

# Prediction of Perinatal and Neurodevelopmental Outcomes in Newborns with a Birth Weight below the 3rd Percentile: Performance of Two International Curves – Prospective Cohort from a Brazilian City

## *Predição de resultados perinatais e de neurodesenvolvimento em recém-nascidos com peso ao nascer abaixo do percentil 3: Desempenho de duas curvas internacionais – coorte prospectiva de uma cidade brasileira*

Marcos Masaru Okido<sup>1</sup>  Ricardo de Carvalho Cavalli<sup>1</sup>  Viviane Cunha Cardoso<sup>2</sup>   
Alessandra Cristina Marcolin<sup>1</sup> 

<sup>1</sup> Department of Obstetrics and Gynecology, University of São Paulo, Ribeirão Preto, Brazil

<sup>2</sup> Department of Puericulture and Pediatrics, University of São Paulo, Ribeirão Preto, Brazil

Address for correspondence Marcos Masaru Okido, Avenida dos Bandeirantes, 3900, 14049-900, Monte Alegre, Ribeirão Preto, SP, Brazil (e-mail: mokido@hcrp.usp.br).

Rev Bras Ginecol Obstet 2023;45(5):225–234.

### Abstract

**Objectives** To evaluate the performance of Intergrowth-21 st (INT) and Fetal Medicine Foundation (FMF) curves in predicting perinatal and neurodevelopmental outcomes in newborns weighing below the 3rd percentile.

**Methods** Pregnant women with a single fetus aged less than 20 weeks from a general population in non-hospital health units were included. Their children were evaluated at birth and in the second or third years of life. Newborns (NB) had their weight percentiles calculated for both curves. Sensitivity, specificity, positive (PPV) and negative predictive value (NPV), and area under the ROC curve (ROC-AUC) for perinatal outcomes and neurodevelopmental delay were calculated using birth weight < 3rd percentile as the cutoff.

**Results** A total of 967 children were evaluated. Gestational age at birth was 39.3 ( $\pm 3.6$ ) weeks and birth weight was 3,215.0 ( $\pm 588.0$ ) g. INT and FMF classified 19 (2.4%) and 49 (5.7%) newborns below the 3rd percentile, respectively. The prevalence of preterm birth, tracheal intubation >24 hours in the first three months of life,

### Keywords

- ▶ fetal growth retardation
- ▶ birth weight
- ▶ neurodevelopmental disorders

received  
May 6, 2022  
accepted  
February 2, 2023

DOI <https://doi.org/10.1055/s-0043-1770131>.  
ISSN 0100-7203.

© 2023. Federação Brasileira de Ginecologia e Obstetrícia. All rights reserved.

This is an open access article published by Thieme under the terms of the Creative Commons Attribution License, permitting unrestricted use, distribution, and reproduction so long as the original work is properly cited. (<https://creativecommons.org/licenses/by/4.0/>)

Thieme Revinter Publicações Ltda., Rua do Matoso 170, Rio de Janeiro, RJ, CEP 20270-135, Brazil

5th minute Apgar <7, admission to a neonatal care unit (NICU admission), cesarean section rate, and the neurodevelopmental delay was 9.3%, 3.3%, 1.3%, 5.9%, 38.9%, and 7.3% respectively. In general, the 3rd percentile of both curves showed low sensitivity and PPV and high specificity and NPV. The 3rd percentile of FMF showed superior sensitivity for preterm birth, NICU admission, and cesarean section rate. INT was more specific for all outcomes and presented a higher PPV for the neurodevelopmental delay. However, except for a slight difference in the prediction of preterm birth in favor of INT, the ROC curves showed no differences in the prediction of perinatal and neurodevelopmental outcomes.

**Conclusion** Birth weight below the 3rd percentile according to INT or FMF alone was insufficient for a good diagnostic performance of perinatal and neurodevelopmental outcomes. The analyzes performed could not show that one curve is better than the other in our population. INT may have an advantage in resource contingency scenarios as it discriminates fewer NB below the 3rd percentile without increasing adverse outcomes.

## Resumo

**Objetivos** Avaliar o desempenho das curvas de Intergrowth-21 st (INT) e Fetal Medicine Foundation (FMF) na predição de resultados perinatais e de neurodesenvolvimento de recém-nascidos com peso abaixo do percentil 3.

**Métodos** Foram incluídas gestantes de feto único com idade inferior a 20 semanas de uma população geral em unidades de saúde não hospitalares. Seus filhos foram avaliados ao nascimento e no segundo ou terceiro anos de vida. Os recém-nascidos tiveram seus percentis de peso calculados para ambas as curvas. Sensibilidade, especificidade, valor preditivo positivo (VPP) e negativo (VPN) e área sob a curva ROC (ROC-AUC) foram calculados para desfechos perinatais e atraso de neurodesenvolvimento considerando o peso ao nascimento menor que o percentil 3 como ponto de corte.

