The Accuracy of Sonographic Assessment for Fetal Weight: Technicians versus Ultrasound-Certified Physicians

Amihai Rottenstreich MD, Nili Yanai MD, Simcha Yagel MD and Shay Porat MD PhD

Department of Obstetrics and Gynecology, Hadassah-Hebrew University Medical Center, Jerusalem, Israel

ABSTRACT: **Background:** Sonographic estimation of birth weight may differ among evaluators due to its operator-dependent nature.

Objectives: To compare the accuracy of estimation of fetal birth weight by sonography between ultrasound-certified physicians and registered diagnostic medical technicians.

Methods: The authors reviewed ultrasound examinations that had been performed by either technicians or ultrasound-certified obstetricians between 2010 and 2017, and within 2 days of delivery. Inclusion criteria were: singleton viable pregnancy, details of four ultrasound measurements (abdominal circumference, bi-parietal diameter, head circumference, and femur length), and known birth weight. The estimated fetal weight (EFW) was calculated according to the Hadlock formula, incorporating the four ultrasound measurements. The mean percentage error (MPE) was calculated by the formula: (EFW-birth weight) x100 / birth weight.

Results: Technicians performed 9741 examinations and physicians performed 352 examinations. The proportion of macrosomic neonates was similar in both groups. Technicians were more accurate than physicians in terms of the MPE, absolute MPE, proportion of estimates that fell within ± 10% of birth weight, and Euclidean distance ($P < 0.0001$ for all comparisons). They were also more accurate in terms of sensitivity, specificity, positive predictive value, negative predictive value, and area under the receiver operating curve. Furthermore, for fetuses weighing more than 4000 grams the technicians had a lower total false prediction rate.

Conclusions: Medical technicians in our institute performed better than physicians in estimating fetal weight. Further studies are warranted to confirm our findings and better delineate the role of repeat physician's examination after an initial estimation by an experienced technician.

IMAJ 2019; 21: 801–805

KEY WORDS: birth weight, sonographic technicians, ultrasound-certified physicians

For editorial see page 831

The antenatal assessment of fetal weight is a critical component of prenatal care as abnormalities in growth are associated with perinatal morbidity and mortality [1]. The evaluation is imperative for decision-making regarding antenatal management, as well as delivery mode and timing. Recognition of both excessive and impaired growth is essential. Fetal growth restriction is the single strongest risk factor for intrauterine demise, which necessitates increased surveillance with the goal of decreasing perinatal mortality [2,3]. Macrosomia is a risk factor for both the fetus (shoulder dystocia, unplanned cesarean delivery, and neonatal complications) [4] and the mother (intrapartum fever and birth trauma including obstetrical anal sphincter injuries).

Although fetal weight can be assessed clinically, sonographic estimation based on biometric measurements and regression equations is preferred for its ease, objectivity, and reproducibility [5,6]. Nevertheless, current evidence indicates significant errors in the sonographic estimation of birth weight due to the operator-dependent nature of this method (e.g., inter-examiner variation) [7]. This assumption led us to hypothesize that birth weight predictions will differ substantially depending on who performed the ultrasound examination. Therefore, we compared the diagnostic accuracies in predicting birth weight between ultrasound-certified physicians and registered diagnostic medical technicians.

PATIENTS AND METHODS

PATIENTS

This study was approved by the local institutional review board of the Hadassah Medical Center Helsinki Committee (IRB approval number No. HMO 0128-17). Institutional review board approval waiving informed consent was obtained for this retrospective study.

Our center is a large tertiary care university hospital with over 10,000 births per year. For the purpose of this study, we reviewed all consecutive deliveries between 2010 and 2017. We included neonates who had a sonographic estimation of fetal weight (EFW) performed by a single experienced examiner (either a technician or a physician) within 2 days prior to delivery and who were delivered at our institution. Exclusion criteria were multifetal gestation, intrauterine fetal demise, missing ultrasound or birth weight data, and congenital anomalies not allowing for accurate assessment of fetal biometric parameters. The examinations were performed by nine registered diagnostic medical technicians and six ultrasound-certified physicians, all of whom work in our medical center. All technicians and physicians included in this study had extensive experience and performed biometric ultrasound evaluations on a daily basis for at least 10 years.

