2013

Sonographic accuracy of estimated fetal weight in twins

Lorie M. Harper
University of Alabama - Birmingham

Kimberly A. Roehl
Washington University School of Medicine in St. Louis

Methodius G. Tuuli
Washington University School of Medicine in St. Louis

Anthony O. Odibo
Washington University School of Medicine in St. Louis

Alison G. Cahill
Washington University School of Medicine in St. Louis

Sonographic Accuracy of Estimated Fetal Weight in Twins

Lorie M. Harper, MD, MSCI, Kimberly A. Roehl, MPH, Methodius G. Tuuli, MD, MPH, Anthony O. Odibo, MD, MSCE, Alison G. Cahill, MD, MSCIs

Objectives—We aimed to determine the accuracy of sonographic formulas for estimating birth weight in twin pregnancies.

Methods—We conducted a retrospective cohort study of consecutive twin pregnancies undergoing sonography at within 1 week of delivery. Pregnancies were included if biometric measurements and birth weight were available and excluded if anomalies were present. Estimated fetal weight was calculated using three sonographic formulas: two derived from singletons (Hadlock and Shepard) and one from twins (Ong). The correlation between estimated fetal weight and birth weight was determined using the Pearson correlation coefficient. The accuracy of each formula (bias) was assessed using the mean percentage error [(estimated fetal weight – birth weight)/birth weight × 100], and the precision (random error) was estimated from the standard deviation of the percentage error. The screening efficiency of each formula for intrauterine growth restriction, defined as below the 10th percentile on the Alexander growth standard, was assessed. The effect of twin presentation was determined using a paired analysis.

Results—Of 1744 consecutive twin pregnancies, 270 (540 infants) met inclusion criteria. The estimated fetal weight of all 3 formulas strongly correlated with the birth weight (Pearson $r = 0.90$ for Hadlock, 0.87 for Shepard, and 0.92 for Ong). Each formula had similar sensitivity (65%–70%) and specificity (85%–90%) for intrauterine growth restriction. For each formula, the correlation coefficient was similar between twins A and B (Pearson $r = 0.85$–0.93); however, the estimated fetal weight for twin A tended to underestimate birth weight, whereas the estimated fetal weight for twin B tended to overestimate birth weight.

Conclusions—Three widely used estimated fetal weight formulas, two derived from singletons and one from twins, perform equally well in estimating birth weight in twin gestations.

Key Words—estimated fetal weight; intrauterine growth restriction; multifetal gestation

Compared to singletons, twin pregnancies are at increased risks of stillbirth and growth restriction. For this reason, sonographic estimation of fetal weight is a ubiquitous antenatal tool for assessing fetal growth and well-being. However, currently used formulas for determining the estimated fetal weight from sonographic biometry are based on singleton gestations.[1] Because of the number of fetuses, placental mass, and fetal positions, sonography of twin gestations presents unique technical challenges compared to singletons, thereby potentially systematically altering crucial measurements such as biparietal diameter, head circumference, and abdominal circumference. The biparietal diameter and head

Received September 4, 2012, from the Department of Obstetrics and Gynecology, University of Alabama, Birmingham, Alabama USA (L.M.H.); and Department of Obstetrics and Gynecology, Washington University, St Louis, Missouri USA (K.A.R., M.G.T., A.O.O., A.G.C.). Revision requested September 5, 2012. Revised manuscript accepted for publication September 11, 2012.

This work was presented as an oral session at the American Institute of Ultrasound in Medicine 2012 Annual Convention; March 31, 2012; Phoenix, Arizona. Dr Harper is supported by grant K12HD001258-13 from the National Institute of Child Health and Human Development (principal investigator, William W. Andrews, MD, PhD).

Address correspondence to Lorie M. Harper, MD, MSCI, Department of Obstetrics and Gynecology, University of Alabama, 1700 Sixth Ave S, Suite 10270, Birmingham, AL 35233-7333 USA.

E-mail: lmharper@uabmc.edu

Abbreviations
ROC, receiver operating characteristic

circumference are of particular concern, as the breech presentation, relatively common in twins, is associated with dolichocephaly and decreased accuracy of estimated fetal weight.[2,3] Consequently, formulas derived from singleton populations, which rely heavily on head measurements, may perform less well in twins by magnifying the effects of an inaccurate or difficult to obtain head measurement. In fact, one study found that two commonly used singleton formulas have reduced sensitivity for detecting growth restriction in twin gestations.[4]

A formula derived from twin gestations, which relies more on the easily obtained femur length, may have improved accuracy and superior detection of fetal growth restriction. Therefore, we aimed to evaluate the accuracy of several commonly used formulas for calculating sonographically estimated fetal weight in twin gestations.

