Original Article

Sonographic fetal weight estimation prior to delivery: analysis of accuracy and interobserver variability between obstetrician and radiologist

Hüseyin Aksoy1, Özkan Özdamar2, Ülkü Aksoy3, Fulya Çağlı4, Özge Idem Karadağ5, Gökhan Açmaz4, Mustafa Alparslan Babayiğit6, Turgut Aydın5

1Department of Obstetrics and Gynecology, Kayseri Military Hospital, Kayseri, Turkey; 2Department of Obstetrics and Gynecology, Istanbul Medeniyet University, Faculty of Medicine, Istanbul, Turkey; 3Department of Obstetrics and Gynecology, Kayseri Memorial Hospital, Kayseri, Turkey; 4Department of Obstetrics and Gynecology, Kayseri Education and Research Hospital, Kayseri, Turkey; 5Department of Obstetrics and Gynecology, Kayseri Acıbadem Hospital, Kayseri, Turkey; 6Department of Public Health and Epidemiology, Gülhane Military Faculty of Medicine, Ankara, Turkey

Received October 19, 2015; Accepted March 15, 2016; Epub June 15, 2016; Published June 30, 2016

Abstract: The aim of this study was to evaluate the interobserver reliability of sonographic fetal weight estimation between experienced obstetrician and radiologist immediately prior to delivery and to compare the sonographic accuracies of estimated fetal weights (EFW) between the sonographers. This was a prospective study of singleton pregnancies that underwent sonographic fetal weight estimation prior to scheduled delivery. One obstetrician and one radiologist, performed all examinations prior to delivery. Sonographic fetal weight estimations were obtained from each ultrasonographic examination. EFWs were compared with the actual birth weight (ABW) and interobserver reliability of sonographic fetal weight estimation between two sonographers was assessed. A total of 245 subjects were included. Inter-observer reliability coefficient for EFW was 0.864 (95% confidence interval: 0.825-0.894, P < 0.001). Mean values of EFWs by radiologist and obstetrician did not significantly differ (P = 0.480) while both did significantly differ from the ABW (P < 0.001, for both comparisons). The mean absolute error for EFW obtained by radiologist and obstetrician was 227 (± 160) and 197 (± 138), respectively (P < 0.01). The mean absolute percentage error for EFW obtained by radiologist and obstetrician was 7.09 (± 5.17) and 6.15 (± 4.38), respectively (P < 0.01). EFWs measured by radiologists and obstetricians are similar but both are significantly distant from the ABW. However, obstetricians measure EFW significantly more accurate than the radiologists do.

Keywords: Interobserver reliability, fetal weight estimation, ultrasonography

Introduction

Sonographic estimation of the fetal weight (EFW) is an integral part of routine antenatal care and has a pivotal role in obstetric practice with a major impact on antenatal management [1]. The optimal practice of obstetrical care requires an optimal estimation of fetal weight. Many important clinical decisions depend upon a precise and accurate assessment of sonographic fetal weight estimation. The accuracy of EFW is also crucial for optimal delivery management and allows obstetrician to make management decisions [1]. Sonographic EFW prior to delivery is part of the daily routine obstetric practice worldwide; however, sonographic EFW has some limitations due to the subjective elements of this measurement [2, 3]. Various clinical factors such as maternal weight, amniotic fluid volume, fetal sex, multiple pregnancy, the presence of gestational diabetes mellitus, time interval from ultrasound examination to delivery, fetal size and gestational age have also potential to affect the accuracy of sonographic EFW [2, 4-8]. The majority of literature on this topic has focused on the maternal and fetal component in the accuracy of sonographic EFW. Technical factors related to the experience of the sonographer may further differentially impact the accuracy of sonographic EFW.
Accuracy and variability of fetal weight estimation

Technical factors regarding the operator himself have not been well investigated. There are quite few studies, which investigate the effect of operator-dependent subjective factors on accuracy of sonographic EFW [7, 9-12].

Although widely used in routine obstetric practice and numerous weight formulas have been suggested to compute the sonographic EFW, the accuracy of fetal weight estimation has not been achieved precisely [1, 13-16]. Despite these efforts to improve the accuracy of fetal weight prediction, ultrasonography provides weight estimates with errors of up to 15% when compared with actual birth weights [15, 17-20]. The operator-dependent subjective components of the method are the main cause of this high error in predicting fetal weight. Moreover, prenatal estimation of fetal weight is performed by either obstetricians or radiologists in many parts of Turkey, depending on the practices of health care provider, however, it has not been investigated whether these examinations do vary with regards to accuracy and reliability.

