Available online at www.sciencedirect.com
SciVerse ScienceDirect
Taiwanese Journal of Obstetrics & Gynecology 51 (2012) 545–553
www.tjog-online.com

Original Article

Efficient fetal size classification combined with artificial neural network for estimation of fetal weight

Yueh-Chin Cheng [a], Gwo-Lang Yan [b], Yu Hsien Chiu [c], Fong-Ming Chang [d], Chiung-Hsin Chang [d,*], Kao-Chi Chung [a,**]

[a] Department of Biomedical Engineering, National Cheng Kung University, Tainan, Taiwan
[b] Department of Computer Science and Information Engineering, Southern Taiwan University of Technology, Tainan, Taiwan
[c] Department of Healthcare Administration and Medical Informatics, Kaohsiung Medical University, Kaohsiung, Taiwan
[d] Department of Obstetrics and Gynecology, National Cheng Kung University, Medical College and Hospital, Tainan, Taiwan

Accepted 12 July 2012

Abstract

Objectives: A novel analysis was undertaken to select a significant ultrasonographic parameter (USP) for classifying fetuses to support artificial neural network (ANN), and thus to enhance the accuracy of fetal weight estimation.
Methods: In total, 2127 singletons were examined by prenatal ultrasound within 3 days before delivery. First, correlation analysis was used to determine a significant USP for fetal grouping. Second, *K*-means algorithm was utilized for fetal size classification based on the selected USP. Finally, stepwise regression analysis was used to examine input parameters of the ANN model.
Results: The estimated fetal weight (EFW) of the new model showed mean absolute percent error (MAPE) of 5.26 ± 4.14% and mean absolute error (MAE) of 157.91 ± 119.90 g. Comparison of EFW accuracy showed that the new model significantly outperformed the commonly-used EFW formulas (all $p < 0.05$).
Conclusion: We proved the importance of choosing a specific grouping parameter for ANN to improve EFW accuracy.

Keywords: artificial neural network; estimated fetal weight; ultrasonographic parameter

1. Introduction

It is very important to assess estimated fetal weight (EFW) accurately in obstetrics. Nowadays, ultrasound (US) is a major tool for EFW. Most published formulas of EFW are derived from 2D US parameters (USPs), such as biparietal diameter (BPD), occipito-frontal diameter (OFD), abdominal circumference (AC), and femur length (FL). Although EFW by regression methods is relatively acceptable in clinical obstetrics [1–9], the accuracy of EFW needs to be improved. As birth weight (BW) varies with gestation age and the distribution of BW was negatively skewed [10], in conflict with the normal distribution hypothesis for the commonly-used regression models.

Sabbagha et al [11] proposed a classification of three subgroups according to the gestational age (GA) and percentiles of AC for EFW in large-, appropriate-, and small-for-gestational-age fetuses. Their errors of EFW were reduced as compared with previous regression formulas. However, their method of fetal classification by GA and AC is prone to the errors of uncertain GA for women with irregular menstrual cycles.

The artificial neural network (ANN) is a nonlinear statistical data modeling tool where the complex relationships between inputs and outputs are modeled [12]. Recently, it was

* Corresponding author. Department of Obstetrics and Gynecology, National Cheng Kung University Hospital, 138 Sheng-Li Road, Tainan 70403, Taiwan.
** Corresponding author. Department of Biomedical Engineering, National Cheng Kung University, No. 1, University Road, Tainan 701, Taiwan.
E-mail addresses: ahsin@mail.ncku.edu.tw (C.-H. Chang), chengyc@mail.ncku.edu.tw (K.-C. Chung).

1028-4559/$ - see front matter
http://dx.doi.org/10.1016/j.tjog.2012.09.009

shown that estimating fetal weight using ANN models was more accurate as compared with regression methods. Nevertheless, Farmer et al [13] focused mainly on macrosomia and did not cover the entire range of fetal weight. Chuang et al [14] reported another ANN model of EFW, covering 200–4400 g. Although their results showed ANN model could provide more accurate EFW than before, their EFW was less accurate in actual fetal body weights below 2500 g and above 4000 g.

In fact, the accuracy of EFW is affected by multiple variables and the variations among USPs in different groups of fetuses. It is necessary to determine significant variables and to classify fetuses for reducing errors in EFW. In medical literature, most articles revealed EFW was less accurate in small- and large-for-gestational groups when using a regression formula or an ANN model only and entirely. In this study, we proposed a novel grouping approach conjoined with an ANN model to improve the accuracy of EFW estimation.

2. Materials and methods

Specifically, the aims of this study were: (1) to use correlation analysis to cross-validate the significance among USPs; and (2) to utilize *K*-means algorithm to classify fetal size with regard to the discriminative classes for the training of the ANN to enhance the EFW accuracy. The accuracy of our novel classification-ANN model was examined by comparing with the regression-based EFW formulas. Fig. 1 depicts the flow chart of overall experiment.

3. Data collection

A retrospective study was undertaken in National Cheng Kung University Hospital. Informed consents were obtained from the pregnant women before US examinations, which included agreement for further data analysis without showing the personal identification data to the analysts. This study was approved by the Institutional Review Board of National Cheng Kung University Hospital (IRB: ER-99-011). All the fetuses were examined by US within 3 days prior to delivery. Fetal US examinations were undertaken by using conventional US scanners (GE Voluson 730 Expert, Milwaukee, WI, USA; Medison Accuvix V20, Seoul, Korea; Aloka SSD-680, Tokyo, Japan), with 3.5–5.0 MHz convex transducers. Anomalous fetuses, multiple gestations and fetuses not delivered within 3 days were excluded from this study. The collected data were further examined to exclude those with missing data or unreasonable code. A total of 2127 consecutive singleton fetuses were used in this study. The data of total fetuses were randomly divided into training and testing groups. According to our study design, 1489 fetuses (70%) were randomly assigned to the training group for the ANN model training, and the other 638 fetuses (30%) were used to validate the ANN model. The concept of v-fold cross validation and the similar criterion of 7:3 of our previous study [14] were adopted to randomly separate the training and the test datasets; the evaluation was performed five times.

