A Comparison of Clinical and Ultrasonic Estimation of Fetal Weight

DAN J. SHERMAN, MD, SHLOMO ARIELI, MD, JOSEPH TOVBIN, MD, GABRIELA SIEGEL, MD, ELIAHU CASPI, MD, AND IAN BUKOVSKY, MD

Objective: To compare the accuracy of routine ultrasonic and clinical birth weight estimation.

Methods: The study sample included 1717 women with singleton pregnancies, admitted in early labor with an ultrasonic estimated fetal weight (EFW) performed during the preceding week. Clinical EFW was obtained before rupture of the membranes by the attending senior resident, who was unaware of the ultrasonic EFW. Accuracy was determined by the percentage error, the absolute percentage error, and the proportion of estimates within 10% of the actual birth weight (birth weight ± 10%). Statistical analysis was done by the paired t test, the comparison of correlated variances, the Wilcoxon sign test, and the χ² test. Actual birth weight in the study sample averaged 3334 ± 607 g (± standard deviation, [SD]) and ranged between 690 and 5320 g.

Results: The means of all error terms of the clinical EFW were significantly smaller than those of the ultrasonic EFW. However, the rates of estimates within 10% of birth weight were not significantly different (72 and 69%, respectively). In birth weights less than 2500 g, both methods overestimated the birth weight, but the mean errors of the ultrasonic EFW were significantly smaller than those of the clinical EFW. The ultrasonic EFW had significantly higher rates of birth weight ± 10% than the clinical EFW (63 compared to 49%, respectively). In the 2500–4000 g birth weight, only the clinical EFW had no systematic error, whereas the ultrasonic EFW underestimated the birth weight. The mean errors of the clinical EFW were significantly smaller and the rate of birth weight ± 10% significantly higher than those of the ultrasonic EFW. In the birth weight greater than 4000 g, both methods underestimated the birth weight, and the mean errors and the rate of estimates within 10% of birth weight were similar for both methods.

Conclusion: Clinical estimation of birth weight in early labor is as accurate as routine ultrasonic estimation obtained in the preceding week. In the lower range of birth weight (less than 2500 g), ultrasonic estimation is more accurate; in the 2500–4000 g range, clinical estimation is more accurate. In the higher range of birth weight (greater than 4000 g), both methods have similar accuracy. (Obstet Gynecol 1998; 91:212–7. © 1998 by The American College of Obstetricians and Gynecologists.)

Accurate estimated fetal weight (EFW) is of paramount importance in the management of labor and delivery. During the last decade, EFW has been incorporated into the standard, routine antepartum evaluation of high-risk pregnancies and deliveries. The accuracy of predicting birth weight by a variety of different formulas, incorporating different ultrasonic measurements, has been studied extensively.1,2 However, no particular formula or biometric measurement has superior accuracy.3,4 In general, the mean absolute error of sonographically predicted birth weight varies between 6 and 12% of the actual birth weight, and 40–75% of estimates fall within 10% of the actual birth weight.3,5–7

Before the introduction of ultrasound, fetal weight was assessed clinically by external palpation of fetal parts and uterine contour. Early studies8–10 showed that 80–85% of clinical estimates are within 500 g of the actual birth weight, and 69% of estimates fall within 10% of the actual birth weight.8 Accuracy is higher in the average range of birth weight, is not related to obstetric training or experience, and is not improved by measurements of uterine diameters with tape or calipers.8–12

Several investigators13,14 have suggested that palpating the uterus to estimate fetal weight is inaccurate. It generally is accepted that the objectivity and reproducibility of sonographic measurements yield more accurate estimates than clinical assessment of birth weight, but only a few studies15–21 have compared both methods.

The present study was undertaken to determine the
accuracy of birth weight estimation by routine antepartum sonogram as compared with clinical examination on admission to the labor and delivery ward.

