Arch Gynecol Obstet (2015) 291:59–66
DOI 10.1007/s00404-014-3384-4

Statistical limits in sonographic estimation of birth weight

Marco Scioscia · Floriano Scioscia · Gaetano Scioscia · Stefano Bettocchi

Received: 10 January 2014 / Accepted: 14 July 2014 / Published online: 29 July 2014

Abstract

Purpose The accuracy of sonographic estimation of birth weight (EBW) is compromised by the precision of the biometrical measurements and the quality of the algorithms. This prospective study was to evaluate technical aspects to derive new equations for the EBW.

Methods Three consecutive phases were carried out (1) to recruit a homogenous population, (2) to derive eight new algorithms using a multiple stepwise mathematical/statistical method, and (3) to test the accuracy of the developed equations. Only those patients with a singleton pregnancy who delivered within 48 h from the scan were considered for the analysis.

Results The study population was made of 473 women. Four polynomial, two square root and two logarithmic algorithms were derived from a balanced study group of 200 women selected from the original study population. These formulas were subsequently applied and compared between them and showed a significant correlation with birth weight ($p < 0.0001$) and satisfactory statistical performances ($r > 0.9$), nevertheless they performed similarly to other equations previously published.

Conclusions The present findings define better the limitations associated with the intrinsic properties of algorithms and highlight that the possibility to improve the precision of sonographic measurements remains the only point at issue to increase the accuracy in the prediction of birth weight.

Keywords Ultrasound · Estimation of fetal weight · Formula · Birth weight · Error

M. Scioscia (✉)
Department of Obstetrics and Gynecology, Sacro Cuore Don Calabria General Hospital, Via Don A. Sempreboni, 5, 37024 Negrar, VR, Italy
e-mail: marcoscioscia@gmail.com; marco.scioscia@sacrocuore.it

F. Scioscia
Department of Electrotechnology and Electronics, Politecnico of Bari, Bari, Italy

G. Scioscia
IBM Innovation Lab, IBM Italy SpA, Bari, Italy

S. Bettocchi
Department of Gynecology, Obstetrics and Neonatology (DIGON), I Clinic, Medical University Policlinico of Bari, Bari, Italy

Introduction

Sonographic estimation of fetal weight (EFW) is widely used in obstetrics as a valuable information for planning the mode of delivery and management of labor. Notwithstanding extensive researches in the past years, very little improvements have been reported in the accuracy of this estimation [1, 2]. Notably, most formulas have been proposed nearly 30 years ago using different combinations of standardized fetal biometric parameters, such as biparietal diameter (BPD), head circumference (HC), abdominal circumference (AC), and femur length (FL). Subsequent efforts have aimed at identifying new parameters such as the sonographic evaluation of fetal fat and lean mass [3–7] mainly to overcome the limited accuracy of ultrasonographic estimations at extremes of birth weight [1].

The assessment of the deviation of EFW from the actual birth weight has been reported to be a combination of the measurement error and the intrinsic properties of the formula [8]. The measurement error is based on the intra- and inter-observer variability that increases at the extremes of fetal weight [9]. For instance, AC is recognized as the most predictive single measurement for the excessive fetal growth, although it is subject to larger variability as it increases [10]. The assessment of the intrinsic properties of the formulas has been poorly studied over the years with general assumptions drawn from inappropriate approaches like either a fully computerized evaluation [11] or the application of formulas derived from different populations (ethnic groups, formulas derived from the general population or specific subgroups) on general study populations. This carries an intrinsic a priori bias that leads to inevitable inaccurate conclusions.

The present study was carried out not to derive a new formula for the EFW, but to evaluate technical aspects of the accuracy of different equations for the estimation of birth weight using a combined mathematical and statistical approach on a real clinical data set. In the first part of the study, a balanced population was used to ensure a homogenous stratification of birth weights to improve the accuracy at extremes of the curve.

