ABSTRACT. Background. Existing fetal growth references all suffer from 1 or more major methodologic problems, including errors in reported gestational age, biologically implausible birth weight for gestational age, insufficient sample sizes at low gestational age, single-hospital or other non-population-based samples, and inadequate statistical modeling techniques.

Methods. We used the newly developed Canadian national linked file of singleton births and infant deaths for births between 1994 and 1996, for which gestational age is largely based on early ultrasound estimates. Assuming a normal distribution for birth weight at each gestational age, we used the expectation-maximization algorithm to exclude infants with gestational ages that were more consistent with 40-week births than with the observed gestational age. Distributions of birth weight at the corrected gestational ages were then statistically smoothed.

Results. The resulting male and female curves provide smooth and biologically plausible means, standard deviations, and percentile cutoffs for defining small- and large-for-gestational-age births. Large-for-gestational age cutoffs (90th percentile) at low gestational ages are considerably lower than those of existing references, whereas small-for-gestational-age cutoffs (10th percentile) postterm are higher. For example, compared with the current World Health Organization reference from California (Williams et al, 1982) and a recently proposed US national reference (Alexander et al, 1996), the 90th percentiles for singleton males at 30 weeks are 1837 versus 2159 and 2710 g. The corresponding 10th percentiles at 42 weeks are 3233 versus 3086 and 2998 g.

Conclusions. This new sex-specific, population-based reference should improve clinical assessment of growth in individual newborns, population-based surveillance of geographic and temporal trends in birth weight for gestational age, and evaluation of clinical or public health interventions to enhance fetal growth.

Acknowledgments. This work was supported by the National Institutes of Health Research. Drs Kramer, Platt, and Abrahamowicz are career scientists of the Canadian Institutes of Health Research. Drs Kuyat, Blumenfeld, and Béatrice Blondel, MD, PhD, for the Fetal/Infant Health Study Group of the Canadian Perinatal Surveillance System.

ABBREVIATIONS. SD, standard deviation; LGA, large for gestational age; SGA, small for gestational age.

For more than half a century, clinicians and investigators have proposed reference data for assessing birth weight for gestational age. These references have been used by clinicians and researchers to assess fetal growth in individual infants and in populations.1 Unfortunately, none of the available references are entirely satisfactory. Some are hospital-based,2–7 giving rise to potential selection bias and problems of generalizability, particularly in view of the low2,6,7 or high4,5 socioeconomic status or high altitude2 that characterizes some of the study hospitals; others are prescriptive rather than descriptive, ie, they are based on infants without known risk factors for impaired fetal growth and thus may not be applicable to populations with mixtures of low- and high-risk pregnancies.8,9 Some are unisex references that fail to account for the known larger birth weight for gestational age in male versus female infants,8,9 others,11,12 go to the opposite extreme and provide curves that are specific for different races, parity, maternal size, and other customizing features for which available data do not permit confident inferences as to whether variations in fetal growth are physiologic or pathologic. Finally, some references are now several decades old and may no longer be pertinent to infants born in more recent years, given the increase in the size of infants born at or near term over the last several decades.

The largest problem with existing references for birth weight for gestational age, however, relates to the measurement of gestational age. Early references rounded gestational age to the nearest week,2–6 rather than truncated to completed weeks. Although this practice makes sense from a biological standpoint, it is not consistent with World Health Organization (WHO; International Classification of Diseases) recommendations to base gestational age on the number of completed weeks, and references based on the nearest week cannot be applied to populations complying with the WHO recommendations. But even when measured in completed weeks, gestational ages in older references are flawed by...
being based on the date of onset of the last menstrual period, which has been shown to grossly overestimate the proportion of infants with postterm gestational ages and underestimate those born preterm when compared with early ultrasound measurements. This leads to artificially wide, or even bimodal, distributions of birth weight at very low gestational ages, and flattening or declining curves postterm, despite the use of various statistical or clinical methods for excluding biologically implausible combinations of birth weight and gestational age. Finally, sample sizes for many references are too small, particularly at low gestational ages, leading to irregularity and even nonmonotonicity in the shape of fetal growth curves at these low gestational ages. Some recent references have not smoothed the curves to correct these irregularities.

