Performance of the Intergrowth-21st and World Health Organization fetal growth charts for the detection of small for gestational age neonates in a population from Latin America

Jezid Miranda[1], Natalia Maestre[2], Angel Paternina[3], Miguel Parra-Saavedra[4], Javier Caradeux[5], Alvaro Sepulveda-Martinez[6], Melisa Pelaez[7], Andrés Torrres[8], Mauro Parra-Cordero[9], Pilar Diaz[5], Dahiana Gallo[10], Dario Santacruz[11], Nicolás Rodriguez[12], Andres Sarmiento[13], Jesus Benavides[14], Sergio Girado[1], Jose Rojas-Suarez[15], Eduard Gratacos[16], and Francesc Figueras[17]

[1]Cartagena University Faculty of Natural Sciences
[2]GRICIO
[3]Universidad de Cartagena
[4]Universidad Simón Bolívar Facultad de Ciencias de la Salud
[5]Clínica Dávila
[6]Hospital Clinico de la Universidad de Chile
[7]Hospital Nacional Docente Madre Niño San Bartolomé
[8]Instituto Mexicano del Seguro Social
[9]University of Chile Hospital
[10]Universidad del Valle
[11]Clínica Versalles
[12]Fundacion Santa Fe de Bogota
[13]Universidad de los Andes
[14]Universidad Tecnológica de Pereira
[15]Universidad de Cartagena Facultad de Medicina
[16]BCNatal - Barcelona Center for Maternal-Fetal and Neonatal Medicine (Hospital Clinic and Hospital Sant Joan de Deu)
[17]Hospital Clinic de Barcelona

April 25, 2022

Abstract

Objective: To evaluate the performance of INTERGROWTH-21st and World Health Organization (WHO) fetal growth charts to identify small-for-gestational-age (SGA) and fetal growth restriction (FGR) neonates as well as their specific risks for adverse neonatal outcomes. Design: Multicenter cross-sectional study. Setting: Ten maternity units across four Latin American countries, 2016-2018. Population: 67,968 singleton live births. Methods: According to each standard, the neonates were classified as SGA and FGR (birthweight <10th and <3rd centiles, respectively). Main Outcomes Measures: The relative risk (RR) and diagnostic performance for the occurrence of a low Apgar score and low ponderal index were calculated for each standard. Results: The WHO charts identified more neonates as SGA than IG-21st (13.9% vs. 7%, respectively). 6.9% babies were considered SGA only by the WHO chart. Compared to normally grown babies, neonates classified as FGRs by both standards had the highest RR for a low Apgar (RR: 5.57; 95% CI: 3.99–7.78), followed by those SGA by both curves (RR:

3.27; 95% CI: 2.52–4.24), while SGAs identified by WHO alone did not have an additional risk (RR: 0.87; 95% CI: 0.55–1.39). Furthermore, the diagnostic odds ratio for a low Apgar was higher when INTERGROWTH-21st was used than when SGA and FGR were defined by WHO charts. Conclusions: In a large population of singleton deliveries from Latin America, the WHO fetal growth charts seem to identify significantly more SGA neonates than the INTERGROWTH-21st charts, but the diagnostic performance of the latter for low Apgar score and low ponderal index is better.

Introduction

Small-for-gestational-age (SGA) neonates are at increased risk of mortality and several morbidities (1,2), suboptimal neurodevelopment (3–7), and susceptibility to cardiovascular disease later in life (8–10). Unrecognized SGA fetuses are at increased risk of perinatal death compared with those who are appropriately followed and managed (11). Its opportune identification allows timely interventions to reduce the risk of adverse perinatal outcomes (APOs) (12–15). Current guidelines recommend the 10^{th} percentile as a cutoff to define SGA and the 3^{rd} percentile to define fetal growth restriction (FGR) (16–18) since several studies have demonstrated an increased risk of perinatal morbidity and mortality beyond these cutoffs (17–23). However, there is disagreement on which charts should be used (16–21,24,25).

Two international standards for fetal growth have been constructed and published as a global effort to reduce the reported variability and the worldwide discrepancy when defining fetal growth restriction. First, the INTERGROWTH-21^{st} (IG-21^{st}) project reported fetal biometry standards constructed with 20,486 low-risk pregnancies delivered between 33 and 42 weeks (26–29). Using a similar concept and methodology, the World Health Organization (WHO) multicenter growth reference study proposed an alternative standard (30). However, previous studies evaluating the diagnostic performance of these fetal growth standards in different populations have reported conflicting results, preventing their worldwide adoption and implementation (31–37).

Latin America represent one of the most unequal regions globally regarding maternal and perinatal health (38–41). The region demonstrates an excess in stillbirths with an estimated rate of 8.2 stillbirths per 1000 births (95% CI 7.5-9.2) (42–44) and approximately, 60% of deaths before the age of five years old in the region occur during the first year of life, with 50% of those during the first 28 days (45). Potential differences in diagnosis of SGA among physicians in Latin America region can exacerbate an inappropriate use of the limited health resources, disadvantaging outcomes of SGA infants. At present, no studies have been performed comparing the performance of both standards to identify SGA neonates in Latin America. Therefore, the objectives of this study were to evaluate the diagnostic performance of INTERGROWTH-21^{st} and WHO fetal growth charts to identify SGA and FGR neonates and to assess the specific risks of adverse perinatal outcomes of SGA and FGR neonates identified by each fetal growth chart in a large cohort of deliveries from Latin America.

Methods

Design and study population

This was a multicenter cross-sectional study including all singleton livebirths between 24^{+0} and 41^{+6}weeks of gestational age occurring in ten obstetric centers from four Latin American countries (Colombia, Peru, Mexico, and Chile) between January 2017 and December 2018. Participating centers included 1) ESE Clínica de Maternidad Rafael Calvo (Colombia) (n=11009); 2) Clínica Santa Cruz de Bocagrande (Colombia) (n=6264); 3) ESE Hospital Local Cartagena de Indias (Colombia) (n=641); 4) ESE Hospital la Divina Misericordia (Colombia) (n=4526); 5) Clínica Versalles (Colombia) (n=16976); 6) Hospital Universitario de la Fundación Santa Fe de Bogotá (Colombia) (n=869); 7) Hospital Materno Nacional San Bartolomé de Lima (Perú) (n=11182); 8) Instituto Mexicano Seguro Social (Hospital General de Zona N014, Unidad de Quemados, Hermosillo, México) (n=4,799); 9) Hospital Clínico Universidad de Chile (Chile) (n=1705) and 10) Clínica Dávila (Chile) (n=9997). The exclusion criteria were as follows: 1) stillbirth; 2) missing data on birth weight, gestational age, infant sex, or maternal country of birth; 3) birth weight below 500 grams; and 4) multiple gestation.

