Frank P. Hadlock, MD • Ronald B. Harrist, PhD • Juan Martinez-Poyer, MD

In Utero Analysis of Fetal Growth: A Sonographic Weight Standard[1]

Regression analysis was used to develop an in utero fetal weight model from a population of 392 predominantly middle-class white patients with certain menstrual histories. There was a gradual increase in fetal weight from 35 g at 10 weeks to 3,619 g at 40 weeks, with uniform variance of ±12.7% (1 standard deviation) throughout gestation. When tested against the estimated weights of 1,771 chromosomally normal fetuses between 14 and 21 weeks, the mean percent difference was 0.8% and the average absolute percent error was 3.3%. When compared with actual delivery data for 163 fetuses in the group, the mean percent difference was 0.8% and the average absolute percent error was 1.1%. These data are compared with other prenatal weight curves obtained at ultrasound and with data from several large postnatal weight studies.

Index terms: Fetus, growth and development • Fetus, US studies, 856.1298

Radiology 1991; 181:129–133

[1] From the Department of Radiology, Baylor College of Medicine, 1 Baylor Plaza, Houston, TX 77030 (F.P.H.); the Department of Biometry, University of Texas School of Public Health, Houston (R.B.H.); and the Department of Obstetrics and Gynecology, University of Oklahoma Medical School, Oklahoma City (J.M.P.). Received April 8, 1991; revision requested May 15; revision received and accepted May 23. Supported in part by a grant from the Woman's Hospital of Texas Education and Research Foundation, Houston. **Address reprint requests to** F.P.H.

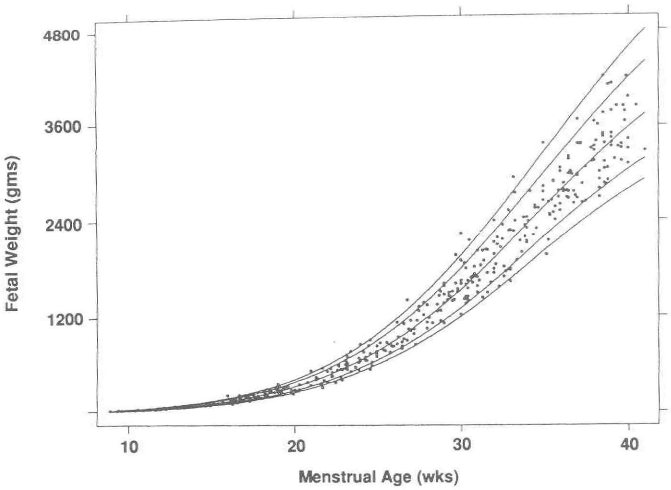
THE prenatal diagnosis of abnormal fetal growth patterns such as growth retardation and growth acceleration is important, since in utero recognition of these patterns may reduce the high perinatal morbidity and mortality associated with their occurrence. A number of postnatal weight standards have been developed for the detection of abnormal fetal growth, but there are problems associated with their use (1–10). In a recent review of the subject, Goldenberg et al outlined inconsistencies in these studies, including variations in population characteristics, sample size, source of data, geographic location, and criteria for exclusion, and in the mathematic methods used to determine normal boundaries (10). Another major limitation of these studies is that, by definition, prematurely born babies form the basis for normal values before 38 weeks, and these fetuses may or may not have been growing normally prior to delivery (11).

To address the issue of normal fetal growth before term, several authors have developed in utero weight standards at ultrasound (US) examination (12–14). Some of these models are based on longitudinal studies of individual fetal growth, with multiple measurements of the same fetus throughout pregnancy (12), while other models are based on cross-sectional analysis of data from each fetus obtained only once during pregnancy (13). Still other studies have been developed with use of the mean ultrasound values for the population under study (14) and therefore do not directly represent the growth of any individual fetus. Given these differences in methods, it is not surprising that the mean values and the confidence limits in these studies differ considerably.

