Invited Commentary

Invited Commentary: The Incremental Value of Customization in Defining Abnormal Fetal Growth Status

Jun Zhang* and Kun Sun

*Correspondence to Dr. Jun Zhang, MOE-Shanghai Key Laboratory of Children’s Environmental Health, Xinhua Hospital, Shanghai Jiao Tong University School of Medicine, 1665 Kong Jiang Road, Shanghai 200092, China (e-mail: junjimzhang@gmail.com).

Initially submitted May 28, 2013; accepted for publication June 17, 2013.

Reference tools based on birth weight percentiles at a given gestational week have long been used to define fetuses or infants that are small or large for their gestational ages. However, important deficiencies of the birth weight reference are being increasingly recognized. Overwhelming evidence indicates that an ultrasonography-based fetal weight reference should be used to classify fetal and newborn sizes during pregnancy and at birth, respectively. Questions have been raised as to whether further adjustments for race/ethnicity, parity, sex, and maternal height and weight are helpful to improve the accuracy of the classification. In this issue of the Journal, Carberry et al. (Am J Epidemiol. 2013;178(8):1301–1308) show that adjustment for race/ethnicity is useful, but that additional fine tuning for other factors (i.e., full customization) in the classification may not further improve the ability to predict infant morbidity, mortality, and other fetal growth indicators. Thus, the theoretical advantage of full customization may have limited incremental value for pediatric outcomes, particularly in term births. Literature on the prediction of short-term maternal outcomes and very long-term outcomes (adult diseases) is too scarce to draw any conclusions. Given that each additional variable being incorporated in the classification scheme increases complexity and costs in practice, the clinical utility of full customization in obstetric practice requires further testing.

birth weight; customization; fetal growth; large for gestational age; small for gestational age

Abbreviations: LGA, large for gestational age; SGA, small for gestational age.

The accurate assessment of fetal growth status is critical for patient care and research. In the field of epidemiology, fetal growth restriction and, to a lesser degree, fetal overgrowth, are 2 of the most commonly used variables, both as exposures and outcomes. However, defining abnormal fetal growth status has long been a challenge. There is still no widely accepted definition. Small for gestational age (SGA) and large for gestational age (LGA), which are defined as fetal or birth weights less than the 10th or greater than the 90th percentile at a given gestational week, have been generally used as proxies for fetal growth restriction and overgrowth, respectively. Yet, it is well recognized that small or large body size does not necessarily represent abnormal body size. The use of different reference tools may result in a different classification status for the same fetus or infant.

Lubchenco et al. (1) first presented the birth weight distribution by gestational week and implicitly introduced the concepts of SGA and LGA to the perinatal field. This year, 2013, marks the 50th anniversary of their landmark paper. Over the years, people realized that newborn sizes vary naturally by race/ethnicity, sex, and parity. In 1976, Brenner et al. (2) were the first to propose an adjustment for these factors. In the 1990s, Gardosi et al. (3) popularized this idea by creating a computer program that was easy to use and free to download. Their method adjusted for race/ethnicity, sex of the newborn, parity, and maternal prepregnancy or early pregnancy weight and height. It was intended to provide more accurate categorization of fetal or newborn size. What was more significant was that the underlying reference they used in the program was not based on birth weight but on ultrasonography-estimated fetal weight (4). For term babies, the 10th percentiles for the birth weight reference and fetal weight reference differ a little. However, for preterm babies, the 10th percentile for the birth weight reference is substantially lower than that of the fetal weight reference, whereas the 90th percentile of the birth weight
reference is much higher than that of the fetal weight reference (Figure 1) (5). This is because preterm babies are more likely to be growth restricted. A reference based on preterm births (i.e., the birth weight reference) is likely to have lower cut-off points than a reference based on all fetuses born and unborn (i.e., the fetal weight reference) (6). On the other hand, most birth weight references in the literature use the mother’s last menstrual period to estimate gestational age (1, 7–9). The well-known errors in last menstrual period data can misclassify term births as preterm and produce artificially high birth weight percentiles at the upper limit (7). Thus, the birth weight reference has a much wider percentile range than the fetal weight reference. The use of a fetal weight reference, therefore, leads to a significant improvement in the classification of abnormal fetal and newborn sizes at both ends of the percentile range, especially in preterm births (5).

A number of studies have demonstrated that the fetal weight reference classifies more preterm babies as SGA (10). Because preterm babies have a higher incidence of adverse perinatal outcomes, the fetal weight reference appears to have improved the SGA classification when using adverse perinatal outcomes as the evaluation criteria (11). But such improvement is significant only in preterm babies (6). Charkaluk et al. (12) further showed that very preterm children classified as SGA by the fetal weight reference had poorer cognitive or academic outcomes than those classified by the birth weight reference. To date, overwhelming evidence indicates that the fetal weight reference should be used to classify fetal and newborn size (SGA, appropriate for gestational age, and LGA) during pregnancy and at birth. Birth weight references should be abandoned in both clinical care and research.

However, fetal and newborn sizes differ quite substantially among countries and races (13). Whether this is due to nature or nurture is yet to be concluded (14, 15). Nevertheless, country- or race-specific fetal weight references are not always available or feasible to construct, especially in resource-poor countries. Indeed, most fetal weight references are based on Caucasian women in North America or Europe. Recently, a global reference that can be easily customized to fit a local population has been proposed (13). This reference has proven useful in better identifying preterm infants who have a higher risk of perinatal and infant mortality (5).