Variables

Maternal baseline characteristics, including demographic details, obstetric history, and anthropometric measures at birth and perinatal outcomes, were collected from the hospital maternity records. Gestational age was calculated using maternal menstrual history or early prenatal ultrasound (before 20 weeks). A low Apgar score was considered below seven at the fifth minute. We used the standardized formula to calculate the ponderal index and applied reference ranges to the entire population (46,47). Similarly, the cephalization index (head circumference [cm]/birth weight [g] $\times$ 100) was calculated and applied to the entire population (48). The study protocol was approved by the ethics committee of the University of Cartagena (Ethics committee N 139, August 31, 2020).

Statistical analysis

According to their distribution, continuous variables were reported as the means or medians using interquartile ranges (IQRs) or standard deviations. Categorical variables are reported as percentages. First, the birth weight percentile was calculated using IG-21st software, and the calculation coefficients derived from the WHO study were used to calculate the WHO birth weight percentile. Then, for each growth chart (IG-21st and WHO standards), we calculated the proportion of live births with a birth weight below the $<10^{th}$ percentile (SGA) and $<3^{rd}$ percentile (FGR). To evaluate the relative validity of each reference growth chart, neonatal outcomes (i.e., low Apgar rate, ponderal, and cephalization indexes) between the "non-overlapping" populations were determined and compared with neonates at or above the 10^{th} percentile using the chi-squared test. Finally, relative risk (RR) was calculated as the ratio of the incidence of adverse perinatal outcomes among SGA and FGR neonates.

To account for a country-specific effect, we further evaluated the association of SGA by different standards with the adverse outcome using multilevel regression analyses, where the subjects were at the lower level and countries at the upper level. The relationships between patient-level and country-level variables and the adverse perinatal outcomes were examined with multilevel linear regression using the R 'lm4' package. Fixed effects were estimated for maternal education and nulliparity. The multilevel analysis was implemented in a stepwise manner. First, an unconditional means model was used to determine the attributable variance explained by the multilevel design. Second, using a backward elimination approach, all selected variables for inclusion were added to the unconditional means model as fixed effects, and nonsignificant variables were removed sequentially until only significant (i.e., $p<0.05$) variables remained. Finally, diagnostic performance (sensitivity; specificity; positive and negative likelihood ratio; and the diagnostic odds ratio) was estimated and used to compare the accuracy of the two fetal growth standards to identify neonates at risk of adverse perinatal outcomes. We compared the likelihood and diagnostic odds ratios by bootstrapping 2000 replicates with replacement. The receiver-operating characteristics (ROC) curve analysis determined the performance for predicting a low APGAR score and ponderal index by each fetal growth standard was determined by the receiver–operating characteristics (ROC) curve analysis. The resulting areas under the ROC curves (AUCs) were compared using the DeLong method, and a p-value <0.05 was considered statistically significant. Data processing was performed using R software. A value of $p<0.05$ was considered statistically significant.

Results

The study included 70,852 pregnant women who delivered live births. A total of 1293 (1.9%) pregnancies were excluded due to multiple gestations (n=1273), birth weight less than 500 grams (n=9), and missing data (n=309).

Population description

Following exclusions, we considered 67,968 deliveries for the analysis. Table S1 summarizes each country's contribution to the overall population. The median maternal age was 26 (IQR: 22 – 31) years, with differences across countries. There were also differences in nulliparity rate, ethnicity, and educational level. The median gestational age at delivery in the study population was 39.5 (IQR: 38.5 – 39.5) weeks. The rate of preterm delivery was 7.9% (5359/67,968). The proportion of neonates classified as SGA was significantly different

between the two standards. The WHO growth standard classified the neonates as follows: at or above the 10^{th} percentile: 58,542 (86.1%) and SGA: 9426 (13.9%), while for IG-21^{st}, 63,244 (93%) neonates were at or above the 10^{th} percentile, and 4724 (7%) neonates were identified as SGA. Thus, the rate of neonates classified as SGA by the IG-21^{st} was almost two times lower than that classified by the WHO (7 vs. 13.9%, $p<0.001$). Similarly, the proportion of neonates classified as FGR was significantly different between the two standards. The WHO growth standard classified the neonates as follows: at or above the 3^{rd} percentile: 63,730 (93.8%) and FGR: 4238 (6.2%), while for IG-21^{st}, 66,517 (97.9%) neonates were at or above the 3^{rd} percentile, and 1451 (2.1%) neonates were classified as FGR. Thus, the rate of neonates classified as FGR by the IG-21^{st} was almost three times fewer than that classified by the WHO (2.1% vs. 6.2%, $p<0.001$).

Figures 1A and 1B are Venn diagrams describing the classification of newborns according to the percentiles of each standard (at or above the 10^{th} percentile vs. SGA and FGR) using both standards simultaneously. Specifically, 86.1% (58,523/67,968) were considered at or above the 10^{th} percentile by both standards, 6.95% (4721/67,968) of neonates were classified as SGA only by the WHO standard (SGA-WHO only), 0.03% (19/67,968) of neonates were classified as SGA only by the IG-21 standard (IG-21^{st}- only), and 6.92% (4705/67,968) were classified as SGA by both standards (Figure 1A). All neonates identified as SGA by IG-21^{st} alone were preterm births. With respect to FGR, 93.7% (63,718/67,968) were considered above the 3^{rd} percentile by both standards, 4.1% (2799/67,968) of neonates were classified as FGR only by the WHO standard (FGR-WHO only), 0.02% (12/67,968) of neonates were classified as FGR only by the IG-21 standard (IG-21 only), and 2.1% (1439/67,968) were classified as SGA by both standards (Figure 1B).