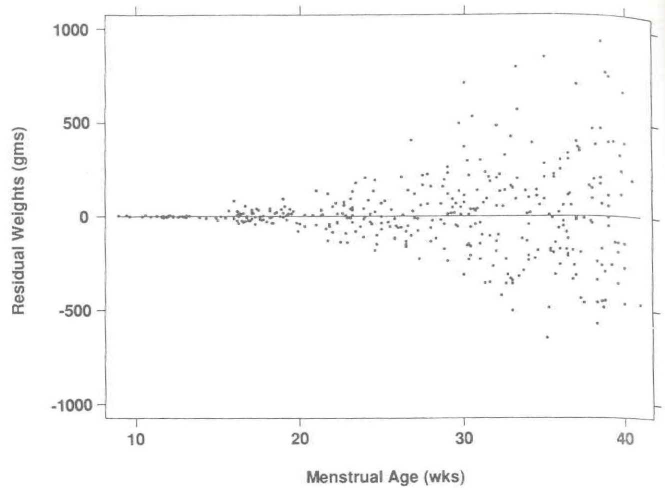
Goldenberg et al (10) noted that differences in study methods may be as or more important than population differences in defining the 10th and 90th percentile boundaries for normal fetal growth. This is consistent with the fact that previously published US studies have failed to demonstrate significant differences in basic in utero fetal measurements (biparietal diameter, head circumference, abdominal circumference, femur length) in fetuses from different racial and socioeconomic backgrounds (15). Goldenberg et al concluded that establishment of a single national standard for in utero growth would enable intercenter comparison among studies of risk factors, diagnostic tests, treatment, and long-term follow-up of fetuses and infants with abnormal growth patterns such as intrauterine growth retardation (10). The purpose of the study reported herein was to develop a national in utero fetal weight standard with US imaging.

MATERIALS AND METHODS

The study population consisted of 392 predominantly middle-class, pregnant, white women with certain menstrual dates, who were seen in our department between menstrual weeks 10 and 41 for evaluation with US. From this group, 361 fetuses formed the basis for a previously published article that focused on evaluation of menstrual age between 14 and 41 weeks (16). The population was expanded for the study reported herein to include uniform representation down to 10 menstrual weeks. The additional patients (n = 31) all had corroboration of menstrual age by measuring crown-rump length. In addition, all patients included in the study met the following criteria: (*a*) a well-known last menstrual period corroborated in the first trimester with either US or clinical evaluation; (*b*) a history of regular menses, with no history of use of oral contraceptives in the 3 months prior to the US study; (*c*) no history of maternal diseases known to affect fetal growth (eg, diabetes mellitus, hypertension); (*d*) no evidence of multiple gestation at US; and (*e*) no evidence of congenital anomalies at the time of the US study.



1.



2.

Figures 1, 2. **(1)** Graph demonstrates distribution of raw data of predicted values (n = 392) superimposed on the fetal weight curve developed from these fetuses. Boundaries are at the 3rd, 10th, 50th, 90th, and 97th percentiles. **(2)** Graph demonstrates the raw residuals across menstrual age for the 392 fetuses in this study. The residual equals the difference between the estimated weight for each fetus and the weight predicted for that fetus based on its age (by using the weight curve developed from these fetuses). While the residual weight in grams increased over time, the residuals were relatively constant at ±12.7% (1 standard deviation) when expressed as a percentage of the predicted weight.

Each fetus underwent US examination only once during gestation; measurements of biparietal diameter, head circumference, abdominal circumference, and femur length were obtained by physicians who used previously published methods (16). The gestation age for each fetus was calculated to the nearest 10th of a week on the basis of the mother's last normal menstrual period. For example, if the fetal age calculated on the basis of the last menstrual period was 39 weeks 3 days, it was coded as 39.4 weeks. The estimated fetal weight was calculated at the time of the US study by using a previously reported model based on measurements of biparietal diameter, head circumference, abdominal circumference, and femur length in combination (17). This model has been demonstrated to be free of systematic bias (mean percentage error, 0.1%) in our population of normally growing fetuses, and the magnitude of the random errors is relatively small (1 standard deviation = 7.4%) compared with that of other published models.