In spite of a long history of utilizing ultrasound to estimate fetal weight and the potential implications of these estimates, it is surprising that few interoperator variability studies have been published. Ultrasound examination is routinely performed in Turkey for all pregnancies prior to delivery. The aim of this study was to assess the reliability in reproducing sonographic EFW between radiologist and obstetricians prior to delivery and to compare the sonographic EFW accuracies of both operators.

Materials and methods

A prospective clinical study was conducted between December 2013 and September 2014 in Kayseri Education and Research Hospital of Medicine, a tertiary center where approximately 7,200 women deliver each year. The study was approved by the institutional ethics committee and all participants signed an informed consent form regarding to participate in the present study.

Singleton pregnant women who had admitted to our department for scheduled labor induction or scheduled elective cesarean section enrolled to the study. Two hundred and forty-five singleton pregnant women who fulfilled the inclusion criteria were included in the study. Inclusion criteria include singleton third-trimester pregnancies with cephalic presentation and admission for scheduled delivery (vaginal or elective cesarean delivery). All pregnant participants were between 37-42th weeks of gestation with singleton cephalic presentation and none had any medical or obstetrical problems.

Our exclusion criteria were as follows: multiple pregnancies, abnormal presentations (breech, transverse and other noncephalic presentations), oligohydramnios, anhydramnios, congenital fetal anomalies, hydrops, intrauterine fetal deaths, active stage of labor, unstable patients, and obstetric emergencies such as antepartum hemorrhage, eclampsia and acute fetal distress. Moreover, patients who did not deliver within 48 hours after the sonographic fetal weight estimation were not included into study. After informed consent, the participant completed an enrollment questionnaire assessing sociodemographic characteristics and medical information. The final study group was composed of 245 subjects.

Gestational age was calculated according to last menstrual period and confirmed by in all cases by crown-rump length measurements done in the first trimester. Two experienced sonographers (one experienced radiologist and one experienced obstetrician with average knowledge in obstetric ultrasound) performed all examinations on the day of the patient’s admission to clinic. Each measurement was performed once by each operator. The initial measurement was recorded by the first sonographer who is an experienced obstetrician (F.C). Subsequently, a second sonographer who is an experienced radiologist (S.T), blinded to the results of the first sonographer, performed the same measurements. The fetal biometric measurements and EFW of the first sonographer were always removed from the ultrasound screen, after a hard copy had been made, before the second sonographer entered the room. The operators were not allowed to present in the ultrasound room during each other’s examinations to remove any possible influence by the second operator when generating the image and measuring the fetal biometry. Each operator was blinded to any pre-existing measurements. All ultrasound examinations were performed in a single room. Fetal biometrical measurements (biparietal diameter, head circumfe-
Accuracy and variability of fetal weight estimation

The accuracy and variability of fetal weight estimation were evaluated in a study of 245 pregnant women. Fetal biometries and EFW were obtained by each ultrasonic examination. Fetal biometries and EFW calculations were performed by using the formula described by Hadlock [15, 21]. For BPD, a transverse section of the head was used with both the lateral ventricles symmetrical in view with a horizontal midline. The fetal biometry measurements were obtained at the appropriate levels described elsewhere. Ultrasonographic fetal weight (EFW), in grams, was estimated with Hadlock formula using a combination of the BPD, AC, and FL using a formula that combines BPD, AC and FL [15]. In all cases, the sonographic fetal weight estimation was done 1-2 days before delivery to eliminate the possible impact of the time elapsed between ultrasound examination and delivery on the accuracy of the EFW. Actual birth weight was measured within 30 min after delivery by an experienced midwife on a weighing scale, which was standardized by monthly calibration.

Ultrasound examinations were performed transabdominally using a commercially available ultrasound system, Toshiba Xario (Toshiba Medical Systems Corporation, Japan) equipped with a 4-7 MHz curved, high frequency, and curved array transducer for all attendants. EFW was compared with the actual birth weight (ABW) and the difference between the EFW and the ABW (simple error) was recorded as error in grams. Interobserver reliability of sonographic fetal weight estimation between two sonographers was assessed. Estimated fetal weight (EFW) was compared with the actual birth weight (ABW) with respect to:

- mean of absolute error (MAE) = absolute value of [EFW-ABW],
- mean of absolute percentage error (MAPE) = absolute value of [EFW-ABW] x 100/ABW.

Statistical analysis

Continuous variables were expressed as mean ± standard deviation and median (minimum-maximum) where appropriate. Kolmogorov-Smirnov Goodness of Fit test was used to test the distribution of data. For normally distributed data, paired samples t-test was used, whereas Wilcoxon signed ranks test was used to compare the non-parametric continuous variables. Inter-observer comparisons were done by reliability tests (Cronbach alpha and intraclass correlation coefficients). Two-way mixed effects model where people effects are random and measures effects are fixed (Absolute Agreement Definition) was used. The accuracy of actual and sonographic fetal weight estimation were compared using Wilcoxon Signed Ranks Test and Pearson's correlation coefficient test for BMI of participants. All the analysis were done by Statistical Package for Social Sciences version 15.0 (SPSS Inc., Chicago, IL, USA). Two-tailed p value less than 0.05 was accepted to be statistically significant.