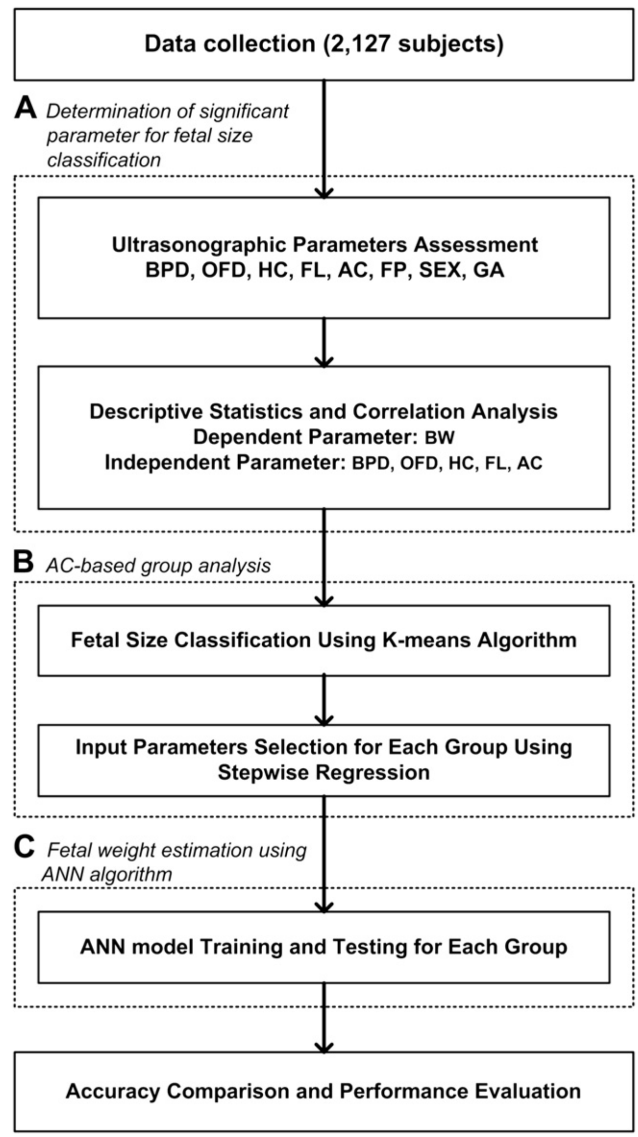

Fig. 1. Flow chart of the method used in this research. AC = abdominal circumference; BPD = biparietal diameter; BW = birth weight; FL = femur length; FP = fetal presentation; GA = gestational age; HC = head circumference; OFD = occipito-frontal diameter; SEX = gender.

4. Determination of significant parameter for fetal size classification

4.1. USP assessment

USP measurement included numerical variables, e.g., BPD, OFD, AC, head circumference (HC), FL, GA, BW, and nominal variables, e.g., gender (SEX) and fetal presentation (FP). The BW for each fetus was recorded for comparing with EFW. Fetal AC was calculated by:

$$AC = \frac{\pi}{2} \times (APD + ATD) \quad (1)$$

where APD is abdominal anteroposterior diameter and ATD is abdominal transverse diameter. GA (unit: weeks) was based on the first day of the last normal menstrual period and confirmed by first-trimester ultrasound scan. In this study, GAs for all cases required both the date of last menstrual period and US confirmation [6,15,16]. FPs were divided into vertex and malpresentation groups according to US examination. Fetuses with breech or transverse presentations were represented in the malpresentation group.

4.2. Descriptive statistics and correlation analysis

To describe the distribution of data, numerical data were checked for normal distribution using the Kolmogorov-Smirnov test [17]. Item analysis containing criterion of internal consistency and correlation analysis were used to observe the characteristics and discrimination of parameters. The critical ratio of each parameter was calculated and tested to determine which parameters were selected for test inclusion or deleted for test exclusion [18]. During the parameter selection processing, item screening was performed first to evaluate the usefulness of all the adopted USPs. In the process, the observations in each parameter were sorted and fractionally categorized into front, middle, and behind parts with predefined levels. In this study, the front part at a level of 27% and the behind part at a level of 27%, named as high score and low score groups [19], respectively, were extracted and tested by using t test. The parameters with a significance of $p < 0.05$ were used for the following experiment.

The distribution property of each parameter showed that the tendency was negatively skewed. It was necessary to classify the fetuses to reduce the effect of data sparseness. Spearman correlation analysis was then used to determine which USP was highly correlated with the BW. A low correlation coefficient suggested that the relationship between two parameters was weak or nonexistent. The AC had a higher correlation coefficient than other USPs, so it was selected as the basis for grouping fetuses.

5. AC-based group analysis

5.1. Fetal size classification using K-means algorithm

K-means algorithm is used to cluster fetuses with similar body size based on AC, and also to reduce the effects of date sparseness: θ_k denotes the mean value of the k^{th} cluster; x_f is the AC of the f^{th} fetus;F denotes the total number of fetuses. The K-means algorithm [20,21] is used to cluster fetuses into K clusters as:

1. Choose arbitrary initial estimates $\theta_k(0)$ from $\{x_f | f = 1, \cdots, F\}$ for θ_k $k = 1, \cdots, K$
2. Repeat
   a. For $f = 1$ to F
      i. Determine the closest representative, say θ_k for x_f.
      ii. Set $b(f) = k$.
   b. End {For}
   c. For $k = 1$ to K
      i. Parameter updating: determine θ_k as the mean value of x_f with $b(f) = k$.
   d. End {For}
3. Until no change occurs for all θ_k between two successive iterations.

K-means is used to classify fetal size to reduce body size heterogeneities between fetuses. Stepwise regression is utilized for each group to select input parameters for the ANN model.

5.2. ANN input parameters selection for each group using stepwise regression

Stepwise estimation [22,23] is used to select input parameters by examining the contribution to the ANN model (Fig. 2). The parameter most highly correlated with the BW is

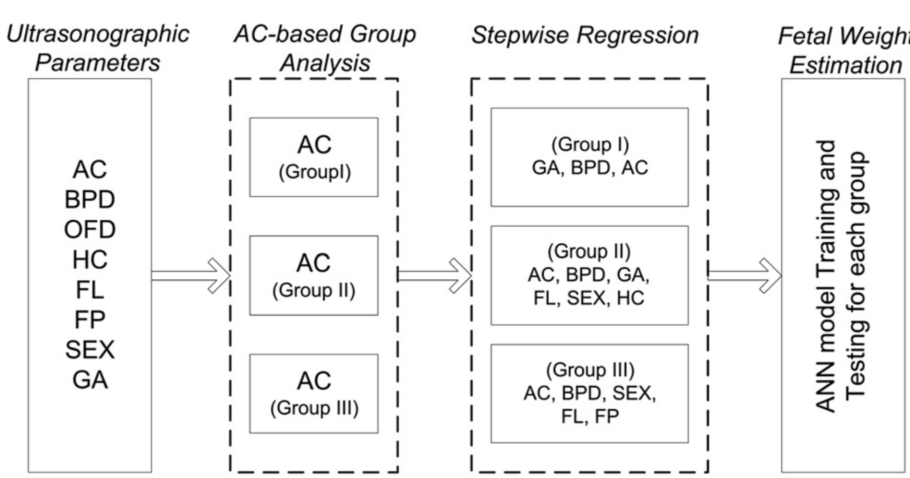

Fig. 2. Input selection for each artificial neural network (ANN) group model. AC = abdominal circumference; BPD = biparietal diameter; FL = femur length; FP = fetal presentation; GA = gestational age; HC = head circumference; OFD = occipito-frontal diameter; SEX = gender.

