, abnormal umbilical cord Doppler findings, diminished fetal fat, stress testing, and close clinical surveillance including fetal heart rate monitoring, are then indicated to optimize delivery timing.[15–21]

References

1. Committee on Obstetric Practice, the American Institute of Ultrasound in Medicine, the Society for Maternal-Fetal Medicine. Committee opinion no 700: methods for estimating the due date. *Obstet Gynecol* 2017; 129:e150–e154.
2. Robinson HP, Fleming JEE. Critical evaluation of sonar crown-rump length measurements. *Br J Obstet Gynaecol* 1975; 82:702–710.
3. Whitworth M, Bricker L, Mullan C. Ultrasound for fetal assessment in early pregnancy. *Cochrane Database Syst Rev* 2015; 7:CD007058.
4. Butt K, Lim K, Diagnostic Imaging Committee. Determination of gestational age by ultrasound. *J Obstet Gynaecol Can* 2014; 36: 171–181.
5. Hadlock FP, Shah YP, Kanon DJ, Lindsey JV. Fetal crown-rump length: reevaluation of relation to menstrual age (5–18 weeks) with high-resolution real-time US. *Radiology* 1992; 182:501–505.
6. Wisser J, Dirschedl P, Krone S. Estimation of gestational-age by transvaginal sonographic measurement of greatest embryonic length in dated human embryos. *Ultrasound Obstet Gynecol* 1994; 4:457–462.
7. 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; 73:243–249.
8. Wilcox M, Gardosi J, Mongelli M, Ray C, Johnson I. Birth weight from pregnancies dated by ultrasonography in a multicultural British population. *BMJ* 1993; 307:588–591.
9. Benson CB, Doubilet PM. Sonographic prediction of gestational age: accuracy of second-trimester and third-trimester fetal measurements. *AJR Am J Roentgenol* 1991; 157:1275–1277.
10. Daya S. Accuracy of gestational-age estimation by means of fetal crown-rump length measurement. *Am J Obstet Gynecol* 1993; 168: 903–908.
11. AIUM-ACR-ACOG-SMFM-SRU practice parameter for the performance of standard diagnostic obstetric ultrasound examinations. *J Ultrasound Med* 2018; 37:E13–E24.
12. Kalish RB, Thaler HT, Chasen ST, et al. First- and second-trimester ultrasound assessment of gestational age. *Am J Obstet Gynecol* 2004; 191:975–978.
13. Pexsters A, Daemen A, Bottomley C, et al. New crown-rump length curve based on over 3500 pregnancies. *Ultrasound Obstet Gynecol* 2010; 35:650–655.
14. Hadlock FP, Deter RL, Harrist RB, Park SK. Computer-assisted analysis of fetal age in the third trimester using multiple fetal growth-parameters. *J Clin Ultrasound* 1983; 11:313–316.
15. Hadlock FP, Deter RL, Harrist RB, Park SK. Estimating fetal age: computer-assisted analysis of multiple fetal growth-parameters. *Radiology* 1984; 152:497–501.
16. Hidaka N, Sato Y, Kido S, Fujita Y, Kato K. Ductus venosus doppler and the postnatal outcomes of growth restricted fetuses with absent end-diastolic blood flow in the umbilical arteries. *Taiwan J Obstet Gynecol* 2017; 56:642–647.
17. Kingdom JC, Audette MC, Hobson SR, Windrim RC, Morgen E. A placenta clinic approach to the diagnosis and management of fetal growth restriction. *Am J Obstet Gynecol* 2018; 218:S803–S817.
18. Seravalli V, Di Tommaso M, Petraglia F. Managing fetal growth restriction: surveillance tests and their interpretation. *Minerva Ginecol* 2019; 71:81–90.
19. Sirico A, Diemert A, Glosemeyer P, Hecher K. Prediction of adverse perinatal outcome by cerebroplacental ratio adjusted for estimated fetal weight. *Ultrasound Obstet Gynecol* 2018; 51: 381–386.

20. Wu L, Richardson ML, Dubinsky T. Predicting adverse neonatal outcome especially when gestational age is uncertain: utility of sonographic measurement of fetal abdominal wall thickness. *Ultrasound Q* 2017; 33:208–212.
21. Dubinsky TJ, O'Regan J, Sonneborn R, Hippe DS, Dighe M, Moshiri M. A nomogram of lateral abdominal wall fat thickness in normal third trimester fetuses. *Ultrasound Q* 2019; 35:30–34.
22. Dubinsky TJ, Parvey HR. Cerebral arterial-umbilical cord arterial systolic/diastolic ratios: an independent predictor of poor neonatal outcome. *Am J Obstet Gynecol* 1995; 172:1943–1944.
23. Bennett KA, Crane JMG, O'Shea P, Lacelle J, Hutchens D, Copel JA. First trimester ultrasound screening is effective in reducing postterm labor induction rates: a randomized controlled trial. *Am J Obstet Gynecol* 2004; 190:1077–1081.
24. Wegienka G, Baird DD. A comparison of recalled date of last menstrual period with prospectively recorded dates. *J Womens Health* 2005; 14:248–252.
25. Mongelli M, Wilcox M, Gardosi J. Estimating the date of confinement: ultrasonographic biometry versus certain menstrual dates. *Am J Obstet Gynecol* 1996; 174:278–281.
26. Savitz DA, Terry JW Jr, Dole N, Thorp JM Jr, Siega-Riz AM, Herring AH. Comparison of pregnancy dating by last menstrual period, ultrasound scanning, and their combination. *Am J Obstet Gynecol* 2002; 187:1660–1666.
27. Wilcox MA, Johnson IR, Maynard PV, Smith SJ, Chilvers CE. The individualised birthweight ratio: a more logical outcome measure of pregnancy than birthweight alone. *Br J Obstet Gynaecol* 1993; 100: 342–347.