American Journal of Epidemiology
Vol. 151, No. 6
Printed in U.S.A.
Copyright © 2000 by The Johns Hopkins University School of Hygiene and Public Health
All rights reserved

Concordance among Measures of Pregnancy Outcome Based on Fetal Size and Duration of Gestation

David A. Savitz,[1] Cande V. Ananth,[2] Gertrud S. Berkowitz,[3] and Robert Lapinski[4]

Epidemiologic and clinical studies of pregnancy outcome often consider a variety of related, overlapping outcome measures. The overlap among these measures was analyzed using data from the Mount Sinai Hospital Perinatal Data Base, New York City, New York. A total of 52,621 births from 1986 through 1996 were included, with information on gender, ethnicity, birth weight, and gestational age assigned based on last menstrual period or early ultrasound. The authors considered very low birth weight (VLBW) (<1,500 g), low birth weight (LBW) (<2,500 g), degrees of preterm delivery (less than 32, 34, and 37 weeks' gestation), and small for gestational age (less than the 10th percentile of weight for gestational age) births. Infants at the extremes of gestational age (<32 or 34 weeks' gestation) were almost always LBW (97.6 and 91.7%, respectively), and those who were VLBW were almost always preterm (99.2%). However, only 69.2% of LBW infants were preterm, and 50.2% of preterm infants were LBW (kappa = 0.54). Only for VLBW and less than 32 weeks' gestation were both measures of overlap at least 70% (kappa = 0.98). The lack of concordance among measures suggests that multiple outcome measures be considered and that results from analyses using disparate measures not be compared directly. *Am J Epidemiol* 2000;151:627–33.

birth weight; fetal growth retardation; gestational age; infant, low birth weight; infant, premature

Epidemiologic and clinical studies of pregnancy outcome often consider a variety of related, overlapping outcome measures (such as birth weight or duration of gestation) or indices derived from those measures (such as term low birth weight (LBW) (infants weighing <2,500 g) or small-for-gestational age (SGA) births). Infants who are small or are born earlier have increased morbidity and mortality, and the more extremely early or small they are, the higher the risk (1). Biologic considerations suggest separating indicators of reduced fetal growth (such as low weight for gestational age) from measures of gestational duration (such as preterm delivery), although even these may be related, for example, if limited fetal growth stimulates parturition (2). LBW, historically the most commonly used measure, represents a mixture of babies who are normal but small, small because of being born too early, and small because of having suffered from stunted fetal growth (3). Pathways leading to small size at birth appear to have distinctive causes (4, 5) and consequences (1, 6, 7) for infant health.

For identification of preventable causes of adverse pregnancy outcome, the categories should be as homogeneous as possible, ideally isolating causal pathways from one another. Although understanding of etiologies and mechanisms is presently insufficiently advanced to define distinctive endpoints for analysis, we can examine the extent to which candidate measures are statistically distinctive. If measures that seem conceptually distinct define the same infants, they are functionally equivalent. Only if measures identify different infants might they be worth distinguishing in epidemiologic studies. For example, if essentially all very low birth weight (VLBW) (<1,500 g) infants were severely preterm (born prior to 32 weeks' completed gestation) and vice versa, there would be little practical value in analyzing those as separate measures. On the other hand, if the infants identified as preterm (born prior to 37 weeks' completed gestation) were often not of LBW or vice versa, then they might be indicative of different etiologic pathways and worthwhile to examine separately. The concordance would

Received for publication January 12, 1999, and accepted for publication June 4, 1999.

Abbreviations: LBW, low birth weight; SGA, small for gestational age; VLBW, very low birth weight.

[1] Department of Epidemiology, School of Public Health, University of North Carolina, Chapel Hill, NC.
[2] Division of Epidemiology and Biostatistics, Department of Obstetrics, Gynecology, and Reproductive Sciences, Robert Wood Johnson Medical School, University of Medicine and Dentistry of New Jersey, New Brunswick, NJ.
[3] Department of Community and Preventive Medicine, Mount Sinai School of Medicine, New York, NY.
[4] Department of Obstetrics, Gynecology, and Reproductive Science, Mount Sinai School of Medicine, New York, NY.
Reprint requests to Dr. David A. Savitz, Department of Epidemiology, CB #7400, University of North Carolina, Chapel Hill, NC 27599–7400.