Materials and Methods

We performed a retrospective cohort study of all patients who underwent routine second-trimester (15–22 weeks) sonography for an anatomic survey at a single tertiary care center. Institutional Review Board approval was obtained from the Washington University School of Medicine. Data were collected prospectively by dedicated nurses from 1990 to 2009. Each patient undergoing sonography in our center receives a standardized handout requesting information regarding pregnancy complications, delivery complications, and neonatal outcomes, to be filled out and returned after delivery. The coordinator called the patient, and in cases in which the patient could not be reached, the physician, if the form was not returned within 4 weeks of the delivery date.

Patients were included in this study if they carried a twin gestation and if sonographically obtained biometric measurements (biparietal diameter, head circumference, abdominal circumference, and femur length) were available within 7 days of delivery and if birth weight was available. Singleton gestations, intrauterine fetal death, higher-order multiple gestations, and pregnancies complicated by anomalies were excluded. All sonographic examinations were performed by trained sonographers specializing in obstetric sonography, and all sonographic examinations were interpreted by an attending physician dedicated to obstetric sonography.

Estimated fetal weight was calculated from biometric measurements using three different sonographic formulas (Table 1).[4–6] Two of these formulas (Hadlock and Shepard) are widely used formulas derived from singleton gestations. The third formula (Ong) was derived from twin gestations.[4] The Hadlock and Shepard formulas were chosen as representative of singleton-derived formulas because they rely on different measurements (head circumference versus biparietal diameter); the Ong formula was selected because it was derived from a twin population and, in contrast to the singleton formulas, does not rely on any cranial measurements.

Maternal and infant characteristics were evaluated using descriptive statistics. The correlation between estimated fetal weight and birth weight was determined using the Pearson correlation coefficient. The accuracy (bias) of each formula was determined by calculating the mean percentage error [(estimated fetal weight – birth weight)/birth weight × 100%], and the precision (random error) of each formula was determined from the standard deviation of the mean percentage error. The mean percentage error of each formula was compared using the Student *t* test. The screening efficiency (sensitivity, specificity, positive predictive value, and negative predictive value) of each formula for intrauterine growth restriction, defined as below the 10th percentile on the Alexander growth standard,[7] was assessed. Exact binomial confidence intervals were calculated for each value. In addition, a receiver operating characteristic (ROC) curve was created to describe each formula's ability to predict growth restriction after adjusting for gestational age. The ROC curves for each formula were compared using the nonparametric method described by DeLong et al.[8]

Because the fetal position in the pelvis can affect biometric measurements (ie, measurement of the biparietal diameter can be difficult when the presenting part is low in the pelvis), the effect of twin order on the mean percentage error of the estimated fetal weight was determined using a paired analysis. Because the study cohort spanned a long period, the effect of the year on accuracy and preci-

Table 1. Sonographic Formulas for Estimated Fetal Weight

Source	Formula
Hadlock	$Log_{10}EFW = 1.326 - 0.00326(AC)(FL) + 0.0107(HC) + 0.0438(AC) + 0.158(FL)$
Shepard	$Log_{10}EFW = -1.7492 + 0.166(BPD) + 0.046(AC) - 0.002546(BPD)$
Ong	$Log_{10}EFW = -2.7606 + 0.0259(AC) + 0.67200(FL) - 0.0475(FL)^2$

AC indicates abdominal circumference; BPD, biparietal diameter; EFW, estimated fetal weight; FL, femur length; and HC, head circumference.

sion of the estimated fetal weight was considered by comparing the mean percentage errors from 1990 to 1999 and 2000 to 2009. The statistical analysis was performed using Stata version 11 Special Edition software (StataCorp, College Station, TX).

Results

Of 1744 twin pregnancies, 270 twin pregnancies (540 infants) met inclusion criteria (109 pregnancies were excluded for fetal anomalies; 192 pregnancies were excluded for undocumented birth weight of one or both infants; and 1173 were excluded for no sonography within 7 days of delivery). The maternal and infant characteristics of the included patients are listed in Table 2; most pregnancies delivered in the third trimester with a median sonography-to-delivery interval of 3 days.