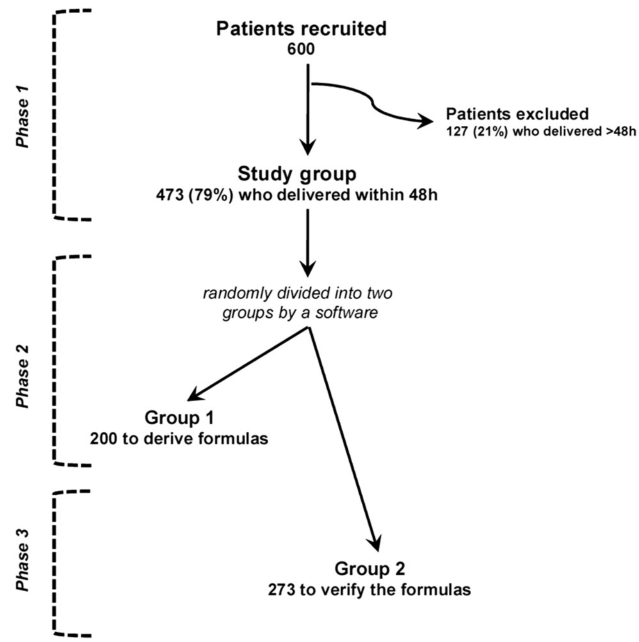

Fig. 1 Study population

Materials and methods

Study design and subjects

The study was planned as a three-step project based on a first stage with the recruitment of the whole study population. Therefore, a prospective recruitment of women with singleton pregnancies likely to give birth within 48 h was carried out at the University Hospital of Bari (Italy) and at the Sacro Cuore General Hospital of Negrar, Verona (Italy) over 18 months. The approval of the local ethics review boards was obtained. Six hundred patients were enrolled consecutively among those admitted for elective cesarean section (maternal request or previous cesarean section), induction of labor (post-term pregnancies, pre-labor rupture of the membranes), or initial spontaneous labor. All patients signed informed consent prior to study participation. Ultrasound measurements of classical fetal biometric parameters (BPD, HC, AC, and FL) [2] were taken by a single consultant (MS) with extensive experience in obstetrical ultrasound (>10 years) using a 3.5–6 MHz or 3.5–5 MHz convex transducer on Aloka ProSound alpha5 or Toshiba Aplio machines. Only those women who succeeded in delivering within 48 h from ultrasound were considered for the analysis. Neonatal birth weights were recorded at delivery.

Subsequently, the all cases who met the main entry criterion (delivery within 48 h from ultrasound) were divided into two groups to derive (phase 2) and to apply those formulas (phase 3) as reported in Fig. 1. A specific software was developed (by GS) to divide the whole study population in two groups: a group of 200 cases for the derivation of the formulas and the rest for their application. The software took into account three factors to select cases for the derivation group: (1) a stratification of actual birth weights (six 500-g intervals ranging from less than 2,000 g to more than 4,000 g) to produce a balanced study population (the same number of cases in the six groups), (2) maternal BMI (± 3 kg/m^2), and (3) gestational age at birth (± 2 weeks). The first point was to avoid a priori bias as a mathematical equation tends to fit the curve to the larger proportion of cases (therefore, the central part of a Gaussian distribution) making the extremes less represented by the formula. The other two points aimed at reducing the intrinsic error linked to the sonographic measurement as obesity and preterm birth were shown to reduce the precision of measurements [8, 9].

According to our previous studies [1, 7], a study population of four hundred cases (patients who actually succeed in delivering within 48 h) were considered sufficient to draw information for the study producing two groups of 200 cases each for phase 2 and 3, respectively (this was based on the assumption that the study population had a Gaussian distribution with a mean birth weight of 3,000 g, thus about 68 % of values would have been within one standard deviation away from the mean, with the need for 411 cases to ensure sufficient numbers in the tails of the distribution for a power of 0.8 and an alpha of 0.05). According to findings of two previous studies of ours [1, 7], about 70 % of all women deemed likely to delivery within 48 h actually succeed in that, therefore, 600 women were required to obtain four hundred cases.

Statistical and mathematical analyses

To compare the performance of new formulas with classical formulas, we used only standard measurements. The vast majority of formulas present as polynomials [1] so we decided to derive four polynomial equations (two formulas with degree two or less and two equations with a higher degree) regardless of the number of variables included but based on the performance of the formulas. Two more formulas with square roots and two with logarithms derived as functions with sublinear growth are less sensitive to parameter measurement errors.