**Resultados** Um total de 967 crianças foram avaliadas ao nascimento e no segundo ou terceiro anos de vida. A idade gestacional ao nascer foi de 39,3 ( $\pm 3,6$ ) semanas e o peso ao nascimento foi de 3.215,0 ( $\pm 588,0$ ) g. INT e FMF classificaram 19 (2,4%) e 49 (5,7%) recém-nascidos abaixo do percentil 3, respectivamente. A prevalência de parto pré-termo, intubação traqueal > 24 horas nos primeiros três meses de vida, Apgar de 5º minuto < 7, internação em unidade de terapia intensiva neonatal (internação em UTIN), taxa de cesariana e atraso de neurodesenvolvimento foi 9,3%, 3,3%, 1,3%, 5,9%, 38,9% e 7,3% respectivamente. Em geral, o percentil 3 de ambas as curvas apresentou baixa sensibilidade e VPP e alta especificidade e VPN. O percentil 3 de FMF mostrou sensibilidade superior para parto prematuro, internação em UTIN e taxa de cesariana. INT foi mais específico para todos os desfechos e apresentou maior VPP para o atraso do neurodesenvolvimento. Entretanto, exceto por uma pequena diferença na predição de parto pré-termo em favor de INT, as curvas ROC não mostraram diferenças na predição de resultados perinatais e de desenvolvimento neurológico.

## Palavras-chave

- ▶ retardo do crescimento fetal
- ▶ peso ao nascer
- ▶ transtornos do neurodesenvolvimento

**Conclusão** O peso ao nascer abaixo do percentil 3 segundo INT ou FMF isoladamente foi insuficiente para um bom desempenho diagnóstico de desfechos perinatais e de neurodesenvolvimento. As análises realizadas não puderam mostrar que uma curva é melhor que a outra em nossa população. INT pode ter vantagem em cenários de contingência de recursos, pois discrimina menos recém-nascidos abaixo do percentil 3 sem aumentar os desfechos adversos.

## Introduction

Fetal growth restriction is associated with adverse perinatal outcomes, neurodevelopmental delay, and the onset of chronic disease in adults.<sup>1-3</sup> Identification of fetuses and newborns (NB) with growth restriction could help improve these results by intensifying prenatal and postnatal care.<sup>4,5</sup>

Several estimated fetal and birth weight (BW) charts have been published worldwide, showing significant differences.<sup>6-12</sup> While some authors suggest that these differences are due to racial and geographic variations, others attribute them to socioeconomic inequalities, nutritional deficits, or methods used in the studies.<sup>7,9,13,14</sup>

One of the healthcare challenges in Brazil is to define which fetal and neonatal growth charts better discriminate children with growth restriction in the Brazilian population. It is unclear whether using North American or European curves could increase false positive diagnoses, as the cut-off points may be too high. Using references that underestimate diagnoses bears even more risk as it could deprive the most vulnerable pregnant women and their children of the necessary care, leading to an increase in the incidence of adverse outcomes.

Therefore, it should be understood that the random choice of a fetal or neonatal weight curve without an in-depth analysis of morbidity and mortality is not recommended. Choosing a particular reference over another is only justified if the reference can better identify the NB with the highest risk of morbidity and mortality without excess diagnoses.<sup>15</sup>

The objective of this study was to evaluate the performance of two international BW curves, Intergrowth-21 st (INT) and Fetal Medicine Foundation (FMF) in predicting perinatal and neurodevelopmental outcomes in newborns based on birth weight below the 3rd percentile in a Brazilian city. This is the first prospective Brazilian study to include neurodevelopmental outcomes in assessing of birth weight curves.

## Methods

This prospective cohort evaluated children at birth and in the second or third years of life.

Data from a BRISA-RP cohort study were used.<sup>16</sup> The preliminary study assessed etiological factors of preterm birth and the consequences of perinatal factors on child health. The research ethics committee approved this study at the University Hospital where it was performed. Ribeirão Preto is located in south-eastern Brazil and has ~710,000 inhabitants. This is a prosperous country region regarding income, consumption, and longevity. Birth data were collected between April 2010 and December 2011, and the neurodevelopmental assessment data in the second or third years of life.

Recruitment of this cohort started during pregnancy. Pregnant women from a general population with a single fetus aged less than 20 weeks were sequentially recruited from selected primary care units in this city. These units are part of the public health system, are not linked to the

University, and generally serve a low- and middle-income population. The first ultrasound determined the gestational age (GA) used in this study. All pregnant women recruited had already undergone a first-trimester ultrasound with gestational age was calculated using the crown-rump length. Furthermore, all pregnant women underwent a new ultrasound to confirm the GA by certified physicians from the research team before 24 weeks. At birth, all NB with BW greater than or equal to 500 g were potentially eligible. NB older than 42 weeks were excluded to reduce the risk of bias in an incorrect recording of gestational age, as it is unlikely that a pregnancy will exceed this limit spontaneously or on medical advice. NB with severe malformations were also excluded. All cohort participants were encouraged to bring their children between the second and third years of life for a neurodevelopment assessment using the Bayley Scales of Infant Development, III edition (BSID-III).<sup>17</sup> All patients gave their informed written consent to participate in the study.

Sociodemographic and clinical data were collected, including maternal age, body mass index, race, parity, education level, smoking, alcohol use, hypertension, and diabetes. Data from the newborns of the included women were collected in their respective maternity hospitals in the city on birth or the following day by the research team.

The predictor variable was BW below the 3rd percentile for GA. The 3rd percentile was chosen because there is a high perinatal morbidity and mortality risk below this thresholds.<sup>18,19</sup> Furthermore, fetal or birth weight below the 3rd percentile is considered an isolated criterion for fetal and neonatal growth restriction, according to the latest expert consensus.<sup>20,21</sup>

The perinatal outcomes were preterm birth, tracheal intubation for more than 24 hours in the first three months of life (Intubation), 5 minute Apgar <7, admission to a neonatal intensive care unit (NICU admission), and cesarean section rate. The long-term outcome variable was the risk of neurodevelopmental delay between the second and third years of life.