selected first. In addition, independent parameters are selected for inclusion based on the incremental contribution over the parameter already in the model. Independent parameters may also be dropped if their predictive power drops to a nonsignificant level when another independent parameter is added to the model [22,23].

6. EFW using the ANN algorithm

This study used a back propagation network to establish the ANN model [12]. After AC group analysis, three back propagation networks were developed for EFW. Fig. 3 illustrates the architecture of the ANN model. Fetuses in each group were randomly divided into a training set and a testing set. In the training set, Group I consisted of 70 samples, Group II 735 samples, and Group III 684 samples, respectively ($n_{11} = 70$, $n_{21} = 735$, $n_{31} = 684$). In the testing set, Group I consisted of 32 samples, Group II 309 samples, and Group III 297 samples, respectively ($n_{12} = 32$, $n_{22} = 309$, $n_{32} = 297$). The numbers of train and testing fetuses for each group are shown in Table 1. The used ANN model has an input layer, a hidden layer, and an output layer. The number of neurons for each layer is described in Table 2. The input vector $\boldsymbol{p}$ in the input layer has $\boldsymbol{R}$ input parameters that are selected by the stepwise estimation method. The transfer function f^1 is the hyperbolic-tangent-sigmoid function in the hidden layer and f^2 is a linear function in the output layer for all three groups.

In the low BW group (Group I), the input layer includes three parameters, the hidden layer has 12 neurons, and the output layer has a neuron. In the normal BW group (Group II), the input layer includes six parameters, the hidden layer has 14 neurons, and the output layer has a neuron. In the high BW group (Group III), the input layer includes five parameters, the hidden layer has 12 neurons, and the output layer has a neuron. The conjugate gradient method is used to train neural network for optimizing

Table 1
The sample sizes in the training set and the testing set for the artificial neural network (ANN) model.

Groups	Sample sizes in the training set	Sample sizes in the testing set	Total
Low BW group (Group I)	$n_{11} = 70$	$n_{12} = 32$	102
Normal BW group (Group II)	$n_{21} = 735$	$n_{22} = 309$	1044
High BW group (Group III)	$n_{31} = 684$	$n_{32} = 297$	981
Total	1489	638	2127

BW = birth weight.

the performance of the ANN model. This study made use of the MATLAB software package to perform the ANN algorithm.

7. Accuracy comparison and performance evaluation

The error of fetal weights estimated by the proposed novel ANN model was compared with regression-based formulas listed in Table 3, including Hsieh's formula 1B [6], Hsieh's formula 2B [6], and Hadlock's formula [5].

Two indexes, mean absolute error (MAE), and mean absolute percent error (MAPE), were calculated from EFW and actual BWs. Equations are represented as follow:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|X_i - \widehat{X}_i\right| \tag{2}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{\left|X_i - \widehat{X}_i\right|}{X_i} \times 100\% \tag{3}$$

where n is the number of fetuses. X_i and $\widehat{X}_i$ are the BW and EFW of the ith fetus.

Because MAE and MAPE were not normally distributed, the Friedman test was used for comparing the performance of the proposed ANN model with three regression-based methods. Significance was defined as $p < 0.05$. Data

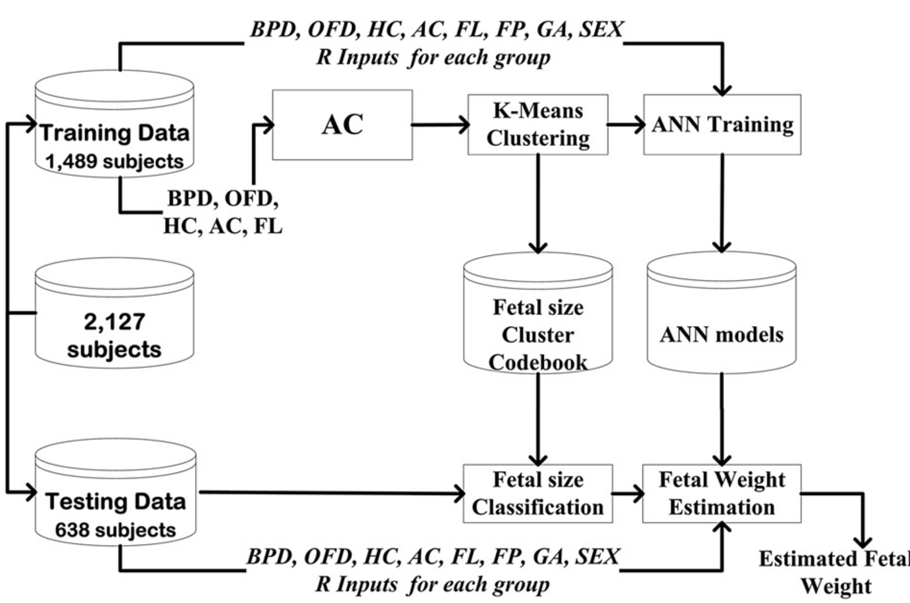

Fig. 3. Illustration of the architecture of the artificial neural network (ANN) model. AC = abdominal circumference; BPD = biparietal diameter; FL = femur length; FP = fetal presentation; GA = gestational age; HC = head circumference; OFD = occipito-frontal diameter; SEX = gender.

Table 2
Input parameters, the number of neurons, and transfer function of the artificial neural network (ANN) model.

