

New Ballard Score, expanded to include extremely premature infants

J. L. Ballard, MD, J. C. Khoury, MSc, K. Wedig, MD, L. Wang, MD,
B. L. Eilers-Walsman, MD, and R. Lipp, MD

From the Division of Neonatology, Department of Pediatrics, and the Division of Epidemiology and Biostatistics, Department of Environmental Health, University of Cincinnati College of Medicine, the Children's Hospital Research Foundation, the Crosley Memorial Nursery of the University of Cincinnati Hospital, the Regional Intensive Care Nursery of the Children's Hospital Medical Center, the Bethesda Hospital Special Care Nursery, and the Perinatal Center of Good Samaritan Hospital, Cincinnati, Ohio

The Ballard Maturation Score was refined and expanded to achieve greater accuracy and to include extremely premature neonates. To test validity, accuracy, interrater reliability, and optimal postnatal age at examination, the resulting New Ballard Score (NBS) was assessed for 578 newly born infants and the results were analyzed. Gestational ages ranged from 20 to 44 weeks and postnatal ages at examination ranged from birth to 96 hours. In 530 infants, gestational age by last menstrual period was confirmed by agreement within 2 weeks with gestational age by prenatal ultrasonography (C-GLMP). For these infants, correlation between gestational age by NBS and C-GLMP was 0.97. Mean differences between gestational age by NBS and C-GLMP were 0.32 ± 1.58 weeks and 0.15 ± 1.46 weeks among the extremely premature infants (<26 weeks) and among the total population, respectively. Correlations between the individual criteria and C-GLMP ranged from 0.72 to 0.82. Interrater reliability of NBS, as determined by correlation between raters who rated the same subgroup of infants, was 0.95. For infants less than 26 weeks of gestational age, the greatest validity (97% within 2 weeks of C-GLMP) was seen when the examination was performed before 12 hours of postnatal age. For infants at least 26 weeks of gestational age, percentages of agreement with C-GLMP remained constant, averaging 92% for all postnatal age categories up to 96 hours. The NBS is a valid and accurate gestational assessment tool for extremely premature infants and remains valid for the entire newborn infant population. (J PEDIATR 1991;119:417-23)

Accurate morbidity and mortality statistics and comparisons between subpopulations of extremely low birth weight infants depend on precise identification of the degree of prematurity, for which a reliable estimate of gestational age

is vital.¹ Many premature infants are born to mothers with an unreliable menstrual history or no prenatal care; the clinical assessment of fetal maturation is thus frequently the only available measure of gestational age. Current litera-

Supported in part by National Institutes of Health grants HD 11725 (Diabetes in Pregnancy), HD 20748 (Perinatal Emphasis Research Center), and HD 07200 (Research Training in Perinatology), and by U.S. Public Health Service grants MCH MCT 174 (Training in Perinatal Care and Research) and MCH ODH 173 (CHD Maternal and Infant Care Project).

Submitted for publication Feb. 20, 1991; accepted April 18, 1991.
Reprint requests: Jeanne L. Ballard, MD, University of Cincinnati College of Medicine, 231 Bethesda Ave., Cincinnati, OH 45267-0541.

9/23/30213

C-GLMP	Gestational age by last menstrual period confirmed (± 2 weeks) by ultrasonography
GLMP	Gestational age by last menstrual period
GUS	Gestational age by ultrasonography
NBS	New Ballard Score

ture states that gestational age is the most significant predictor of survival and calls for an accurate tool for assessment of gestational age in very low birth weight neonates.² To this end, we have expanded the score most widely used

Neuromuscular Maturity

	-1	0	1	2	3	4	5
Posture							
Square Window (wrist)	>90°	90°	60°	45°	30°	0°	
Arm Recoil		180°	140°-180°	110°-140°	90°-110°	<90°	
Popliteal Angle	180°	160°	140°	120°	100°	90°	<90°
Scarf Sign							
Heel to Ear							

