Customized vs INTERGROWTH-21st standards for the assessment of birthweight and stillbirth risk at term

Andre Francis, MSc; Oliver Hugh, BSc (Hons); Jason Gardosi, MD, FRCOG

BACKGROUND: Fetal growth abnormalities are linked to stillbirth and other adverse pregnancy outcomes, and use of the correct birthweight standard is essential for accurate assessment of growth status and perinatal risk.

OBJECTIVE: Two competing, conceptually opposite birthweight standards are currently being implemented internationally: customized gestation-related optimal weight (GROW) and INTERGROWTH-21st. We wanted to compare their performance when applied to a multiethnic international cohort, and evaluate their usefulness in the assessment of stillbirth risk at term.

STUDY DESIGN: We analyzed routinely collected maternity data from 10 countries with a total of 1.25 million term pregnancies in their respective main ethnic groups. The 2 standards were applied to determine small for gestational age (SGA) and large for gestational age (LGA) rates, with associated relative risk and population-attributable risk of stillbirth.

The customized standard (GROW) was based on the term optimal weight adjusted for maternal height, weight, parity, and ethnic origin, while INTERGROWTH-21st was a fixed standard derived from a multiethnic cohort of low-risk pregnancies.

RESULTS: The customized standard showed an average SGA rate of 10.5% (range 10.1-12.7) and LGA rate of 9.5% (range 7.3-9.9) for the set of cohorts. In contrast, there was a wide variation in SGA and LGA rates with INTERGROWTH-21st, with an average SGA rate of 4.4% (range 3.1-16.8) and LGA rate of 20.6% (range 5.1-27.5). This variation in INTERGROWTH-21st SGA and LGA rates was correlated closely (R = ±0.98) to the birthweights predicted for the 10 country cohorts by the customized method to derive term optimal weight, suggesting that they were mostly due to physiological variation in birthweight. Of the 10.5% of cases defined as SGA according to the customized standard, 4.3% were also SGA by INTERGROWTH-21st and had a relative risk of 3.5 (95% confidence interval, 3.1-4.1) for stillbirth. A further 6.3% (60% of the whole customized SGA) were not SGA by INTERGROWTH-21st, and had a relative risk of 1.9 (95% confidence interval, 3.1-4.1) for stillbirth. An additional 0.2% of cases were SGA by INTERGROWTH-21st only, and had no increased risk of stillbirth. At the other end, customized assessment classified 9.5% of births as large for gestational age, most of which (9.0%) were also LGA by the INTERGROWTH-21st standard. INTERGROWTH-21st identified a further 11.6% as LGA, which, however, had a reduced risk of stillbirth (relative risk, 0.6; 95% confidence interval, 0.5-0.7).

CONCLUSION: Customized assessment resulted in increased identification of small for gestational age and stillbirth risk, while the wide variation in SGA rates using the INTERGROWTH-21st standard appeared to mostly reflect differences in physiological pregnancy characteristics in the 10 maternity populations.

Key words: birthweight, customized growth charts GROW, epidemiology, ethnicity, fetal growth, INTERGROWTH-21st, large for gestational age, pregnancy risk, small for gestational age, stillbirth

Introduction
Fetal growth restriction and low birthweight are closely linked to risk of stillbirth and other indicators of adverse perinatal outcome. As these associations have become ever clearer, the focus has shifted to prevention, which requires adequate tools and standards.

Many reference curves and tables have been produced in various settings for the assessment of fetal growth and birthweight. They can vary because of the methods used, the quality of the data they originated from, and whether they were based on longitudinal or cross-sectional, fetal, or neonatal data. They also vary with the physiological and pathological characteristics of the population. Therefore, an approach that has gained traction in recent years is not to base reference curves on the whole population, but to set a standard that seeks to represent the optimal growth and birthweight that can be achieved in the absence of any complications, and that therefore should be better able to detect abnormalities in fetal growth.