The derivation of a formula was based on a mathematical and statistical approach for finding the best-fitting curve to a given set of points, thus a quality control of data was performed using Grubb's test. Although deletion of outlier data is a controversial practice in statistics, outlier values play an important role in mathematics [12]. Extreme outliers (observations beyond 1.96 standard deviations from the mean in phase 2, $n = 5/200$) were rejected. In fact, outliers can occur by chance in any distribution, but they are often indicative either of measurement error or that the population has a heavy-tailed distribution. Outliers were not omitted in phase 3 to assess the robustness of formulas. Birth weight estimation equations were obtained using multiple linear regression analysis technique, with actual birth weight as the dependent variable. Classical biometric parameters (BPD, HC, AC, and FL) were considered as explanatory (independent) variables, as well as their square root, natural logarithm, square and cube transformations to derive the eight equations that were both accurate and simple to use. Models exhibiting multicollinearity were discarded, as hidden mutual dependence between explanatory variables could cause unreliability of results. The possibility of multicollinearity was evaluated by calculating Dillon and Goldstein condition number C from eigenvalues of the correlation matrix [13] and testing the threshold condition $C > 30$. Adjusted coefficient of determination R_a^2 was used to assess the compromise between accuracy and simplicity of different models. Higher coefficient of determination R^2 corresponds to lower standard error of a regression formula. However, R^2 always increases when a further explanatory variable is added, no matter how much or how little explanatory power it has. R_a^2 corrects this behavior by adjusting the coefficient for the number of variables in the model, therefore, taking into account the loss of freedom degrees that results from using more variables [14].

For each of the four formula types that were considered, all formulas obtained by all possible combinations of parameters were considered as candidates. Formulas were then grouped by the number of contained variables; in each group, the five candidates with highest adjusted coefficient of determination R^2 were examined in detail as described in the paper. In detail: (1) for formulas with degree 1 and 2: BPD, HC, AC, FL, their squares and pairwise cross-products were considered, for a total of 14 possible variables, hence all regression formulas with 1,2,…,14 variables were computed; (2) for formulas with degree 3 and 4: the third and fourth powers of BPD, HC, AC, FL, as well as products of any three of them were considered, for a total of 12 possible variables; (3) for formulas with logarithms: BPD, HC, AC, FL, and their logarithms were considered, for a total of eight possible variables; (4) for formulas with square roots: BPD, HC, AC, FL, and their square roots were considered, for a total of eight possible variables.

Statistical significance of candidate models was verified using the phase 2 data by means of an ANOVA table and F test for the overall fit, as well as t tests of individual parameters. Furthermore, for best candidates the regression function plot was audited visually to detect anomalies in the behavior at the extremes of the data set (very low or very high values) and the histogram of residual values over the two-hundreds data set was examined through visual inspection to confirm the hypothesis of normality [12].

Subsequently, the performance of the different equations (Phase 3) was assessed by calculating the mean absolute and signed "potential error" of the estimate weights using the following formula (EFW − ABW) × 100/EBW expressed in percentage terms (where ABW stands for actual birth weight) with the SD and SE of the percent differences representing the variability and the precision of the mean, respectively. This method was previously proposed by other Authors [1, 15, 16] as the sonographic estimation represents the actual relevant information to clinicians for decision making (clinical management), thus EBW not ABW was used as the denominator.

The general tendency of each formula to over or underestimate birth weights was assessed using the Bland–Altman method [17] and reported as signed biases (negative values indicate that there was an overall tendency of that algorithm to overestimations). This method assesses the agreement between two measurements (EBW and ABW), not the strength of a relationship as the correlation coefficient does. The bias (mean difference between the paired measurements) and 95 % limits of agreement (the two values within which 95 % of the differences between paired measurements will lie) were calculated for all algorithms. To run the limits of agreement analysis (Bland–Altman), a logarithmic transformation of weights was necessary because of the variability in the difference between EBW and ABW as birth weight increased. The performance of all algorithms was calculated as the ability to successfully predict the ABW within 10 and 15 % of absolute error.

Continuous variables were assessed using a t test if a normal distribution was confirmed by the method of

Kolmogorov and Smirnov. Data were analyzed using the GraphPad Prism software system (version 4.00 for Windows, GraphPad Software, San Diego California USA) with significance set at $p < 0.05$. For the derivation of the new formula for estimation of fetal weight, KyPlot data analysis package (version 2.0, KyensLab Inc., Tokyo, Japan, 2002) was used.