Patients and Methods

The study population consisted of 1717 women admitted for delivery between January 1990 and April 1995 to the Assaf Harofeh Medical Center, within 1 week of an ultrasonic EFW. Inclusion criteria were 1) singleton pregnancy with unruptured membranes, 2) admission for planned delivery or in early labor (latent phase), and 3) a routine sonographic EFW obtained in the preceding week. More than 80% of ultrasonic evaluations were carried out by level I–II certified sonographers in our facility (a tertiary care center) or in one of several imaging institutes in our region. Because routine ultrasonic reports did not indicate the specific formula used for EFW, this information was not available for each estimate. However, a later investigation revealed that during the study period, all participating institutes used the formulas devised by Shepard et al,5 Hadlock et al,22 or Sabbaghia et al6 for carrying out their routine obstetric sonograms. If more than one report was available from measurements done in the preceding week, the latest estimate was chosen. Clinical EFW was carried out by the attending senior resident, using abdominal palpation, after the admission nurse identified a case that met the inclusion criteria. No special training was undertaken nor was a standardized method used to clinically estimate the birth weight. Neither the prenatal records nor previous ultrasonic estimations were available to the resident before the clinical estimation. Both estimates were entered into the patient’s chart. The actual birth weight was determined by electronic scales and recorded in the chart immediately after delivery. Of 1849 women who met inclusion criteria, 132 (7%) were excluded because of two or more ultrasonic EFW per sonogram or because of incomplete or inconsistent records.

Accuracy of birth weight estimation was determined by calculating the percentage error ([estimated birth weight – actual birth weight] × 100/actual birth weight, by percentage), the absolute percentage error (absolute value [estimated birth weight – actual birth weight] × 100/actual birth weight, by percentage) and the ratio (by percentage) of estimates within 10% of the actual birth weight. Each of these error terms was averaged for each method of estimation in the entire study group and in three strata of birth weights (less than 2500 g, 2500–4000 g, and greater than 4000 g). The percentage error expresses the deviation as a percentage (or g/kg) of the actual birth weight and is comparable across samples or studies. The mean percentage error represents the sum of the positive (overestimation) and negative (underestimation) deviations from the actual birth weight, approximating zero in a method with very low or no systematic error. The Student t test was used to determine if this mean was significantly different from zero (ie, the presence of a systematic error in each method). Differences between both methods in the mean percentage error (ie, the size of the systematic error) were assessed by the paired t test. The standard deviations (or the variances) of the mean percentage errors represent the random error of each method. Because the standard deviations of the mean errors were calculated on the same patients, differences in random errors between both methods were assessed by the comparison of two correlated variances (a modification of the standard F test).23 The mean absolute percentage error is the sum of the absolute deviations (regardless of their direction), reflecting the size of the overall predictive error in terms of the actual birth weight. Because the absolute errors are not normally distributed, the Wilcoxon sign test (nonparametric) was used to test for differences between clinical and ultrasonic estimations. Differences in ratios of estimates that were within 10% of the actual birth weight were assessed by the χ² test. Values are mean ± standard deviation (SD), with P < .05 considered statistically significant.

Results

The actual birth weight in the study population averaged 3334 ± 607 g and ranged between 690 and 5320 g. The distribution of birth weights is given in Table 1. Mean maternal age was 28.6 ± 5.3 years and ranged between 18 and 46 years. Mean gravidity and parity were 1.6 ± 1.8 pregnancies and 1.1 ± 1.3 deliveries, respectively; 41% of the gravidas were nulliparous, and 2.2% had 5 or more prior births. Spontaneous onset of labor occurred in 73% of the women, 17% had induction of labor, and 10% were admitted for cesarean delivery. On admission, vertex presentations were noted in 94% of the pregnancies, whereas breech presentations and transverse lies were observed in 5.5 and 0.5%, respec-

Table 1. Distribution of Birth Weights in the Study Population

<table>
<thead>
<tr>
<th>Birth weight strata (g)</th>
<th>n</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤1500</td>
<td>13</td>
<td>0.8</td>
</tr>
<tr>
<td>1501–2499</td>
<td>121</td>
<td>7.0</td>
</tr>
<tr>
<td>2500–4000</td>
<td>1389</td>
<td>80.9</td>
</tr>
<tr>
<td>4001–4499</td>
<td>155</td>
<td>9.0</td>
</tr>
<tr>
<td>≥4500</td>
<td>39</td>
<td>2.3</td>
</tr>
</tbody>
</table>

...
tively. Gestational age at delivery averaged 39.5 ± 1.9 weeks and ranged between 24 and 43 weeks. Term deliveries (ie, 37–41 weeks) made up 84.6% of the study group, 6.2% were preterm, and 9.2% were post-term deliveries.