Table 1 describes clinical characteristics and perinatal outcomes for pregnancies assigned as SGA and FGR by WHO standard alone, by both standards, and those classified as at or above the 10^{th} percentile for both curves. The rate of preterm delivery was higher in the newborns classified as FGR by WHO alone and by both standards than in those classified as above the 10^{th} percentile by the two standards (all p values <0.001). However, the rate of preterm delivery was significantly higher in those classified as SGA by the two curves than in those classified as AGA (16.3% vs. 7.32%, $p<0.05$). In addition, there were significant differences in the cesarean section rate among the groups, being significantly higher in neonates classified as SGA and FGR by both standards compared to those classified as above the 10^{th} percentile by the two standards (59.4% vs. 46.2%, $p<0.001$) (Table 1).

The rate of a low APGAR score was significantly higher for neonates classified as SGA and FGR by both standards (1.51% and 2.64%, respectively), followed by neonates classified as FGR only by the WHO (0.82%), being significantly lower in neonates classified as at or above the 10^{th} percentile by both curves (0.46%). Notably, there were no significant differences in the rate of low APGAR scores between those neonates classified as SGA only by WHO and neonates classified as at or above the 10^{th} percentile by both curves (0.40% vs. 0.46%, $p=0.64$, Table 1). Figures 2A and 2B show the RRs for a low APGAR score or ponderal index, respectively, in neonates identified as SGA and FGR. Neonates classified as SGA and FGR by both standards exhibited the most significant RR for an APGAR score below seven at five minutes (RR: 3.27; [95% CI: 2.52 – 4.24], and 5.57 [3.99 – 7.78], respectively). Importantly, neonates classified as SGA only by WHO alone did not have a significantly higher risk of a low APGAR score (RR: 0.87; 95% CI: 0.55 – 1.39) (Figure 2A).

The median ponderal index was significantly lower in the group of neonates classified as SGA and FGR by both standards than in those classified as above the 10^{th} percentile by both standards (FGR by both standards: 22.5 [IQR: 20.6 – 24.7] & SGA by both standards: 23.5 [IQR: 21.7 – 25.4] vs. 26.7 [24.9 – 28.6]; all p values <0.001). Similarly, the rate of a ponderal index below the 5^{th} percentile was significantly higher in these groups. Neonates classified as SGA and FGR by both standards exhibited the most significant RR for a low ponderal index (RR: 11.95; [95% CI: 10.7 – 13.4], and 14.9 [13.2 – 16.8], respectively) (Figure 2). Furthermore, neonates classified as SGA only by WHO alone also had a significantly higher risk of a low ponderal index (RR: 4.75; 95% CI: 4.1 – 5.53) (Figure 2B). Finally, the cephalization index was significantly higher in neonates classified as SGA by WHO alone and in those classified as SGA and FGR by both standards, displaying, in addition, a trend toward worst values in the latter groups (Table 2). Table S2

shows the odds ratios of SGA by each standard for neonatal outcomes under a hierarchical logistic regression model. In brief, we found that SGA babies only by WHO had an OR of 0.98 (95% CI: 0.61 – 1.57) for a low APGAR score at five minutes and 4.14 (95% CI: 3.52 – 4.86) and a ponderal index below the 5th percentile, respectively.

Table 2 displays the diagnostic performance of the WHO and IG-21st for identifying an APGAR score below seven at 5 minutes and a ponderal index below the 5thpercentile. Both charts exhibited low sensitivities for low Apgar scores (below 30%) and high specificity. We next assessed the diagnostic effectiveness of both fetal growth charts for specific obstetric outcomes, demonstrating that the IG-21st had the highest diagnostic odds ratios (DORs) (Table 2). As an overall measure of diagnostic performance for a low Apgar score, the diagnostic odds ratio was higher when SGA (3.70 vs. 2.02, mean difference: 0.61, 95% CI: (0.45 – 76.6), p <0.001) and FGR (6.22 vs. 3.01, mean difference 0.72, 95% CI: (0.48 – 0.96), p <0.001) were defined by IG-21st than by WHO charts. Similarly, the diagnostic odds ratio for a low ponderal index was also higher when SGA (10.4 vs. 9.01, mean difference 0.14, 95% CI: (0.06 – 0.23) p-value = 0.001) and FGR (14.6 vs. 10.6, mean difference 0.32, 95% CI: (0.2 – 0.42) p<0.001) were defined by IG-21st than by WHO charts. When we applied both fetal growth charts for the identification of a low APGAR score and ponderal index, the IG-21st fetal growth charts marginally improved the prediction of a low APGAR score based on the area under the receiver operating characteristic (ROC) curve (AUC), estimated using 2,000-fold bootstrapping to account for overfitting (Table S3). Specifically, for low APGAR scores, the AUC of the identification of SGA neonates for WHO fetal growth charts were 55.3 (95% CI: 53.1 – 57.5) vs. 57.3 (55.2 – 59.4) for IG-21st, p *=0.005* (two-sided) (Table S3).

Discussion

Main findings

In this large multicenter study, including an unselected population of singleton pregnancies from four countries in Latin America, the proportion of SGA and FGR neonates identified by the WHO fetal growth standard was significantly higher than that obtained using the IG-21st standards. Nevertheless, the overall diagnostic performance for the adverse neonatal outcome and the low ponderal index was better when IG-21st defined SGA and FGR.

Comparison with results of previous studies and interpretation of results

The ability of the IG-21st standard to identify fetuses and neonates at risk of adverse outcomes has been recently challenged by several studies worldwide. Those studies have consistently reported that the use of IG-21st resulted in a lower prevalence of SGA compared with reference (31,32,34,36,49) or customized charts (31). Moreover, undiagnosed SGA fetuses are at increased risk of adverse perinatal outcomes and stillbirth (31,37,50). Importantly, similar to other reports from developed countries (37,42–44), we reported that the IG-21st chart identified fewer neonates as SGA and that the Latin American population turns to the right in the distribution percentiles within the IG-21st standard.

Another important finding is that the WHO identified an additional group of 4721 SGA babies who were not at significant risk of a low APGAR score. However, they have anthropometric features resembling intrauterine growth restriction. There have been several explanations for the discrepancy between the two standards. One explanation is that the calculation of EFW in the WHO study was based on the Hadlock formula (53), while IG-21st created a new formula based only on HC and AC.(54) On the other hand, IG-21st assumed a parametric distribution of the fetal growth trajectories under a linear mixed model. Researchers in the WHO project have used quantile regression to estimate percentiles directly and have fewer assumptions. It would be rational to assume, then, that compared to IG-21st, the aim of WHO was to be more of a reference, including pregnancies with complications. A previous study including 9409 women from the US reported limited accuracy of the IG-21st, NICHD, and WHO standards for identifying neonates at risk of adverse perinatal outcomes (including death) (49). IG-21st has been compared to customized charts reporting that IG-21st failed to detect SGA neonates, particularly among ethnic groups with larger maternal size.(31) Similar to our results, IG-21st classified fewer newborns as being below the 5th and 10thpercentiles by birth

weight than WHO and NICHD standards (49).