Percentiles of BW were obtained for each NB. INT (specific gender) and FMF calculators were used to predict BW.<sup>22,23</sup> These charts were chosen because both include fetal and neonatal charts and are the only ones whose calculators were found on official open access Web sites.

The INT standards were constructed from a prescriptive population of over 4,500 healthy pregnancies in a study of over 59,000 total pregnancies. The project involved 8 countries from 4 continents and included only highly selected women with optimal nutrition and low risk of pregnancy complications. Fetal anthropometric data were prospectively collected every 5 weeks starting at 14 weeks. The aim of that approach was to create charts that could be used worldwide.<sup>9,24</sup>

Nicolaides et al. (FMF) used a heterogeneous sample of unselected pregnant women, most of them being white women from the United Kingdom. Data were collected from two sources. The first comprised 5163 paired measurements of EFW and BW, and the second of 95,579 pregnancies with EFW obtained by routine fetal ultrasound biometry. In

this study, the authors proposed to consider that all babies of the same gestational age, even intrauterine babies, could be included for BW references. Thus, the construction of the curves considered that in a given population with a defined gestational age, the median fetal weight and the median birth weight are similar, with different degrees of deviations from the median for fetal weight and birth weight, depending on the gestational age.<sup>10</sup>

Maternal, gestational, and childbirth data were obtained by filling out previously prepared questionnaires with interview data and medical records.

Neurodevelopment was assessed by ten psychologists who received identical, simultaneous, and group training. Children were assessed in three domains: cognitive, language (receptive and expressive), and motor (gross and fine), with each domain including specific tests for each age. For each age group (13–24 months; 25–42 months), there is a corresponding starting point. The child's performance on each test item was scored 0 or 1. The points obtained in each domain were summed, and children were classified as competent, emerging, or at risk according to the cut-off points provided by the test. A score that resulted in a risk for any domains was considered positive for neurodevelopmental risk.

There was no interference from researchers in prenatal care, labor and delivery, and postnatal care of children. All NB of pregnant women with conditions potentially associated with fetal growth deviations such as hypertension, diabetes, smoking, and preterm deliveries were included in the analysis.

Statistical analysis was performed using the SAS System for Windows (Statistical Analysis System), version 9.2., SAS Institute Inc, 2002–2008, Cary, NC, USA. Comparisons of descriptive variables were performed using the generalized estimating equations (GEE analysis) (numerical variables). Sensitivity, specificity, and positive and negative predictive values (PPV and NPV) of perinatal outcomes and risk of neurodevelopmental delay were estimated, with differences being determined by the McNemar test and GEE analysis considering a BW below the 3rd percentile as the cut-off point. The discriminatory ability of each curve was assessed using the AUC of the ROC curve. To be significant, were considered results with  $p$ -value  $<0.05$  with a confidence interval of 95%.

## Results

### Participants

A total of 1417 pregnant women were recruited. Seventy-three dropouts were reported, and 17 NB were excluded (6 with major malformations, 3 without weight records, 2 weighing less than 500 g, and 6 with a gestational age of 42 weeks or more). The total number of children with no or incomplete neurodevelopmental tests was 360 (27.1%). The final number of cases for analysis was 967 (► Fig. 1).

### Descriptive Data

The median age of pregnant women was 26.0 ( $\pm 12.0$ ) years. The majority (79.9%) were nuligest or secundigest, and 90.7%

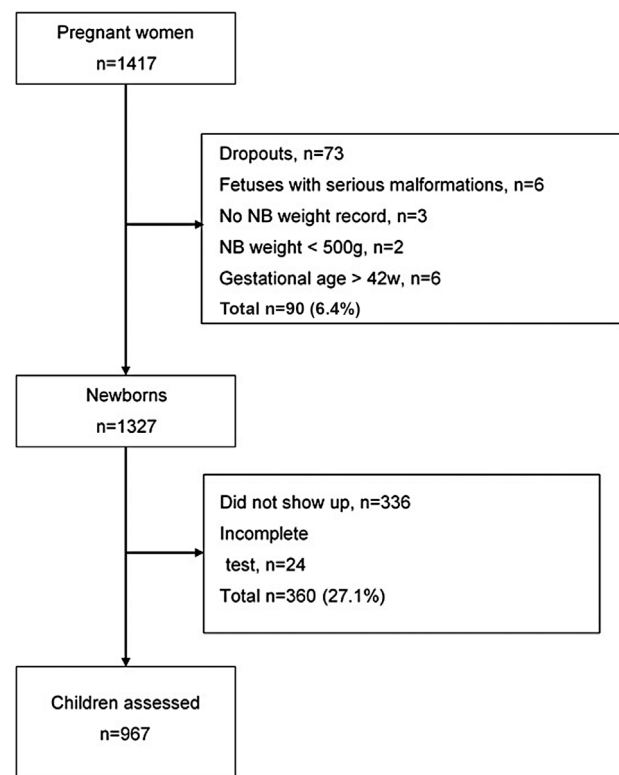


Fig. 1 Flowchart of the participants included in the study.

had full-term delivery. The median GA at birth, and the NB weight was 39.3 ( $\pm 3.6$ ) weeks, and 3,215.0 ( $\pm 588.0$ ) g respectively. The prevalence of hypertension was 15.4%, and smoking, 12.4%. The total prevalence of diabetes in the sample was 6.1%. In the INT and FMF groups, the diabetes rate was 10.5 and 4.1%, respectively. The overall prevalence of prematurity was 9.3%; however, among NB below the 3rd percentile, it was 46.2% for INT and 32.4% for FMF. For neurodevelopment, 741 children were evaluated between 13 and 24 months and 226 between 25 and 36 months (► Table 1). The prevalence of intubation, 5 minute Apgar  $<7$ , NICU admission, cesarean section and neurodevelopment delay was 3.3%, 1.3%, 5.9%, 38.9% and 7.3%, respectively (► Table 2).