Regression analysis was used to evaluate the relationship between estimated weight in grams and menstrual age in weeks. The models tested included both log and non-log functions of estimated fetal weight on menstrual age, menstrual age squared, and menstrual age cubed. To be included in the equation, the coefficient for each variable had to be statistically significant at the .05 level. The optimal model was chosen on the basis of the largest coefficient of determination (R squared) and the smallest standard deviation and by inspection of the residuals for uniformity of variance. This model was used to calculate predicted normal weight values between 10 and 41 weeks.

To evaluate the accuracy of our predicted values at the lower and upper ends of the weight curve (where one might expect accuracy to be weakest), two further calculations were made. Data from 1,771 chromosomally normal patients seen in our department for amniocentesis were used to calculate estimated weights for each of those fetuses. The mean values for each week in gestation were calculated for that group (14–21 weeks), and those values were compared with the predicted 50th percentile values for the population in this study. To evaluate the weights at the other extreme of the weight curve, we retrieved birth weight data for those fetuses in the study who were delivered at term (38–42 weeks) and whose US examination had been performed prior to the 35th week of pregnancy and who therefore had no direct influence on the predicted values of the curve beyond 35

Table 1
In Utero Fetal Weight Standards at US

Menstrual Week	Percentiles (g)				
	3rd	10th	50th	90th	97th
10	26	29	35	41	44
11	34	37	45	53	56
12	43	48	58	68	73
13	55	61	73	85	91
14	70	77	93	109	116
15	88	97	117	137	146
16	110	121	146	171	183
17	136	150	181	212	226
18	167	185	223	261	279
19	205	227	273	319	341
20	248	275	331	387	414
21	299	331	399	467	499
22	359	398	478	559	598
23	426	471	568	665	710
24	503	556	670	784	838
25	589	652	785	918	981
26	685	758	913	1,068	1,141
27	791	876	1,055	1,234	1,319
28	908	1,004	1,210	1,416	1,513
29	1,034	1,145	1,379	1,613	1,724
30	1,169	1,294	1,559	1,824	1,649
31	1,313	1,453	1,751	2,049	2,189
32	1,465	1,621	1,953	2,285	2,441
33	1,622	1,794	2,162	2,530	2,703
34	1,783	1,973	2,377	2,781	2,971
35	1,946	2,154	2,595	3,036	3,244
36	2,110	2,335	2,813	3,291	3,516
37	2,271	2,513	3,028	3,543	3,785
38	2,427	2,686	3,236	3,786	4,045
39	2,576	2,851	3,435	4,019	4,294
40	2,714	3,004	3,619	4,234	4,524

Table 2
Comparison of Prenatal Weight Curves at US

Menstrual Week	Predicted 50th Percentile Weight (g)		
	Deter et al (12)*	Hadlock†	Ott (13)†
14	82	93	93
16	122	146	162
18	204	223	259
20	327	331	392
22	492	478	562
24	699	670	773
26	948	913	1,026
28	1,239	1,210	1,321
30	1,572	1,559	1,655
32	1,947	1,953	2,027
34	2,363	2,377	2,431
36	2,821	2,813	2,864
38	3,321	3,236	3,318
40	3,863	3,619	3,788

*Longitudinal study.
†Cross-sectional study.

Table 3
Comparison of Prenatal and Postnatal Weight Curves

Menstrual Week	Fetal Weight (g)		
	Williams et al (1) (postnatal)	Hadlock (prenatal)	Babson et al (2)* (postnatal)
22	513 (30)	478 (13)	NA
24	675 (30)	670 (13)	NA
26	882 (29)	913 (13)	NA
28	1,143 (28)	1,210 (13)	1,118 (30)
30	1,484 (27)	1,559 (13)	1,458 (23)
32	1,920 (25)	1,953 (13)	1,861 (21)
34	2,394 (21)	2,377 (13)	2,298 (19)
36	2,849 (17)	2,813 (13)	2,697 (15)
38	3,227 (14)	3,236 (13)	3,171 (14)
40	3,462 (13)	3,619 (13)	3,448 (13)

Note.—Numbers in parentheses are standard deviations in percentages. All charts are data of a white population (analysis of data of Williams et al is restricted to only white patients for purposes of comparison).
*NA = no available data in this time period.

weeks. The mean delivery weights in this group were calculated for each week, rounded to the nearest week, for comparison with the predicted values.