Human body proportions are thought to be the product of environmental and gene interactions, and they are notable differences across different races/ethnicities and countries (55). The ponderal index is an indicator of leanness in neonates. Previous studies have shown that asymmetric fetal growth (characterized by a low ponderal index) reflects fetal malnutrition (56), is associated with cerebral palsy (57,58), and increases the risk of perinatal morbidity and mortality (59). Developing countries might welcome using the cephalization index due to its low cost. Based on this index, some recommendations postulated that the greater the brain weight: body ratio (the more severe the intrauterine malnourishment), the higher the chances are for suboptimal brain development despite compensatory mechanisms such as brain sparing (48). Since intrauterine adverse events might not be clinically relevant until late in child development, it is crucial to identify as early as possible those small neonates at risk for neurodevelopmental disabilities who need early life interventions. This is especially relevant in deprived environments, where these interventions can improve cognitive performance and reduce antisocial behavior at a young age (60,61).

Strength and limitations

The strength of this study is that this birth dataset is the most extensive compilation to date from Latin America, including data from four countries and more than 67,000 births. In addition to the increased data quantity, we simultaneously evaluated the two current prescriptive international fetal growth standards to adjust the risk estimation of adverse perinatal outcomes and anthropometric measures associated with FGR. Differences in maternal age and antenatal care across countries might be due to population characteristics, culture, and obstetric practice. However, non-black Hispanics are currently used to agglomerate the Latin American population worldwide, so we did not consider ethnic differences within our population. Study limitations include the retrospective nature of this study. Another limitation of our study is that we only reported APGAR scores. However, this is an objective measure used to identify babies with a high risk of perinatal morbidity and poor neurological development. In addition, stillbirths were excluded because of uncertainty regarding their classification as SGA by birth weight. Other large series have shown that IG21st standards miss a fraction of babies at risk for this complex event (62). It is also a limitation of our study that we could not compare the performance of customized standards due to the lack of published and validated coefficients for all the participating countries. Although customized curves have been proposed (63–65), their superiority in identifying adverse perinatal outcomes has not been supported by more recent literature (66,67). Finally, although WHO detected a significantly higher proportion of SGA fetuses, this fraction of small fetuses likely contains instances of adverse outcomes that the data available (only Apgar) could not reveal.

Interpretation

There is a trade-off between maximizing sensitivity (few false negatives) and specificity (few false positives) in the chart selection. For SGA screening, using data from a previous large cohort study conducted in France, a false negative conferred an adjusted 2.1-increased risk for stillbirth (68). In absolute terms (according to a prevalence of stillbirth among detected SGA of 1%), this means one additional fetal death for each 87 non-detected SGA. The WHO charts exhibited higher sensitivity for SGA-associated adverse outcomes and a low ponderal index. However, false positives are also an issue to consider. A false positive for SGA means unnecessary follow-up and planned delivery, which should be at term in adherence to the international guidelines. A large cohort study in the UK showed that two otherwise normal small babies are picked up for every SGA fetus with complications identified (69). There is evidence from nationwide studies that compared with true negatives, iatrogenic preterm deliveries were 4.6 times higher than false positives. Thus, the ideal chart for fetal growth assessment should combine a good capacity to rule in and rule out SGA-associated complications. Under the assumption that the same weight is given to false negatives and positives, the diagnostic odds ratio [DOR] (+LHR/-LHR) estimates the performance. Especially for the definition of FGR, the IG-21st charts exhibited a better overall performance in predicting low Apgar scores. Furthermore, the diagnostic performance for a low ponderal index (a surrogate of the thirty phenotypes) was better when SGA and FGR were defined using the IG-21st charts.

Research implications

Although, in the past years, the field's focus has been to answer which chart we should use, one alternative might be to identify cutoffs for each standard in which the perinatal morbidity increases, which might be not necessarily the 3rd or the 10thpercentile. Such perinatal risk-based cutoffs can be an opportunity to provide personalized care (52,70). This strategy might enhance the clinical applicability and use of the two standards while adapting to local scenarios. Environmental constraints are also a well-known factor influencing fetal growth, and usually, growth percentiles are not controlled adequately. Therefore, in addition to the argument about which chart should be used, the debate should move on to integrate functional parameters that enhance the fundamental objective in antenatal care, which is to assess placental function rather than fetal size.

Conclusion

In a large population in Latin America, the WHO fetal growth standard increases the identification of SGA and FGR neonates compared to the IG-21st project standard. Nevertheless, the former resulted in a lower overall diagnostic performance for a low Apgar score and low neonatal ponderal index.

Acknowledgments

We thank our medical and nursing colleagues of all participating centers who agreed to participate and contribute to this study.

Disclosure of interest

The authors report no conflict of interest.

Contribution to Authorship

JM, FF, and AP analyzed the data and drafted the manuscript. NM, MPS, JC, AS, MP, AT, MP, PDC, DG, DS, NR, AS, JAB, SG, JAR, EG, and FF interpreted the results and revised the manuscript critically for important intellectual content. FF contributed to the design of the study. All authors approved the final version of the manuscript.

Details of Ethics Approval

The ethics committee of the University of Cartagena (Ethics committee N 139, August 31, 2020) approved the study protocol.

Funding

This project has been funded with the support of Universidad de Cartagena (Resolución 00526 15 de Marzo de 2019).

References

1. Lawn JE, Blencowe H, Oza S, You D, Lee ACC, Waiswa P, et al. Every newborn: Progress, priorities, and potential beyond survival. Lancet. 2014;384(9938):189–205.

2. Lees C, Marlow N, Arabin B, Bilardo CM, Brezinka C, Derks JB, et al. Perinatal morbidity and mortality in early-onset fetal growth restriction: Cohort outcomes of the trial of randomized umbilical and fetal flow in Europe (TRUFFLE). Ultrasound Obstet Gynecol. 2013;42(4):400–8.