## Main Results

INT and FMF classified 19 (1.9%) and 49 (5.1%) NB below the 3rd percentile, respectively. As a rule, high specificity was observed, but low sensitivity and positive predictive value. Some results were statistically superior for FMF, such as sensitivity for preterm delivery, NICU admission, and cesarean, respectively 13.3% (95%CI 7.4–22.5) versus 6.7% (95%CI 2.7–15.5) ( $p = .014$ ), 13.8% (95%CI 6.6–22.9) versus 3.5% (95%CI 0.6–13.0) ( $p = .014$ ) and 6.7% (95%CI 4.4–9.8) versus 2.1% (95%CI 1.0–4.3) ( $p < .001$ ). On the other hand, INT was superior in all outcomes when specificity was evaluated ( $p < .001$ ). In addition, the NPV for neurodevelopment delay was 21.1% (95%CI 7.0–46.1) for INT versus 10.2% (95%CI 3.8–23.0) for FMF ( $p = .005$ ) (► Table 3). However, analysis using

**Table 1** Cohort demographic characteristics

	Total n = 967	INT NB weight $p < 3$ n = 19	FMF NB weight $p < 3$ n = 49	p-value <sup>d</sup>
Maternal age, years (IQR)	26.0 (12.0)	26.0 (12.0)	26.0 (12.0)	0.718
BMI, kg/m <sup>2</sup> (IQR)	26.5 (6.8)	27.8 (8.5)	27.0 (6.6)	0.218
Study US, weeks IQR)	22.9 (1.7)	23.1 (2.3)	22.9 (1.7)	0.919
Ethnicity				
- White (%)	510 (52.4)	11 (57.9)	27 (55.1)	0.756
- Non-white (%)	457 (47.3)	8 (42.1)	22 (44.9)	
Schooling				
- $\geq 12$ years (%)	77 (8.0)	1 (5.3)	4 (8.2)	0.595
- $< 12$ years (%)	890 (92.0)	18 (94.7)	45 (91.8)	
Marital status				
- Married or cohabiting (%)	778 (80.5)	14 (73.7)	37 (75.5)	0.810
- No partner (%)	189 (19.5)	5 (26.3)	12 (24.5)	
Parity				
- 0–1 (%)	773 (79.9)	16 (84.2)	43 (87.8)	0.515
- $\geq 2$ (%)	194 (19.5)	3 (15.8)	6 (12.2)	
NB sex				
- Male (%)	474 (49.0)	11 (57.9)	23 (46.9)	0.228
- Female (%)	493 (51.9)	8 (42.1)	26 (53.1)	
Hypertension <sup>a</sup> (%)	149 (15.4)	4 (21.1)	12 (24.5)	0.667
Diabetes <sup>a</sup> (%)	59 (6.1)	2 (10.5)	2 (4.1)	$< 0.001$
Smoking <sup>b</sup> (%)	120 (12.4)	6 (31.6)	11 (22.5)	0.189
Alcohol <sup>b</sup> (%)	238 (24.6)	6 (31.6)	14 (28.6)	0.706
Delivery, weeks (IQR)	39.3 (3.6)	39.6 (4.6)	39.3 (3.4)	0.220
$\geq 37$ weeks	39.4 (1.7) <sup>e</sup>	40.0 (1.6) <sup>g</sup>	39.7 (1.4) <sup>i</sup>	
$< 37$ weeks	35.7 (2.4) <sup>f</sup>	35.9 (1.0) <sup>h</sup>	34.8 (2.3) <sup>j</sup>	
NB weight, g (IQR)	3215.0 (588.0)	2425.0 (660.0)	2455.0 (465.0)	0.464
$\geq 37$ weeks	3250.0 (570) <sup>e</sup>	2500.0 (135) <sup>g</sup>	2500.0 (295) <sup>i</sup>	
$< 37$ weeks	2645.0 (877.5) <sup>f</sup>	1842.5 (178.8) <sup>h</sup>	1735.0 (381.3) <sup>j</sup>	
Breastfeeding $\geq 1$ month (%)	863 (89.3)	18 (94.7)	41 (83.7)	0.187
Day care center (%)	469 (48.5)	4 (21.1)	12 (24.5)	0.667
Neurodevelop. Assessm <sup>c</sup> , years (IQR)	1.9 (0.3)	1.8 (0.3)	1.8 (0.3)	0.428

Abbreviations: BMI, body mass index; FMF, Fetal Medicine Foundation; INT, Intergrowth-21st; IQR, interquartile interval; NB, newborn; p, percentile; US, ultrasound.

<sup>a</sup>Types of hypertension included chronic, gestational, or preeclampsia. Types of diabetes included: type 1, 2 our gestational; <sup>b</sup>Any amount of consumption; <sup>c</sup>Child's age at evaluation of neurodevelopment by Bayley III test; <sup>d</sup>Comparisons of categorical and numerical variables considering only newborns with  $< p3$  in each curve, through GEE (Generalized Estimating Equations) analysis; <sup>e</sup>n = 877; <sup>f</sup>n = 90; <sup>g</sup>n = 13; <sup>h</sup>n = 6; <sup>i</sup>n = 37; <sup>j</sup>n = 12.