RESULTS

The optimal model was a natural log model of weight in grams on menstrual age (in weeks) and menstrual age squared: Log n weight (g) = 0.578 + 0.332 MA − 0.00354 MA^2, where MA is menstrual age (standard deviation = 0.12, R^2 = 99.1%).

In Table 1, the predicted values for the 3rd, 10th, 50th, 90th, and 97th percentiles are listed for each week in gestation; these are presented in graphic form with the raw data superimposed in Figure 1. The distribution of the differences among the individual weight estimations and those predicted with the model are demonstrated in Figure 2; as one would expect, there is greater variability (in grams) as pregnancy advances, but this variability was uniform when expressed as a percentage of the weight predicted by the model (1 standard deviation = ±12.7%). In Tables 2 and 3, the predicted values for each week in gestation are compared with data from prenatal and postnatal charts from institutions with a similar population base. The potential reasons for the observed differences will be addressed in the Discussion section below.

The comparison of predicted weights obtained in this study in the middle trimester (14–21 weeks) with estimated weights for the 1,771 chromosomally normal fetuses demonstrated remarkable similarities (Fig 3). The mean percentage difference between the predicted and the observed weights was 0.8% (range, 0.06%–2.7%), and the average absolute percentage error was 3.3%. The comparison of predicted term weights with those actually observed at delivery was also impressive. The average age at the time of the initial US study in this group (n = 163) was 23 weeks (1 standard deviation = 5.9 weeks), and the average age at delivery was 39.9 weeks (1 standard deviation = 1.07 weeks). The predicted 50th percentile weights at term (38–41 weeks) compared favorably with the observed mean birth weights at term (Fig 3). The mean percentage difference between the predicted and observed weights was 0.8% (range, 0.06%–2.7%), and the average absolute percentage error was 1.1%. The predicted and observed term birth weights reported herein are higher than those published in earlier postnatal studies (1–9) but are comparable to data published in the most recent postnatal studies reported in the literature (18) (Table 4).

DISCUSSION

Among the many problems associated with comparing current postnatal weight charts, as outlined by Goldenberg et al (10), the most important is probably the proper assignment of menstrual age, since inaccurate dating will result in the inappropriate entry of data at the wrong menstrual week. These errors, which frequently have a magnitude of 2–4 weeks (since implantation bleeding or early first trimester bleeding can be mistaken for a menstrual period), are commonly found in large postnatal studies because certain dates are usually known in less than 90% of patients (1–2). Ultimately, errors of this type will have some impact on the mean value for the week in question but will more likely expand the variability around the mean. Our study, in which dates were known with certainty, should not be affected by these errors.

Another inconsistency that can affect the mean values at each menstrual week is the way in which menstrual age is reported (ie, rounded to the nearest week, reported in completed weeks, or reported to the nearest 10th of a week). An example of this problem would be the fetus that is 40 weeks 4 days on the basis of a well-known last menstrual period. If the age were rounded to the nearest week, the fetus would be classified as a 41-week fetus. If the age were reported in completed weeks, as recommended by the American Academy of Pediatrics (19), the fetus would be classified as a 40-week fetus.

If the age were reported to the nearest 10th of a week, the fetus would be properly classified as a 40.6-week fetus. We chose the latter approach not because of the implied precision of the last menstrual period but simply because we thought dilution of the age assignment might make the regression equation and its predicted values at each week less meaningful. When comparing our predicted values with those from postnatal studies, these differences in the way menstrual age is reported should be kept in mind.