3. Arcangeli T, Thilaganathan B, Hooper R, Khan KS, Bhide A. Neurodevelopmental delay in small babies at term: a systematic review. Ultrasound Obs Gynecol. 2012;40(3):267–75.

4. Eixarch E, Meler E, Iraola A, Illa M, Crispi F, Hernandez-Andrade E, et al. Neurodevelopmental outcome in 2-year-old infants who were small-for-gestational age term fetuses with cerebral blood flow redistribution. Ultrasound Obstet Gynecol. 2008;32(7):894–9.

5. Baschat AA. Neurodevelopment following fetal growth restriction and its relationship with antepartum parameters of placental dysfunction. Ultrasound Obstet Gynecol. 2011;37(5):501–14.

6. Monteith C, Flood K, Pinnamaneni R, Levine TA, Alderdice FA, Unterscheider J, et al. An abnormal cerebroplacental ratio (CPR) is predictive of early childhood delayed neurodevelopment in the setting of fetal growth restriction. Am J Obstet Gynecol. 2019;221(3):273.e1-273.e9.

7. Paules C, Miranda J, Policiano C, Crovetto F, Youssef L, Hahner N, et al. Fetal neurosonography detects differences in cortical development and corpus callosum in late-onset small fetuses. Ultrasound Obstet Gynecol. 2021;58(1):42–7.

8. Gluckman PD, Hanson M a., Cooper C, Thornburg KL. Effect of in utero and early-life conditions and adult health and disease. N Engl J Med. 2008;359:1523–4; author reply 1524.

9. Barker DJP, Osmond C, Forsén TJ, Kajantie E, Eriksson JG. Trajectories of Growth among Children Who Have Coronary Events as Adults. N Engl J Med. 2005;353(17):1802–9.

10. Crispi F, Miranda J, Gratacós E. Long-term cardiovascular consequences of fetal growth restriction: biology, clinical implications, and opportunities for prevention of adult disease. Am J Obstet Gynecol. 2018;218(2):S869–79.

11. Gardosi J, Madurasinghe V, Williams M, Malik A, Francis A. Maternal and fetal risk factors for stillbirth: Population based study. BMJ. 2013;346(7893).

12. Richardus JH, Graafmans WC, Verloove-Vanhorick SP, Mackenbach JP, Masuy-Stroobant G, Alderdice F, et al. Differences in perinatal mortality and suboptimal care between 10 European regions: Results of an international audit. BJOG An Int J Obstet Gynaecol. 2003;110(2):97–105.

13. Nohuz E, Rivière O, Coste K, Vendittelli F. Prenatal identification of small-for-gestational age and risk of neonatal morbidity and stillbirth. Ultrasound Obstet Gynecol. 2020;55(5):621–8.

14. Flenady V, Koopmans L, Middleton P, Frøen JF, Smith GC, Gibbons K, et al. Major risk factors for stillbirth in high-income countries: A systematic review and meta-analysis. Lancet. 2011;377(9774):1331–40.

15. Gardosi J, Madurasinghe V, Williams M, Malik A, Francis A. Maternal and fetal risk factors for stillbirth: Population based study. BMJ. 2013;346(7893):f108.

16. Melamed N, Baschat A, Yinon Y, Athanasiadis A, Mecacci F, Figueras F, et al. FIGO (international Federation of Gynecology and obstetrics) initiative on fetal growth: best practice advice for screening, diagnosis, and management of fetal growth restriction. Int J Gynecol Obstet. 2021;152(1):3–57.

17. Martins JG, Biggio JR, Abuhamad A. Society for Maternal-Fetal Medicine Consult Series #52: Diagnosis and management of fetal growth restriction. Am J Obstet Gynecol. 2020;223(4):B2–17.

18. Lees CC, Stampalija T, Baschat AA, da Silva Costa F, Ferrazzi E, Figueras F, et al. ISUOG Practice Guidelines: diagnosis and management of small-for-gestational-age fetus and fetal growth restriction. Ultrasound Obstet Gynecol. 2020 Aug 1;56(2):298–312.

19. Lee ACC, Kozuki N, Cousens S, Stevens GA, Blencowe H, Silveira MF, et al. Estimates of burden and consequences of infants born small for gestational age in low and middle income countries with INTERGROWTH-21 st standard: Analysis of CHERG datasets. BMJ. 2017;358.

20. ACOG Practice Bulletin No. 204: Fetal Growth Restriction. Obstet Gynecol. 2019 Feb;133(2):e97–109.

21. Beune IM, Bloomfield FH, Ganzevoort W, Embleton ND, Rozance PJ, van Wassenaer-Leemhuis AG, et al. Consensus Based Definition of Growth Restriction in the Newborn. J Pediatr. 2018;196:71–6.

22. Francis JH, Permezel M, Davey MA. Perinatal mortality by birthweight centile. Aust New Zeal J Obstet Gynaecol. 2014;54(4):354–9.

23. Pilliod RA, Cheng YW, Snowden JM, Doss AE, Caughey AB. The risk of intrauterine fetal death in the small-for-gestational-age fetus. Am J Obstet Gynecol. 2012;207(4):318.e1-318.e6.

24. Salomon LJ, Alfirevic Z, Da Silva Costa F, Deter RL, Figueras F, Ghi T, et al. ISUOG Practice Guidelines: ultrasound assessment of fetal biometry and growth. Ultrasound Obstet Gynecol. 2019;53(6):715–23.

25. McCowan LM, Figueras F, Anderson NH. Evidence-based national guidelines for the management of suspected fetal growth restriction: comparison, consensus, and controversy. Vol. 218, American Journal of Obstetrics and Gynecology. 2018. p. S855–68.

26. Papageorghiou AT, Ohuma EO, Altman DG, Todros T, Cheikh Ismail L, Lambert A, et al. International standards for fetal growth based on serial ultrasound measurements: the Fetal Growth Longitudinal Study of the INTERGROWTH-21st Project. Lancet. 2014;384(9946):869–79.

27. Villar J, Cheikh Ismail L, Victora CG, Ohuma EO, Bertino E, Altman DG, et al. International standards for newborn weight, length, and head circumference by gestational age and sex: the Newborn Cross-Sectional Study of the INTERGROWTH-21st Project. Lancet. 2014;384(9946):857–68.