In the entire study population (Table 2), only the mean percentage error of the clinical estimation of birth weight was not significantly different from zero, suggesting no systematic error. However, the ultrasonic evaluation underestimated systematically the actual birth weight. In addition, both means of the percentage and absolute percentage errors of the clinical EFW were significantly smaller than those of the ultrasonic EFW (Table 2). There were no significant differences between both methods in the random errors (standard deviations of the mean errors) and in the rate of estimates which were within 10% of the actual birth weight (Table 2). In the low birth weight group (less than 2500 g), both methods overestimated systematically the actual birth weight (Table 2). In this group, all mean errors of the ultrasonic estimation were significantly smaller than those of the clinical estimation, and the proportion of estimates within 10% of actual birth weight was higher for the ultrasonic EFW than for the clinical EFW. Similarly, the random error of the ultrasonic estimation was significantly smaller than that of the clinical estimation. In the middle range of birth weight (2500–4000 g), the clinical estimation had no systematic error, but the ultrasonic method underestimated the actual birth weight (Table 2). In this group, clinical estimation of birth weight resulted in significantly smaller mean errors and higher rates of estimates within 10% of birth weight than the ultrasonic estimation. In addition, the random error of the clinical estimation was significantly smaller than that of the ultrasonic estimation.

In the high birth weight group (greater than 4000 g), both methods underestimated systematically the actual birth weight, but the mean errors were not significantly different and similar proportions of estimates were within 10% of birth weight (Table 2). However, the random error (the SD or the variance of the mean error) of the clinical estimation was significantly smaller than that of the ultrasonic estimation.

Discussion

As fetal weight cannot be measured directly, it must be estimated from fetal or maternal anatomic characteristics. Early methods to predict birth weight with an ultrasound technique were based on measurements of fetal abdominal circumference and biparietal diameter.5,13 The accuracy of EFW was later improved by the incorporation of femur length and head circumference.
into the formulas. Sabbagha et al subsequently reported weight formulas targeted to large-, appropriate-, and small-for-gestational age fetuses. However, despite the great variety of different formulas, no single formula provides uniformly accurate results, particularly when applied to populations that are different from the index population from which they were derived. Also, when applied in large series of fetuses varying widely in their weight range, there appears to be a fixed success and limitation to ultrasonic EFW. Systematic errors (mean percentage error) of ultrasonic estimates in unselected populations usually span ±2% of the actual birth weight, and the random errors (SD of the mean percentage error) vary between 7 and 16% of birth weight. This suggests that the inaccuracies of the ultrasonic estimates are biologic variance primarily and not methodologic bias.

Sonography is widely used for weight estimation because it is objective, reproducible, and involves a well-defined measurement procedure. However, clinical EFW is subjective, poorly defined, and does not require any measurement; but is it less accurate? Only a few studies have compared the accuracy of birth weight estimation by abdominal palpation (ie, clinical) and ultrasonic measurements of fetal dimensions (Table 3). The majority of these studies had relatively small sample sizes and included only term pregnancies with a narrow range of birth weight. The present study encompasses a wide range of gestational ages and birth weights. However, there are some basic differences in design between the present and the earlier studies. In the majority of the previous studies, the ultrasonic and clinical estimates were obtained by a single examiner throughout the study or by the same examiner when more than one investigator was involved. In addition, the clinical estimation preceded the ultrasonic measurements, which raises a concern of potential bias in the latter. In the current study and one other, both estimates were obtained independently by different observers, precluding the possibility that one esti-