Table 2. Maternal and Infant Characteristics

Characteristic	Value
Maternal age, y	30.9 (26.3–35.3)
Gravidity	2 (1–3)
Maternal race, n (%)	
White	159 (63.6)
Black	59 (23.6)
Maternal body mass index, kg/m^2	25.6 (22.5–30.7)
Sonographic examination performed 1990–1999	210 (38.9)
Sonographic examination performed 2000–2009	330 (61.1)
Sonographic age at delivery, wk	35.2 (33.3–36.6)
Sonography-to-delivery interval, d	3 (1–5)
Birth weight, A, g	2263 (1800–2639)
Birth weight, B, g	2185 (1702–2525)
Male fetus	270 (50.0)
Birth weight <10th percentile	157 (29.1)

Data presented as median (interquartile range) and as number (percent).

All three formulas showed a strong correlation between the sonographically estimated fetal weight and actual birth weight (Table 3), with Pearson *r* correlation coefficients between 0.87 and 0.92 and R^2 values of 0.98 or higher. Overlapping confidence intervals of the Pearson correlation coefficients indicate no difference in the correlation of the formulas with actual birth weights. The mean percentage errors of the Shepard and Hadlock formulas were similar, although the mean percentage error of the Ong formula was statistically different from that of the Hadlock formula. However, the mean percentage errors of all three formulas were within 2% of the actual birth weight, with a random error of each formula of ±13% to 15%. The random error of the Shepard formula was higher than that of the other two formulas ($P < .01$). The Hadlock and Ong formulas were within 10% of the actual birth weight in 67% to 69% of cases, although the Shepard formula was within 10% of the actual birth weight in only 61% of cases ($P < .01$ versus either the Hadlock or Ong formula).

All three formulas detected fetal growth restriction with approximately 67% to 70% sensitivity and greater than 85% specificity (Table 4). All three had positive predictive values greater than 65% and negative predictive values of at least 85%. Overlapping confidence intervals suggest no difference in the performance of these formulas for detecting growth restriction. In addition, we calculated the area under the ROC curve after adjusting for the gestational age at delivery. All three formulas had similar areas under the curve for predicting growth restriction ($P = .58$).

The correlation between the sonographically estimated fetal weight and the actual birth weight remained high for each formula when calculated separately for the presenting and nonpresenting twins (Table 5). The sonographically estimated fetal weight for the presenting twin tended to be slightly lower than the actual birth weight (mean percentage error, –2.39% to –0.52%), whereas the

Table 3. Correlation, Accuracy, and Precision of Sonographically Estimated Fetal Weight and Birth Weight

	Correlation of Estimated Fetal Weight With Birth Weight		Accuracy and Precision for Estimating Fetal Weight				Estimates Within 10% of Birth Weight	
Formula	Pearson *r*	Coefficient of Determination (R^2)	Mean % Error (Accuracy)	*P*	SD (Precision)	*P*	%	*P*
Hadlock	0.90 (0.89–0.92)	0.99	–0.34	Ref	13.5	Ref	69.1	Ref
Shepard	0.87 (0.85–0.89)	0.98	–0.39	.88	15.6	<.01	61.7	<.01
Ong	0.92 (0.91–0.93)	0.99	1.68	<.01	13.6	.92	67.2	<.01

Values in parentheses are 95% confidence intervals. Ref indicates reference value.

sonographically estimated fetal weight for the nonpresenting twin tended to be slightly higher (mean percentage error, 0.44% to 3.88%), although the actual mean percentage errors were small for both twins. The precision of the Hadlock and Shepard formulas was slightly better for the nonpresenting twin compared to the presenting twin, although the differences in random error between presenting and nonpresenting twins were not statistically significant. The differences between the mean percentage errors calculated for each twin by each formula were statistically significant, although the absolute differences between formulas were small. Differences in random error were not statistically significant between formulas.

When the cohort was segregated by year of examination (1990–1999 and 2000–2009), the Pearson correlation coefficient for each formula was unchanged (data not shown). The mean percentage errors of the Shepard and Ong formulas were similar in each decade ($P = .16$ and .10, respectively). The mean percentage error of the Hadlock formula was different by decade ($P = .01$), although the absolute difference was clinically insignificant (<3%). When random error was compared by the year of examination, no statistical difference existed for the Hadlock or Shepard formulas ($P = .25$ and .23, respectively). The random error of the Ong formula was statistically different ($P < .01$) but again not clinically significant (<2%).

Discussion

All three sonographic formulas evaluated performed equally well in predicting the actual birth weight of twins and in diagnosing fetal growth restriction. Although all three formulas consistently underestimated the birth weight of the presenting twin and overestimated the birth weight of the second twin, the absolute value of the mean percentage error was relatively small for both twins. Although the Hadlock and Shepard formulas were derived from singleton populations, their use in twin gestations may be appropriate when the biometric measurements necessary to use the calculation can be obtained. Because the Shepard formula performed slightly less well in estimating the fetal weight within 10% of the actual birth weight, the Hadlock formula may be preferred. The Ong formula derived from a twin population did not perform significantly better than the Hadlock and Shepard formulas in our twin population.