Downloaded from http://aje.oxfordjournals.org/ at University of North Texas on January 5, 2017

need to be symmetrical for the outcomes to be functionally equivalent; it is not sufficient for one measure to subsume the other unless the reverse is also true.

Since there is presently no standardization of methods for classifying pregnancy outcomes in the literature, any attempt to integrate results across studies, whether through formal meta-analysis or conventional literature review, requires judgment of which measures are statistically interchangeable and which are not. To provide guidance to investigators who are selecting and constructing pregnancy outcome measures and to help in the interpretation of published studies that use diverse indices, we have examined the concordance among an array of commonly used measures of gestational duration, infant size, and their combination.

MATERIALS AND METHODS

Births occurring between January 1986 and December 1996 in the Mount Sinai Hospital Obstetric and Perinatal Data Base, New York City, New York, were analyzed. Starting with 53,675 singleton pregnancies, we excluded 312 fetal deaths; 43 infants with missing data on race/ethnicity; 685 of ethnicity other than Black, Hispanic, Asian, or White; six of ambiguous gender; three with missing data on gender; and five with missing data on birth weight (1,054 in total), leaving 52,621 in the analysis. For the analysis of SGA births, 430 infants born prior to 25 weeks or after 42 weeks were omitted as well. Information was compiled on race/ethnicity (Black, Hispanic, Asian, or White), gender, birth weight (g), and gestational age at birth.

Commonly used measures of pregnancy outcome were constructed from the information on birth weight and gestational age. We considered two indices of birth weight: VLBW (<1,500 g) and LBW (<2,500 g). Gestational age was estimated based on last menstrual period if it was within 2 weeks of the estimate based on ultrasound before the third trimester, as described elsewhere (8). If the woman was uncertain of her last menstrual period or if the disparity was greater than 2 weeks, then the gestational age was based on ultrasound assessment. When neither last menstrual period nor ultrasound was available, the estimate was based on the clinical evaluation by the obstetrician. The final assignment was based on last menstrual period dates in 69.1 percent of births, ultrasound in 29.3 percent of births, and clinical estimate in 1.6 percent of births in the study population. Gestational age was categorized as less than 32, less than 34, less than 37, and 37 or more weeks. Birth weight and gestational age were combined to define SGA based on US national standards (9) specific to gender and race (with Black norms applied to Hispanics and White norms applied to Asians).

Concordance was examined by cross-tabulating outcome measures for the total population as well as separately by ethnicity and gender. We quantified the proportion of births in one group that also met the criterion for another outcome group, e.g., the proportion of births less than 37 weeks gestation that were less than 2,500 g and the proportion of births less than 2,500 g that were less than 37 weeks gestation. We examined VLBW, LBW, and size for gestational age in relation to duration of gestation and vice versa. To summarize the relation between selected pairs of measures, we calculated kappa coefficients and 95 percent confidence intervals (10), conceptualizing the two measures as alternative indicators of the same event and quantifying the extent to which the two outcomes coincide after taking into account the agreement expected by chance alone.

RESULTS

The study population is a mix of women who are from the neighborhoods around Mount Sinai Hospital and patients from a broader area of Manhattan. As described in a previous report (8), among term births most of the patients were between 20 and 34 years old, with only 9 percent less than age 20 years and 22 percent age 35 or older, two thirds were private patients, and one-third were public clinic patients. Slightly over half were nulliparous and only 4 percent were of parity four or higher, 7 percent smoked during pregnancy, 61 percent began prenatal care prior to 13 weeks' gestation, and diabetes and hypertension were reported in 6 and 5 percent, respectively.

Black and Hispanic patients were represented in sizable numbers (tables 1–3), although over half of the patients were White. Among all births, 7.1 percent were LBW, with higher proportions found among Black and Hispanic patients, and among females compared with males (table 1). The same excess risk among Blacks and Hispanics was found for preterm birth as well and was most pronounced for the earliest preterm births (table 2).