Such a standard has been developed as the computer-generated customized GROW chart, which uses coefficients derived from large birthweight databases to predict optimal growth for each mother in each pregnancy. Physiological variables such as ethnic origin, maternal size, and parity are adjusted for, and the standard is set at a level that is free from pathology, so that the effect adverse influences such as smoking, hypertension, or diabetes, are better recognized. Because the construction of the standard combines a term optimal weight (TOW) with a proportionality fetal weight curve for all gestations, the same chart can be used for the assessment of fetal growth as well as birthweight. Customized charts have been shown to be internationally applicable, are recommended by the Royal College of Obstetricians and Gynecologists, and are now increasingly in clinical and international research use. The GROW (Gestation Related Optimal Weight) application has recently been updated with additional coefficients to represent over 100 ethnic or country-of-origin groups.

An alternative approach to derive a standard is that taken by the INTERGROWTH-21st (IG21) project, which selected low-risk, well-nourished mothers with uncomplicated pregnancies. Data were combined from cohorts in 8 countries to produce a single, prescriptive, multiethnic standard for...
birthweight\textsuperscript{10,11} and fetal growth\textsuperscript{12,13} to be used universally. The recently published World Health Organization fetal growth project,\textsuperscript{14} based on data from 10 countries, used similar methodology, but concluded that there were significant differences between populations in maternal characteristics that affected growth. Similarly, the \\textit{Eunice Kennedy Shriver} National Institute of Child Health and Human Development Fetal Growth Studies\textsuperscript{15} and other studies\textsuperscript{16-18} demonstrated ethnic differences in fetal growth in low-risk pregnancies. Nevertheless, the IG21 standards are being actively promoted and have begun to be implemented in many settings.

We therefore set out to compare the IG21 birthweight standard with the individually customized (GROW) standard in an international cohort based on maternity datasets from 10 countries, to assess how well they were able to associate birthweight with stillbirth risk. We focused our analysis on term data, as preterm birthweight ought to be assessed with a fetal rather than a neonatal weight standard in light of the known associations between prematurity and fetal growth restriction.\textsuperscript{19-21}

\section*{Materials and Methods}

\subsection*{Data source}

The Perinatal Institute administers the Gestation Network (www.gestation.net), which is a portal for provision of free software tools including customized centile calculators for local, national, and international use. The applications contain coefficients for adjustment of the growth and weight standard according to maternal characteristics, derived from anonymized databases submitted from clinicians and researchers who wish to have an application suitable for their own local population. To date, datasets from 23 countries have been received totaling 3.2 million births. Based on this database, the first global customized centile calculator was recently released, which can adjust for over 100 ethnic groups or countries of origin as well as the mother’s height, weight and parity, and the sex of the baby.

Ten of these Gestation Network data sets, totaling 2,140,543 cases, also contained stillbirth as a pregnancy outcome and represented the overall cohort used in this analysis. The origins of the data ranged from hospital-based collections to wider population-based registers, and included, in alphabetical order, datasets from Bhutan (national referral hospital), China (randomly selected births from 150 hospitals), Germany (State of Hesse birth register), India (large private tertiary maternity hospital in Hyderabad), Ireland (6 hospitals in the Perinatal Ireland network), The Netherlands (96 independent Dutch midwifery practices), Slovenia (national perinatal information system), Sweden (national medical birth registry), United Kingdom (83 maternity hospitals within the national growth assessment protocol (GAP) program), and United States (14 hospitals in the Washington State Obstetrics Clinical Outcome Assessment Program). The collaborators providing the data are listed under the Acknowledgment. All data were fully anonymized before receipt, and no institutional review board approval was required for this study.

Each dataset originated in settings with established routine ultrasound dating scans and these had been used to calculate gestational age at birth unless not available, in which case the last menstrual period was used. Maternal height and weight was measured at the beginning of pregnancy and ethnicity was recorded according to mother-declared ethnic origin or country of birth. Multiple pregnancies, congenital anomalies, and preterm births (<37 weeks) were excluded and only the predominant ethnic group from each country was included in the analysis, with complete data on maternal and pregnancy variables required for customized adjustment. This resulted in a study cohort of 1,251,289 cases. The stepwise exclusions are summarized in Table 1.