ROC curves did not show adequate performance in predicting perinatal and neurodevelopmental outcomes (**► Fig. 2**).

## Discussion

For both curves evaluated, the cutoff point for birth weight below the **3rd percentile** alone did not prove to be a good predictor of adverse perinatal outcomes and risk of neurodevelopmental delay. FMF classified more than double NB with BW below the 3rd percentile. This increased the sensitivity but did not improve the other parameters.

The sample size was one of the limitations of the study because some outcomes such as intubation ( $n = 32$ ), 5 minute Apgar score  $< 7$  ( $n = 13$ ), and NICU admission ( $n = 58$ ) had a low incidence, resulting in lower reliability of the results. Until birth, few cases were excluded or dropped out of the study ( $n = 90$ ; 6.4%); however, many children did not attend the neurodevelopment test ( $n = 360$ ; 27.1%). We were unable to investigate the reasons for the withdrawal of pregnant women and the absence of children. Data on the quality of management of fetal growth restriction pregnancies, perinatal deaths, the onset of diseases, and quality of infant

**Table 2** Performance of birth weight below the 3rd percentile for perinatal, and infant neurodevelopment outcomes according to Intergrowth-21 st and fetal medicine foundation curves

	n (%)	Sensitivity% (95%CI)	p-value <sup>a</sup>	Specificity% (95%CI)	p-value <sup>a</sup>	PPV (95%CI)	p-value <sup>b</sup>	NPV (95% CI)	p-value <sup>b</sup>
Preterm birth n = 90 (9.3%)	INT 6 (31.6%) FMF 12 (24.5%)	6.7 (2.7–14.5) 13.3 (7.4–22.5)	0.014	98.5 (97.4–99.2) 95.8 (94.2–97.0)	<0.001	31.6 (13.6–56.5) 24.5 (13.8–39.2)	0.334	91.1 (89.1–92.8)	0.136
Intubation n = 32 (3.3%)	INT 2 (10.5%) FMF 5 (10.2%)	6.3 (1.1–22.2) 15.6 (5.9–33.6)	0.083	98.2 (97.0–98.9) 95.3 (93.7–96.5)	<0.001	10.5 (1.8–34.5) 10.2 (3.8–23.0)	0.952	96.8 (95.5–97.8)	0.217
5 minute Apgar <7 n = 13 (1.3%)	INT 0 FMF 1 (2.0%)	0.0 (0.0–28.4) 7.7 (0.4–37.9)	0.317	98.0 (96.8–98.8) 95.1 (93.4–96.3)	<0.001	0.0 (0.0–20.9) 2.1 (0.1–12.5)	1.00	98.6 (97.6–99.2)	0.537
NICU admission n = 58 (5.9%)	INT 2 (10.5%) FMF 8 (16.3%)	3.5 (0.6–13.0) 13.8 (6.6–25.9)	0.014	98.1 (97.0–98.9) 95.5 (93.9–96.7)	<0.001	10.5 (1.8–34.5) 16.3 (7.8–30.2)	0.429	94.1 (92.4–95.5)	0.065
Cesarean n = 376 (38.9%)	INT 8 (42.1%) FMF 17 (34.7%)	2.1 (1.0–4.3) 6.7 (4.4–9.8)	<0.001	98.1 (96.6–99.0) 95.9 (93.9–97.3)	<0.001	42.1 (21.1–66.0) 51.0 (36.5–65.4)	0.327	61.2 (58.0–64.3)	0.060
Neurodevelopment n = 71 (7.3%)	INT 4 (21.1%) FMF 5 (10.2%)	5.6 (1.8–14.5) 7.0 (2.6–16.4)	0.317	98.3 (97.2–99.0) 95.1 (93.4–96.4)	<0.001	21.1 (7.0–46.1) 10.2 (3.8–23.0)	0.005	92.9 (91.1–94.4)	0.257

Abbreviations: FMF, Fetal Medicine Foundation (n = 49); INT, Intergrowth-21<sup>st</sup> (n = 19); Intubation, tracheal intubation for more than 24 hours in the first three months of life; Neurodevelopment, neurodevelopment delay (assessment in the second and third years of life by the Bayley/III test); NICU, neonatal intensive care unit; VPN, negative predictive value; VPP, positive predictive value.  
<sup>a</sup>Mc Nemar test, <sup>b</sup>GEE (generalized estimating equations).

stimulation are scarce. The high number of absences may cause some bias in the results, as perinatal losses usually occur among those with the lowest weight percentile, which would probably influence the INT group more because it concentrates the smallest NBs. However, it is essential to remember that absences interfere in analyzing both patterns since the study group is the same.

It is important to note that this study only included newborns weighing below the 3rd percentile. We did not include the restricted fetuses because we did not have estimated fetal weight and Doppler data. Therefore, probably some NBs above the 3rd percentile but at risk were not included.

The prospective design add advantages to the study, as it was possible to obtain a sample of pregnant women of non-hospital origin and were still in the first half of pregnancy. Including pregnant women with complications such as hypertension and smoking was purposeful, as low birth weight can result from multiple maternal and gestational conditions. Thus, we intended to obtain a sample that well represented the general population with its proportion of healthy women and others with prevalent diseases during pregnancy.