Concordance among measures is presented in figure 1 for the total population, aggregated across gender and ethnicity groups. The arrows across the rows indicate the proportion of those meeting the row attribute who also meet the column attribute, e.g., of those infants who weigh less than 1,500 g (VLBW), 88.4 percent were born prior to completing 32 weeks' gestation, 95.5 percent prior to completing 34 weeks' gestation, etc. Similarly, the column arrows indicate the proportion of those who fall into the column who also meet the row attributes, e.g., of those infants who were born prior to completing 32 weeks' gestation, 71.0 percent weighed less than 1,500 g

Downloaded from http://aje.oxfordjournals.org/ at University of North Texas on January 5, 2017

TABLE 1. Number of births by ethnicity, gender, and birth weight, Mt. Sinai Hospital, New York City, NY, 1986–1996

	Birth weight						Total no. of births
	<1,500 g		1,500–<2,500 g		≥2,500 g		
	No.	%	No.	%	No.	%	
Total	621	1.2	3,095	5.9	48,905	92.9	52,621
White							
Male	100	0.7	448	3.1	13,903	96.2	14,451
Female	84	0.6	614	4.5	12,886	94.9	13,584
Black							
Male	116	2.6	424	9.4	3,990	88.1	4,530
Female	106	2.4	488	10.9	3,900	86.8	4,494
Hispanic							
Male	93	1.4	479	7.4	5,909	91.2	6,481
Female	108	1.7	535	8.4	5,701	89.9	6,344
Asian							
Male	5	0.4	51	3.7	1,333	96.0	1,389
Female	9	0.7	56	4.2	1,283	95.2	1,348

TABLE 2. Number of births by ethnicity, gender, and gestational age, Mt. Sinai Hospital, New York City, NY, 1986–1996

	Gestational age (weeks)							
	<32		32–33		34–36		≥37	
	No.	%	No.	%	No.	%	No.	%
Total	773	1.4	546	1.0	3,727	7.1	47,575	90.5
White								
Male	118	0-.8	94	0.7	787	5.4	13,452	93.1
Female	103	0.8	80	0.6	634	4.7	12,767	94.0
Black								
Male	144	3.2	92	2.0	499	11.0	3,795	83.8
Female	131	2.9	84	1.9	451	10.0	3,828	85.2
Hispanic								
Male	131	2.0	97	1.5	638	9.8	5,615	86.5
Female	132	2.1	89	1.4	578	9.1	5,545	87.4
Asian								
Male	5	0.4	7	0.5	75	5.4	1,302	93.7
Female	9	0.7	3	0.2	65	4.8	1,271	94.3

(VLBW), 97.6 percent weighed less than 2,500 g (LBW), etc.

The magnitude of overlap among pregnancy outcome measures depends strongly on the severity of prematurity and size (figure 1). Those at the extremes of gestational age (<32 or <34 weeks) are almost always LBW (<2,500 g), and those who are VLBW are predominantly preterm (<37 weeks' gestation). However, with less restrictive criteria for preterm delivery (<37 weeks' gestation) and LBW (<2,500 g), the overlap is far more modest, with 69.2 percent of LBW infants also preterm and 49.8 percent of preterm infants also LBW. There are marked asymmetries in the extent to which one measure subsumes the other when an extreme criterion for size or gestational duration is used but a less restrictive criterion is used for the other. For example, among VLBW infants, 99.2 percent are preterm, whereas only 12.2 percent of preterm infants are VLBW. Only for the measures of VLBW and birth at less than 32 weeks' gestation are

Downloaded from http://aje.oxfordjournals.org/ at University of North Texas on January 5, 2017

TABLE 3. Number of births by ethnicity, gender, and size for gestational age, New York City, NY, 1986–1996*

	SGA†		Not SGA	
	No.	%	No.	%
Total	5,288	10.1	46,903	89.9
White				
Male	1,547	10.8	12,811	89.2
Female	1,500	11.1	11,995	88.9
Black				
Male	450	10.1	4,026	89.9
Female	430	9.7	4,021	90.3
Hispanic				
Male	484	7.6	5,917	92.4
Female	485	7.7	5,802	92.3
Asian				
Male	209	15.1	1,174	84.9
Female	183	13.7	1,157	86.3

*A total of 430 births outside the range of 25–42 weeks' completed gestation were excluded.
† SGA, small for gestational age.

both measures of overlap at least 70 percent, with none of the other cells defined by the intersection of LBW and reduced gestational age both greater than 50 percent.