Another major problem with most postnatal studies that has been addressed in the literature is the question of whether prematurely born babies can be assumed to have grown normally prior to their delivery (11). It is currently believed that prematurely born babies have a tendency to be undergrown (11), although clearly fetuses that are either small or large for gestational age can undergo premature birth (but for different reasons). The expected end result of this phenomenon would be lower mean values at each menstrual week and significantly broadened variability about the mean. Indeed, the latter finding can be observed in virtually all large postnatal growth studies in that the variability (when expressed as a percentage of the predicted value) is considerably larger preterm than at term (Table 3). For example, at 28 weeks the standard deviations of the predicted value for the large postnatal studies of Williams et al (1) and Babson et al (2) are 28% and 30%, respectively, while the standard deviation at term in both studies is 13%. On the other hand, the standard deviation in our study was uniform at 13% throughout pregnancy. The fact that these differences are due to inaccurate assignment of menstrual age is supported by the recent work of Secher et al (20), who demonstrated uniform variance in a group of fetuses whose menstrual dates had been confirmed at early US.

For purposes of comparison, our data should be most comparable with the data in the prenatal US studies of Deter et al (12) and Ott (13) (Table 2) and the postnatal studies of Williams et al (1) and Babson et al (2) (Table 3). The predicted values from Deter et al are based on a quadratic function derived from longitudinal data of 20 white fetuses seen at US performed at sea level; the predicted values are very similar to those obtained in this study, with a mean percentage difference of −1.1%. These differences may be the result in part of the fact that Deter et al used the weight prediction model of Warsof et al (21) that was based on measurement of biparietal diameter and abdominal circumference; this measurement has been demonstrated to result in systematically underestimated fetal weight (22). The difference in predicted birth weight at 40 weeks between our study and that of Deter et al is probably related to the fact that Deter et al had very little data beyond 39 weeks when their model was developed.

The study by Ott (13) was a cross-sectional analysis of 402 white patients seen in St Louis; these data are remarkably similar to those for our population. When looking at the predicted values for each week of gestation, it is clear that the values obtained by Ott are typically larger than ours, averaging approximately 6.7% larger throughout gestation (15–40 weeks). The only significant difference in Ott's methods compared with ours is his use of a model for predicting fetal weight that has been demonstrated to result in systematically overestimated weight in utero by 3% in fetuses weighing less than 3,000 g (23) and at term by 6.1% in fetuses seen at US imaging at 24 weeks. We believe this is the most likely explanation for the differences observed among our estimates of normal fetal weight at each week of gestation prior to term. We cannot readily explain the difference in variances observed in our study and Ott's, which is considerable, but the higher random errors reported by Ott (8.9%) may have played some role.

We chose the studies of Williams et al (1) and Babson et al (2) as postnatal studies for comparison with our study because both of those studies exam-

Table 4
Comparison of Fetal Birth Weights at Term

Menstrual Week	Weights (g) Hadlock Predicted*	Hadlock Observed	McLean et al (18) Observed†
38	3,236	3,234	ND
39	3,435	3,469	3,438
40	3,619	3,598	3,560
41	ND	3,686	3,674

Note.—ND = No data or insufficient data at these time points.
*The 50th percentile weight for age based on the equation developed in this study.
†The actual mean weights at birth.