\begin{table}
\centering
\begin{tabular}{|l|c|c|c|}
\hline
Exclusions from original data submitted from 10 countries (2,140,543) resulting in cohort used in this study (1,251,289) & Excluded, n & Remaining, n & Remaining, \% \\
\hline
Congenital anomalies and multiple pregnancies & 57,322 & 2,083,221 & 97.3 \\
Missing or invalid gestational age or birthweight & 41,581 & 2,041,640 & 95.4 \\
Preterm deliveries (<259 d) & 121,676 & 1,919,964 & 89.7 \\
Minority ethnic group or missing ethnic origin data & 490,406 & 1,429,558 & 66.8 \\
Missing or invalid sex or maternal height, weight or parity & 178,269 & 1,251,289 & 58.5 \\
\hline
\end{tabular}
\caption{Exclusions from original data submitted from 10 countries (2,140,543) resulting in cohort used in this study (1,251,289)}
\end{table}

\section*{Standards for calculating centiles}

Small for gestational age (SGA) was defined as <10th, and large for gestational age (LGA) as >90th weight for gestational age centile, according to 2 methods:

1. Customized centiles were determined using the global centile calculator, entering the birthweight and gestational age at delivery, sex of the neonate, and information about maternal height, early pregnancy weight, parity (as it was at beginning of pregnancy), and ethnic origin. Coefficients for all predominant ethnic groups and associated maternal variables were available within the global centile calculator (GROW v.8.0.1).

2. IG21 centiles were based on the published IG21 neonatal weight-for-gestational age standard\textsuperscript{19} and included birthweight and gestational age at delivery as well as adjustment for neonatal sex.

Centiles for stillborn babies were also calculated according to the above
<table>
<thead>
<tr>
<th>Characteristics of 10 country cohorts</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>No.</strong></td>
</tr>
<tr>
<td>Maternal height, cm</td>
</tr>
<tr>
<td>Mean (SD)</td>
</tr>
<tr>
<td>Median (IQR)</td>
</tr>
<tr>
<td>Maternal weight, g</td>
</tr>
<tr>
<td>Mean (SD)</td>
</tr>
<tr>
<td>Median (IQR)</td>
</tr>
<tr>
<td>Body mass index</td>
</tr>
<tr>
<td>Mean (SD)</td>
</tr>
<tr>
<td>Median (IQR)</td>
</tr>
<tr>
<td>Nullipara %</td>
</tr>
<tr>
<td>Sex male %</td>
</tr>
<tr>
<td>Gestational age</td>
</tr>
<tr>
<td>Mean (SD)</td>
</tr>
<tr>
<td>Median (IQR)</td>
</tr>
<tr>
<td>Birthweight, g</td>
</tr>
<tr>
<td>Mean (SD)</td>
</tr>
<tr>
<td>Median (IQR)</td>
</tr>
<tr>
<td>Stillbirths</td>
</tr>
<tr>
<td>Frequency</td>
</tr>
<tr>
<td>Rate/1000</td>
</tr>
<tr>
<td>TOW1 standardized</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>TOW2 adjusted</td>
</tr>
<tr>
<td>Median gestation</td>
</tr>
<tr>
<td>GROW—SGA %</td>
</tr>
<tr>
<td>GROW—LGA %</td>
</tr>
<tr>
<td>IG21—SGA %</td>
</tr>
<tr>
<td>IG21—LGA %</td>
</tr>
</tbody>
</table>

GROW, gestation-related optimal weight; IG21, INTERGROWTH-21st birthweight standard; IQR, interquartile range; LGA, large for gestational age; SGA, small for gestational age; TOW, term optimal weight.

a Adjusted for para 0, height 163 cm, weight 64 kg, sex averaged, delivery at 280 d; b Adjusted for average parity, height and weight, sex averaged, delivery at median gestation.

methods, but with 2 days deducted from the gestational age at delivery, as an approximation of the gestational age at the time of intrauterine demise.22-24

SGA and LGA numbers, rates, and relative risk (RR) of stillbirths were presented according to the GROW and IG21 methods, as well as in subgroups according to whether they overlapped, ie, the birthweight was SGA or LGA by both standards, or was SGA or LGA by GROW or IG21 only.