Results

Delivery within 48 h was observed in 473 cases (78.8 %) that were included in the analysis. The characteristics of the study populations are reported in Table 1. The birth weights ranged from 760 to 4,630 g with a Gaussian distribution as confirmed statistically (K–S > 0.1). None of the newborns had physical or chromosomal malformations. More than three-quarters of the scans (76.7 %) were performed within 24 h of delivery.

According to the study protocol, 200 cases were selected randomly by the developed software with the application of stratifying criteria and used to derive eight formulas (Fig. 1). Mathematical and statistical characteristics of the formulas are reported in Table 2.

A strong correlation between the estimates and the ABW was found for all formulas being $r > 0.9$ in each case (Table 3). The mean absolute percent error of the eight algorithms is shown in Fig. 2. All of them provided accurate estimates of birth weight that was defined as ≤10 % of mean absolute error. The signed potential error was very low for all formulas demonstrating a good distribution of errors (Table 3). Although the mean error was similar, the standard error (that represents the precision of the mean) produced by polynomial formulas was lower than the other algorithms and Bartlett's test suggests that the differences among the standard deviations was

Table 1 Demographic and clinical data of the study population. All characteristics refer to patients who delivered within 48 h, thus considered for the analysis

	Phase 2 $n = 200$	Phase 3 $n = 273$
Age (years; mean and SD)	32.3 ± 4.7	31.7 ± 5.1
BMI (kg/m^2; mean and SD)	23.7 ± 4.5	22.6 ± 5.8
Gestation (weeks; median and range)	38; 30–41	38; 29–41
Fetuses in cephalic presentation (n, %)	191 (95.5)	257 (94.1)
Patients with intact membranes at ultrasound (n, %)	162 (81.0)	232 (85.0)
Birth weight (g; median and range)	3,180; 760–4,600	3,250; 780–4,630
Ultrasound within 24 h of delivery (n, %)	156 (78.0)	207 (75.8)

Table 2 Mathematical and statistical characteristics of the derived formulas on phase 2 data

Classification	Formula	Dillon–Goldstein	Coefficient determination Adjusted R^2	ANOVA F	ANOVA p	Regressogram evaluation
Polynomial	$EFW = -893.105 + 0.027 \times AC^2 + 0.226 \times FL^2$	23.5	0.93	694.3	<0.001	Regular with a normal distribution
	$EFW = -1093.569 + 0.0257 \times AC^2 + 0.2312 \times BPD \times FL$	31.6	0.93	689.0	<0.001	Quite regular
	$EFW = 114.351 + 5.744 \times 10^{-5} \times AC^3 + 2.655 \times 10^{-3} \times FL^3$	19.3	0.94	762.9	<0.001	Quite regular, not so good for higher values
	$EFW = -196.792 + 1.577 \times 10^{-3} \times BPD \times AC \times FL$	19.7	0.91	664.0	<0.001	Satisfactory, especially on higher values
Square root	$EFW = -5070.205 + 17.126 \times AC + 315.798 \times \sqrt{FL}$	29.6	0.91	510.2	<0.001	Good for low and medium, less for higher values
	$EFW = -4342.382 + 19.895 \times AC + 56.916 \times \sqrt{HC}$	13.7	0.91	488.2	<0.001	Quite regular with underestimation at extremes
Logarithm	$EFW = -6476.316 + 18.241 \times AC + 867.973 \times \ln FL$	29.7	0.91	496.7	<0.001	Tendency to underestimation of medium values and overestimation at extremes
	$EFW = -5461.534 + 19.877 \times AC + 474.878 \times \ln BPD$	19.5	0.91	487.0	<0.001	Tendency to overestimation of medium values and underestimation at extremes

The second polynomial formula was considered satisfactory notwithstanding the Dillon–Goldstein condition number >30 as provided a good coefficient of determination and a quite regular regressogram. The regressogram evaluation is based on a visual interpretation of the distribution of the differences

Table 3 Correlation between predicted and actual weight using three classes of the derived formulas on patients of phase 3 ($n = 273$)