Since a modest proportion of births overall are either preterm or LBW, being classified as normal on one criterion is strongly predictive of being classified as normal on the other. Of infants born at term, 97.5 percent weighed 2,500 g or more, and of those born weighing 2,500 g or more, 94.8 percent had completed at least 37 weeks gestation. For this reason, quantifying the concordance in a single measure must take into account the level of agreement expected by chance; hence, the use of kappa statistics (table 4). Based on these measures, the agreement between VLBW and birth at less than 32 or less than 34 weeks is very high, with moderate agreement of LBW with births at less than 32, less than 34, and less than 37 weeks' gestation and lower agreement between LBW and SGA.

Classification of SGA is intended to be independent of gestational age, and there is only a small tendency for SGA births to be preterm to a greater extent than are non-SGA births (12.3 vs. 9.0 percent) (data not shown), with deviations from exactly 10 percent reflecting an imperfect correspondence between the distributions observed in the study population compared with the reference population. The predictiveness of SGA for VLBW and especially for LBW is much greater, with 38.5 percent of SGA births being LBW compared with 4.1 percent of non-SGA births (figure 1). Similarly, LBW is predictive of being SGA, with 45.7 percent of LBW births qualifying as SGA. Only when the hybrid index of "term LBW" is constructed are the measures nearly equivalent, with 96.0 percent of such infants classified as SGA (data not shown).

Another hybrid index sometimes used is "preterm LBW" (11), combining shortened gestation with small size in order to exclude erroneously classified term infants. Preterm LBW births represent 67.7 percent of all LBW infants and 49.8 percent of all preterm infants

Birth Weight/SGA	<32 wks ↓	<34 wks ↓	<37 wks ↓	37+ wks ↓	SGA ↓	Not SGA ↓
<1500 gm →	88.4 71.0	95.5 45.0	99.2 12.2	0.8 0.0	24.8 3.1	75.2 0.8
<2500 gm →	20.2 97.6	32.6 91.7	69.2 49.8	32.3 2.5	45.7 38.5	54.3 4.1
2500+ gm →	0.0 2.9	0.2 8.3	5.2 50.2	94.8 97.5	5.4 61.5	94.6 95.9

FIGURE 1. Predictions of low birth weight, preterm birth, and SGA (less than 10th percentile of weight for gestational age) with each predicted by the other categories. Results are based on 52,621 births from Mt. Sinai Hospital, New York City, New York, 1986–1996. Arrows designate the proportion of those who have the attribute indicated in the column or row heading and who also have the attribute in the intersecting column or row. For example, of all births less than 2,500 g (row 2), 32.6% are also less than 34 weeks gestation; of all births less than 34 weeks gestation (column 2), 91.7 percent are less than 2,500 g.

Downloaded from http://aje.oxfordjournals.org/ at University of North Texas on January 5, 2017

TABLE 4. Kappas and 95% confidence intervals for concordance among low-birth weight, preterm, and small-for-gestational age births, Mt. Sinai Hospital, New York City, NY, 1986–1996

Gestational age (weeks) and SGA* births	Birth weight (g) and SGA births					
	<1,500		<2,500		SGA	
	Kappa	95% CI*	Kappa	95% CI	Kappa	95% CI
<32	0.98	0.97, 0.99	0.54	0.52, 0.56	0.01	0.00, 0.01
<34	0.91	0.90, 0.93	0.64	0.62, 0.65	0.01	0.01, 0.02
<37	0.31	0.29, 0.33	0.54	0.52, 0.55	0.03	0.02, 0.04
SGA	0.06	0.05, 0.08	0.37	0.36, 0.39		

* SGA, small for gestational age; CI, confidence interval.