Groups	Items	ANN		
		Input layer	Hidden layer	Output layer
Low BW group (Group I)	Parameters for inclusion (R)	GA, BPD, AC (R = 3)		
	Number of neurons (S)		S^1=12	S^2=1
	Transfer function (f)		Hyperbolic tangent sigmoid (f^1)	Linear (f^2)
Normal BW group (Group II)	Parameters for inclusion (R)	AC, BPD, GA, FL, SEX, HC (R = 6)		
	Number of neurons (S)		S^1=14	S^2=1
	Transfer function (f)		Hyperbolic tangent sigmoid (f^1)	Linear (f^2)
High BW group (Group III)	Parameters for inclusion (R)	AC, BPD, SEX, FL, FP (R = 5)		
	Number of neurons (S)		S^1=12	S^2=1
	Transfer function (f)		Hyperbolic tangent sigmoid (f^1)	Linear (f^2)

AC = abdominal circumference; BPD = biparietal diameter; BW = birth weight; FL = femur length; FP = fetal presentation; GA = gestational age; HC = head circumference; SEX = gender.

management and statistical analysis were performed using SPSS for Windows version 15.0 (SPSS Inc., Chicago, IL, USA) and STATISTICA version 8 (StatSoft Inc., USA).

8. Results

8.1. Data description

The range of BWs of 2127 babies was 500–4736 g. GAs at birth ranged from 21 weeks to 43 weeks. The mean ± standard deviation (SD) for each parameter: BW 3062.7 ± 604.7 g, AC was 32.6 ± 3.3 cm, BPD was 9.1 ± 0.7 cm, OFD was 11.0 ± 0.9 cm, FL was 6.8 ± 0.6 cm, HC was 31.6 ± 2.3 cm and GA was 38.5 ± 3.1 weeks. The Kolmogorov-Smirnov test ($p <$ 0.05) revealed that all numeric data had non-normal distributions. Moreover, the distribution for each parameter showed that the tendency was negatively skewed ($\text{Skewness}_{BW} = -1.72$, $\text{Skewness}_{AC} = -2.13$, $\text{Skewness}_{BPD} = -3.13$, $\text{Skewness}_{OFD} = -1.48$, $\text{Skewness}_{FL} = -2.96$, $\text{Skewness}_{HC} = -2.69$, $\text{Skewness}_{GA} = -3.14$) and the peakedness was a leptokurtosis ($\text{Kurtosis}_{BW} = 5.14$, $\text{Kurtosis}_{AC} = 7.24$, $\text{Kurtosis}_{BPD} = 13.23$, $\text{Kurtosis}_{OFD} = 5.10$, $\text{Kurtosis}_{FL} = 12.13$, $\text{Kurtosis}_{HC} = 11.18$, $\text{Kurtosis}_{GA} = 12.16$), as illustrated in Fig. 4. In the process of preliminary parameter selection, the numerical parameters with significant discrimination include AC, BPD, OFD, HC, FL, and GA. The t values for each parameter were as follows: AC was $t_{1164, 0.95} = 41.65$*, BPD was $t_{1468, 0.95} = 32.44$*, OFD was $t_{1254, 0.95} = 50.29$*, HC was $t_{1179, 0.95} = 37.28$*, FL was $t_{1387, 0.95} = 35.44$*, and GA was $t_{1269, 0.95} = 30.73$*. The results of the item analysis were shown in Table 4.

8.2. Strong correlation between USP and BW

The Spearman correlation between each parameter and BW has a positive coefficient, as shown in Table 5. The scatter plots of BW and USP are illustrated in Fig. 5. In particular, AC showed strong positive correlation to BW (r = 0.81, $p < 0.01$). The AC has a higher correlation coefficient than other USPs, so it is selected as the basis for grouping fetuses.

8.3. Fetal size grouping with their respective discriminative parameters

The scatter plots of BW and AC are illustrated in Fig. 6: low BW group (Group I), $r = 0.92$, $p < 0.01$, $n = 102$; normal BW group (Group II), $r = 0.61$, $p < 0.01$, $n = 1{,}044$; high BW group (Group III), $r = 0.64$, $p < 0.01$, $n = 981$. The mean ± SD (range) of fetal AC in Groups I, II, and III were 21.8 ± 3.5 cm (15.7–26.5 cm), 31.5 ± 1.4 cm (26.7–33.1 cm), and 34.8 ± 1.2 cm (33.3–39.9 cm), respectively. The BWs for three groups classified by AC were 1124.1 ± 524.4 g for Group I, 2919.7 ± 336.0 g for Group II, and 3416.5 ± 326.6 g for Group III.

Table 2 shows the selected parameters of each group. In Group I, the input layer includes three parameters of GA, BPD, and AC ($R^2 = 0.95$, $n = 102$). In Group II, the input layer includes six parameters of AC, BPD, GA, FL, SEX, and HC ($R^2 = 0.72$, $n = 1{,}044$). In Group III, the input layer includes five parameters of AC, BPD, SEX, FL, and FP ($R^2 = 0.61$, $n = 981$).

Table 3
Three published regression methods.

References	Formulas
Hsieh's formula 1B (1987)	$\log_{10}BW = 5.6541 \times 10^{-3} \times AC \times BPD - 1.5515 \times 10^{-4} \times AC^2 \times BPD + 1.9782 \times 10^{-5} \times AC^3 + 5.2594 \times 10^{-2} \times BPD + 2.1315$
Hsieh's formula 2B (1987)	$\log_{10}BW = 9.4962 \times 10^{-3} \times AC \times BPD - 0.1432 \times FL - 7.6742 \times 10^{-4} \times AC \times BPD^2 + 1.7450 \times 10^{-3} \times BPD^2 \times FL + 2.7193$
Hadlock's formula (1985)	$\log_{10}BW = 1.304 + 0.05281 \times AC + 0.1938 \times FL - 0.004 \times AC \times FL$

AC = abdominal circumference; BPD = biparietal distance; BW = birth weight; FL = Femur length.