Physical Maturity

Skin	sticky friable transparent	gelatinous red, translucent	smooth pink, visible veins	superficial peeling &/or rash, few veins	cracking pale areas rare veins	parchment deep cracking no vessels	leathery cracked wrinkled
Lanugo	none	sparse	abundant	thinning	bald areas	mostly bald	
Plantar Surface	heel-toe 40-50mm: -1 <40mm: -2	>50mm no crease	faint red marks	anterior transverse crease only	creases ant. 2/3	creases over entire sole	
Breast	imperceptible	barely perceptible	flat areola no bud	stippled areola 1-2mm bud	raised areola 3-4mm bud	full areola 5-10mm bud	
Eye/Ear	lids fused loosely: -1 tightly: -2	lids open pinna flat stays folded	sl. curved pinna; soft; slow recoil	well-curved pinna; soft but ready recoil	formed & firm instant recoil	thick cartilage ear stiff	
Genitals male	scrotum flat, smooth	scrotum empty faint rugae	testes in upper canal rare rugae	testes descending few rugae	testes down good rugae	testes pendulous deep rugae	
Genitals female	clitoris prominent labia flat	prominent clitoris small labia minora	prominent clitoris enlarging minora	majora & minora equally prominent	majora large minora small	majora cover clitoris & minora	

Maturity Rating

score	weeks
-10	20
-5	22
0	24
5	26
10	28
15	30
20	32
25	34
30	36
35	38
40	40
45	42
50	44

Fig. 1. Expanded NBS includes extremely premature infants and has been refined to improve accuracy in more mature infants.

for estimating gestational age³ to include gestational ages of less than 26 weeks. This study was undertaken to validate the new tool, to test its interrater reliability, and to determine the optimal postnatal ages for gestational examination of extremely premature infants and of the entire infant population.

METHODS

Expansion of the gestational assessment tool involved careful examination of extremely premature infants, including many who were determined to be nonviable by their attending physicians. We sought neurologic and physical characteristics that might differentiate extremely premature infants from more mature infants. It was noted that flexibility was virtually absent at the wrist and that passive

flexor tone was virtually absent at the knee, shoulder, and hip. This information enabled us to expand four of the neuromuscular criteria by adding a score of -1. On careful physical examination, the extremely immature infants were noted to have sticky, transparent skin, to be free of any lanugo, and to have imperceptible breast markings and nearly undifferentiated genitalia. These observations in turn allowed expansion of the relevant physical items with the addition of a score of -1 (Fig. 1).

Two physical maturational criteria of the existing score, "plantar surface" and "ear," were expanded. The relationship between foot length and gestational age has been shown by Streeter⁴ and confirmed by Hern.⁵ According to these authors, fetuses between 20 and 26 weeks have foot lengths between 30 and 50 mm. Foot length was measured from the

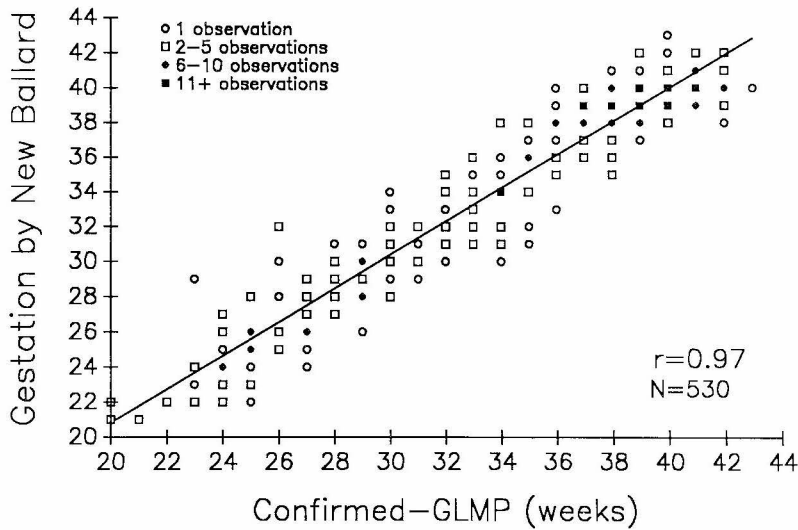


Fig. 2. Regression for gestational age by NBS versus C-GLMP.

Table I. Population demographic data (N = 578)

Parameter	Value	Range
Gender (% male)	50.4	
Race (% white)	78.1	
Apgar score <7 at 1 min (%)	33.6	
Apgar score <7 at 5 min (%)	15.8	
Small for gestational age (%)	7.0	
Large for gestational age (%)	16.5	
Gestational age by date (wk)	35.2 ± 5.8*	20-44
Gestational age by ultrasonography (wk)	35.0 ± 5.8*	18-48
Gestational age by NBS (wk)	35.5 ± 5.5*	21-43
Weight (gm)	2626 ± 1224*	230-55/70
Gestational week at ultrasound examination†	14.3 ± 5.7*	5-32
Postnatal age at examination (hr)	21.3 ± 17.4*	0.2-96

*Values are expressed as mean ± SD.