(data not shown). Among preterm LBW infants, 21.8 percent were classified as SGA.

Because the size distribution of births differs by gender and ethnicity, the pattern of overlap between measures of size and duration of gestation differs as well according to these characteristics (table 5). Percent LBW provides an indication of the overall distribution of birth weights, with the lowest proportion among Asians and Whites and the highest among Blacks. Males are less often LBW (i.e., are larger on average) than are females. The proportion preterm among LBW births does not show a pattern by ethnicity, but there is a pronounced tendency for the proportion to be higher among males. A LBW male is more likely to be preterm than is a LBW female, which would be predicted since females are smaller at each gestational age. The proportion LBW among preterm births follows a much more consistent pattern: With the exception of Asian births, the gender and ethnic groups that are largest have the lowest proportions (table 5). Those ethnic and gender groups that tend to be larger tend to have larger infants at all gestational ages, including those who are preterm, so that a smaller proportion of the preterm infants are LBW. For VLBW among preterm births, there is no consistent pattern by gender, but the excess among Black infants is clear. Even among preterm births, extreme small size and prematurity are more common than among Whites.

TABLE 5. Concordance among low-birth weight, preterm, and small-for-gestational age births by gender and ethnicity, Mt. Sinai Hospital, New York City, NY, 1986–1996

	LBW* (%)	Preterm among LBW (%)	LBW among preterm (%)	VLBW* among preterm (%)
Total	7.1	67.7	49.8	12.2
White	4.4	65.4	44.9	10.1
Male	3.8	72.5	39.7	9.9
Female	5.1	59.9	51.2	10.3
Black	12.6	70.1	56.8	15.6
Male	11.9	75.7	55.7	15.8
Female	13.2	65.0	58.0	15.5
Hispanic	9.5	69.4	50.6	12.0
Male	8.8	72.0	47.6	10.7
Female	10.1	67.0	53.9	13.4
Asian	4.4	51.2	37.8	8.5
Male	4.0	51.8	33.3	5.8
Female	4.8	50.8	42.9	11.7

* LBW, low birth weight; VLBW, very low birth weight.

DISCUSSION

These data suggest that the extent of overlap or concordance among commonly used measures of pregnancy outcome is insufficient to create effective equivalence among them. On the basis of these patterns, no index of birth weight or gestational age is functionally equivalent to the other, with the most severely preterm VLBW infants coming closest. Even for this group, however, severe growth retardation and extreme prematurity have different consequences (7). Certainly, for purposes of formal meta-analyses, only studies that use identical measures can be considered homogeneous, and even in less formal efforts to reconcile findings across the literature, disparate findings across studies may well result solely from use of different measures. Standardization of measures across studies would be optimal for the goal of being able to integrate findings.

LBW has long been recognized to be a heterogeneous category, given variation in gestational age among LBW infants. Yerushalmy (12) proposed a scheme for isolating reduced fetal growth that was argued to be more practical than applying percentiles to classify SGA. In his approach, group I is less than 1,500 g (regardless of gestational age), group II is 1,500–2,500 g and less than 37 weeks' gestation, group III is 1,500–2,500 g and greater than or equal to 37 weeks' gestation, group IV is greater than or equal to 2,500 g and less than 37 weeks' gestation, and group V is greater than or equal to 2,500 g and greater than or equal to 37 weeks' gestation. Such an approach iso-

Downloaded from http://aje.oxfordjournals.org/ at University of North Texas on January 5, 2017

lates the extremely premature (group I), those who are small for dates (group III), and the healthiest infants (group V). The "preterm LBW" infants (group II) and "preterm normal birth weight" (group IV) combine to include all but the most severely preterm infants, with the former likely to be at the lower end and the latter at the upper end of the gestational age spectrum, close to 37 weeks' gestation. Not surprisingly, these groups have distinctive mortality rates, but whether this scheme is more effective than one based solely on gestational age with finer gradations is unclear.