<table>
<thead>
<tr>
<th>Reference</th>
<th>No. of patients</th>
<th>Gestational age (wk)</th>
<th>Birth weight (g)</th>
<th>Method of estimation</th>
<th>Mean absolute % error (BW ± 10%)</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patterson 1985</td>
<td>43/62</td>
<td>&gt;37</td>
<td>3330 ± 445</td>
<td>Clinical</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Warsof et al 13</td>
<td>100</td>
<td>&gt;37</td>
<td>2280–4650</td>
<td>Clinical</td>
<td>8.2%</td>
<td>66%</td>
</tr>
<tr>
<td>Raman et al 1992</td>
<td>50</td>
<td>NA</td>
<td>1800–4500</td>
<td>Clinical</td>
<td>7.9%</td>
<td>67%</td>
</tr>
<tr>
<td>Chauhan et al 1992</td>
<td>106</td>
<td>&gt;37</td>
<td>2440–5225</td>
<td>Clinical</td>
<td>9.0%</td>
<td>66%</td>
</tr>
<tr>
<td>Chauhan et al 1993</td>
<td>200</td>
<td>&gt;37</td>
<td>2440–5225</td>
<td>Clinical</td>
<td>9.1%</td>
<td>65%</td>
</tr>
<tr>
<td>Shamley et al 1994</td>
<td>223</td>
<td>35–42</td>
<td>2028–4678</td>
<td>Clinical</td>
<td>8.4%</td>
<td>66%</td>
</tr>
<tr>
<td>Chauhan 1995</td>
<td>602</td>
<td>&gt;37</td>
<td>2302–5225</td>
<td>8 U/S formulas</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Chauhan 1995</td>
<td>67†</td>
<td>&gt;37</td>
<td>NA</td>
<td>Clinical</td>
<td>9.9%</td>
<td>54%</td>
</tr>
<tr>
<td>Present study</td>
<td>1717</td>
<td>24–43</td>
<td>690–5320</td>
<td>3 U/S formulas</td>
<td>7.9%</td>
<td>72%</td>
</tr>
</tbody>
</table>

BW = birth weight; NA = not available; Clin = clinical; U/S = ultrasound.
Clin > U/S—Clinical estimation significantly more accurate than ultrasonic estimations.
Clin ≥ U/S—Clinical estimation significantly more accurate, or not significantly different from ultrasonic estimations.
Clin = U/S—Accuracy of clinical estimation not significantly different from ultrasonic estimations.
Clin ≤ U/S—Clinical estimation significantly less accurate, or not significantly different from ultrasonic estimations.
Clin < U/S—Clinical estimation significantly less accurate than ultrasonic estimations.
* Prediction of macrosomia.
† BW >4 kg.
mate may influence the other. The previous studies15–21 usually involved only one or few estimators in a single institute, whereas the present study involved several institutes and many certified sonographers and senior residents with varying levels of experience. The use of different ultrasonic estimation formulas might pose a major drawback in the present study. Because a large number of ultrasonic estimates were obtained at imaging institutes outside our own facility, we could not control for the ultrasonic method of measurement nor the choice of formula used to estimate fetal weight. Our study population represents a heterogeneous group of ultrasonic methods and formulas for EFW. The inclusion of a number of formulas might be considered to be inappropriate and, in a small sample size, may lead to biased conclusions. However, authors2,4,24 who compared the accuracy of conventionally used formulas suggested that no single formula estimates birth weight more accurately to a significant degree than any other. In addition, the ultrasonic accuracy in our study is similar and compares favorably with that reported by other weight estimation studies (Table 3).1–4,14 in which individual formulas were compared. Our results do not pertain to the accuracy of any specific ultrasonic weight estimation formula but rather they determined the accuracy of routine, noninvestigational ultrasonic EFW as measured by one of several conventional formulas relative to clinical EFW.