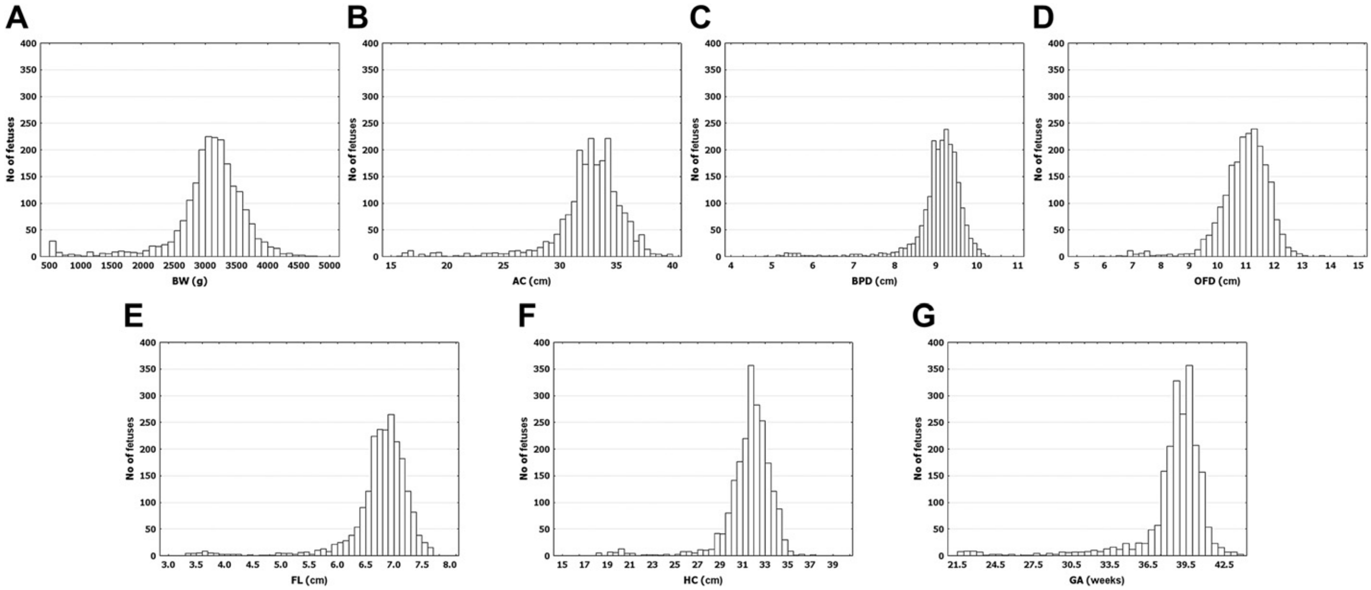

Fig. 4. The distribution for each parameter. (A) BW = birth weight; (B) AC = abdominal circumference; (C) BPD = biparietal diameter; (D) OFD = occipito-frontal diameter; (E) FL = femur length; (F) HC = head circumference; (G) GA = gestational age; $n = 2,127$.

8.4. The accuracy comparison of fetal weight estimation

As shown in Table 6, the MAPE and MAE of the proposed method were 7.1 ± 6.1% and 83.5 ± 82.8 g for Group I, 5.4 ± 4.6% and 157.2 ± 131.2 g for Group II, and 4.9 ± 3.2% and 166.7 ± 107.9 g for Group III, respectively. By lumping total estimation errors together, the MAPE and MAE were 5.3 ± 4.1% and 157.9 ± 119.9 g for the proposed method, 6.0 ± 4.6% and 173.2 ± 120.3 g for the Hsieh 1B model, 6.5 ± 7.2% and 175.1 ± 120.4 g for the Hsieh 2B model, and 7.4 ± 5.3% and 224.6 ± 169.0 g for the Hadlock model, respectively. The results of the Friedman test showed that the MAPE and MAE had statistical differences among the four methods, as shown in Table 7. The results of the multiple-comparisons procedure showed that the MAPE and MAE of three pairs had statistical differences. The three pairs were compared as follows: (1) the proposed method and the Hsieh's formula 1B method; (2) the proposed method and the Hsieh's formula 2B method; and (3) the proposed method and the Hadlock's method. Our proposed approach is significantly better than the three regression methods using the Friedman test. The highly significant correction between the actual BW and the EFW by the proposed method of the testing set, the scatter plot of actual BW, and fetal weight estimation by this study ANN model ($r = 0.95$, $R^2 = 0.89$, $n = 638$) is illustrated in Fig. 7. It shows that the input–output relationship was a straight line.

9. Discussion

We have shown an efficient classification framework combined with an ANN model for improving EFW accuracy in the wider range of gestation period and fetal weight from the database of 2127 babies. Full term infants (GAs from 37–42 weeks) are in the majority in this study. When applying the regression models, researchers have to assume that the distribution of their subjects is "normally" distributed. However, BWs are not normally distributed and might be inappropriate to use the regression models derived EFW

Table 4
Item analysis of numerical parameters.

Item	Group	N	Mean	SD	df	*t* value	*p*
AC	High score	576	35.60	1.07	1164	41.65*	<0.05
	Low score	590	28.92	3.74			
BPD	High score	750	9.59	0.18	1468	32.44*	<0.05
	Low score	720	8.53	0.86			
OFD	High score	667	11.86	0.35	1254	50.29*	<0.05
	Low score	589	9.96	0.86			
HC	High score	595	33.57	0.69	1179	37.28*	<0.05
	Low score	586	29.10	2.82			
FL	High score	677	7.25	0.15	1387	35.44*	<0.05
	Low score	712	6.26	0.73			
GA	High score	640	40.56	0.70	1269	30.73*	<0.05
	Low score	631	35.45	4.11			

AC = abdominal circumference; BPD = biparietal distance; df = degree of freedom; FL = femur length; GA = gestational age; HC = head circumference; OFD = occipito-frontal diameter; SD = standard deviation.
* The difference was statistically significant using the *t* test.

Table 5
The correlation coefficient of each parameter and birth weight (BW).

Groups	Correlation coefficients (γ)				
	AC	BPD	FL	HC	OFD
Low BW group (Group I)	0.92**	0.92**	0.92**	0.92**	0.89**
Normal BW group (Group II)	0.61**	0.53**	0.44**	0.47**	0.31**
High BW group (Group III)	0.64**	0.46**	0.24**	0.38**	0.25**
Total (non-grouped)	0.81**	0.58**	0.51**	0.54**	0.43**

AC = abdominal circumference; BPD = biparietal distance; BW = birth weight; FL = femur length; HC = head circumference; OFD = occipito-frontal diameter.
** Indicate that the correlation coefficient has statistical significance ($p < 0.01$).