†Gestational week at ultrasound examination was available for 201 infants.

tip of the great toe to the back of the heel. Values of -1 and -2 were assigned to foot lengths of 40 to 50 mm and less than 40 mm, respectively. This information was incorporated into the criterion labeled "plantar surface." Eyelid separation normally occurs toward the end of fetal life or during the seventh month of human gestation.^{6,7} We defined loosely fused eyelids as closed but able, in the case of one or both lids, to be partly separated by gentle traction, and we defined tightly fused eyelids as bilaterally inseparable by gentle traction. Values of -1 and -2 were assigned for loosely and tightly fused eyelids, respectively. The criterion originally labeled "ear" was renamed "eye/ear" (Fig. 1). The remainder of the criteria were refined and clarified in an attempt to correct reported inaccuracies of the original score.^{8,9}

To validate the new tool, we examined a convenience sample of 681 live-born infants of all gestational ages,

without knowledge of menstrual history or ultrasonographic results. Only the 578 infants whose mothers had both undergone prenatal ultrasound examinations and provided menstrual histories were used for analysis. Assessments were performed before the infants were 96 hours of postnatal age by neonatologists specifically trained in performing the examination. Birth weight percentiles were recorded according to Usher and McLean's intrauterine growth curves.¹⁰ Demographic data are shown in Table I. The study was approved by the institutional review boards of University Hospital, Children's Hospital Medical Center, Bethesda Hospital, and Good Samaritan Hospital of Cincinnati.

Hypotheses to be tested were as follows: the New Ballard Score (1) validly and accurately reflects the gestational age of the extremely premature infant (less than 26 weeks of gestation) and (2) validly and accurately reflects the gesta-

Table II. Validity and accuracy of New Ballard Score at various gestational ages

GLMP (wk)	Validity (agreement with GLMP)			Accuracy			
	NBS (%)	GUS (%)	n	p	NBS minus C-GLMP (wk)	n	p
<26	84	93	61	0.09	0.32 ± 1.58	57	0.14
26-31	87	93	89	0.12	0.34 ± 1.68	83	0.07
32-37	79	90	127	0.02	0.60 ± 1.70	114	0.001
38+	94	92	301	0.42	-0.13 ± 1.19	276	0.08
TOTAL	88	92	578	0.06	0.15 ± 1.46	530	0.02

Validity is percentage of infants with NBS/GUS within 2 weeks of agreement with GLMP; accuracy is mean (±SD) difference (in weeks) between NBS and C-GLMP.

Table III. Correlations with C-GLMP for individual criteria and for total score

	r
Neuromuscular maturity	
Posture	0.82
Square window (wrist)	0.79
Arm recoil	0.71
Popliteal angle	0.74
Scarf sign	0.82
Heel to ear	0.81
Total neuromuscular	0.94*
Physical maturity	
Skin	0.75
Lanugo	0.81
Plantar surface	0.72
Breast	0.80
Eye/ear	0.73
Genitals	0.82
Total physical	0.95*
TOTAL SCORE	0.97*

r, Spearman rank correlations between individual criteria and C-GLMP (n = 530).

Significance: $p \leq 0.0001$ for all correlations.

*Pearson correlation.

tional age of the entire population of live-born infants. Validity was defined as follows: (1) the correlation between gestational age by NBS and gestational age by last menstrual period must be at least as high as the correlation between gestational age by ultrasonography and GLMP; and (2) the percentage of infants for whom gestational age by NBS is within 2 weeks of GLMP must be no different from the percentage of infants for whom GUS is within 2 weeks of GLMP. Accuracy was defined as a mean difference between gestational age by NBS and gestational age by confirmed GLMP that is equivalent to zero. To assess interrater reliability, a subgroup of 67 infants, examined consecutively within a 2-month period and ranging in gestational age from 24 to 41 weeks, was tested by two independent observers who made their assessments within 4 hours of one another. Postnatal age was recorded at examination.

Statistical analysis. Gestation was measured in completed weeks. Correlation coefficients and the Fisher Exact Test were used for the validation defined previously. Accuracy was analyzed with a paired *t* test. Interrater reliability of NBS was determined by using the κ statistic and correlation. Individual criteria were correlated with confirmed menstrual dates and with GUS by using the Spearman rank correlation. Optimal postnatal ages at examination for both the extremely premature and the entire infant population were examined by using the Fisher Exact Test and chi-square test. Data were managed and analyzed with the SAS statistical package (SAS Institute Inc., Cary, N.C.).¹¹ A *p* value less than 0.05 was considered significant. Data are expressed as mean ± standard deviation unless otherwise stated.