Phase 3	Polynomial	Square root	Logarithm
Correlation with birth weight (r, 95 % CI)	0.91 (0.90–0.92)	0.91 (0.89–0.92)	0.90 (0.88–0.92)
Signed percent difference (%, mean, SD and SE)	−0.45	0.09	0.22
	9.78	20.78	21.24
	0.32	0.97	0.99
Absolute mean error <5 %	44.57 %	46.52 %	45.22 %
Absolute mean error <10 %	75.54 %	72.39 %	72.61 %
Absolute mean error <15 %	91.09 %	89.13 %	88.04 %

Data are presented also as signed mean percent difference that was calculated as EBW minus ABW divided by EBW times 100. A comparison of the estimates within 5, 10 and 15 % of the actual birth weight was reported

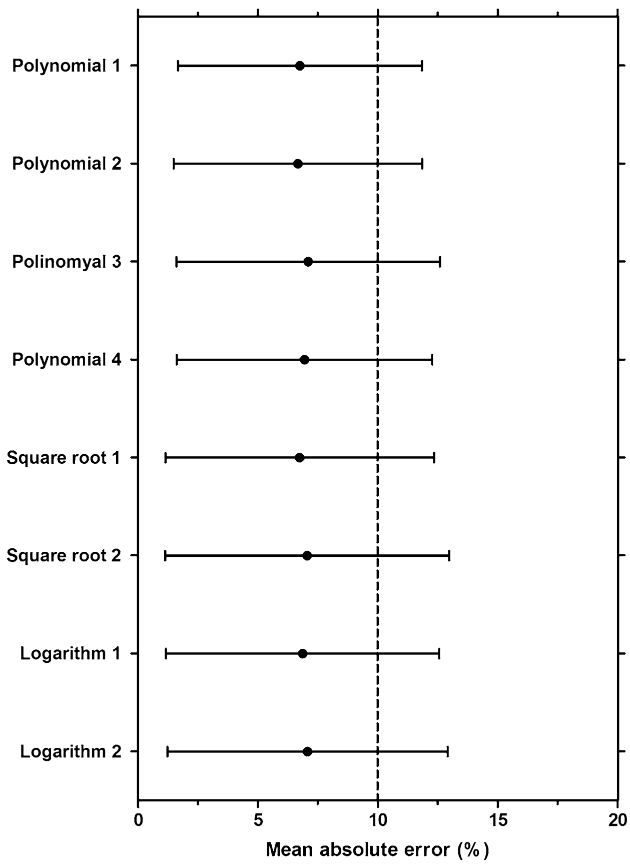

Fig. 2 Mean of absolute error (±1 SD) of the studied formulas

extremely significant ($p < 0.0001$). The percentage of birth weight predictions within ±10 and ±15 % of ABW of the studied formulas were on average 73 and 89 %, respectively (Table 3). No significant differences were found for accurate predictions (within 10 % of mean absolute error) between the group of women who underwent ultrasound with intact and ruptured membranes or between cephalic versus non-cephalic presentations in each group of categorized algorithms (data not shown).

Table 4 shows the agreement between the EBW and the ABW assessed by the limits of agreement method (Bland–Altman). All algorithms tended to overestimate birth weight (negative values) with a general good accuracy showing a bias ≤0.15 and low variability, mainly for the polynomial equations (Fig. 3).

The analysis of algorithms after stratification for ABW was made assessing the signed mean error and the fraction of EBW within ±10 of the ABW (Fig. 4a, b, respectively). The three categorized groups showed a parabolic trend in six birth weight groups (Fig. 4a) with a tendency to underestimate large fetuses with about 7 % of mean error and about 60 % of estimates with ±10 % of error (Fig. 4b). The mean discrepancy was definitively acceptable (within 150 g) till 4,000 g, although the deviation (SD) of about 300 g has also to be considered. In fact, the analysis of standard deviations (using Bartlett's test) suggested that the differences among SDs were extremely significant ($p < 0.001$) with higher variations as the ABW increased (from 229 g for ABW <2,000 g to 398 g for ABW >4,000 g).The performance of all groups was very high for ABWs between 3,000 and 4,000 g with more than 80 % of predictions within 10 % of error (Fig. 4b). It is interesting that the polynomial group showed the highest accuracy for newborns weighing less than 2,000 g ($p < 0.01$). In fact, although the mean error for the three categorized groups was similar (Fig. 4a), the square root and logarithm groups provided the poorest results with less than 40 % of predictions within the 10 % difference (Fig. 4b) and this can be explained by a higher standard deviation (the differences between estimated and actual weights were more spread out).