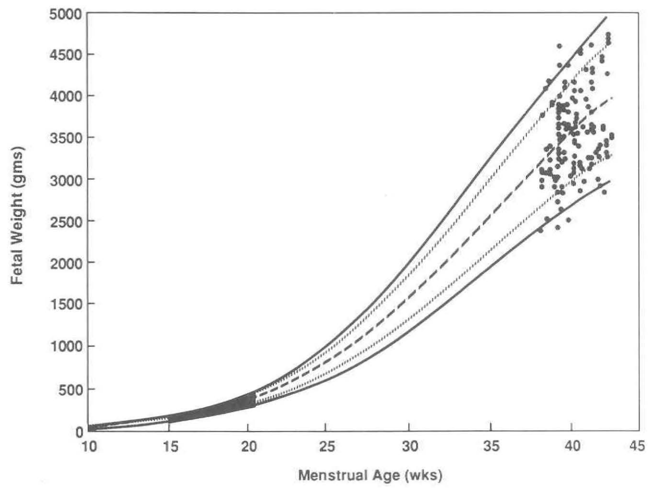


Figure 3. Graph demonstrates the predicted weights for 1,771 chromosomally normal fetuses (15–21 weeks) and the actual birth weights of 163 fetuses plotted on the weight curve developed from the study reported herein. The actual birth weight of 2.5% of babies was below the 3rd percentile and above the 97th percentile; the actual birth weight of 8.5% of the neonates fell below the 10th percentile and above the 90th percentile. Boundaries are at the 3rd, 10th, 50th, 90th, and 97th percentiles.

ined a large, predominantly white population seen at sea level (Table 3). Between 28 and 40 weeks, the mean data of Babson et al are consistently lower than our data by an average of 2%, and the mean data of Williams et al are lower than our data by approximately 1%. These relatively small differences may be related in the preterm period to subtle growth disturbances in those babies born prematurely (11). We believe that errors in assignment of fetal age account for the large differences in birth weights at term when our data are compared with the data of Babson et al and Williams et al. This issue has been addressed in a recently published article by McLean et al (18), who reported much higher birth weights at term (in comparison with their previously published data and those of others) in a population of fetuses whose menstrual age was confirmed early in pregnancy at US. Predicted normal weight values at term in our study were comparable to those observed in our population at term and are quite similar to the data recently reported by McLean et al (Table 4).

The differences in variances (percentage of standard deviation) observed among our data and those of Babson et al and Williams et al (Table 3) are probably the result of abnormal premature growth and errors in assignment of fetal age. For example, in the study of Williams et al (1) only 80% of the population had certain menstrual dates, while in the study of Babson et al (2) only 60% of the population knew the precise day of the beginning of the last menstrual period. The fact that the differences in variance are the result of inaccurate assignment of menstrual age is further supported by findings reported in the postnatal study of Secher et al (20), which to our knowledge represents the first postnatal weight study in which assignment of gestation age was determined precisely at an early US examination. In that study, there was a rather uniform variance of approximately ±16% at the 10th and 90th percentiles, which is very similar to the variance observed in our study (±17%). This is further supported by the observation that the variance in virtually all postnatal studies is relatively uniform after 38 weeks if very large numbers of fetuses were included. Indeed, in most postnatal studies, over 90% of the population base comprises fetuses born at term.

We believe that prenatal US studies can enable accurate measurement of the boundaries of normal fetal growth. Moreover, we believe that the weight curve presented herein should be equally applicable to populations of different racial and socioeconomic origins, since our study and those of others suggest that there are no significant population differences in fetal growth prior to term (15). This is not to say, however, that this common genetic potential for preterm growth cannot be altered in a systematic way by factors that are unique to a specific population (eg, extremely high altitude [5]). Indeed, given the very large numbers of term fetuses included in most postnatal studies, the differences observed among studies at term may represent true population differences if the accuracy of the last menstrual period and the method of reporting menstrual age are common to both studies.

In summary, we attempted to define normal predicted values and confidence limits for fetal weight as estimated in utero at US. It is our hope that proper use of these data will lead to earlier recognition of abnormal growth patterns such as growth retardation and/or acceleration. In view of the distribution of our normal data on this weight curve (Figs 3, 4), we believe that the 3rd and 97th percentile boundaries should be used as confidence limits for normal growth. Because these confidence limits are relatively narrow in the preterm period, as compared with limits reported in postnatal studies, it is imperative that menstrual age be known before these data are used to diagnose abnormal fetal growth. Indeed, we cannot recommend use of these limits for fetuses whose ages have not been verified at US during early pregnancy, since minor errors in assignment of fetal age may lead to false diagnosis of altered fetal growth and may result in errors of management. ■