Around the same time as Yerushalmy's proposal, the classification of infants based on gestational age and percentiles of birth weight was made much easier to implement in clinical settings through nomograms (13). By plotting the birth weight against the week of gestation, infants could be classified simultaneously as preterm, term, or postterm, and as small, appropriate, or large for gestational age. Over the intervening 30 years, the only important refinement may well be recognition of the need to subdivide preterm deliveries on the basis of extent of prematurity within the preterm range, given the marked gradients of perinatal mortality by week of gestation (1). Duration of gestation is a strong predictor of infant survival, with additional information provided by considering relative birth weight.

Being born smaller and being born earlier each confer a disadvantage for survival, with the extremes most disadvantageous (14, 15). In fact, the sharp gradients in survival in the lower range of the gestational age or birth weight distribution argue for either a focus on the most severe cases or at least a stratification by severity (1). Previous studies have indicated that measures of weight or gestational age that isolate infants at the extreme are effective in predicting the group with very high mortality, encouraging use of measures that isolate preterm infants of varying levels of severity.

Another relevant question in choosing among measures is the extent to which predictors of one versus another differ empirically. Secular trends over the decade 1970–1980 show a marked decline in term LBW (effectively, SGA) (16) and modest declines for preterm LBW (which captures only half of preterm births, according to our results) (16), but a lack of decline in preterm births in the aggregate (17). These data suggest that infants are getting larger but not through extending the duration of gestation.

Analyses of risk factors for specific pregnancy outcomes defined by gestational age, birth weight, or their combination yield mixed findings with regard to the distinctiveness of the predictors. In a study of nutritional supplements and pregnancy outcome in Guatemala in the 1970s, different predictors were found for preterm births (lack of calorie or protein supplementation) versus SGA births (low maternal head circumference, low maternal weight) (4). Lang et al. (5) divided births into spontaneous preterm deliveries (labeled "preterm labor") and term SGA births to examine the patterns of association among a set of 23 potential risk factors. Preterm births were associated with an array of obstetric factors, which also predicted SGA births. However, additional predictors of SGA births were race, maternal height, and low prepregnancy weight, in addition to a stronger effect of smoking. When the array of findings is examined, the similarities are more impressive than the differences, suggesting either that many conditions affect growth and duration of gestation through independent pathways or that restriction of fetal growth might tend to result in preterm delivery or a small, term delivery (term SGA).

Limitations of the data used for these analyses should be recognized, including the inherent uncertainties in assessing duration of gestation (18) not fully remedied by use of ultrasound (19). Recent evidence suggests that the magnitude of error in gestational age assessment may be substantial—often a full menstrual cycle off (20). Therefore, independent of the true biologic discrepancy between growth and duration of gestation, there is a component due to error in measurement, principally in the gestational age estimate.

There is a potential influence of size on the clinical decision regarding the timing of delivery, through labor induction or cesarian section. For example, if the smallest of the infants of a given gestational age are preferentially delivered early, then the concordance between small and early will be artificially inflated. The decisions made by the obstetric care providers may have influenced some of the observed patterns.

Application of norms for Whites to classify SGA of Asian infants and application of norms for Blacks to classify Hispanic infants make the evaluation of size for gestational age among these ethnic groups somewhat suspect. As noted previously, despite efforts to isolate infants who have not realized their growth potential through the measure of SGA, the mixing of truly growth-retarded infants with those who are normal, but small, is an inherent problem (3).

On the basis of our analyses and those of other investigators (1, 3, 5), the authors would discourage the analysis of LBW and measures that integrate birth weight and gestational age (e.g., term LBW, preterm LBW). Use of birth weight to correct for erroneous gestational age measurements may be beneficial in removing some term infants erroneously classified as preterm, but also undoubtedly excludes truly preterm

Downloaded from http://aje.oxfordjournals.org/ at University of North Texas on January 5, 2017

infants who are relatively large. Independent measures of gestational age (using differing degrees of cutpoints for prematurity, perhaps 32, 34, and 37 weeks) and SGA would be sufficient and would render the measures based solely on birth weight unnecessary. The more extreme categories of preterm birth would isolate groups increasingly certain to be truly preterm. Only at the extreme of VLBW, which corresponds closely to birth at less than 32 weeks' gestation, is there an unambiguous interpretation of birth weight as an indication of severe prematurity, and studies that are large enough to focus on the smallest, most premature infants can avert the error in gestational age estimation by focusing on birth weight. Since the biologic basis for choosing among measures remains elusive, further research to determine the relation of various measures to infant mortality and to evaluate the distinctiveness of predictors of pregnancy outcome is needed.