The main strength of this study was the inclusion of neurodevelopmental outcomes, as the predictive capacity of perinatal outcomes is low.<sup>25,26</sup> Long-term results are essential to assess the role of gestational complications such as low birth weight in the onset of permanent neurological damage, which is difficult to assess in the neonatal period.<sup>27</sup> This is the first prospective Brazilian study to include neurodevelopmental outcomes in assessing birth weight curves.

We believe that the assessment of the curves made in this study was timely, as they represent two distinct types of population samples. One of them (INT) is based on intercontinental and multi-ethnic populations, including the Brazilian population, while the other (FMF) is predominantly based on the population of European women.<sup>9,10</sup>

Both curves were poor predictors of perinatal and neurodevelopmental outcomes, probably because these outcomes are influenced by multiple factors beyond the birth weight, such as prematurity, birth conditions, neonatal care, breastfeeding, and infant stimulation. This has already been demonstrated in other studies with fetal and neonatal patterns.<sup>28,29</sup> The most remarkable difference between the curves in this study was the highest cut-off point on the FMF graph. INT discriminated 19 NB below the 3rd percentile while FMF 49. In practice, we had more than twice as many newborns classified as having restricted growth according to the new consensus for one standard (FMF) compared with the other (INT). To exemplify, if we had a hypothetical 39-week NB weighing 2,580 g, this NB would be in the fifth (male chart) or seventh (female chart) percentiles of the INT, but by FMF references, it would be in the second percentile. This difference may seem insignificant, but it can place many additional newborns in the growth-restricted group in a population context. Changes in cutoff points could make any growth curve potentially suitable for any population. We

**Table 3** Comparison of area under receptor operating characteristics curve for prediction of perinatal and infant neurodevelopmental outcomes between intergrowth-21 st and fetal medicine foundation using birth weight percentiles

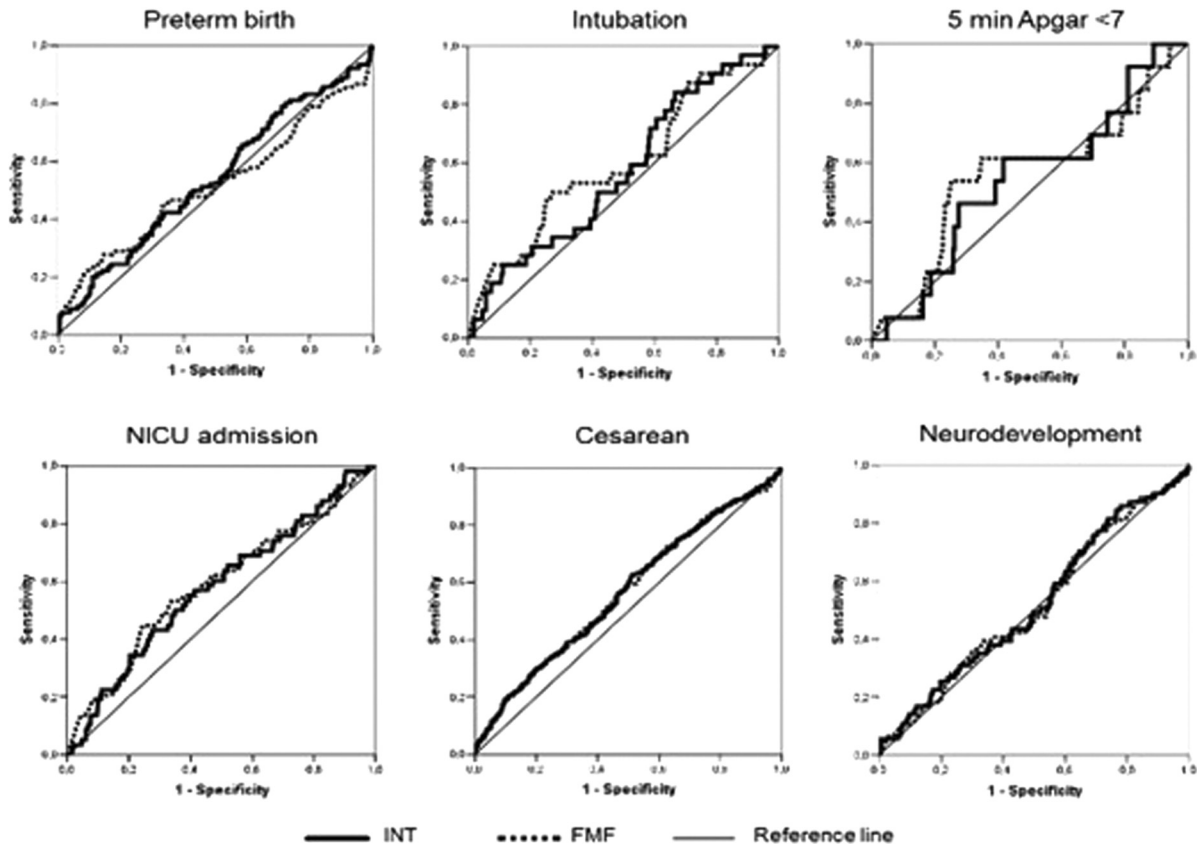
	INT	FMF	p-value
	ROC curve AUC (95%CI)	ROC curve AUC (95%CI)	
Preterm birth	0.54 (0.47–0.60)	0.52 (0.45–0.59)	0.032
Intubation	0.57 (0.48–0.67)	0.60 (0.49–0.70)	0.090
5 minute Apgar <7	0.54 (0.39–0.70)	0.56 (0.39–0.73)	0.437
NICU admission	0.58 (0.50–0.66)	0.59 (0.51–0.67)	0.319
Cesarean	0.57 (0.53–0.60)	0.56 (0.53–0.60)	0.887
Neurodevelopment	0.52 (0.45–0.59)	0.52 (0.45–0.59)	0.453