28. Villar J, Altman DG, Purwar M, Noble JA, Knight HE, Ruyan P, et al. The objectives, design and implementation of the INTERGROWTH-21 st Project. BJOG An Int J Obstet Gynaecol. 2013;120(SUPPL. 2):9–26.

29. Papageorghiou AT, Kennedy SH, Salomon LJ, Altman DG, Ohuma EO, Stones W, et al. The INTERGROWTH-21(st) fetal growth standards: toward the global integration of pregnancy and pediatric care. Am J Obstet Gynecol. 2018 Feb;218(2S):S630–40.

30. Kiserud T, Piaggio G, Carroli G, Widmer M, Carvalho J, Neerup Jensen L, et al. The World Health Organization Fetal Growth Charts: A Multinational Longitudinal Study of Ultrasound Biometric Measurements and Estimated Fetal Weight. PLoS Med. 2017;14(1).

31. Anderson NH, Sadler LC, McKinlay CJD, McCowan LME. INTERGROWTH-21st vs customized birthweight standards for identification of perinatal mortality and morbidity. Am J Obstet Gynecol. 2016;214(4):509.e1-509.e7.

32. Pritchard N, Lindquist A, Siqueira I dos A, Walker SP, Permezel M. INTERGROWTH-21st compared with GROW customized centiles in the detection of adverse perinatal outcomes at term. J Matern Neonatal Med. 2020;33(6):961–6.

33. Kabiri D, Romero R, Gudicha DW, Hernandez-Andrade E, Pacora P, Benshalom-Tirosh N, et al. Prediction of adverse perinatal outcome by fetal biometry: comparison of customized and population-based standards. Ultrasound Obstet Gynecol. 2020;55(2):177–88.

34. Poon LCY, Tan MY, Yerlikaya G, Syngelaki A, Nicolaides KH. Birth weight in live births and stillbirths. Ultrasound Obstet Gynecol. 2016;48(5):602–6.

35. Choi SKY, Gordon A, Hilder L, Henry A, Hyett JA, Brew BK, et al. Performance of six birthweight and estimated fetal weight standards for predicting adverse perinatal outcomes: a 10-year nationwide population-based study. Ultrasound Obstet Gynecol. 2021;58(2):264–77.

36. Cheng YKY, Leung TY, Lao TTH, Chan YM, Sahota DS. Impact of replacing Chinese ethnicity-specific fetal biometry charts with the INTERGROWTH-21st standard. BJOG An Int J Obstet Gynaecol. 2016;123:48–55.

37. Francis A, Hugh O, Gardosi J. Customized vs INTERGROWTH-21 st standards for the assessment of birthweight and stillbirth risk at term. Am J Obstet Gynecol. 2018;218(2):S692–9.

38. Barcena A, Byanyima W. Latin American is the world's most unequal region. Here's how to fix it. Economic Commission for Latin America and the Caribbean (CEPAL). January . 2016.

39. Regional Task Force for the reduction of maternal mortality. Overview of the Situation of Maternal Morbidity and Mortality: Latin America and the Caribbean. 2017.

40. Observatory GE, Caribbean for LA and the. Maternal mortality. 2017.

41. WHO, UNICEF, UNFPA WBG and the UNPD. Trends in Maternal Mortality : 2000 To 2017. WHO, UNICEF, UNFPA, World Bank Group and the United Nations Population Division. 2019.

42. Pingray V, Althabe F, Vazquez P, Correa M, Pajuelo M, Belizan JM. Stillbirth rates in 20 countries of Latin America: an ecological study. BJOG An Int J Obstet Gynaecol. 2018;

43. Blencowe H, Cousens S, Jassir FB, Say L, Chou D, Mathers C, et al. National, regional, and worldwide estimates of stillbirth rates in 2015, with trends from 2000: A systematic analysis. Lancet Glob Heal. 2016;

44. Conde-Agudelo A, Belizan JM, Diaz-Rossello JL. Epidemiology of fetal death in Latin America. Acta Obstet Gynecol Scand. 2000;

45. Belizan JM, McClure EM, Goudar SS, Pasha O, Esamai F, Patel A, et al. Neonatal death in low- to middle-income countries: A global network study. Am J Perinatol. 2012;

46. Rohrer F. Eine neue Formel zur Bestimmung der Korperfulle. Ges Anthr. 1908;39(5).

47. Landmann E, Reiss I, Misselwitz B, Gortner L. Ponderal index for discrimination between symmetric and asymmetric growth restriction: Percentiles for neonates from 30 weeks to 43 weeks of gestation. J Matern Neonatal Med. 2006;19(3):157–60.

48. Harel S, Tomer A, Barak Y, Binderman I, Yavin E. The cephalization index: A screening device for brain maturity and vulnerability in normal and intrauterine growth retarded newborns. Brain Dev. 1985;7(6):580–4.

49. Hua X, Shen M, Reddy UM, Buck Louis G, Souza JP, Gulmezoglu AM, et al. Comparison of the INTERGROWTH-21st, National Institute of Child Health and Human Development, and WHO fetal growth standards. Int J Gynecol Obstet. 2018;143(2):156–63.

50. Liu S, Metcalfe A, Leon JA, Sauve R, Kramer MS, Joseph KS. Evaluation of the INTERGROWTH-21st project newborn standard for use in Canada. PLoS One. 2017;12(3).

51. Grantz KL, Hediger ML, Liu D, Buck Louis GM. Fetal growth standards: the NICHD fetal growth study approach in context with INTERGROWTH-21st and the World Health Organization Multicentre Growth Reference Study. Am J Obstet Gynecol. 2018 Feb;218(2S):S641-S655.e28.

52. Vieira MC, Relph S, Persson M, Seed PT, Pasupathy D. Determination of birth-weight centile thresholds associated with adverse perinatal outcomes using population, customised, and Intergrowth charts: A Swedish populationbased cohort study. PLoS Med. 2019;16(9).

53. Hadlock FP, Harrist RB, Sharman RS, Deter RL, Park SK. Estimation of fetal weight with the use of head, body, and femur measurements-A prospective study. Am J Obstet Gynecol. 1985;151(3):333–7.

54. Stirnemann J, Villar J, Salomon LJ, Ohuma E, Ruyan P, Altman DG, et al. International estimated fetal weight standards of the INTERGROWTH-21st Project. Ultrasound Obstet Gynecol. 2017;49(4):478–86.