In this study, we have attempted to overcome these previous deficits by constructing a sex-specific reference of birth weight for gestational age that has several advantages over existing references. It is based on a linked live birth-infant death file for all infants born in all of the Canadian provinces and territories (with the exception of Ontario) born between 1994 and 1996, during which time gestational ages have reflected the increasing use of early ultrasound for gestational age estimation. Gestational age is measured in completed weeks, and we have used a new mixture distribution method to exclude infants with implausible gestational ages. This results in smooth, monotonic curves with biologically sensible distributions at all gestational ages. We present both a graphical depiction and a tabular representation of means, standard deviations (SDs), and the 3rd, 5th, 10th, 50th (median), 90th, 95th, and 97th percentiles in the hope that they will be useful to clinicians in classifying fetal growth of newborn infants under their care and of stillbirths, and to researchers and public policy makers in comparing geographic differences and temporal trends in birth weight for gestational age in populations.

METHODS

The references we constructed were based on all births contained in the linked file of live births and infant deaths occurring in the provinces and territories of Canada (with the exception of Ontario) born between January 1, 1994, and December 31, 1996. (The linked file was used, rather than the entire Canadian Birth Database, because the linkage procedure allowed elimination of several duplicates contained in the Birth Database.) Ontario was excluded from the study population base for constructing these references because of documented problems with data quality.

Canadian birth certificates include gestational age recorded in completed weeks. Although no information is contained on the birth certificate about the method used to assess gestational age, nor documentation of the dates and results of ultrasound estimates, careful evaluation of these data over several decades strongly suggests that early ultrasound has increasingly been the basis for gestational age assessments in recent years. No information on racial origin is contained on Canadian birth certificates.

Our proposed reference is based on singletons with recorded gestational ages of 22 to 43 weeks. The reference sample comprises 347,370 males and 329,035 females. We used a mixture distribution method for correcting gestational age errors, a modification of the procedure described by Oja et al. The model of Oja et al assumes a log normal distribution of birth weight at each gestational age and errors of +4 weeks or -4 weeks. As assumed in previous studies and as confirmed by ultrasound-based estimated fetal weights, our method assumes that the true distribution of birth weight at each gestational age is normal (Gaussian). We also assume that the observed gestational ages are a mixture of correct values and of term (40-week) gestations erroneously measured at the observed (recorded) gestational age. This latter assumption is consistent with the frequently observed curve "bump" at the upper percentiles for extremely preterm infants.

We used the estimation-maximization algorithm to derive maximum likelihood estimates of the mean and SD and of the probability that the true gestational age of an individual infant was the observed (recorded) gestational age rather than 40 weeks. Starting values for the mean and SD parameters were estimated from the lower half of the observed distribution of birth weight for gestational age (which should not be affected by misclassifying infants born at term as extremely preterm). We then resampled all observations. This resampling resulted in either maintaining the infant in the dataset with the recorded gestational age or deleting that infant, with the probability of deletion equal to the estimated probability that the true gestational age was 40 weeks. Finally, the percentiles of the birth weight distribution at the corrected gestational ages were generated using a smoothing spline with 7 degrees of freedom, weighted by the square root of the (corrected) number of infants at each gestational age.

Tables and graphs were created separately for males and females for the 3rd, 5th, 10th, 50th (median), 90th, 95th, and 97th percentiles at 22 to 43 completed weeks, based on the smoothed estimated curves, and for the mean and SD calculated from the empirical distribution of birth weights after correction. The means and SDs are also tabulated to allow calculation of z scores rather than (or in addition to) percentiles, where

$$z = \frac{\text{observed birth weight} - \text{mean birth weight}}{\text{SD}}$$