Discussion

A critical evaluation of the methodological approach to derive a formula for the estimation of fetal weight shows that intrinsic limitations set the cut-off of accuracy around 10 % of error. Notably, this value is similar to that reported for classical formulas [1, 2]. Although this limit is widely recognized as acceptable in clinical practice, a technical

Table 4 Measures of bias and 95 % limits of agreement from Bland–Altman analyses after log transformation (ABW–EBW) and expressed as percentage in estimates. Data set from phase 3

Classification	Formula	Bias	SD	95 % limit of agreement	
Polynomial	$EFW = -893.105 + 0.027 \times AC^2 + 0.226 \times FL^2$	−0.11	1.24	−2.54	2.32
	$EFW = -1093.569 + 0.0257 \times AC^2 + 0.2312 \times BPD \times FL$	−0.13	1.26	−2.60	2.35
	$EFW = 114.351 + 5.744 \times 10^{-5} \times AC^3 + 2.655 \times 10^{-3} \times FL^3$	−0.04	1.11	−2.23	2.14
	$EFW = -196.792 + 1.577 \times 10^{-3} \times BPD \times AC \times FL$	−0.13	1.15	−2.38	2.11
Square root	$EFW = -5070.205 + 17.126 \times AC + 315.798 \times \sqrt{FL}$	−0.11	1.98	−3.99	3.77
	$EFW = -4342.382 + 19.895 \times AC + 56.916 \times \sqrt{HC}$	−0.09	1.89	−3.78	3.61
Logarithm	$EFW = -6476.316 + 18.241 \times AC + 867.973 \times \ln FL$	−0.10	2.01	−4.04	3.84
	$EFW = -5461.534 + 19.877 \times AC + 474.878 \times \ln BPD$	−0.08	1.93	−3.86	3.69

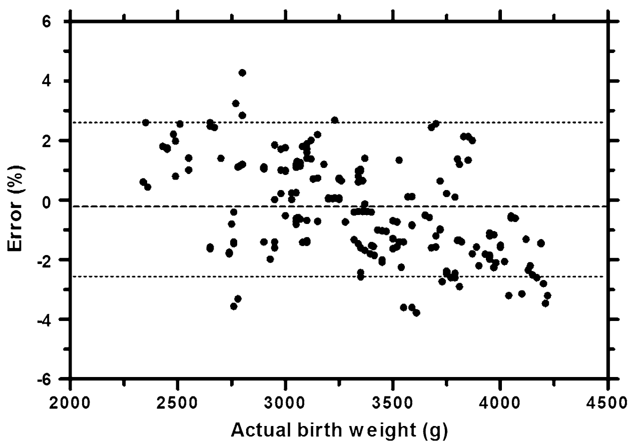

Fig. 3 Example of limits of agreement between birth weight and sonographically estimated fetal weight by polynomial number 2. Data set from Phase 3