Results

A total of 468 patients who admitted to our department for scheduled delivery were evaluated for eligibility and 223 patients were excluded, of which 17 refused to participate in the study and 206 did not meet the inclusion criteria. The final study group was composed of 245 subjects.

Some demographic and clinical characteristics of patients are illustrated in Table 1. The mean of fetal weight estimated by radiologist and obstetrician were 3356 ± 375 g and 3344 ± 365 g, respectively. The mean actual birth weight was 3266 ± 366 g. Significant differences were found with respect to the comparison of EFW obtained by each operator and ABW (P < 0.001, for both observer). Mean EFW

Table 1. Demographic characteristics of the patients included in the study (n = 245)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Mean ± SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, y</td>
<td>25.84 ± 5.46</td>
</tr>
<tr>
<td>BMI, kg/m²</td>
<td>28.29 ± 5.23</td>
</tr>
<tr>
<td>Gravidity, n</td>
<td>2.59 ± 1.35</td>
</tr>
<tr>
<td>Parity, n</td>
<td>1.21 ± 1.01</td>
</tr>
<tr>
<td>Gestational age, w</td>
<td>39.59 ± 1.23</td>
</tr>
</tbody>
</table>

All data were expressed as mean ± standard deviation (SD). BMI, body mass index; y, years; n, numbers; w, weeks.
Accuracy and variability of fetal weight estimation

Table 2. Descriptive statistics for EFW by observers and ABW and inter-observer reliability coefficients for EFW (n = 245)

<table>
<thead>
<tr>
<th></th>
<th>Radiologist</th>
<th>Obstetrician</th>
<th>P*</th>
<th>Inter-observer reliability coefficients</th>
<th>95% CI</th>
<th>P**</th>
</tr>
</thead>
<tbody>
<tr>
<td>EFW-(Mean ± SD)</td>
<td>3356.20 ± 375.13</td>
<td>3344.59 ± 365.97</td>
<td>0.480</td>
<td>0.864</td>
<td>0.825</td>
<td>0.894</td>
</tr>
<tr>
<td>EFW-Med (Min-Max)</td>
<td>3365 (2083-4391)</td>
<td>3356 (2345-4410)</td>
<td>0.660</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Actual Birth Weight ± SD</td>
<td>3266.64 ± 366.48</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p*</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


Table 3. Comparisons of observers for mean absolute and percentage errors (n = 245)

<table>
<thead>
<tr>
<th></th>
<th>Radiologist Mean ± SD</th>
<th>Obstetrician Mean ± SD</th>
<th>P*</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAE</td>
<td>227.32 ± 160.98</td>
<td>197.51 ± 138.54</td>
<td>&lt; 0.01*</td>
</tr>
<tr>
<td>MAPE</td>
<td>7.09 ± 5.17</td>
<td>6.15 ± 4.38</td>
<td>&lt; 0.01*</td>
</tr>
</tbody>
</table>


obtained by each operator and ABW of the participants are presented in Table 2. There was no significant difference between mean EFW obtained by radiologist and obstetrician.

Inter-observer reliability coefficients with 95% confidence intervals for sonographic EFW was 0.864 (0.825-0.894, P < 0.001). The reliability coefficient for sonographic fetal weight estimation for interclass observations was summarized in Table 2. Interobserver reliability and agreement in sonographic EFW measurement was high.

Table 3 demonstrates the median absolute and median absolute percentage errors of sonographic EFW for two operators. Significant difference was found between operators regarding absolute errors and absolute percentage errors (P < 0.01). The median absolute error for sonographic EFW obtained by radiologist and obstetrician was 192 (5-862) and 160 (0-590) g, respectively (P < 0.01). The median absolute percentage error for sonographic EFW obtained by radiologist and obstetrician was 5.7 (0.1-26.9) and 5.0 (0-19.4), respectively (P < 0.01). There was significant difference in the magnitude of the median absolute errors and absolute percentage errors between two operators.

Discussion

Prediction of fetal birth-weight by ultrasound in third-trimester is integral to the antenatal care and the accuracy of EFW is also crucial for optimal delivery management [1]. In this present study, we evaluated the interobserver reliability of sonographic fetal weight estimation between two experience sonographers in the third-trimester shortly before delivery. We also compared the sonographic EFW obtained by both sonographers with the ABW using the absolute error and percentage error data. Our study demonstrated that the interobserver reliability and agreement in sonographic EFW measurement was high. Despite this high interobserver agreement for measurement of EFW; absolute error and absolute percentage error between two operators were found significantly different. In parallel with high interobserver agreement, clinical evaluation of this significant difference between operators regarding absolute error and absolute percentage error makes this finding clinically insignificant.