Analysis
Statistical analyses were performed using software programs Excel (2016; Microsoft, Redmond, WA) and Stata (Version 14.2; StataCorp, College Station, TX). A descriptive table was constructed including data from each of the 10 country cohorts, showing mean, SD, median, and interquartile range for birthweight; gestation; and maternal height, weight, and body mass index (BMI), as well as listing gender, parity, and stillbirths rates. The multiple regression-derived ethnic constants were based on median gestation at delivery, and expressed as term optimal weight TOW in 2 ways: TOW1 adjusted to 280 days and standardized for maternal characteristics (para 0, height 163 cm, weight 64 kg, sex averaged, delivery at 280 days); and TOW2 adjusted for average maternal characteristics within that group, and at the respective median gestational age. RR with 95% confidence intervals (CI) and population attributable risk values were calculated.

To assess how IG21 standards relate to birthweight variation in the 10 country cohorts, IG21 SGA and LGA rates were compared with the predicted weight (TOW2) adjusted for average maternal height, weight, and parity and the median length of pregnancy in that cohort. The correlation between predicted and actual mean birthweights for the 10 country cohorts was high: \( R = 0.979 \) (Figure 1).

Results
Details of the 10 datasets are listed in Table 2. Averages with measures of dispersion are provided to illustrate the wide variation in maternal characteristics between country cohorts. Maternal height ranged from 155-170 cm, early pregnancy weight from 54-69 kg, median gestational age at delivery from 273-282 days, and median birthweight from 3040-3610 g.

Table 2 also shows the variation between cohorts in SGA and LGA rates according to the customized GROW and the IG21 methods of assessment. Average GROW SGA rate was 10.5% with a range of 10.1-12.7, while LGA averaged 9.5% with a range of 7.3-9.9. In contrast, average IG21 values were lower for SGA: 4.4%, ranging from 3.1-16.8, while average LGA rates were much higher: 20.6%, with a range of 5.1-27.5.

In Figures 2 and 3, the adjusted TOW constants (TOW2) for each of the 10 country cohorts are plotted against the IG21 SGA and LGA rates, respectively. The relationship with IG21 SGA rates (Figure 2) is curvilinear, with the 2 cohorts with the lowest predicted weights assigned very high SGA rates of 16.8 (India) and 14.4 (Bhutan). For LGA (Figure 3), there is a direct, linear relationship between the predicted weight and IG21-determined LGA rates. These significant correlations suggest that the varying proportions of cases identified as SGA or LGA with IG21 merely reflect normal variation in birthweight between these country cohorts. Tables 3 and 4 detail the association between stillbirths and SGA and LGA.
The highest RR for stillbirth (3.5; 95% CI, 3.1–4.1) was observed for babies SGA by both standards. IG21 adds only another 2087 SGA cases including 3 stillbirths, which do not represent an elevated stillbirth risk. In contrast, using GROW, a further 60% of births (78,703 of 131,950) are categorized as SGA and add another 185 (45%) to the 226 stillbirths identified by both methods, with RR 1.9 (95% CI, 1.6–2.2).

At the other end of the spectrum, being LGA by both standards was not associated with stillbirth (RR, 0.9; 95% CI, 0.8–1.1) and according to GROW there were another 6792/118,954 or 5.7% LGA with 13 stillbirths and no effect on risk. IG21 however classified a further 56.5% (145,570 of 257,732) as LGA but these cases had in fact a lower RR for stillbirth.

**Comment**

This is, to our knowledge, the first multinational comparison of the IG21 and customized birthweight standards. It shows firstly that using IG21, there are wide differences in SGA and LGA rates across the 10 cohorts studied, ranging from 3.1–16.8% for SGA and 5.1–27.5% for LGA rates. As Figures 2 and 3 show, these values are strongly correlated with the TOW calculated by GROW for each cohort, suggesting that IG21 SGA and LGA rates vary mostly due to physiological differences between different populations.

For example, the high IG21 SGA rate for India (16.8%) (Table 2) most likely represents physiological variation due to ethnic origin and small maternal size, as the data represent a mostly middle-class Indian population that has a GROW SGA rate of just 11.3%. GROW centiles adjust only between normal BMI limits; for example if a mother’s BMI is 17, the GROW software will limit downward adjustment to 18.5 when calculating the predicted term weight. This means that GROW adjustments do not extend to birthweights that might reflect undernutrition. The IG21 SGA rate is high because physiological maternal characteristics are not taken into account, and the correlation with the GROW TOW confirms that IG21 SGA is mostly dictated by physiological variation. Furthermore, this exaggerated IG21 SGA rate does not correspond to the stillbirth rate, which was not elevated in this predominantly middle-class maternity population receiving high standard of care.