Table 4. Test Characteristics of Each Formula for Predicting Fetal Growth Restriction

Formula	Sensitivity, %	Specificity, %	Positive Predictive Value, %	Negative Predictive Value, %	Area under the ROC Curve[a]
Hadlock	70.1 (62.2–77.1)	86.4 (82.5–89.6)	67.9 (60.1–75.0)	87.5 (83.7–90.7)	.8694
Shepard	70.7 (62.9–77.7)	85.3 (81.3–88.7)	66.5 (58.8–73.6)	87.6 (83.8–90.8)	.8585
Ong	66.9 (58.9–74.2)	89.0 (85.4–91.9)	71.4 (63.4–78.6)	86.7 (82.9–89.9)	.8699

Values in parentheses are 95% confidence intervals.
[a]$P = .58$.

Table 5. Comparison of Formulas by Presenting Versus Nonpresenting Twin

	Twin A					Twin B					
Formula	Pearson r	Mean % Error (Accuracy)	P^a	SD (Precision)	P^a	Pearson r	Mean % Error (Accuracy)	P^a	SD (Precision)	P^a	P^b
Hadlock	0.89 (0.86–0.91)	–2.39	Ref	14.3	Ref	0.92 (0.90–0.94)	1.71	Ref	12.4	Ref	<.01
Shepard	0.85 (0.82–0.88)	–1.22	<.01	17.0	.38	0.90 (0.87–0.92)	0.44	<.01	14.0	.92	.15
Ong	0.91 (0.89–0.93)	–0.52	<.01	13.6	.39	0.93 (0.91–0.95)	3.88	.02	13.2	.44	<.01

Values in parentheses are 95% confidence intervals. Ref indicates reference value.
[a]P compares mean error and SD between formulas using Hadlock as the reference.
[b]P compares mean error of twin A versus twin B.

Although several studies have examined the most effective way to predict birth weight discordance, relatively few have evaluated the most effective formula to use to estimate fetal weight in twins. The estimated fetal weight is essential in the management of twin pregnancies; it is used to determine the mode of delivery and when or if antenatal testing is indicated. In monochorionic gestations, discordance in the estimated fetal weights can also be used to aid in the diagnosis of twin-twin transfusion. Therefore, it is crucial that the formulas used to translate sonographic measurements of fetal parts into an estimated fetal weight are accurate calculations of the birth weight, rather than magnifying the effects of difficult-to-obtain measurements in twins.

Prior studies have shown that the Hadlock formula has a higher mean percentage error in twins compared to the Ong formula. Danon et al[9] performed a case-control study showing that the Hadlock formula had a higher mean percentage error in twins than in singletons, although the mean absolute error in twins remained low at 174 g. The formula evaluated by Ong et al[4] was derived from a population of 73 twins and validated in a population of 152 twins. In this population, the Ong formula performed better than Hadlock or Shepard formula in detecting growth restriction below the 10th percentile, although there were no statistically significant differences between the Ong formula and the other formulas in the mean percentage error. In a cohort study of 283 twin pregnancies, the Hadlock formula was superior to the Ong and Shepard formulas at estimating fetal weight in twins, although the Ong formula had an mean percentage error value of less than 2%.[10] In another cohort study that examined the factors influencing the accuracy of sonographically estimated fetal weight, the Hadlock formula was found to have a mean percentage error of approximately 8%.[11]

In contrast, we found that the Hadlock formula performed slightly better in our population of twins, with a mean percentage error of less than 1%, although the Ong formula had a mean percentage of error of less than 2%. Both the Hadlock and Ong formulas were similar in the percentage of time that estimated fetal weight was within 10% of the actual birth weight. The Shepard formula did not perform as well as the Hadlock or Ong formula in the percentage of cases in which the estimated fetal weight was within 10% of the actual birth weight. These findings suggest that the Hadlock formula can be used in twin gestations. This finding is important for a busy sonography or labor and delivery unit, where it may be impractical and time-consuming to change the formula used to calculate the estimated fetal weight between patients.