RESULTS

The total number of infant examinations used for analysis was 578. Correlations between GLMP and both the gestational age by NBS and the GUS were 0.96. The percentage of infants for whom gestational age by NBS was within 2 weeks of agreement with GLMP was compared with the percentage of infants for whom GUS was within 2 weeks of agreement with GLMP. There was no significant difference between these percentages for either the extremely premature or the total infant populations. Infants examined at 32 to 37 weeks of gestational age had a greater percentage of 2-week agreement of GLMP with GUS than with NBS. The NBS overestimated the gestational age of these infants by 0.6 ± 1.7 weeks ($p = 0.001$; Table II). The difference between gestational age by NBS and C-GLMP was 0.32 ± 1.58 weeks (p not significant) for extremely premature infants and 0.15 ± 1.46 weeks ($p = 0.02$) for all gestational ages (Table II).

In the 530 infants whose mothers provided confirmed menstrual history, the correlation between NBS and C-GLMP was 0.97 (Fig. 2). Correlations between individual criteria and C-GLMP ranged from 0.71 to 0.82 (Table III).

Correlation between raters for the total score was 0.95. Spearman correlation coefficients for interrater agreement

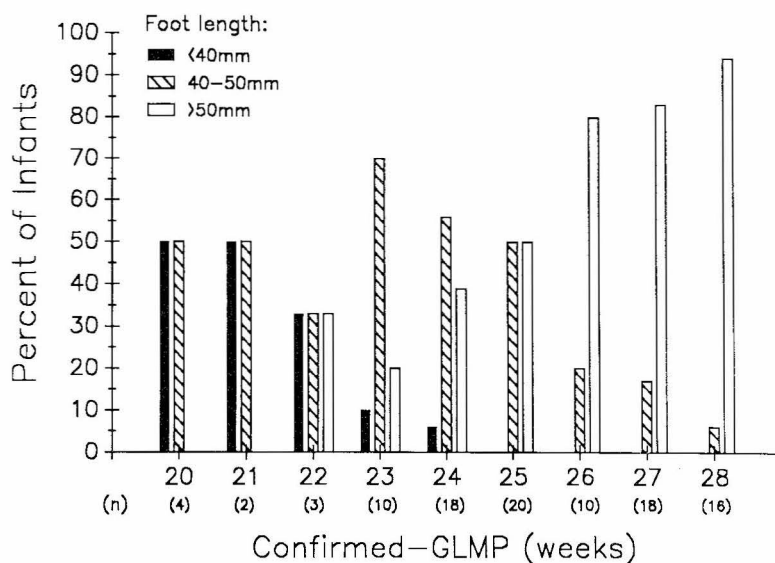


Fig. 3. Distribution of foot length at gestational weeks 20 through 28 by C-GLMP (confirmed-GLMP).

Table IV. Interrater reliability for individual criteria (n = 67)*

	Spearman correlation	Agreement (% ± 1 score point)	Kappa (± 1 score point)
Posture	0.81	86	0.93
Square window	0.84	98	0.98
Arm recoil	0.61	92	0.89
Popliteal angle	0.72	79	0.77
Scarf sign	0.80	84	0.80
Heel to ear	0.74	85	0.85
Skin	0.68	98	0.98
Lanugo	0.82	94	0.89
Plantar surface	0.90	100	1.00
Breast	0.91	97	0.96
Eye/ear	0.72	100	1.00
Genitals	0.87	100	1.00
TOTAL	0.95	86†	0.93†

Significance: $p < 0.01$ for all correlations and κ statistics. There was significant agreement between two independent observers for all criteria.

*Score was performed on a subset of infants representing gestational ages from 24 through 41 weeks.

†Two-week agreement was used for the total score.

on individual criteria ranged from 0.61 to 0.91. The percentage of agreement within 1 score point ranged from 79% to 100%, with associated κ values ranging from 0.79 to 1.00 (Table IV).

The two new criteria, foot length and eyelids status, were further analyzed. Distribution of foot length is illustrated at individual weeks 20 through 28 in Fig. 3. We analyzed the eyelid fusion criterion in infants of up to 28 completed weeks of gestational age. All infants less than 22 weeks of gestational age had tightly fused eyelids and all those at least 28 weeks of gestational age had unfused eyelids. Proportions of tightly fused, loosely fused, and unfused eyelids at gestational ages between 20 and 28 weeks are shown in Fig. 4.