RESULTS

Tables 1 and 2 show the 3rd, 5th, 10th, 50th (median), 90th, 95th, and 97th percentile; mean; and SD birth weights for male and female infants, respectively. Figures 1 and 2 show the crude percentile graphs, and Figs 3 and 4 show the curves produced after correcting the gestational ages and smoothing the resulting birth weight distributions. The crude curves show a convex "bump" in the upper percentiles at 27 to 33 weeks, consistent with our model’s assumption that some truly term births are misclassified at these (grossly underestimated) gestational ages. The corrected curves are smooth and evenly spaced and show no biologically implausible bumps preterm or flattening or declines postterm. Moreover, each of the percentile curves shows the expected sigmoid shape, with birth weight increasing monotonically as gestational age advances. The smoothed curves closely follow (within limits of random error) the empiric birth weight distribution at each (corrected) gestational age. For example, at 30 weeks, the 90th percentile curve cuts off 91.4% of males and 87.7% of females in our data set. At 42 weeks, the 10th percentile curve cuts off 10.4% of males and 10.2% of females.

Figure 5 illustrates how our procedure has corrected the crude curves at 30 weeks’ gestation (for males). At low percentiles, the curves before and after correction are virtually identical, but starting at the 90th percentile, and especially at the 95th and 97th percentiles, the cutoffs after the correction are substantially lower. Table 3 compares our new proposed reference (based on the corrected curves) with 4 population-
based references: 1 from Canada (Arbuckle et al20) and 3 from the United States (Williams et al16 [white non-Hispanics only], Zhang and Bowes12 [whites only], and Alexander et al19). Tabulated values for the Canadian reference were provided by Statistics Canada (Russell Wilkins, Health Analysis and Modeling Group, Statistics Canada, Ottawa, Ontario, personal communication). The comparison focuses on the 2 sensitive regions mentioned earlier: the 90th percentile (the conventional large-for-gestational-age [LGA] cutoff) at 30 weeks, and the 10th percentile (the conventional small-for-gestational-age [SGA] cutoff) at 42 weeks. Each of the previous references uses a correction procedure to reduce errors in pret-
DISCUSSION

The reference curves and tables presented herein differ in several important respects from those currently available. Their main advantages include their recent vintage; population base; increased reliance on early ultrasound-based estimates of gestational age; statistical modeling of gestational ages to correct biologically implausible values (based on the observed and expected birth weights); sex-specificity; and adaptability to the use of either percentiles or z scores. This reference differs from others currently available in showing lower means and particularly upper (90th, 95th, and 97th) percentile birth weights at low gestational ages, as well as higher birth weights (particularly at the 10th, 5th, and 3rd percentiles) for gestational age in the postterm period. The absence of a downturn in the curves in the postterm period is consistent with evidence based on early ultrasound-based gestational ages.29,30 Our reference also differs from older references by reflecting the recent temporal trend toward increasing birth weight at or near term.

This new reference should prove helpful for several types of users in developed country settings. For clinicians, it enables better classification of individual infants as small, appropriate, or large for their gestational age. It should prove useful for clinicians caring for extremely preterm infants in whom unexpectedly high birth weights should lead to skepticism about the reported gestational age. It should also be reassuring to prenatal care providers by eliminating the impression that fetuses fail to gain, or
even lose weight postterm. For researchers, this new reference should allow better classification of groups of infants for determination of geographic differences, temporal trends, etiologic determinants, and short- and long-term prognosis. For public health policy makers, the reference should be useful in surveillance, ie, in tracking population differences by geographic location, socioeconomic status, race/ethnicity, and other relevant factors, as well as in tracking trends over time and in response to educational or other public health interventions.

Despite the advantages of our new reference over its predecessors, several limitations should be acknowledged. First, like any purely statistical correction procedure, our method compares poorly with early ultrasound for valid estimation of the true gestational age. In particular, early ultrasound corrects last menstrual period-based estimates by a systematic shift to the left (ie, toward lower gestational ages) across the entire gestational age distribution, primarily because of the effect of delayed ovulation. Our method, however, corrects only gross errors by excluding preterm (and some postterm) infants whose birth weights are more compatible with 40-week births. We have also experimented with more complex statistical corrections (eg, ±1, 2, 4, and 8 weeks), but the more complex algorithms are far more computer-intensive, result in many more exclusions, and do not substantially alter the slopes or SGA or LGA cutoffs of the reference curves.