References

1. Williams RL, Creasy RK, Cunningham GC. Fetal growth and perinatal viability in California. Obstet Gynecol 1982; 59:624–632.
2. Babson SG, Behrman RE, Lessel R. Fetal growth: live-born birth weights for gestational age of white middle class infants. Pediatrics 1970; 45:937–944.
3. Gruenwald P. Growth of the human fetus. Am J Obstet Gynecol 1966; 94:1112–1119.
4. Usher R, McLean F. Intrauterine growth of live-born Caucasian infants at sea level: standards obtained from measurements in seven dimensions of infants born between 25 and 44 weeks of gestation. Pediatrics 1969; 74:901–910.
5. Lubchenco LO, Hansman C, Dressler M, et al. Intrauterine growth as estimated from live-born birth-weight data at 24 to 42 weeks of gestation. Pediatrics 1963; 32:793–800.
6. Miller HC, Merritt TA. Fetal growth in humans. Chicago: Year Book, 1979; 31–57, 127–141.
7. Brenner WE, Edelman DA, Hendricks CH. A standard of fetal growth for the United States of America. Am J Obstet Gynecol 1976; 126: 555–564.
8. Naeye RL, Dixon JB. Distortions in fetal growth standards. Pediatr Res 1978; 12:987–991.
9. Carr-Hill RA, Pritchard CW. Reviewing birthweight standards. Br J Obstet Gynaecol 1983; 90:718–725.
10. Goldenberg RL, Cutter GR, Hoffman HJ, Foster JM, Nelson KG, Hauth JC. Intrauterine growth retardation: standards for diagnosis. Am J Obstet Gynecol 1989; 161:271–277.
11. Weiner CP, Sabbagha RE, Vaisrub N, Depp R. A hypothetical model suggesting suboptimal intrauterine growth in infants delivered preterm. Obstet Gynecol 1985; 65:323–326.
12. Deter RL, Harrist RB, Hadlock FP, et al. Longitudinal studies of fetal growth with the use of dynamic image ultrasonography. Am J Obstet Gynecol 1982; 143:545–554.
13. Ott WJ. The diagnosis of altered fetal growth. Obstet Gynecol Clin North Am 1988; 15:237–263.
14. Jeanty P, Cantraine F, Romero R, et al. A longitudinal study of fetal weight growth. J Ultrasound Med 1984; 3:321–325.
15. Hadlock FP, Harrist RB, Shah YP, et al. Sonographic fetal growth standards: are current data applicable to a racially mixed population? J Ultrasound Med 1990; 9:157–160.
16. Hadlock FP, Deter RJ, Harrist RB, Park SK. Estimating fetal age: computer-assisted analysis of multiple fetal growth parameters. Radiology 1984; 152:497–501.
17. Hadlock FP, Harrist RB, Sharman RS, Deter RL, Park SK. Estimating of fetal weight with the use of head, body and femur measurements: a prospective study. Am J Obstet Gynecol 1985; 151:333–337.
18. McLean FH, Boyd ME, Usher RH, Kramer MS. Postterm infants: too big or too small? Am J Obstet Gynecol 1991; 164:619–624.
19. Committee on Fetus and Newborn. Nomenclature for duration of gestation, birth weight and intrauterine growth. Pediatrics 1967; 39: 935–939.
20. Secher NJ, Hansen PK, Lenstrup C, Pedersen-Bjergaard L, Eriksen PS. Birthweight-for-gestational age charts based on early ultrasound estimation of gestational age. Br J Obstet Gynaecol 1986; 93:128–134.
21. Warsof SL, Gohari B, Berkowitz RL, Hobbins JC. The estimation of fetal weight by computer assisted analysis. Am J Obstet Gynecol 1977; 128:881–892.
22. 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.
23. Ott WJ, Doyle S, Flamm S, Wittman J. Accurate ultrasonic estimation of fetal weight. Am J Perinatol 1986; 3:307–310.
24. Ott WJ. Defining altered fetal growth by second-trimester sonography. Obstet Gynecol 1990; 75:1053–1059.