GROW centiles applied to the population cohorts had an overall narrower range of values (SGA 10.1–12.7, LGA 7.3–9.9) than that obtained with IG-21. The actual SGA (<10th centile) rate tends to be above 10%, as in any maternity population it is more likely that fetuses do not fulfill their predicted growth potential due to pathological influences than exceed it. The overall SGA rate tends to be higher if the cohort includes preterm deliveries due to their association with growth restriction, or if they are derived from high-risk referral centers with an elevated risk level for fetal growth problems and associated perinatal morbidities. None of the cohorts showed a GROW SGA rate as low as the average 4.4% displayed by the IG21 standard.

Our findings are consistent with that of Anderson and colleagues, who also compared the IG21 birthweight standard with customized GROW centiles, applying them to their Auckland database. The authors reported similarly low SGA rates with IG21, and substantial variation within their main ethnic groups, which did not reflect pathological outcomes.

Lee and colleagues recently applied the IG21 standard to the Child Health Epidemiology Reference Group (CHERG) dataset of 14 birth cohorts from low- and middle-income countries, and reported SGA rates ranging from 5% or 6% in Eastern Asia and Northern Africa, to 34% in India. While countries in South Asia and sub-Saharan Africa had high SGA rates as well as high neonatal mortality rates, causality was not demonstrated, and the association was contradicted by other countries or
regions with elevated mortality risk that had low SGA rates according to IG21.

Misclassification with a one-size-fits-all model that does not adjust for physiological variation can have considerable unwanted clinical effects. Antenatal overdiagnosis of SGA may lead to unnecessary investigations, intervention, and anxiety for babies with appropriate growth and size for their population. At the same time, with the falsely low IG21 SGA rate, many at-risk babies will go unrecognized because they are being classified as not <10th centile, and we have shown this to be the case in 60% of the at-risk population (Table 3). Postnatally, babies falsely considered SGA may receive unnecessary supplementary feeding to compensate for an imagined deficit. The wrong standard may result in misdirection of available resources in the target population, and a loss of focus on identifying babies truly at risk.

A weakness of this study is that we had only stillbirth as recorded outcome measure in these 10 cohorts, which is not the only relevant outcome measure to assess fetal growth abnormalities. In particular, LGA babies may have complications such as shoulder dystocia and associated morbidities due to birth trauma. There have, however, been several studies with such outcome data that compared customized and population-based centiles for macrosomia and found the customized definition of LGA to be superior and able to identify an additional group in the population that is also at risk of complications. A standard such as IG21 that consistently classifies >20% of cases as LGA is likely to lead to excessive maternal anxiety and unnecessary interventions.

Inclusion of LGA in our analysis allowed a look at both ends of the spectrum. It shows that the low IG21 SGA rate is accompanied by a high LGA rate, indicating that the standard not only ignores physiological variation but is overall too low for this population. According to the published IG21 formula, the predicted weight at 40 weeks is 3380 g (boys) and 3260 g (girls), while the optimal 40.0-week weight predicted by the customized standard for this cohort (3561 g) as well as the actual weight reached (3490 g) were substantially higher. The datasets we studied recorded birthweight and not fetal weight and had the benefit of being routinely collected, whereas fetal weight measurements at term are likely to represent a smaller, selected population that had indications for ultrasound scans. Nevertheless, our cohort (3561 g) as well as the actual weight reached (3490 g) were substantially higher.

The datasets we studied recorded birthweight and not fetal weight and had the benefit of being routinely collected, whereas fetal weight measurements at term are likely to represent a smaller, selected population that had indications for ultrasound scans. Nevertheless, our

| FIGURE 3 | Predicted customized birthweight vs INTERGROWTH-21st LGA rate |

Predicted birthweight customized for average maternal height, weight, parity and sex, and controlled for gestational age at delivery. \( R = 0.9775; P < .01 \).