ACKNOWLEDGMENTS

Dr. Ananth was supported in part through grant 029553 from the Robert Wood Johnson Foundation awarded to the Center for Perinatal Health Initiatives.

REFERENCES

1. Wilcox AJ, Skjaerven R. Birth weight and perinatal mortality: the effect of gestational age. Am J Public Health 1992;82: 378–82.
2. Secher NJ, Hansen PK, Thomsen BL, et al. Growth retardation in preterm infants. Br J Obstet Gynaecol 1987;94:115–20.
3. Wilcox AJ. Intrauterine growth retardation: beyond birthweight criteria. (Editorial). Early Hum Dev 1983;8:189–93.
4. Villar J, Khoury MJ, Finucane FF, et al. Differences in the epidemiology of prematurity and intrauterine growth retardation. Early Hum Dev 1986;14:307–20.
5. Lang JM, Lieberman E, Cohen A. A comparison of risk factors for preterm labor and term small-for-gestational-age birth. Epidemiology 1996;7:369–76.
6. Rush RW, Keirse MJNC, Howat P, et al. Contribution of preterm delivery to perinatal mortality. Br Med J 1976;2: 965–8.
7. Arnold CC, Kramer MS, Hobbs CA, et al. Very low birth weight: a problematic cohort for epidemiologic studies of very small or immature neonates. Am J Epidemiol 1991;134: 604–13.
8. Berkowitz GS, Blackmore-Prince C, Lapinski RH, et al. Risk factors for preterm birth subtypes. Epidemiology 1998;9: 279–85.
9. Zhang J, Bowes WA Jr. Birth-weight-for-gestational-age patterns by race, sex, and parity in the United States population. Obstet Gynecol 1995;86:200–8.
10. Fleiss JL. Statistical methods for rates and proportions. 2nd ed. Chap. 13. The measurement of interrater agreement. New York, NY: John Wiley & Sons, 1981:212–36.
11. Hillier SL, Nugent RP, Eschenbach DA, et al. Association between bacterial vaginosis and preterm delivery of a low-birth-weight infant. N Engl J Med 1995;333:1737–42.
12. Yerushalmy J. The classification of newborn infants by birth weight and gestational age. J Pediatr 1967;71:164–72.
13. Battaglia FC, Lubchenco LO. A practical classification of newborn infants by weight and gestational age. J Pediatr 1967; 71:159–63.
14. Wilcox AJ. Birth weight, gestation, and the fetal growth curve. Am J Obstet Gynecol 1981;139:863–7.
15. Hertz-Picciotto I, Din-Dzietham R. Comparisons of infant mortality using a percentile-based method of standardization for birthweight or gestational age. Epidemiology 1998;9:61–7.
16. Kessel SS, Villar J, Berendes HW, et al. The changing pattern of low birth weight in the United States, 1970 to 1980. JAMA 1984;251:1978–82.
17. Berkowitz GS, Papiernik E. Epidemiology of preterm birth. Epidemiol Rev 1993;15:414–43.
18. Kramer MS, McLean FH, Boyd ME, et al. The validity of gestational age estimation by menstrual dating in term, preterm, and postterm gestations. JAMA 1988;250:3306–8.
19. Henriksen TB, Wilcox AJ, Hedegaard M, et al. Bias in studies of preterm and postterm delivery due to ultrasound assessment of gestational age. Epidemiology 1995;6:533–7.
20. Gjessing HK, Skaerven R, Wilox AJ. Errors in gestational age: evidence of bleeding early in pregnancy. Am J Public Health 1999;89:213–18.

Downloaded from http://aje.oxfordjournals.org/ at University of North Texas on January 5, 2017