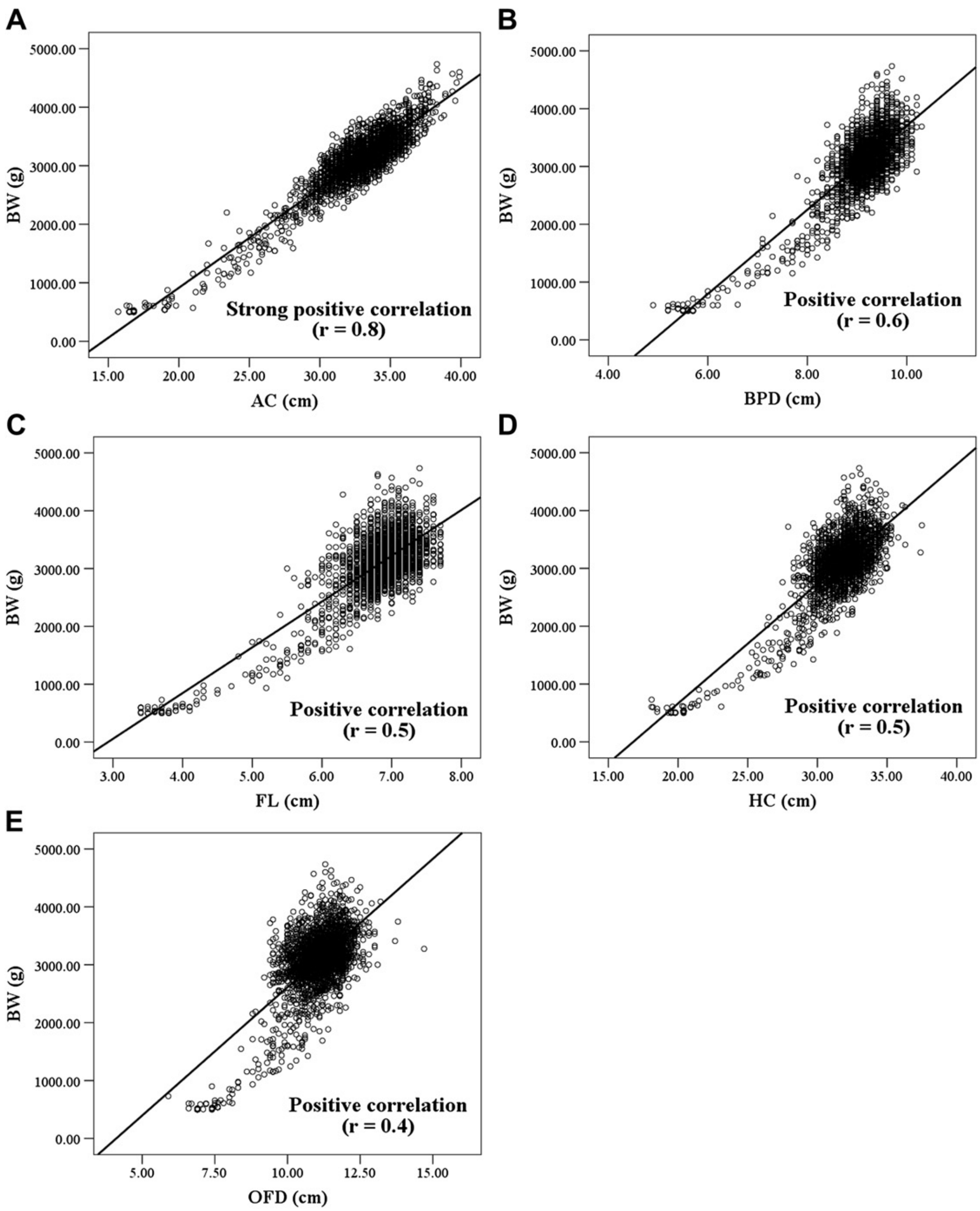

Fig. 5. The scatter plot of birth weight (BW) and ultrasonographic parameters. (A) AC = abdominal circumference (cm), r = 0.8; (B) BPD = biparietal diameter (cm), r = 0.6; (C) FL = femur length (cm), r = 0.5; (D) HC = head circumference (cm), r = 0.5; (E) OFD = occipito-frontal diameter (cm), r = 0.4; BW = birth weight; n = 2,127.

formulas. In contrast, the prerequisite of a normal distribution was not necessary in the ANN model.

In our ANN model in this series, back-propagation algorithm is an adaptive system that changes the weight by an amount proportional to the difference between the predicted output and the actual output [12,24]. It is composed of a large number of highly interconnected processing elements working in unison to solve nonlinear problems, which is based on the concept of gradient descent. This concept minimizes the error function between the EFW and the BW. The results of the neural network model for fetal weight estimation are found to be superior to the regression models. Farmer et al [13] and Chuang et al [14] proposed the back propagation network with the steepest descent algorithm that uses an orthogonal search direction to find the minimum point of the error function. The algorithm tends to coverage most slowly for small learning rate. If the gradient of the error function is changing too fast for making the learning rate large, the algorithm will become unstable. The advantage of using the ANN method is to reduce the heterogeneity among fetuses and among parameters on the EFW based on scientific statistical analysis.

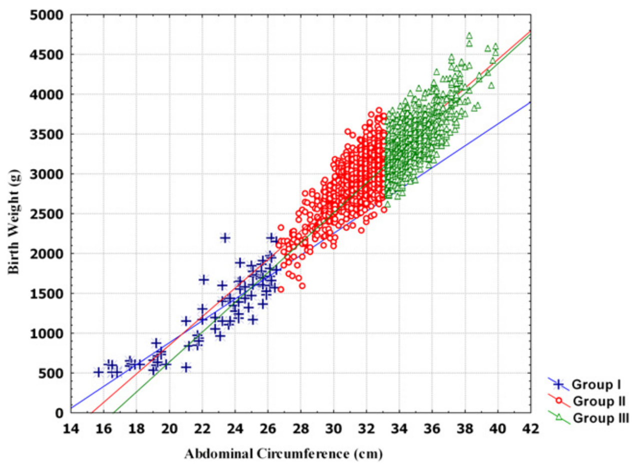

Fig. 6. The scatter plot of birth weight and abdominal circumference. Group I: the low birth-weight group, r = 0.92, n = 102. Group II: the normal birth-weight group, r = 0.61, n = 1,044. Group III: the high birth-weight group, r = 0.64, n = 981.

As AC is the most sensitive estimator for fetal size in this series, AC is selected as the basis of grouping fetuses. As Sabbagha et al [11] pointed out in their report, it is necessary to classify fetuses for reducing errors of EFW in small- and large-for-GA fetuses. For example, there were fetuses with an average AC measurement, but their BWs exceeded the 90th percentile. In that situation, the single predicting variable of AC is incapable and apparently insufficient to define all the heterogeneities generated from the gathered data. Therefore, the combination of USPs had been claimed to be better than the single parameter in the prediction of EFW [6, 25,26].