It was noted that the random errors, or precisions, of all formulas were 13% to 15%. This finding is similar to that in other publications about twins, suggesting that the difficulty in obtaining sonographic measurements in twins is not systematic and may lead to random errors rather than bias.[4,9]

Our study had several strengths. First, we strictly limited the sonography-to-delivery interval to 7 days or less, and most women delivered within 5 days of their last sonographic examination. This restriction enhanced the calculated accuracy of our sonographic estimations by limiting the amount of fetal growth that occurred between the sonographic examination and the birth weight measurement. Because fetal growth can be as high as 200 g per week in the third trimester,[7] or almost 10% of the infant body weight, it could be a potential source of substantial error when determining the sonographic accuracy of fetal weight. In addition, we had a fairly large sample size of 270 pregnancies and 540 infants, limiting any individual contribution to the observed mean percentage error.

One limitation of the study was that most of our patients were delivered between 33 and 36 weeks' gestation, limiting the generalizability of the study. We are unable to evaluate how well these sonographic formulas perform at substantially earlier gestations, and thus with lower estimated fetal weights, because of small numbers of these patients. However, the late second and third trimesters are the periods during which growth restriction is most likely to be diagnosed, and we are able to determine that these formulas perform well during those periods. In addition, we used a singleton growth curve for defining growth restriction, which may have led to overdiagnosis of growth restriction in a twin population; however, it is common practice to use a singleton growth curve because it generally leads to more conservative management of these high-risk twin pregnancies.[12] To accumulate a large number of twin pregnancies, our study spanned 20 years, with changes in sonographers and ultrasound technology. However, minimal differences existed in the accuracy of the estimated fetal weight compared to the birth weight over time. Finally, all sonographic examinations were performed by sonographers specializing in obstetric sonography, and all of the examinations were interpreted by attending physicians dedicated to obstetric sonography. Although this process provides high-quality images that adhere to guidelines set for obstetric sonography by the American Institute of Ultrasound in Medicine,[13] it may limit the generalizability of our findings to institutions with similar sonography practices and may not be applicable in all settings.

In conclusion, although derived from singleton populations, the Hadlock and Shepard formulas performed well in estimating the fetal weight of twins. The use of these formulas may be appropriate even in twin gestations as long as the biometric measurements necessary for the calculations can be obtained. The use of a formula specific to twin gestations may not be necessary to improve the sonographic estimation of fetal weight and the antenatal detection of fetal growth restriction.

References

1. Callen PW. *Ultrasonography in Obstetrics and Gynecology*. 5th ed. Philadelphia, PA: Saunders Elsevier; 2008.
2. Chauhan SP, Magann EF, Naef RW III, Martin JN Jr, Morrison JC. Sonographic assessment of birth weight among breech presentations. *Ultrasound Obstet Gynecol* 1995; 6:54–57.
3. Melamed N, Ben-Haroush A, Meizner I, Mashiach R, Yogev Y, Pardo J. Accuracy of sonographic fetal weight estimation: a matter of presentation. *Ultrasound Obstet Gynecol* 2011; 38:418–424.
4. Ong S, Smith AP, Fitzmaurice A, Campbell D. Estimation of fetal weight in twins: a new mathematical model. *Br J Obstet Gynaecol* 1999; 106:924–928.
5. Hadlock FP, Harrist RB, Sharman RS, Deter RL, Park SK. Estimation of fetal weight with the use of head, body, and femur measurements: a prospective study. *Am J Obstet Gynecol* 1985; 151:333–337.
6. Shepard MJ, Richards VA, Berkowitz RL, Warsof SL, Hobbins JC. An evaluation of two equations for predicting fetal weight by ultrasound. *Am J Obstet Gynecol* 1982; 142:47–54.
7. Alexander GR, Himes JH, Kaufman RB, Mor J, Kogan M. A United States national reference for fetal growth. *Obstet Gynecol* 1996; 87:163–168.
8. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988; 44:837–845.
9. Danon D, Melamed N, Bardin R, Meizner I. Accuracy of ultrasonographic fetal weight estimation in twin pregnancies. *Obstet Gynecol* 2008; 112:759–764.
10. Diaz-Garcia C, Bernard JP, Ville Y, Salomon LJ. Validity of sonographic prediction of fetal weight and weight discordance in twin pregnancies. *Prenat Diagn* 2010; 30:361–367.
11. Ocer F, Aydin Y, Atis A, Kaleli S. Factors affecting the accuracy of ultrasonographical fetal weight estimation in twin pregnancies. *J Matern Fetal Neonatal Med* 2011; 24:1168–1172.
12. Creasy RK, Resnik R, Iams JD. *Creasy and Resnik's Maternal-Fetal Medicine: Principles and Practice*. 6th ed. Philadelphia, PA: Saunders Elsevier; 2009.
13. American Institute of Ultrasound in Medicine. AIUM practice guideline for the performance of an antepartum obstetric ultrasound examination. *J Ultrasound Med* 2003; 22:1116–1125.