The most important limitation of all population-based references, including ours, is their cross-sectional nature, ie, they are based on the birth weights of different infants born at different gestational ages,
rather than longitudinal measurements of the same infants over the course of gestation.\(^3\) Growth is defined as an increase in size over time, and documentation of increasing size therefore requires 2 or more serial measurements. Unfortunately, serial anthropometric measurements during fetal life are feasible only with ultrasound, and these have not proved to be sufficiently valid or precise to serve as a reference.\(^3,33–35\) The substitution of cross-sectional for longitudinal data on fetal growth is a problem, because evidence suggests that preterm infants are somewhat smaller than fetuses of the same gestational age who remain in utero.\(^36–39\) This problem may partly reflect the fact that some of the determinants of fetal growth and gestational duration overlap. Pregnancy-induced hypertension and preeclampsia, for example, are risk factors for both intrauterine growth restriction and both spontaneous and induced preterm delivery.\(^39–42\) At or after term, fetuses who remain unborn may not have grown at the same rate as those born earlier, because fetal size is considered to be one of the determinants of the onset of labor, cross-sectional data may reflect the earlier birth of faster-growing fetuses.\(^1\)

If and when ultrasound technology improves to permit more valid and reliable measurements of fetal weight and other anthropometric measurements in utero, a truly longitudinal fetal growth reference will be feasible. Other future changes may also require modification or replacement of our reference. Temporal trends toward increasing maternal prepregnancy weight and weight gain during pregnancy, maternal stature, and reductions in cigarette smoking will probably continue to increase the size of infants born at or near term. Finally, the increasing trend toward obstetric intervention to hasten delivery for pathologic pregnancies in the preterm period, and in response to signs of slow growth or other fetal compromise in the postterm period, may continue to affect the shape of the growth curve at these periods of gestation.

### REFERENCES


Table 3. Comparison of Key Birth Weight Cutoffs (g). Proposed New Reference Versus Four Previous References

<table>
<thead>
<tr>
<th>Reference</th>
<th>LGA (90th Percentile) at 50 Weeks</th>
<th>SGA (10th Percentile) at 42 Weeks</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
<td>Female</td>
</tr>
<tr>
<td>Proposed reference</td>
<td>1837</td>
<td>1783</td>
</tr>
<tr>
<td>Arbuckle et al(^20)</td>
<td>1870</td>
<td>1800</td>
</tr>
<tr>
<td>Williams et al(^16)</td>
<td>2159</td>
<td>2113</td>
</tr>
<tr>
<td>Zhang and Bowes(^12)</td>
<td>2355*</td>
<td>2288*</td>
</tr>
<tr>
<td>Alexander et al(^19)</td>
<td>2710†</td>
<td>2710†</td>
</tr>
</tbody>
</table>

* Average of values for primiparas and multiparas.
† Sex-specific value not reported.
estimation by menstrual dating in term, preterm, postterm gestations. *JAMA.* 1988;260:3306–3308
A New and Improved Population-Based Canadian Reference for Birth Weight for Gestational Age

Michael S. Kramer, Robert W. Platt, Shi Wu Wen, K.S. Joseph, Alexander Allen, Michal Abrahamowicz, Béatrice Blondel, Gérard Bréart and for the Fetal/Infant Health Study Group of the Canadian Perinatal Surveillance System

*Pediatrics* 2001;108;e35
DOI: 10.1542/peds.108.2.e35
A New and Improved Population-Based Canadian Reference for Birth Weight for Gestational Age

Michael S. Kramer, Robert W. Platt, Shi Wu Wen, K.S. Joseph, Alexander Allen, Michal Abrahamowicz, Béatrice Blondel, Gérard Bréart and for the Fetal/Infant Health Study Group of the Canadian Perinatal Surveillance System

*Pediatrics* 2001;108:e35
DOI: 10.1542/peds.108.2.e35

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://pediatrics.aappublications.org/content/108/2/e35