55. Grantz KL, Kim S, Grobman WA, Newman R, Owen J, Skupski D, et al. Fetal growth velocity: the NICHD fetal growth studies. Am J Obstet Gynecol. 2018;219(3):285.e1-285.e36.

56. Nieto A, Matorras R, Villar J, Serra M. Neonatal morbidity associated with disproportionate intrauterine growth retardation at term. J Obstet Gynaecol J Inst Obstet Gynaecol. 1998 Nov;18(6):540–3.

57. Streja E, Miller JE, Wu C, Bech BH, Pedersen LH, Schendel DE, et al. Disproportionate fetal growth and the risk for congenital cerebral palsy in singleton births. PLoS One. 2015;10(5):e0126743.

58. Dahlseng MO, Andersen GL, Irgens LM, Skranes J, Vik T. Risk of cerebral palsy in term-born singletons according to growth status at birth. Dev Med Child Neurol. 2014;56(1):53–8.

59. Cheung YB, Yip PSF, Karlberg JPE. Size at birth and neonatal and postneonatal mortality. Acta Paediatr. 2002;91(4):447–52.

60. Barnett WS. Long-Term Effects of Early Childhood Programs on Cognitive and School Outcomes. Futur Child. 1995;5(3):25.

61. Yoshikawa H. Long-term effects of early childhood programs on social outcomes and delinquency. Vol. 5, The Future of children / Center for the Future of Children, the David and Lucile Packard Foundation. 1995. p. 51–75.

62. Anderson NH, Sadler LC, McKinlay CJD, Mccowan LMEE. INTERGROWTH 21st versus customized birthweight standards for identification of perinatal mortality and morbidity. Am J Obstet Gynecol. 2015;214(4):509.e1-509.e7.

63. Gardosi J, Chang A, Kalyan B, Sahota D, Symonds EM. Customised antenatal growth charts. Lancet. 1992;339(8788):283–7.

64. McCowan LME, Harding JE, Stewarf AW. Customised birthweight centiles predict SGA pregnancies with perinatal morbidity. BJOG An Int J Obstet Gynaecol. 2005;112(8):1026–33.

65. Gardosi J, Francis A. Adverse pregnancy outcome and association with small for gestational age birthweight by customized and population-based percentiles. Am J Obstet Gynecol. 2009;201(1):28.e1-28.e8.

66. Carberry AE, Gordon A, Bond DM, Hyett J, Raynes-Greenow CH, Jeffery HE. Customised versus population-based growth charts as a screening tool for detecting small for gestational age infants in low-risk pregnant women. Cochrane Database Syst Rev. 2014;2014(5).

67. Chiossi G, Pedroza C, Costantine MM, Truong VTT, Gargano G, Saade GR. Customized vs population-based growth charts to identify neonates at risk of adverse outcome: systematic review and Bayesian meta-analysis of observational studies. Ultrasound Obstet Gynecol. 2017;50(2):156–66.

68. Nohuz E, Riviere O, Coste K, Vendittelli F. Prenatal identification of small-for-gestational age and risk of neonatal morbidity and stillbirth. Ultrasound Obstet Gynecol. 2020 Apr;55(5):621–8.

69. Sovio U, White IR, Dacey A, Pasupathy D, Smith GCS. Screening for fetal growth restriction with universal third trimester ultrasonography in nulliparous women in the Pregnancy Outcome Prediction (POP) study: A prospective cohort study. Lancet. 2015;386(10008):2089–97.

70. Ganzevoort W, Thilaganathan B, Baschat A, Gordijn S, Gardosi J. Fetal growth and risk assessment : is there an impasse? Am J Obstet Gynecol. 2019;220(1):80.

Table 1. Clinical characteristics and perinatal outcomes for pregnancies assigned as SGA and FGR by each standard and those classified as above the 10^{th} percentile for both curves.

	Above or equal the 10^{th} percentile by both curves n=58,523	**Below the 10^{th} percentile (SGA)**	**Below the 10^{th} percentile (SGA)**	**Above or equal the 3^{rd} percentile by both curves** n=63,718	**Below the 3^{rd} percentile (FGR)**	**Below the 3^{rd} percentile (FGR)**
		By WHO Alone n=4721 **(IQR or %)**	**By both standards** n=4705 **(IQR or %)**	**(IQR or %)**	**By WHO Alone** n=2799 **(IQR or %)**	**By both standards** n=1439 **(IQR or %)**
Gestational age (weeks)	39 (SD ± 1.66)	39.4 (SD ± 1.7)*	38.5 (SD ±2.7)	39 (SD ±1.7)	38.9 (SD ±2)	37.5 (SD ± 3.5)

Preterm delivery (<37 weeks)	4282 (7.3)	289 (6.1)*	769 (16.3)	4618 (7.25)	317 (11.3)	412 (28.6)
Route of delivery						
Cesarean	27031 (46.2)	1779 (37.7)	2247 (47.8)	28978 (45.5)	1228 (43.9)	854 (59.4)
Instrumented	751 (1.3)	41 (0.9)	25 (0.5)	797 (1.3)	15 (0.5)	5 (0.4)
Vaginal	30730 (52.5)	2900 (61.4)	2432 (51.7)	33931 (53.3)	1556 (55.6)	579 (40.3)
Birth weight (grams)	3340 (3100 - 3610)	2866* (2690 - 3004)	2514 (2230 - 2710)	3300 (3020 - 3595)	2620 (2420 - 2790)	2180 (1783 - 2438)
WHO centile	52.4 (29.9 - 77.3)	8.16* (6.75 - 9.06)	0 (0 - 0)	48.3 (24.5 - 75.4)	0 (0 - 0)	0 (0 - 0)
IG-21st centile	63 (42.1 - 83)	14.6* (12.2 - 17.3)	5.1 (2.3 - 7.5)	59.1 (36.5 - 81.7)	6.2 (4.7 - 9.5)	1.2 (0.4 - 2.2)
Apgar score <7 at 5 minutes	270 (0.46)	19 (0.4)	71 (1.5)	302 (0.5)	23 (0.8)	38 (2.6)
Ponderal Index	26.7 (24.9 - 28.6)	24.5* (23 - 26)	23.5 (21.7 - 25.4)	26.5 (24.7 - 28.5)	23.7 (22.2 - 25.3)	22.5 (20.6 - 24.7)
Ponderal index <5th centile	584 (1)	224 (4.7)*	561 (11.9)	841 (1.3)	245 (8.7)	283 (19.7)
Cephalization index	1 (0.9 - 1.1)	1.2 (1.1 - 1.2)*	1.3 (1.2 - 1.5)	1.04 (0.9 - 1.1)	1.26 (1.2 - 1.4)	1.48 (1.34 - 1.8)

Data are median (interquartile range) or N (%). Abbreviations: FGR: fetal growth restriction; IG-21st: INTERGROWTH-21st; SGA: small-for-gestational age neonates; WHO: World Health Organization. In this table, there are not included IG-21st only cases.