DATA COLLECTION

The ultrasound database was cross-referenced with the births database of all women who delivered within 2 days of the ultrasound examination during the study period. The following data were extracted: patient characteristics, sonographic measurements of biparietal diameter (BPD), head circumference (HC), abdominal circumference (AC), femur length (FL), neonatal sex, gestational age (GA) at delivery, birth weight, and the calculation of the standard estimated sonographic fetal weight using the Hadlock formula, which incorporates the fetal BPD, HC, AC, and FL [8]. In all cases, first-trimester sonographic examination was available. GA was based on the last menstrual period (LMP); however, if first-trimester ultrasound dating differed by more than 7 days from the LMP dating, the GA was changed to correspond with the ultrasound dating. All sonographic examinations were obtained on one of the following ultrasound machines: Voluson E6, Voluson E8, Voluson E10, Voluson 730 (GE Healthcare, WI, USA) or Accuvix A30 (Samsung Medison, Republic of Korea), with standard techniques.

To compare the accuracy of predicting neonatal birth weight between the technicians and physicians, the systematic error, absolute systematic error, random error, absolute random error of the calculated versus birth weight, percentage of predictions within 10% of birth weight, and the Euclidean distance were compared between the two groups. The mean percentage error (MPE), expressed as the percentage of error out of the birth weight (BW) was calculated by the formula (EFW-BW) x100/BW, which represents the systematic error. The absolute form of the MPE represents the distance of the estimation from the true value, regardless of whether the birth weight was higher or lower than the estimation. The random error is the standard deviation (SD) of the MPE (MPE SD). The Euclidean distance, a recently described measure of accuracy [9], was calculated using the formula $\sqrt{(MPE^2 + MPESD^2}$. The same calculations were performed after stratifying by birth weight: ≤ 2500 grams, > 2500 grams to < 4000 grams, and ≥ 4000 grams.

To characterize the overall accuracy of the two groups of examiners in predicting low-birth weight (≤ 2500 grams) and macrosomia (≥ 4000 grams) sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), overall accuracy [defined as (true negative + true positive cases) / all cases], and the area under the receiver operating characteristic (ROC) curve (AUC) were calculated. In addition, for fetuses with birth weight ≥ 4000 grams, the total false prediction rate, defined as the arithmetic sum of false positive and false negative prediction rates, was calculated. The false positive prediction rate was defined as the false positive rate multiplied by the true negative proportion of the cohort. The false negative prediction rate was defined as the false negative rate multiplied by the true positive proportion of the cohort.

STATISTICAL ANALYSIS

Patient characteristics are described as proportions for categorical variables, and as medians and interquartile ranges for continuous variables without a normal distribution. Sensitivity, specificity, PPV, NPV, and overall accuracy were determined. Statistical significance between groups was assessed by Chi-square and Fisher's exact test for categorical variables, and the Mann–Whitney U test for continuous variables. Correlations for the number of examinations performed and the percentage of predictions within 10% of birth weight were calculated using the Pearson test with the correspondent r and *P* values. A *P* value < 0.05 was considered statistically significant. Statistical analyses were performed using IBM Statistical Package for the Social Sciences statistics software, version 22 (SPSS, IBM Corp, Armonk, NY, USA).

RESULTS

The dataset comprised 10,093 ultrasound examinations, of which 9741 were performed by registered diagnostic medical technicians and 352 by physicians. Pregnancy and patient characteristics are summarized in Table 1. In the entire cohort,