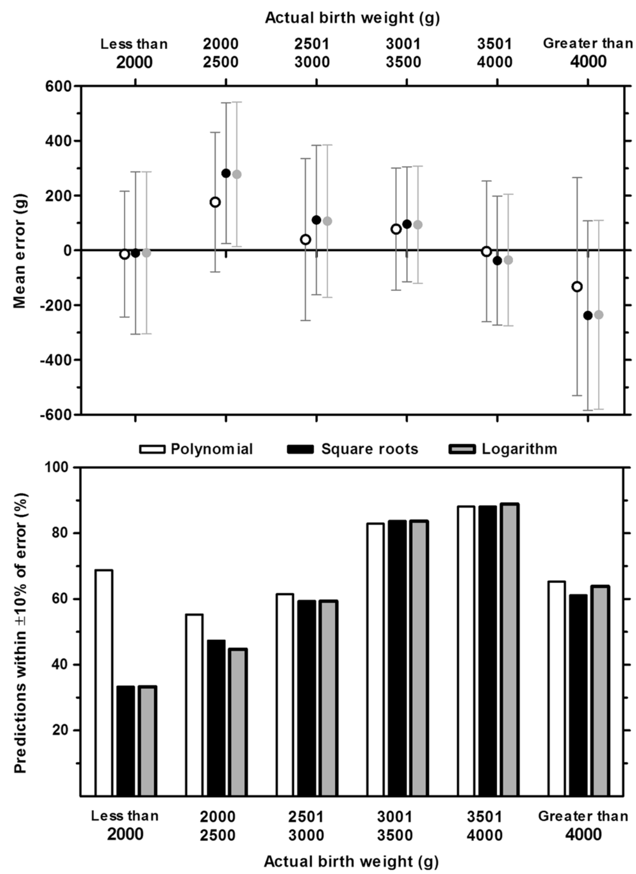

Fig. 4 Performance of the three groups of algorithms after stratification for ABW. **a** Mean error (±1 SD) of birth weight estimations in different groups of formulas as birth weight increases. Data are reported in grams to clearly visualize the increasing standard deviations. **b** Predictions within 10 % of absolute error according to the class of algorithm and the ABW interval

appraisal was required as many studies have tried to improve the accuracy.

This study adds important information to this issue demonstrating that the limitations linked to the formula cannot be overcome using standard sonographic parameters only. The error (deviation of EBW from ABW) has been reported to be half due to the measurement error and half arising from the intrinsic properties of the formula [8]. As for the first point, the general tendency to underestimate large fetuses and to overestimate the small ones was shown to be related to technical difficulties to obtain reproducible/accurate biometric parameters for large/small fetuses (observational error) [9], while for the second point, it was not linked to the sonographic parameters that the formula rely on [1]. Besides, the possibility of multicollinearity (mutual dependency between variables that enhances the intrinsic error of each measurement) and interdependency (a hidden linear correlation between biometric parameters) reduces the strength of formulas that incorporate multiple parameters. Therefore, the error due to the equations is a definite source of disagreement between predictions and actual birth weights, although not the largest one. In fact, since the accuracy has not been improved using a combined mathematical/statistical approach, a homogenous study population and a single scan operator (to reduce the inter-observer variability), we may assume that the largest bias is linked to the ultrasonographic measurement.

The aim of this study was to test the ability of eight formulas belonging to three different classes of algorithms and derived on purpose to accurately predict birth weight. The strength of our study is that (1) all scans were made by an experienced physician; (2) the study population for the derivation of formulas and their application was homogenous; in fact we used a balanced population in the study phase-1 for the derivation of the new algorithms to represent a wide range of EFWs. This reflects what happens in clinical practice where we do not know the actual birth weight, so the application of formulas for a specific subpopulation (i.e. SGA or LGA fetuses) is a personal choice of the operator and is an additional bias. (3) Only fetuses born within 48 h from the ultrasound were considered for the study; (4) the prospective design allowed us to use EBW instead of ABW as reference (independent variable) to assess the accuracy of the studied formulas as ABW is clinically less useful than the sonographic estimation, because the birth weight is unknown until after birth; (5) each class of algorithms included at least two formulas. The findings of this study support a slightly better performance of polynomial equations with consistent good accuracy in the different classes of birth weight.

In conclusion, definite limitations are associated with the intrinsic properties of algorithms as also shown using postpartum biometric parameters [18]. In fact, directly measured neonatal biometry overcomes some limitations of ultrasonography at term (due to reduced amniotic fluid, breech presentation, head engagement,…) that reduce the accuracy of sonographic measurements. Given the wide availability of computed machines, a different mathematical-statistical approach may be proposed for the derivation of new formulas using natural cubic spline equations or multivariate adaptive regression splines (MARS). These produce far more complex formulas as they are an extension of linear models that automatically model non-linearities by joining hinge functions. Furthermore, these functions may also add novel variables such as parental data (BMI, ethnicity,…) that have no multicollinearity and non-standard fetal measurements like lean and fat mass [7] to improve the accuracy of predictions. Since these equations are generally used to merge variables of different origins (i.e. sonographic with non-sonographic parameters), we decided not to use them for this study that focuses on 2D sonographic parameters only.

Recent researches based on non-standard biometric parameters such as 3D fractional volume calculation of fetal limbs [19] have shown little improvements but carry the significant disadvantage to be time consuming (for scanning and data processing) and by the need for a 3D ultrasound machine. Nevertheless, it is our opinion that the possibility to improve the quality and precision of sonographic measurements remains the main point at issue. New strategies should be implemented to make ultrasonography more accurate such as an automated individual measurement error evaluation as proposed by recent studies [20].