Previous studies have mostly focused on the impact of different weight equations and various clinical factors on the accuracy of estimated fetal weight [2, 4-8, 18, 19, 22]. Despite the clinical importance of sonographic EFW in obstetrical care and delivery management, few interoperator variability studies have been published [9-12]. In 1993, the first in this issue was the study by Chang [12], which was a prospective study to assess intra and interobserver variability in a series of 40 third-trimester pregnancies. In that study, EFWs were obtained by two experienced sonographers and it was demonstrated that measurements of sonographic EFW were reproducible. They reported interobserver differences of less than 85 g, or approximately 3.5%, in EFW [12]. Our results were in agreement with the prior report of Chang [12]. The median absolute percentage error for sonographic EFW obtained by radiologist and obstetrician in our study was 5.7 (0.1-26.9) and 5.0 (0-19.4), respectively (P < 0.01). This finding was also comparable with mean absolute
percentage errors of 6.1 ± 4.2, 5.9 ± 3.8 and 6.3 ± 5.1% reported in a previous study [23]. In that study, the investigators analyzed the biometric measurements of 39 fetuses at term, which were independently measured by three experienced sonographers [23]. They investigated the contribution of additional examiners to prediction of fetal weight by ultrasound and it was found that averaging the measurement results of two sonographers reduced the mean absolute error in EFW by approximately 17% (from 6.1% to 5.1%). The overall accuracy of sonographic EFW, in terms of median absolute error, obtained by two operators in our study was 192 (5-862) and 160 (0-590) g, respectively and was consistent with a previously published report, in which 840 women undergoing ultrasound assessment for EFW shortly before induction of labor over a 2-year period had been evaluated and a mean absolute error of 227 ± 197 was reported [24].

Several studies compared the accuracy of different sonographic models for fetal weight estimation [22, 23]. In a retrospective cohort study, 26 different models using 3705 sonographic weight estimations performed less than 3 days before delivery were analyzed [22]. They found that mean absolute percentage error for EFW, obtained by using the combination of AC, FL and BPD weight equations, was 6.6%. This finding was also consistent with our study results. In 2003, the accuracies of 25 ultrasonic algorithms for the estimation of term fetal weight in 82 nondiabetic pregnant women at 35 to 42 weeks of gestation were investigated [24]. Mean absolute percentage error for EFW obtained by using the combination of AC, FL and BPD weight equations was reported to be approximately 8.1%. As compared to the results of the study by Nahum [26], the mean absolute percentage error for EFW in our study is lower. The main discrepancy between the results of two studies may be related to differences in the study protocols. Nahum [26] reviewed the medical records retrospectively and the mean interval between sonographic assessment and delivery was 8.2 ± 6.4. In our study, the sonographic assessment was done 1-2 days prior to delivery to eliminate the possible impact of the time elapsed between ultrasound examination and delivery on the accuracy of the EFW.

Precise estimation of fetal weight is critical in determining adequate and accurate obstetrical care to pregnant women. Accurate prediction of birth weight prior to delivery is a valuable tool for determining further obstetric and delivery
Accuracy and variability of fetal weight estimation

management. Inaccurate measurements shortly before the delivery adversely affects delivery management performance and may lead to inappropriate or untimely intervention and perinatal compromise. Despite considerable technical advances in ultrasound equipment have been achieved and numerous weight formulas have been suggested over the past 30 years, there is still a lack of accuracy in sonographic fetal weight estimation. The importance of ascertaining the interobserver reliability of sonographic EFW is necessary to further improve the accuracy of sonographic EFW during pregnancy. Despite the widespread use of ultrasound to estimate fetal weight and the possible implications of these estimates, there is a relative paucity of literature evaluating interobserver reliability of sonographic EFW.

We suggest that sonographic EFW in third-trimester shortly before the delivery is highly reliable between the radiologists and obstetricians. EFWs measured by radiologists and obstetricians are similar but both are significantly distant from the actual birth weight. However, obstetricians measure EFW significantly more accurate than the radiologists do. Further studies are needed to clarify the reasons of the discrepancy between the two measurements and to guide clinicians with respect to sonographic estimation of fetal weight.

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Özkan Özdamar, Department of Obstetrics and Gynecology, Faculty of Medicine, Istanbul Medeniyet University, Istanbul 34750, Turkey. Tel: +90532 4766744; Fax: +90 262 414 11 45; E-mail: ozkan_ozdamar35@hotmail.com

References

Accuracy and variability of fetal weight estimation