Table 1. Baseline characteristics of patients

Characteristics	Technicians n=9741	Physicians n=352	*P* value
Maternal age (years) (%)			0.03
15–20	197 (2.0)	5 (1.4)	
20–30	4502 (46.2)	142 (40.3)	
30–40	4553 (46.7)	178 (50.6)	
› 40	489 (5.0)	27 (7.7)	
Parity, median [interquartile range]	1 [0–15]	2 [0–15]	0.52
Nulliparous (%)	2885 (29.6)	101 (28.7)	0.71
Previous cesarean section (%)			0.14
0	7571 (77.7)	283 (80.4)	
1	1360 (14.0)	50 (14.2)	
2	523 (5.4)	9 (2.6)	
≥ 3	287 (2.9)	10 (2.8)	
Diabetes (gestational and pre-gestational) (%)	134 (1.4)	6 (1.7)	0.60
Gestational age (weeks) (%)			‹ 0.001
‹ 28	13 (0.1)	3 (0.9)	
$28^{0/6}$–$32^{6/6}$	50 (0.5)	21 (6.0)	
$33^{0/6}$–$36^{6/6}$	566 (5.8)	77 (21.9)	
$37^{0/6}$–$39^{6/6}$	5381 (55.2)	199 (56.5)	
≥ 40	3731 (38.3)	52 (14.8)	
Ultrasound to delivery interval (days) median [interquartile range]	0.80 [0-2.0]	0.82 [0.30-1.97]	0.90
Birth weight ≥ 4000	640 (6.6)	25 (7.1)	0.69
Male fetus (%)	5082 (52.2)	181 (51.4)	0.78
Presentation (%)			‹ 0.001
Vertex	9223 (94.7)	310 (88.1)	
Breech	469 (4.8)	38 (10.8)	
Transverse/oblique	49 (0.5)	4 (1.1)	

Continuous variables are expressed as median [interquartile range]

Table 2. Accuracy of estimated weight determination in relation to the examiner and final birth weight

	All examinations (n=10,093)			Birth weight ≤ 2500 grams (n=750)			Birth weight > 2500 to < 4000 grams (n=8678)			Birth weight ≥ 4000 grams (n=665)		
	Technicians n=9741	Physicians n=352	*P* value	Technicians n=616	Physicians n=134	*P* value	Technicians n=8485	Physicians n=193	*P* value	Technicians n=640	Physicians n=25	*P* value
MPE	-2.29	-3.90	< 0.001	-7.75	-5.90	0.03	-2.40	-3.31	0.08	4.38	2.16	0.08
MPE SD	7.49	8.41	0.015	9.02	8.74	0.59	7.07	7.94	0.14	6.29	6.61	0.53
Absolute MPE	6.11	7.33	< 0.001	9.26	8.59	0.33	5.87	6.67	0.04	6.16	5.75	0.66
Absolute MPE SD	4.90	5.67	< 0.001	7.45	6.10	0.07	4.61	5.42	0.001	4.56	3.74	0.13
Percent within ≤ 10% absolute percentage error	7910 (81.2%)	251 (71.3%)	< 0.001	373 (60.6%)	85 (63.4%)	0.54	7020 (82.7%)	145 (75.1%)	0.006	517 (80.8%)	21 (84.0%)	1.0
Euclidean distance	7.83	9.27		11.89	10.54		7.47	8.60		7.66	6.95	

MPE = mean percentage error, SD = standard deviation

the percentage of predictions within 10% of birth weight was similar regardless of fetal presentation (81.0% in vertex-presenting fetuses vs. 79.3% in non-vertex presenting fetuses, $P = 0.35$).

When the entire cohort was assessed, MPE, MPE-SD, absolute MPE, and absolute MPE-SD were all significantly lower for the technicians compared to the physicians, with higher proportions of examinations within 10% of birth weight [Table 2]. The Euclidean distance for the entire cohort was also better for the technicians than the physicians group (7.83 vs. 9.87). For fetuses weighing ≤ 2500 grams (n=750), the physicians had a significantly lower MPE and Euclidean distance. However, the percentage of predictions within 10% of birth weight was similar between the physicians and technicians. For the fetuses weighing between 2500 grams and 4000 grams (n=8678), which constituted the majority of the cohort (86.0%), the technicians had a significantly lower absolute MPE and absolute MPE SD and higher proportion of estimates within 10% of birth weight. The Euclidean distance was also lower for the technicians than physicians regarding this subset of fetuses. For the fetuses weighing ≥ 4,000 grams, the groups did not differ in any of the compared variables, except for the Euclidean distance, which was slightly better for the physicians than the technicians [Table 2]. Among the technician group, the number of examination performed was directly correlated with the percentage of predictions within 10% of birth weight ($r = 0.60$, $P = 0.02$).