When the extremely premature infants were examined

before 12 hours of postnatal age, the percentage for whom gestational age by NBS was within 2 weeks of agreement of C-GLMP was significantly higher than that of those examined after 12 hours of postnatal age (97% vs 78%; $p = 0.043$). For babies at least 26 weeks of gestational age, an optimal postnatal age at examination was not apparent (Table V).

DISCUSSION

Before the mid-1980s, infants less than 26 weeks of gestational age rarely survived beyond delivery room care, but there is increasing optimism for survival of very low birth weight infants.¹²⁻¹⁴ The infants born at the earliest gestational ages tend to be those of young mothers¹⁵ who may

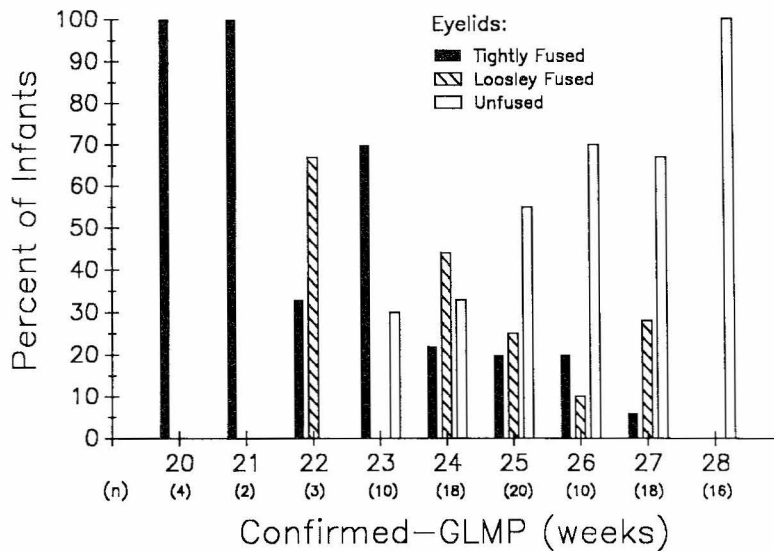


Fig. 4. Eyelid fusion status at gestational weeks 20 through 28 by C-GLMP (confirmed-GLMP).

Table V. Proportion of 2-week agreement between gestational age by New Ballard Score and C-GLMP

Gestation (wk)	Age at examination (hr)*								p
	0.2-12		13-24		25-48		46-90		
	%	n	%	n	%	n	%	n	
<26	97	33/34	78	7/9	75	3/4	80	4/5	0.043†
26-31	96	26/27	80	20/25	100	16/16	86	12/14	NS‡
32-37	93	26/28	79	34/43	78	21/27	93	13/14	NS‡
≥38	94	80/85	98	119/122	98	48/49	91	10/11	NS‡

NS, Not significant.

The optimal postnatal age at examination for infants >26 weeks of gestational age, compared with C-GLMP, was less than 12 hours. For more mature infants, age at examination did not affect validity of the score up to 96 hours of postnatal age.

*Postnatal age at examination was recorded in 513 infants.

†Fisher Exact Test of ≤12 hours versus ≥13 hours.

‡By Fisher Exact Test of ≤12 hours versus ≥13 hours and by chi-square test using age at examination categories for each gestational age group.

have uncertain menstrual histories and may not have undergone prenatal ultrasound assessment or received other prenatal care. Morbidity and mortality statistics and long-term developmental follow-up studies of these very low birth weight infants will be more meaningful if accurate gestational ages are recorded along with birth weights.¹ For these reasons, an expansion of the gestational-maturational assessment tool has become essential.

The newly expanded NBS provides a valid and accurate assessment of gestational age for extremely premature infants that was not previously available while maintaining validity and interrater reliability for the entire infant population. The NBS overestimates gestational age by 0.3 to 0.6 weeks (2 to 4 days) at gestational ages of less than 37 weeks. At gestational ages of between 32 and 37 weeks, this error is statistically significant for validity and accuracy.

We speculate that infants born prematurely at these gestational ages may have had intrauterine stresses that accelerated fetal maturation,^{16,17} resulting in higher maturational scores. Fetal growth, the basis of ultrasonography, may not be affected by intrauterine stresses to the same degree or in the same direction as is fetal maturation, the basis of NBS. These biologic influences may explain certain discrepancies between GUS and gestational age by NBS. Further studies are in progress to test this hypothesis.