Despite differences in study design, our findings and those reported by others15–21 (Table 3) strongly suggest that the accuracy of clinical estimation of birth weight is similar or higher than ultrasonic estimation. In only two15,18 of the seven earlier studies was clinical estimation significantly less accurate than ultrasonic formulas. Shamley and Landon18 noted that the error of clinical estimation was statistically higher than that for the Hadlock et al1 and Shepard et al2 equations, but was comparable to the other two equations studied (Sabbagh et al3 and Rose and McCallum14). However, for the subset of patients in labor, both the formula by Hadlock et al1 and clinical examination “proved to be most accurate.” Patterson et al19 also noted that clinical estimation is less accurate than the Campbell formula, but is comparable to the Warsof et al formula.13 Both formulas were more accurate than clinical estimation in the presence of oligohydramnios or engagement of the fetal vertex. The studies of Raman et al17 and Chauhan et al21 showed that clinical estimation was significantly more accurate than the sonographic prediction. This also was demonstrated in the present study among fetuses with birth weight between 2500 and 4000 g. In this range of birth weight, clinical estimation was more accurate and showed no systematic error. This may be explained by the prior knowledge and expectations of the human observer, which play a major role in “clinical” estimation but only a minor role in ultrasonographic measurement. For example, in a population with a mean (±SD) birth weight of 3300 ± 500 g, approximately 50% of birth weight will be within ±10% of the mean birth weight (ie, 3000–3600 g). This represents the probability that a variate drawn at random from the standard normal distribution has a value lying in the interval mean ± 0.66 of the SD.25 In other words, by knowing the mean birth weight, a relatively accurate estimate of birth weight may be obtained 50% of the time, without even examining the patient. This together with the assumption of a term pregnancy and comparison of the uterine size to the “average term uterus,” an even higher degree of accuracy may be obtained. This is compatible with our findings that approximately 70% of clinical estimates were within ±10% of the actual birth weight in the present study. Another explanation for the somewhat lower accuracy and systematic underestimation of birth weight by the ultrasonic method is the inclusion of estimates that were obtained up to 1 week before delivery. Some studies4,26 suggest that correction for fetal weight gain (28 g/day) taking place during the interval between the scan and delivery may yield more accurate results. However, most ultrasonic weight estimation formulas originally were derived on fetuses scanned within 1 week of delivery.1,3,4,14,17,22

Both methods of EFW demonstrate diminished accuracy at both extremes of fetal weight, ie, the very large or very small fetuses.3,16 Existing formulas for ultrasonic estimation generally underestimate the weight of the macrosomic fetus, and there is a tendency toward overestimation in cases of low birth weight.1,2,4 This pattern was observed consistently by studies9–12 of clinical estimation and in the present study by both methods of weight estimation. Our findings and those of others16,17 also suggest that both methods have similar accuracy in macrosomic (greater than 4000 g) fetuses. However, the study of Chauhan et al21 showed that the prediction of macrosomia and the accuracy of clinical EFW among macrosomic fetuses are significantly better than or similar to sonographic equations. Finally, our data demonstrate that in the low range of birth weight (less than 2500 g), ultrasonic estimation is significantly more accurate than clinical estimation. This finding had not been demonstrated in earlier studies,15–21 which focused entirely on term pregnancies.

There is clearly a role for clinical estimation of birth weight as a diagnostic tool. The present study suggests that a clinical estimate of birth weight is sufficient to manage labor and delivery in a term or near-term pregnancy. Even in estimating the weight of the macrosomic fetus for making decisions regarding trial of
labor, there is no benefit in obtaining a routine sonographic birth weight, while targeted ultrasound or formulas focusing on fetal body dimensions and proportions may be useful in these cases. However, our study suggests that ultrasound estimation is more accurate in preterm gestations or when restricted fetal growth is suspected. Although we propose that clinical estimation can be a reliable tool, like any diagnostic method, it has limitations. The efficacy of the clinical method needs to be determined in situations that may alter birth weight assessment, such as premature rupture of the membranes, or in women with atypical physical characteristics (ie, in women who are obese or especially tall or small). Finally, several reports have suggested that the accuracy of clinical estimation is not related to the experience of the examiner. This aspect was beyond the scope of the present study, and the minimum of experience necessary to make an appropriate estimate has yet to be determined.

References


Address reprint requests to: Dan J. Sherman, MD
Department of Obstetrics and Gynecology
Assaf-Harofeh Medical Center
Zerifin 70300
Israel

Received June 10, 1997.
Received in revised form October 7, 1997.
Accepted October 22, 1997.

Copyright © 1998 by The American College of Obstetricians and Gynecologists. Published by Elsevier Science Inc.