The sensitivity, specificity, PPV, NPV, overall accuracy, and AUC are presented in Table 3 for the birth weight categories of under 2500 grams, 2500 to 4000 grams, and over 4000 grams, according to the physician and technician groups. Compared to the physicians, the technicians had lower sensitivities and higher specificities for the extreme birth weight of ≤ 2500 grams and ≥ 4000 grams. For the fetuses with birth weight between those extremes, the technicians had higher sensitivity but much lower specificity.

The total false prediction rate (the arithmetic sum of false positive and false negative prediction rates) was calculated for fetuses with birth weight ≥ 4000 grams for each group of examiners and is shown in Table 4. The proportion of fetuses with birth weight ≥ 4000 grams was 6.6% of the entire cohort (true positives), while the rest of the cohort (93.4%) were below this weight. In total, the technician group had almost 2% less erroneous predictions for the 4000 gram cutoff, which comes to 193 cases of the entire cohort.

Table 3. Detection of low birth weight and macrosomic infants

	Birth weight ≤ 2500 grams (n=750)		Birth weight > 2500 to < 4000 grams (n=8,678)		Birth weight ≥ 4000 grams (n=665)	
	Technicians n=616	Physicians n=134	Technicians n=8485	Physicians n=193	Technicians n=640	Physicians n=25
Sensitivity (%)	64.0	80.6	96.5	88.6	53.8	56.0
Specificity (%)	99.4	97.2	59.0	76.7	97.3	95.1
PPV (%)	88.1	94.7	94.1	82.2	58.6	46.7
NPV (%)	99.4	89.1	71.2	84.7	96.8	96.6
Accuracy (%)	97.2	90.9	91.6	83.2	94.5	92.3
AUC	0.981	0.976	0.509	0.831	0.947	0.960

AUC = area under the curve, NPV = negative predictive value, PPV = positive predictive value

Table 4. False prediction rate analysis of birth weight ≥ 4000 grams

	False positive prediction rate	False negative prediction rate	Total false prediction rate
Technicians	2.52%	3.05%	5.57%
Physicians	4.58%	2.90%	7.48%

DISCUSSION

This study addressed the diagnostic accuracies of ultrasound-certified physicians compared to registered diagnostic medical technicians in predicting birth weight. The accuracy of predicting birth weight was significantly higher among technicians compared to certified physicians, as evident by lower values of error measures for the technician group and a higher proportion of examinations within 10% of birth weight.

Accurate prediction of fetal weight is an integral part of obstetric practice as it can provide useful information regarding antenatal management, labor methods, and mode of delivery for the prevention of maternal and perinatal complications. Examiner experience is important in achieving accurate fetal weight estimation, as the most common sources of inaccuracy of sonographic EFW were reported to be operator dependent [7]. This was previously demonstrated by the learning curve among residents for whom the best performance of sonographic weight estimation was achieved after at least 2 years of training [10]. However, even among experienced examiners, inter-observer differences were shown [11], with 50% of the random error in fetal weight estimation attributed to observer variability [12]. In our study we observed differences between groups even though all fetal weight estimations were performed by experienced examiners.

Overall, when the entire cohort was assessed, technicians performed better than physicians in terms of the degree of systematic and random error as well as the proportion of examinations within 10% of birth weight. This finding is in accordance with a previous report that included 365 patients who had their fetal biometric measurements taken by technicians, and immediately followed by physicians [13]. This finding is in contrast to early publications that demonstrated similar accuracy rates regardless of the level of training [14,15]. The better performance of the technicians may be accounted for by their rigorous training. This is further confirmed by the positive correlation found between the number of examinations performed by the technicians and the percentage of predictions within 10% of birthweight.