Our results suggest that prompt examination is necessary for optimal validity of NBS estimation of gestational age for extremely premature infants. This may be explained in part by the relative unavailability of such premature infants at older postnatal ages, a result of their brief survival times. More studies with larger numbers of extremely premature infants will be necessary to confirm that a postnatal age at

examination of less than 12 hours is essential to ensure validity of NBS maturational examination.

Although early ultrasound examination provides the most useful prenatal estimate of gestational age,^{18, 19} NBS provides a valid postnatal maturational assessment for infants at all gestational ages of more than 20 weeks.

We acknowledge the cooperation and understanding of the neonatal nursing staffs of Cincinnati University Hospital, Children's Hospital Medical Center, Bethesda Hospital, and Good Samaritan Hospital of Cincinnati. We thank Drs. Reginald Tsang and Bonny Specker for their review and critique.

REFERENCES

1. Kline J, Stein Z, Susser M. Conception to birth: epidemiology of prenatal development. Monographs in epidemiology and biostatistics, vol 14. New York: Oxford University Press, 1989.
2. Phelps DL, Brown DR, Tung B, et al. 28-day survival rates of 6676 neonates with birth weights of 1250 grams or less. *Pediatrics* 1991;87:7-17.
3. Ballard JL, Novak KK, Driver MA. A simplified score for assessment of fetal maturation of newly born infants. *J PEDIATR* 1979;95:769-74.
4. Streeter GL. Weight, sitting height, head size, foot length, and menstrual age of the human embryo. *Contributions to Embryology* 1920;11:143-70.
5. Hern WM. Correlation of fetal age and measurements between 10 and 26 weeks of gestation. *Obstet Gynecol* 1984;63:26-32.
6. Pearson AA. The development of the eyelids. I. External features. *J Anat* 1980;130:33-42.
7. Harris MJ, McLeod MJ. Eyelid growth and fusion in fetal mice: a scanning electron microscope study. *Anat Embryol (Berl)* 1982;164:207-20.
8. Constantine NA, Kraemer HC, Kendall-Tackett KA, Bennett FC, Tyson JE, Gross RT. Use of physical and neurologic observations in assessment of gestational age in low birth weight infants. *J PEDIATR* 1987;110:921-8.
9. Alexander GR, Hulsey TC, Smeriglio VL, Comfort M, Lenkoff A. Factors influencing the relationship between a newborn assessment of gestational maturity and the gestational age interval. *Paediatr Perinat Epidemiol* 1990;4:133-46.
10. Usher R, McLean F. Intrauterine growth of live-born Caucasian infants at sea level: standards obtained from measurements in 7 dimensions of infants born between 25 and 44 weeks of gestation. *J PEDIATR* 1969;74:901-10.
11. SAS user's guide basics and statistics, version 5 editions. Cary, N.C.: SAS Institute, 1985.
12. Horwood SP, Boyle MH, Torrence GW, Sinclair JC. Mortality and morbidity of 500- to 1499-gram birth weight infants live-born to residents of a defined geographic region before and after neonatal intensive care. *Pediatrics* 1982;69:613-20.
13. Hack M, Fanaroff AA. Changes in the delivery room care of extremely small infant (<750 gm): effects on morbidity and mortality. *N Engl J Med* 1983;308:1330-6.
14. Seel EJ. Outcome of very very low birth weight infants. *Clin Perinatol* 1986;13:451-9.
15. Scholl TO, Miller LK, Salmon RW, Cofsky MC, Shearer J. Prenatal care adequacy and the outcome of adolescent pregnancy effects on weight gain, preterm delivery, and birth weight. *Obstet Gynecol* 1987;69:312-6.
16. Amiel-Tison C. Possible acceleration of neurological maturation following high-risk pregnancy. *Am J Obstet Gynecol* 1980;138:303-6.
17. Gould JB, Gluck L, Kulovich MV. The relationship between accelerated pulmonary maturity and accelerated neurological maturity in certain chronically stressed pregnancies. *Am J Obstet Gynecol* 1977;127:181-6.
18. Deter RL, Harrist RB, Hadlock FP et al. The use of ultrasound in the assessment of normal fetal growth: a review. *J Clin Ultrasound* 1981;9:481-93.
19. Siddiqi TA, Miodovnik M, Mimouni F, Clark EA, Khoury JC. Biphasic intrauterine growth in insulin-dependent diabetic pregnancies. *J Am Coll Nutr* 1989;8:225-34.