Conflict of interest None.

References

1. Scioscia M, Vimercati A, Ceci O, Vicino M, Selvaggi LE (2008) Estimation of birth weight by two-dimensional ultrasonography: a critical appraisal of its accuracy. Obstet Gynecol 111:57–65
2. Dudley NJ (2005) A systematic review of the ultrasound estimation of fetal weight. Ultrasound Obstet Gynecol 25:80–89
3. Bernstein IM, Goran MI, Amini SB, Catalano PM (1997) Differential growth of fetal tissues during the second half of pregnancy. Am J Obstet Gynecol 176:28–32
4. Bernstein IM (2005) Fetal body composition. Curr Opin Clin Nutr Metab Care 8:613–617
5. Larciprete G, Valensise H, Vasapollo B, Novelli GP, Parretti E, Altomare F, Di PG, Menghini S, Barbati G, Mello G, Arduini D (2003) Fetal subcutaneous tissue thickness (SCTT) in healthy and gestational diabetic pregnancies. Ultrasound Obstet Gynecol 22:591–597
6. Larciprete G, Valensise H, Di PG, Vasapollo B, Casalino B, Arduini D, Jarvis S, Cirese E (2005) Intrauterine growth restriction and fetal body composition. Ultrasound Obstet Gynecol 26:258–262
7. Scioscia M, Scioscia F, Vimercati A, Caradonna F, Nardelli C, Pinto LR, Selvaggi LE (2008) Estimation of fetal weight by measurement of fetal thigh soft-tissue thickness in the late third trimester. Ultrasound Obstet Gynecol 31:314–320
8. Mongelli M, Tambyraja R (2003) Ultrasonic fetal weight estimation and tolerance to measurement error: a comparative analysis. Australas Radiol 47:389–392
9. Kurmanavicius J, Burkhardt T, Wisser J, Huch R (2004) Ultrasonographic fetal weight estimation: accuracy of formulas and accuracy of examiners by birth weight from 500 to 5,000 g. J Perinat Med 32:155–161
10. Chang TC, Robson SC, Spencer JA, Gallivan S (1993) Ultrasonic fetal weight estimation: analysis of inter- and intra-observer variability. J Clin Ultrasound 21:515–519
11. Mongelli M, Biswas A (2002) Menstrual age-dependent systematic error in sonographic fetal weight estimation: a mathematical model. J Clin Ultrasound 30:139–144
12. Rousseeuw PJ, Leroy AM (1987) Robust regression and outlier detection. Wiley-Interscience, New York (Series in Applied Probability and Statistics)
13. Dillon WR, Goldstein M (1984) Multivariate analysis: methods and applications. Wiley, New York
14. Toutenburg H (2002) Statistical analysis of designed experiments, 2nd edn. Springer, Berlin
15. Anderson NG, Jolley IJ, Wells JE (2007) Sonographic estimation of fetal weight: comparison of bias, precision and consistency using 12 different formulae. Ultrasound Obstet Gynecol 30:173–179
16. Edwards A, Goff J, Baker L (2001) Accuracy and modifying factors of the sonographic estimation of fetal weight in a high-risk population. Aust N Z J Obstet Gynaecol 41:187–190
17. Bland JM, Altman DG (1986) Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1:307–310

18. Kehl S, Schmidt U, Spaich S, Schild RL, Sütterlin M, Siemer J (2012) What are the limits of accuracy in fetal weight estimation with conventional biometry in two-dimensional ultrasound? A novel postpartum study. Ultrasound Obstet Gynecol 39(5):543–548
19. Pagani G, Palai N, Zatti S, Fratelli N, Prefumo F, Frusca T (2014) Fetal weight estimation in gestational diabetic pregnancies: comparison between conventional and three-dimensional fractional thigh volume methods using gestation-adjusted projection. Ultrasound Obstet Gynecol 43(1):72–76
20. Cevenini G, Severi FM, Bocchi C, Petraglia F, Barbini P (2008) An informative probability model enhancing real time echobiometry to improve fetal weight estimation accuracy. Med Biol Eng Comput 46:109–120