Abbreviations: FMF, Fetal Medicine Foundation; INT, Intergrowth-21st; Intubation, tracheal intubation for more than 24 hours in the first three months of life; Neurodevelopment, neurodevelopment delay (assessment in the second and third years of life by the BayleyIII test); NICU, neonatal intensive care unit; ROC curve AUC, area under receptor operating characteristics curve.

could have tested other thresholds to define the one that best discriminates the children at the most significant risk in each pattern.<sup>30</sup> However, we did not do this because it would be of little practical use as most hospitals have their chosen fetal and neonatal growth reference charts and well-known thresholds such as the 3rd or 10th percentile. Some expected differences were found with the applied statistical tests due to differences in the cutoff points of the evaluated curves. Higher sensitivity for FMF and higher specificity for INT. However, the ROC curves, except for a small difference in the

prediction of preterm birth in favor of INT, did not show consistent differences.

The birth weight of our sample was similar to that of other national studies. Barros et al. conducted a study in Pelotas, a city in southern Brazil, with a cohort of 4,558 newborns. Most pregnant women came from urban areas, and 61.7% were white. The BW was 3,149.6 g, slightly lower than in this study (3,215.0 g). However, the sample by Barros et al. included a higher proportion of pregnancies with complications such as smoking (27.5%), premature birth (15.3%), and

**Fig. 2** Receiver operating characteristic curve of Intergrowth-21 st and Fetal Medicine Foundation to predict perinatal and infant neurodevelopmental outcomes using birth weight percentiles.

hypertension (23.7%).<sup>31</sup> Kiserud et al., similarly to the INT project, used a multinational sample that included a Brazilian city (Campinas;  $n = 150$ ) to create EFW and BW standards for the World Health Organization.<sup>7</sup> Interestingly, the GA at the birth of the Brazilian sample was 39 weeks, similar to this study (39.3 weeks), and the BW was 3290.0 g, also similar to this study (3243.0 g) when growth-restricted NB are excluded.

The prevalence of premature births was 9.3%, lower than that found in the study by Passini et al. (12.3%).<sup>32</sup> The study cited above included 20 referral hospitals and more than 33,000 deliveries in Brazil. This difference can be explained by the characteristics of the samples obtained in each study. Our study recruited pregnant women in non-hospital units, while the study by Passini et al. included pregnant women from referral hospitals. Although some of the pregnant women in our study had complications during pregnancy, pregnant women coming from referral hospitals are more likely to have risk factors for preterm delivery, whether spontaneous, due to rupture of membranes, or therapeutic. In addition, the study by Passini et al. showed that Brazilian regions presented slightly different prevalences, with a lower prevalence for the Southeast region, where our study was performed.

Among the maternal diseases that can negatively affect fetal growth and bias the analyses, arterial hypertension had a slightly higher prevalence in this study (15.2%) compared with the general population.<sup>20,33</sup> We postulate that the recruitment may have been biased because the hypertensive pregnant woman seeks the health unit more frequently. The prevalence of smoking was similar to other studies.<sup>34</sup> The prevalence of diabetes was low in this study, despite the indistinct inclusion of type 1, 2, and gestational diabetes, probably due to underreporting and the use of old diagnostic criteria for gestational diabetes used at the time. The International Association of Diabetes in Pregnancy Study Group criteria, which are more sensitive, were adopted from 2015 onwards in our country. The statistical difference observed between the INT and FMF groups is probably due to the small number of diabetes diagnoses in these groups (the same 2 cases for both groups).

The INT and FMF curves were tested in different populations. Kajdy et al., in Poland, obtained a BW reference curve of 39,092 single births and compared their percentiles with 6 published charts, including the INT. In that study, the 50th percentile at 40 weeks was 3645.8 g and 3486.7 g for male and female NB, respectively. The authors obtained 3.2% of NB below the 3rd percentile by the local chart and only 0.6% by INT.<sup>35</sup> Anderson et al. obtained data from 53,484 NB in Auckland, New Zealand, and compared small-for-gestational-age (SGA) new-born outcomes between INT versus a customised standard using maternal characteristics of height, weight, parity, and ethnicity. The GA was 39.4 weeks, and the weight was 3433 g, with a higher weight associated with Pacific ethnicity (3585 g) and a lower weight associated with Indian ethnicity (3130 g). The incidence of SGA was 4.5% when INT was used and 11.6% when the customised standard was used. The authors concluded that customised curves

identified more NB SGA at risk for perinatal morbidity and mortality than INT standards.<sup>36</sup> Francis et al. analyzed data from 1.25 million full-term pregnancies from 10 countries. INT was compared with a customised standard to determine stillbirth rates in SGA and large for gestational age (LGA) groups. Significant differences in SGA rates would be found between countries using INT. The most significant differences were observed between Sweden (10.7% for the customised standard and 3.1% for INT) and India (11.3% for the customised standard and 16.8% for INT). In Sweden ( $n = 257,924$ ), the GA at birth was 40.8 weeks, and the BW was 3623.0 g, while in India ( $n = 6436$ ) the GA at birth was 39.0 weeks, and the BW was 3055.5 g.<sup>37</sup>