Concerning the impacts of FP and gender on EFW accuracy, previously published formulas [27] showed that EFW in breech presentation appears to be less accurate than EFW in vertex presentation. In addition, several studies [28–30] have reported that gender-specific and weight-range-specific method of EFW prediction provided greater accuracy. Furthermore, EFW is more accurate for male than female fetuses [31]. In our study, we utilized the stepwise estimation to extract the contribution of each independent parameter to the ANN. Our results showed that FP was a matter of affecting EFW prediction in the high BW group. In contrast, fetal gender was a contributive parameter for fetal weight estimation in the normal BW group and high BW group. Although the proposed procedures of grouping fetuses and parameters extraction, AC, BPD, SEX, FL, FP are included with the input of the ANN in the high BW group.

Table 6
The mean and standard deviation of MAPE and MAE of each group (n = 638).

Groups	MAPE ± SD (%)	MAE ± SD (g)
Low BW group (Group I)	7.1 ± 6.1%	83.5 ± 82.8
Normal BW group (Group II)	5.4 ± 4.6%	157.2 ± 131.2
High BW group (Group III)	4.9 ± 3.2%	166.7 ± 107.9
Total	5.3 ± 4.1%	157.9 ± 119.9

BW = birth weight; MAE = mean absolute error; MAPE = mean absolute percent error; SD = standard deviation.

Table 7
Comparison of mean absolute percent error (MAPE) and mean absolute error (MAE) of estimated fetal weight (EFW) methods (in testing set, n = 638).

Methods	MAPE ± SD (%)	p	MAE ± SD (g)	p
Hsieh's formula 1B (1987)	6.0 ± 4.6%	<0.01	173.2 ± 120.3	<0.01
Hsieh's formula 2B (1987)	6.5 ± 7.2%	<0.01	175.1 ± 120.4	<0.01
Hadlock's formula (1985)	7.4 ± 5.3%	<0.01	224.6 ± 169.0	<0.01
Proposed method	5.3 ± 4.1%	—	157.9 ± 119.9	—

EFW = estimated fetal weight; MAE = mean absolute error; MAPE = mean absolute percent error; SD = standard deviation.
The estimation error of the proposed method has a significant difference from three regression methods using the Friedman test and multiple-comparisons procedure.

Farmer et al focused mainly on the macrosomia group and used 13 variables as the input of an ANN model [13]. Their report showed that the MAPE of EFWs of 102 fetuses was 4.7 ± 3.9%. In our study, the MAPE of EFWs of 74 fetuses in the high BW group was 4.9 ± 3.2%. The MAPE of EFW estimation in our study was similar to that in Farmer's study. The accuracy of EFW using the proposed approach outperformed that using the Taiwanese conventional regression analysis with Hsieh's formula 1B, or that using Hsieh's formula 2B [6], and even that using the American conventional regression analysis [5]. In Chuang's study, when a BW was below 2500 g, the MAPE of EFW was raised to 9.6%. When a BW was above 4000 g, the MAPE of EWF was raised to 13.0%. In Table 6, the MAPE was 7.1 ± 6.1% in Group I and 4.9 ± 3.2% in Group III. The accuracy of our novel approach in the high BW group (Group III) was significantly improved in comparison with Chuang's study [14], while the accuracy of our novel approach in the low BW group (Group I) was similar.

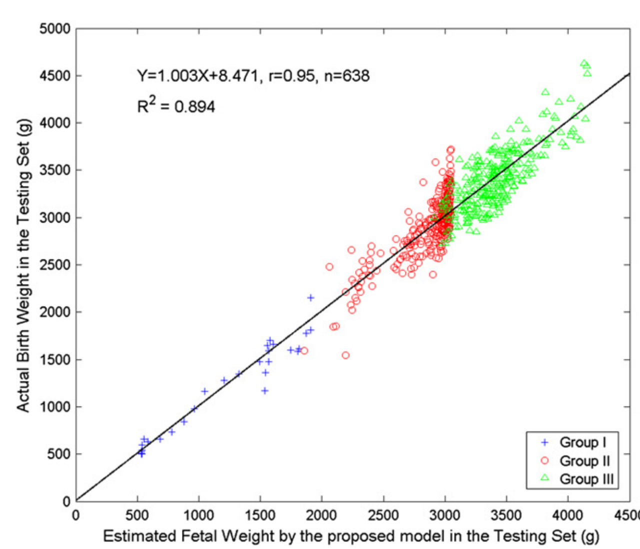

Fig. 7. The scatter plot of actual birth weight and estimated fetal weight by the proposed model in the testing set (n = 638).

In conclusion, our study is an attempt to develop scientific and statistical approaches to the selection and verification of parameters, as well as to develop an ANN model, to improve the accuracy of EFW applicable in a wider range of GAs and weights. The importance of our study is to consider and control the heterogeneity among the high variability and broad ranged parameters by statistics, and to choose the best parameter as a reasonable classified group to improve the accuracy. According to our results, this study has proved that the accuracy of the proposed approach outperforms those of the commonly used EFW formulas based on regression models. Our study attempted to investigate the significant assessment criteria and eliminate noise factors between ultrasound parameters and actual BW before delivery. The risk analysis affecting fetal outcomes, especially on the 3 days prior to delivery was investigated for providing valuable information in the clinical decision-making and management for parturition. In addition, fetal growth is very crucial to prenatal diagnosis and genetic consultation [32−37]. We believe our novel classification−ANN model of EFW estimation may assist clinicians to assess fetal growth in daily practice.

Acknowledgments

We are grateful for the assistance of all the staff of the Prenatal Ultrasound Lab of the Department of Obstetrics and Gynecology, National Cheng Kung University Hospital. This study was supported in part by a grant from the National Science Council, Executive Yuan, Taipei, Taiwan.