UK, United Kingdom; LGA, United States; LGA, large for gestational age.


| TABLE 3 | SGA by GROW and INTERGROWTH 21st and stillbirth risk |

<table>
<thead>
<tr>
<th>SGA</th>
<th>SGA by GROW</th>
<th>SGA by IG21</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classified as SGA, n (%)</td>
<td>131,950 (10.5)</td>
<td>55334 (4.4)</td>
</tr>
<tr>
<td>Stillbirths, n (/1000)</td>
<td>411 (3.1)</td>
<td>229 (4.1)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>SGA by GROW only</th>
<th>SGA by both standards</th>
<th>SGA by IG21 only</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classified as SGA, n (%)</td>
<td>78,703 (6.3)</td>
<td>53,247 (4.3)</td>
<td>2087 (0.2)</td>
</tr>
<tr>
<td>Stillbirths, n (/1000)</td>
<td>185 (2.4)</td>
<td>226 (4.2)</td>
<td>3 (1.4)</td>
</tr>
<tr>
<td>Relative risk (95% CI)</td>
<td>1.9 (1.6–2.2)</td>
<td>3.5 (3.1–4.1)</td>
<td>1.1 (0.4–3.4)</td>
</tr>
<tr>
<td>Population attributable risk %</td>
<td>5.1</td>
<td>9.7</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Total N = 1,251,289.

CI, confidence interval; GROW, gestation-related optimal weight; IG21, INTERGROWTH-21st birthweight standard; SGA, small for gestational age.

Statistically significant relative risk values are shown in bold.

findings here are also relevant for the IG21 fetal weight standard,13 which is being proffered for international use under the same one-size-fits-all assumption. The equivalent 40-week value of the IG21 fetal growth formula13 is 3338 g and is therefore also unlikely to represent growth curves that are suitable for this multicountry cohort.

Our results confirm doubts about the one-size-fits-all approach10 and improve our understanding of the reported difficulties in local implementation of IG21 fetal and neonatal standards in various environments.25,31,32 We demonstrate the substantial variation in maternal and physiological pregnancy characteristics across population cohorts, and present evidence that the varied SGA and LGA rates using the IG21 formula mostly reflect physiological variation, which blunts the standard’s ability to identify pathology. Finally, we have shown that GROW as a globally applicable but individually adjustable standard improves the strength of association with stillbirth as an adverse outcome, and identifies 60% more SGA cases at increased risk.

Acknowledgment

We would like to thank the collaborators and their organizations for supplying data for this multinational cohort study: Prof Phurb Dorji and A. Yangden, Jigme Dorji Wangchuck National Referral Hospital, Thimphu, Bhutan; Prof Jun Jim Zhang, Jiao Tong University, Shanghai, China; Prof Birgit Arabin and Dr Bjorn Miel- selwitz, Gesellschaftstelle Qualitaetsicherung Hessen, Eschborn, Germany; Dr Susieela Vavilala and Dr Nuzhat Aziz, Fernandez Hospital, Hyderabad, India; Prof Fergal Malone, Perinatal Ireland, Dublin, Ireland; Joost Faassen, LinkORB BV, Oirschot, The Netherlands; Prof Tanja Premru-Srsen, University Medical Center, Ljubljana, Slovenia; Dr Britt Claussen, Socialstyrelsen Medical Birth Register, Stockholm, Sweden; Kristin Sticov and Dr Vivienne Souter, Foundation for Health Care Quality, Seattle, WA.

References

21. Groom K, North R, Poppe K, Sadler L, McCowan L. The association between customized small for gestational age infants and pre-eclampsia or gestational hypertension
varies with gestation at delivery. BJOG 2007;114:478-84.

Author and article information
From the Perinatal Institute, Birmingham, United Kingdom.

Received Nov. 20, 2017; revised Dec. 7, 2017; accepted Dec. 8, 2017.

Disclosure: All authors are employees of the Perinatal Institute, a not-for-profit organization that derives income from the provision of support services to health care organizations, which may include the GROW software for customized growth charts mentioned in this article.

Corresponding author: Jason Gardosi, MD, FRCOG. jgardosi@perinatal.org.uk