Our findings may have several important implications. The accuracy of assessment of technicians compared to the ultrasound-certified physicians may argue against the practice of fetal weight estimation by a physician following an initial examination by a technician. In this regard, it is worth noting that there are no evidence-based data to support repeated sonographic fetal weight estimations. Our results may contribute to cost-effectiveness and time-saving. Furthermore, our results may also apply to under-resourced areas, where physician specialists are not readily available, and support for decision-making is based only on ultrasound examinations performed by technicians. Our results concur with a previous report that showed correct fetal weight estimation by residents in training among women who delivered within 2 days of the ultrasound examination [16].

For small fetuses, the percentage of erroneous predictions made by physicians was lower than that of technicians, although statistically significant only for the MPE value. This result could be interpreted to mean that physicians' prediction for this subset of patients is potentially more accurate and may affect patient outcomes. However, since differences were small between the groups, the clinical significance is probably trivial. Moreover, this study was not intended to measure outcomes and was underpowered for such results.

For the fetuses with birth weight in the range of 2500 to 4000 grams, the technicians had higher sensitivity but much lower specificity. For the extremes of ≤ 2500 grams and ≥ 4000 grams, the technicians had lower sensitivities and higher specificities compared to the physicians group, with similar percentage of predictions within 10% of the birth weight. Nevertheless, the overall false prediction rate for the macrosomic fetuses was still lower for the technician group. Thus, from a clinical perspective, their prediction was superior to that of the physicians. The differences in the percentage of erroneous predictions of small and macrosomic fetuses between the groups may relate to physicians' higher concern for fetal growth restriction and its potential hazardous consequences, as compared to misidentifying the presence of macrosomia.

LIMITATIONS

The retrospective design of this study raises the possibility of biases inherent to such investigations. A selection bias may be due to the referral to a physician for patients with more difficult examinations. Furthermore, we could not exclude the possibility that additional unknown factors (e.g., different ultrasound equipment used, anthropometric parameters, amniotic fluid volume, and placental location) could explain the differences observed between the groups.

STRENGTHS

The study has a large number of examinations made exclusively by highly experienced examiners and has full maternal characteristics; therefore, this cohort is representative of the general population of pregnant women for whom birth weight is estimated.

CONCLUSIONS

In this observational study of determining EFW accuracy between technicians and ultrasound-certified physicians, the technicians demonstrated overall better performance. Our study shows that when a qualified sonographer estimates the fetal weight between 2500 and 4000 grams, there is no need in re-estimation by physicians. Examination by a physician is needed when the estimated weight is outside of these weights. Further studies are needed to confirm our findings and better delineate the role of subsequent examination by an ultrasound-certified physician following an initial estimation made by an experienced technician.

Acknowledgements

The authors thank Cindy Cohen for her editorial assistance

Correspondence

Dr. S. Porat

Dept. of Obstetrics and Gynecology, Hadassah-Hebrew University Medical Center, Jerusalem 91120, Israel