A question has been raised as to whether the customized reference by Gardosi et al. (3) further improves the classification compared with the unadjusted fetal weight reference (16). Several studies have shown that if one uses perinatal or infant mortality or severe morbidity as the evaluation criterion, these 2 methods produce similar results in term babies (i.e., full customization offers little to modest incremental value in term babies) (6, 17–20). Mikolajczyk et al. (13) demonstrated that after adjustment for race, further customization did not substantially improve the prediction of perinatal mortality.

Notwithstanding, mortality and severe morbidity are uncommon, dramatic outcomes. If the developmental origin of health and diseases is of concern for babies subject to fetal growth restriction, milder morbidities or health indices may be more sensitive and relevant to chronic adult diseases. In this regard, Carberry et al. (21), in this issue of the Journal, used newborn body fat percentage measured by air displacement plethysmography and a neonatal morbidity composite index (including hypothermia, prolonged hospital stay, and poor feeding) as an outcome measure. They found that, among term births, the fully customized reference does not improve the prediction of low or high body fat percentage or neonatal morbidity. This conclusion is consistent with a study by Charkaluk et al. (12), which showed that customization did not meaningfully improve the identification of very preterm infants at risk of poor cognitive or academic outcomes. Carberry et al. speculate that adjustment for maternal weight and height may overcontrol for any intergenerational growth stunting of the mother (21), which could be a nontrivial contribution (22). This issue might be particularly pertinent in emerging countries, where rapidly improved nutrition status in recent years may lead to an overestimation of LGA status after adjustment for maternal weight and height.

In theory, adjustment for maternal weight, height, and other factors should better classify fetal growth status. But in reality, several factors may limit its impact. First, the relationship between these factors (e.g., weight, height, race, sex, and parity) and outcomes (morbidity and mortality) may be too weak to show any noticeable effect. Hutcheon et al. (23) demonstrated through data simulation that, in order for customization to have real incremental value, the association between maternal characteristics (customizing factors) and birth weight must be “unrealistically” strong. Second, the greater the variation in customizing factors, the larger the adjustment effect. For example, the mean birth weight of black and white infants differs by about 200 g (5). An adjustment for race can lead to significant improvement. On the other hand, in a racially homogenous population, further customization may have limited incremental value. Finally, most reclassification of fetal/newborn growth status occurs in borderline cases (e.g., infants with birth weights around the 10th or 90th percentile). In term births, such infants generally have very low morbidity and mortality. Therefore, fine tuning the classification may not substantially increase the sensitivity and specificity of the classification.

![Figure 1](https://academic.oup.com/aje/article-abstract/178/8/1309/84368/1310-Zhang-and-Sun)
As more evidence suggests that fetal overgrowth may be associated with childhood obesity and adult metabolic disorders (24), the value of customization in LGA classification is drawing more attention. The study by Carberry et al. (21) showed no difference in term newborn body fat percentage between customized and population-based references. Several studies also failed to find important differences in neonatal morbidity and infant mortality between customized and noncustomized classifications of LGA (5, 25).

It should be noted that the vast majority of studies on this issue addressed a research question and used neonatal or pediatric outcomes as the “gold standard” to assess the utility of customization. However, in the clinical management of pregnant women, a more accurate classification of fetal growth status may also have important clinical and cost implications. For example, decisions about the close monitoring of fetal growth and the timing and route of delivery could be influenced by the accuracy of classification. In this regard, adjustment for the sex of the fetus may be made to allow more precise classification, whereas adjustment should not be made for parity (26). Larkin et al. (27) found that pregnancies with infants classified as LGA by the customized reference carried an increased risk of shoulder dystocia, third- or fourth-degree laceration, and cephalopelvic disproportion than those classified by a birth weight reference. Nevertheless, the obstetric implications of customization have not been well addressed in the literature. Errors in ultrasonography-based estimates of fetal weight may be another factor that counter-balances the advantage of fine tuning by full customization (28). Ultimately, randomized controlled trials comparing customized versus noncustomized references for fetal growth assessment may be necessary to provide a definitive answer for obstetric implications (29).

In conclusion, the incremental value of the customization of fetal growth references depends on several factors. The theoretical advantage of full customization may have limited practical value for most women, particularly those with term births that have a low incidence of adverse pediatric outcomes. Country-specific fetal growth references may be necessary to reflect large differences in race/ethnicity and morbidity and mortality among countries. Further fine tuning in a racially/ethnically homogeneous population may have limited incremental value for pediatric outcomes, although adjustment for maternal body size may still be useful for a few very large or small women. Literature on the prediction of short-term maternal outcomes and very long-term outcomes (adult diseases) is too scarce to draw any conclusions. Given that each additional variable being incorporated in the classification scheme increases complexity and costs in practice, the clinical utility of full customization in obstetric practice requires further testing.

acknowledgments

author affiliations: MOE-Shanghai Key Laboratory of Children’s Environmental Health, Xinhua Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China (Jun Zhang); and Department of Pediatric Cardiology, Xinhua Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China (Kun Sun).

J.Z. is supported by a grant from the National Natural Science Foundation of China (no. 81273091).

Conflict of interest: none declared.

REFERENCES