References

[1] Campbell S, Wilkin D. Ultrasonic measurement of the fetal abdominal circumference in the estimation of fetal weight. Br J Obstet Gynaecol 1975;82:689−97.
[2] Warsof SL, Gohan P, Berkowitz RL, Hobbins JC. The estimation of fetal weight by computer-assisted analysis. Am J Obstet Gynecol 1977;128:881−92.
[3] Shepard MJ, Richards VA, Berkowitz RL, Warsof SL, Hobbins JC. An evaluation of two equations for predicting fetal weight by ultrasound. Am J Obstet Gynecol 1982;142:47−54.
[4] Weiner CP, Sabbagha RE, Vaisrub N, Socol ML. Ultrasonic fetal weight prediction: role of head circumference and femur length. Obstet Gynecol 1985;65:812−7.
[5] Hadlock FP, Harrist RB, Shearman RS, Deter RL, Park SK. Estimation of fetal weight with the use of head, body and femur measurements. A prospective study. Am J Obstet Gynecol 1985;151:333−7.
[6] Hsieh FJ, Chang FM, Huang HC, Lu CC, Ko TM, Chen HY. Computer-assisted analysis for prediction of fetal weight by ultrasound: comparison of biparietal diameter, abdominal circumference and femur length. J Formos Med Assoc 1987;86:957−64.
[7] Benacerraf BR, Gelman R, Frigoletto Jr FD. Sonographically estimated fetal weights: accuracy and limitation. Am J Obstet Gynecol 1988;159:1118−21.
[8] Dudley NJ. A systematic review of the ultrasound estimation of fetal weight. Ultrasound Obstet Gynecol 2005;25:80−9.
[9] Hart NC, Hilbert A, Meurer B, Schrauder M, Schmid M, Siemer J, et al. Macrosomia: a new formula for optimized fetal weight estimation. Ultrasound Obstet Gynecol 2010;35:42−7.
[10] Gardosi J, Mongelli M, Wilcox M, Chang A. An adjustable fetal weight standard. Ultrasound Obstet Gynecol 1995;6:168−74.
[11] Sabbagha RE, Minogue J, Tamura RK, Hungerford SA. Estimation of birth weight by use of ultrasonographic formulas targeted to large-, appropriate-, and small-for-gestational-age fetuses. Am J Obstet Gynecol 1989;160:854−62.
[12] Baxt WG. Application of artificial neural networks to clinical medicine. Lancet 1995;346:1135−8.
[13] Farmer RM, Medearis AL, Hirata GI, Platt LD. The use of a neural network for the ultrasonographic estimation of fetal weight in the macrosomic fetus. Am J Obstet Gynecol 1992;166:1467−72.
[14] Chuang L, Hwang JY, Chang CH, Yu CH, Chang FM. Ultrasound estimation of fetal weight with the use of computerized artificial neural network model. Ultrasound Med Biol 2002;28:991−6.
[15] Hadlock FP, Deter RL, Harrist RB, Park SK. Fetal biparietal diameter: a critical re-evaluation of the relation to menstrual age using real-time ultrasound. J Ultrasound Med 1982;1:97−104.
[16] Parker AJ, Davies P, Newton JR. Assessment of gestational age of the Asian fetus by the sonar measurement of crown-rump length and biparietal diameter. Br J Obstet Gynaecol 1982;89:836−8.
[17] Wilcox M, Gardosi J, Mongelli M, Ray C, Johnson I. Birth weight from pregnancies dated by ultrasonography in a multicultural British population. BMJ 1993;307:588−91.
[18] Gorsuch RL. Exploratory factor analysis: its role in item analysis. J Pers Assess 1997;68:532−60.
[19] Lei PW, Dunbar SB, Kolen MJ. A comparison of parametric and nonparametric approaches to item analysis for multiple-choice tests. Educ Psychol Meas 2004;64:565−87.
[20] Hartigan JA, Wang MA. A K-means clustering algorithm. Appl Stat 1979;28:100−8.
[21] Modha DS, Spangler WS. Feature weighting in k-Means clustering. Machine Learning 2003;52:217−37.
[22] Billings SA, Voon WSF. A prediction-error and stepwise-regression estimation algorithm for non-linear systems. Int J Control 1986;44:803−22.
[23] Henderson DA, Denison DR. Stepwise regression in social and psychological research. Psychol Rep 1989;64:251−7.
[24] Cross SS, Harrison RF, Kennedy RL. Introduction to neural networks. Lancet 1995;346:1075−9.
[25] Chen CP, Chang FM, Chang CH, Lin YS, Chou CY, Ko HC. Prediction of fetal macrosomia by single ultrasonic fetal biometry. J Formos Med Assoc 1993;92:24−8.
[26] Chang FM, Ko HC, Lin YS, Yao BL, Wu CH, Kuo PL, et al. Clinical validation of two equations in antenatal prediction of Chinese fetal weight by ultrasonography. J Formos Med Assoc 1991;90:1086−92.
[27] Chauhan SP, Magann EF, Naef 3rd RW, Martin Jr JN, Morrison JC. Sonographic assessment of birth weight among breech presentations. Ultrasound Obstet Gynecol 1995;6:54−7.
[28] Davis RO, Cutter GR, Goldenberg RL, Hoffman HJ, Cliver SP, Brumfield CG. Fetal biparietal diameter, head circumference, abdominal circumference and femur length. A comparison by race and sex. J Reprod Med 1993;38:201−6.
[29] Siemer J, Hilbert A, Wolf T, Hart N, Müller A, Schild RL. Gender-specific weight estimation of fetuses between 2,501 and 3,999 g − new regression formulae. Fetal Diagn Ther 2008;24:304−9.
[30] Schild RL, Sachs C, Fimmers R, Gembruch U, Hansmann M. Sex-specific fetal weight prediction by ultrasound. Ultrasound Obstet Gynecol 2004;23:30−5.
[31] Melamed N, Ben-Haroush A, Meizner I, Mashiach R, Glezerman M, Yogev Y. The accuracy of sonographic weight estimation as a function of fetal sex. Ultrasound Obstet Gynecol 2011;38:67−73.
[32] Chen CP. Prenatal sonographic features of fetuses in trisomy 13 pregnancies (I). Taiwan J Obstet Gynecol 2009;48:210−7.
[33] Chen CP. Prenatal sonographic features of fetuses in trisomy 13 pregnancies (II). Taiwan J Obstet Gynecol 2009;48:218−24.
[34] Chen CP. Prenatal sonographic features of fetuses in trisomy 13 pregnancies (III). Taiwan J Obstet Gynecol 2009;48:342−9.
[35] Chen CP. Prenatal sonographic features of fetuses in trisomy 13 pregnancies (IV). Taiwan J Obstet Gynecol 2010;49:3−12.
[36] Chen CP. Prenatal diagnosis and genetic counseling for mosaic trisomy 13. Taiwan J Obstet Gynecol 2010;49:13−22.
[37] Chen CP. Prenatal diagnosis, fetal surgery, recurrence risk and differential diagnosis of neural tube defects. Taiwan J Obstet Gynecol 2008;47:283−90.