Phone: (972-2) 677-7180, **Fax:** (972-2) 677-6489

email: shay.porat@gmail.com

References

1. Royal College of Obstetricians and Gynaecologists. The investigation and management of the small-for-gestational age fetus. No. 31, 2nd edn. London: Royal College of Obstetricians and Gynaecologists, 2013.
2. Gardosi J, Madurasinghe V, Williams M. Maternal and fetal risk factors for stillbirth: population based study. *Br Med J* 2013; 346: 1-10.
3. Chauhan SP, Cole J, Sanderson M, Magann EF, Scardo JA. Suspicion of intrauterine growth restriction: use of abdominal circumference alone or estimated fetal weight below 10%. *J Matern Fetal Neonatal Med* 2006; 19: 557-62.
4. Haragan A, Hulsey T, Hawk A, et al. Diagnostic accuracy of fundal height and handheld ultrasound-measured abdominal circumference to screen for fetal growth abnormalities. *Am J Obstet Gynaecol* 2015; 820: 1-8.
5. Benacerraf BR, Gelman R, Frigoletto FD Jr. Sonographically estimated fetal weights: accuracy and limitation. *Am J Obstet Gynecol* 1988; 159: 1118-21.
6. Hirata GI, Medearis Al, Horenstein J, et al. Ultrasonographic estimation of fetal weight in the clinically macrosomic fetus. *Am J Obstet Gynecol* 1990; 162: 238-42.
7. Milner J, Arezina J. The accuracy of ultrasound estimation of fetal weight in comparison to birth weight: a systematic review. *Ultrasound* 2018; 26 (1): 32-41.
8. Hadlock F, Harrist R, Sharman R, et al. Estimation of fetal weight with the use of head, body and femur measurements – a prospective study. *Am J Obstet Gynaecol* 1985; 151: 333-7.
9. Hammami A, Mazer Zumaeta A, Syngelaki A, Akolekar R, Nicolaides KH. Ultrasonographic estimation of fetal weight: development of new model and assessment of performance of previous models. *Ultrasound Obstet Gynecol* 2018; 52 (1): 35-43.
10. Predanic M, Cho A, Ingrid F, Pellettieri J. Ultrasonographic estimation of fetal weight: acquiring accuracy in residency. *J Ultrasound Med* 2002; 21: 495-500.
11. Gull I, Fait G, Har-Toov J, et al. Prediction of fetal weight by ultrasound: the contribution of additional examiners. *Ultrasound Obstet Gynecol* 2002; 20: 57-60.
12. Dudley NJ. A systematic review of the ultrasound estimation of fetal weight. *Ultrasound Obstet Gynecol* 2005; 25: 80-9.
13. Humphries J, Reynolds D, Bell-Scarbrough L, Lynn N, Scardo JA, Chauhan SP. Sonographic estimate of birth weight: relative accuracy of sonographers versus maternal–fetal medicine specialists. *J Matern Fetal Neonatal Med* 2002; 11: 108-12.
14. Field NT, Piper JM, Langer O. The effect of maternal obesity on the accuracy of fetal weight estimation. *Obstet Gynecol* 1995; 86: 102-7.
15. Chauhan SP,Cowan BD,Magann EF, et al. Intrapartum detection of a macrosomic fetus: clinical versus eight sonographic models. *Aust NZ J Obstet Gynecol* 1995; 35: 266-70.
16. Platek DN, Divon MY, Anyaegbunam A, Merkatz IR. Intrapartum ultrasonographic estimates of fetal weight by the house staff. *Am J Obstet Gynecol* 1991; 165 (4 Pt 1): 842-5.

Capsule

A convergence for: RNA and DNA sensing

In response to cytosolic DNA, cyclic GMP-AMP synthase (cGAS) initiates a type I interferon response. **Hu** et al. found that endogenous cGAS bound to the nucleotide helicase G3BP1, which is involved in stress-granule formation. After DNA stimulation of human cells, cGAS associated in an RNA-dependent manner with G3BP1 and was found in cytoplasmic foci that also contained messenger RNA and the RNA-dependent kinase PKR. Formation of cytoplasmic cGAS condensates necessary for DNA-stimulated type I interferon production required G3BP1 and PKR activity.

Sci Signal 2019; 12: eaav7934

Eitan Israeli

Capsule

Putting mice on a keto diet

Our immune responses to infections are influenced by extrinsic factors, including weather, social interactions, and diet. **Goldberg** and colleagues found that feeding mice a high-fat, low-carbohydrate ketogenic diet confered protection in the context of lethal influenza infection. By characterizing the immune response in the lungs, the authors identified that a ketogenic diet promoted the expansion of γδ T cells in the lung. Experiments with mice lacking γδ T cells confirmed the functional importance of these cells in conferring protection. The γδ T cells appear to improve barrier function in the lungs by modifying differentiation and function of the airway epithelial cells.

Sci Immunol 2019; 4: eaav2026

Eitan Israeli

Capsule

Microbiota influence vaccine responses

Responses to vaccines can be variable. Recent findings suggest that a potential explanation may lie in the influence of gut microbiota. In a perspective, **Pulendran** discussed how systematic analyses of vaccination responses, for example to the influenza vaccine, are revealing insights into vaccine responses and the importance of the gut microbiota in forming immunogenicity to infectious disease. This finding has implications for vaccine design and for stratifying individuals according to likely vaccine response.

Science 2019; 366; 1074

Eitan Israeli