Table 2. Diagnostic effectiveness of each standard for specific obstetric and perinatal outcomes.

Outcomes	**Predictors**	**Positive LR (95% CI)**	**Negative LR (95% CI)**	**Sensitivity (95% CI)**	**Specificity (95% CI)**	**PPV (95% CI)**	**NPV (95% CI)**	**DOR (95% C**
APGAR score <7 at 5 minutes	WHO <10th percentile	1.77 (1.48 – 2.12)	0.88 (0.83 – 0.93)	24 (20 – 29)	86 (86 – 86)	1 (1 – 1)	100 (99 – 100)	2.02 (1.59 – 2.57)
	WHO <3rd percentile	2.68 (2.13 – 3.38)	0.89 (0.85 – 0.93)	18 (14 – 22)	93 (93 – 93)	1 (1 – 2)	100 (99 – 100)	3.01 (2.29 – 3.98)
	IG-21st <10th percentile	3.12 (2.56 – 3.81)	0.84 (0.80 – 0.89)	21 (17 – 26)	93 (93 – 93)	2 (1 – 2)	100 (99 -100)	3.70 (2.88 – 4.76)
	IG-21st <3rd percentile	5.61 (4.22 – 7.46)	0.90 (0.87 – 0.84)	12 (9 – 15)	98 (98 – 98)	3 (2 – 4)	100 (99 – 100)	6.22 (4. – 8.59)
Ponderal Index below 5th centile	WHO <10th percentile	4.42 (4.21 – 4.64)	0.49 (0.46 – 0.52)	57 (55 – 60)	87 (87 – 87)	8 (8 – 9)	99 (99 – 99)	9.01 (8. – 10.1)

Outcomes	Predictors	Positive LR (95% CI)	Negative LR (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	DOR (95% C
	WHO <3rd percentile	6.92 (6.43 – 7.45)	0.65 (0.62 – 0.68)	39 (36 – 41)	94 (94 – 95)	12 (11 – 13)	99 (99 – 99)	10.6 (9. – 11.9)
	IG-21st <10th percentile	6.56 (6.11 – 7.03)	0.63 (0.60 – 0.66)	41 (38 – 44)	94 (94 – 94)	12 (11 – 13)	99 (99 – 99)	10.4 (9. – 11.6)
	IG-21st <3rd percentile	11.79 (10.5 – 13.3)	0.81 (0.79 – 0.83)	21 (19 – 23)	98 (98 – 98)	20 (17 – 22)	98 (98 – 98)	14.6 (12.6 – 16.8)

The proportions of sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) are given in percentages (%). DOR, diagnostic odds ratio, FN, False-negative; FP, false positive; LR, likelihood ratio; TN, true negative; TP, true positive.

Figure legends

Figure 1. Venn diagrams describe newborns' classification according to the centiles of each standard (at or above the 10th percentile vs. SGA and FGR) using both standards simultaneously.

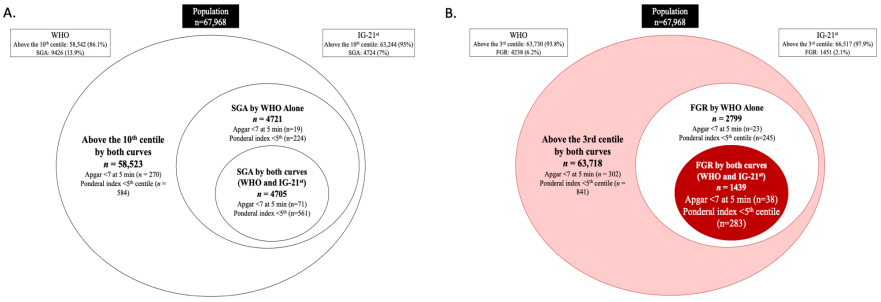
Figure 2. Risk ratios of low APGAR scores and ponderal indexes in SGA and FGR neonates according to the WHO and IG-21st standards. A. Neonates classified as SGA and FGR by both standards exhibited the most significant RR for a low APGAR score. However, neonates classified as SGA only by WHO alone do not have a significantly higher risk of a low APGAR score. B. In terms of anthropometric measures, neonates classified as SGA and FGR by both standards exhibited the most significant RR for a low Ponderal Index. RRs were also increased in neonates identified as SGA by WHO fetal growth standard alone (SGA-WHO only), characterized by anthropometric measures resembling true FGR as reflected by a thrifty phenotype.

Hosted file

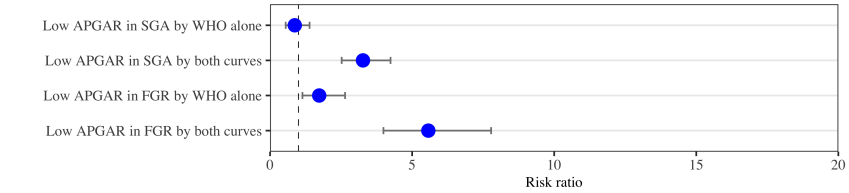
`Table 1.docx` available at https://authorea.com/users/478474/articles/566701-performance-of-the-intergrowth-21st-and-world-health-organization-fetal-growth-charts-for-the-detection-of-small-for-gestational-age-neonates-in-a-population-from-latin-america

Hosted file

`Table 2.docx` available at https://authorea.com/users/478474/articles/566701-performance-of-the-intergrowth-21st-and-world-health-organization-fetal-growth-charts-for-the-detection-of-small-for-gestational-age-neonates-in-a-population-from-latin-america



A. Risk Ratio of Low APGAR score at five minutes.



B. Risk Ratio of low Ponderal Index.