Regarding FMF patterns, Duncan et al. compared the detection of SGA in preterm prelabor rupture of membranes by Hadlock versus the FMF charts. A sample of 106 patients from a university hospital in Tennessee with 84.9% African American was assessed. The cutoff point adopted was the 10th percentile. In this study, the FMF and Hadlock patterns discriminated respectively 48 (45%) and 22 (21%) of NB below the 10th percentile. Both patterns had similar accuracy in predicting SGA and were equally poor in predicting severe adverse neonatal outcomes. The FMF chart resulted in a 2-fold increase in positive cases, potentially increasing surveillance.<sup>38</sup>

Based on the studies cited above, it is evident that different populations can provide different proportions of NB below the 3rd or 10th percentile when using the same curve. Given its miscegenation, it is plausible that the INT provides standards more suited to the Brazilian population as the multiethnic sample is one of the main features of the INT standards. Compared with FMF, INT is less sensitive but appears safe as it does not increase adverse outcomes. This can be advantageous in resource contingency scenarios.

## Conclusion

Although BW below the 3rd percentile is associated with adverse perinatal and neurodevelopmental outcomes, it was insufficient for a good diagnostic performance when evaluated alone. The analyses performed in this study could not show that one curve is unequivocally better than the other in our population. The apparent excess of newborns classified below the 3rd percentile by FMF may mean that it is not advisable to use references imported from countries of different racial composition, such as European countries or the United States. It is plausible that INT is more suitable for the Brazilian population due to its mixed racial composition. Further prospective studies are needed in Brazil to compare global standards, such as INT, with locally customized curves.

## Contributions

All authors participated in the concept and design of the present study; analysis and interpretation of data; draft or revision of the manuscript, and they have approved the manuscript as submitted. All authors are responsible for the reported research.



### Conflicts to Interest

The authors inform that there are no conflicts of interest, whether political, economic, resources to carry out the research or intellectual property.

### References

- McIntire DD, Bloom SL, Casey BM, Leveno KJ. Birth weight in relation to morbidity and mortality among newborn infants. *N Engl J Med.* 1999;340(16):1234–1238. Doi: 10.1056/NEJM199904223401603
- Sacchi C, Marino C, Nosarti C, Vieno A, Visentin S, Simonelli A. Association of intrauterine growth restriction and small for gestational age status with childhood cognitive outcomes: a systematic review and meta-analysis. *JAMA Pediatr.* 2020;174(08):772–781. Doi: 10.1001/jamapediatrics.2020.1097
- de Mendonça ELSS, de Lima Macêna M, Bueno NB, de Oliveira ACM, Mello CS. Premature birth, low birth weight, small for gestational age and chronic non-communicable diseases in adult life: A systematic review with meta-analysis. *Early Hum Dev.* 2020;149:105154. Doi: 10.1016/j.earlhumdev.2020.105154
- Lees C, Marlow N, Arabin B, et al; TRUFFLE Group. 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(04):400–408. Doi: 10.1002/uog.13190
- Boivin MJ, Kakooza AM, Warf BC, Davidson LL, Grigorenko EL. Reducing neurodevelopmental disorders and disability through research and interventions. *Nature.* 2015;527(7578):S155–S160. Doi: 10.1038/nature16029
- Hadlock FP, Harrist RB, Martinez-Poyer J. In utero analysis of fetal growth: a sonographic weight standard. *Radiology.* 1991;181(01):129–133. Doi: 10.1148/radiology.181.1.1887021
- Kiserud T, Piaggio G, Carroli G, 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(01):e1002220. Doi: 10.1371/journal.pmed.1002220
- Mikolajczyk RT, Zhang J, Betran AP, et al. A global reference for fetal-weight and birthweight percentiles. *Lancet.* 2011;377(9780):1855–1861. Doi: 10.1016/S0140-6736(11)60364-4
- Papageorghiou AT, Ohuma EO, Altman DG, et al; International Fetal and Newborn Growth Consortium for the 21st Century (INTERGROWTH-21st) International standards for fetal growth based on serial ultrasound measurements: the Fetal Growth Longitudinal Study of the INTERGROWTH-21st Project. *Lancet.* 2014;384(9946):869–879. Doi: 10.1016/S0140-6736(14)61490-2
- Nicolaides KH, Wright D, Syngelaki A, Wright A, Akolekar R. Fetal Medicine Foundation fetal and neonatal population weight charts. *Ultrasound Obstet Gynecol.* 2018;52(01):44–51
- Fenton TR, Kim JH. A systematic review and meta-analysis to revise the Fenton growth chart for preterm infants. *BMC Pediatr.* 2013;13(01):59. Doi: 10.1186/1471-2431-13-59
- Alexander GR, Himes JH, Kaufman RB, Mor J, Kogan M. A United States national reference for fetal growth. *Obstet Gynecol.* 1996; 87(02):163–168. Doi: 10.1016/0029-7844(95)00386-X
- Cheng YKY, Lu J, Leung TY, Chan YM, Sahota DS. Prospective assessment of INTERGROWTH-21<sup>st</sup> and World Health Organization estimated fetal weight reference curves. *Ultrasound Obstet Gynecol.* 2018;51(06):792–798. Doi: 10.1002/uog.17514
- Ioannou C, Talbot K, Ohuma E, et al. Systematic review of methodology used in ultrasound studies aimed at creating charts of fetal size. *BJOG.* 2012;119(12):1425–1439. Doi: 10.1111/j.1471-0528.2012.03451.x
- Andreasen LA, Tabor A, Nørgaard LN, Rode L, Gerds TA, Tolsgaard MG. Detection of growth-restricted fetuses during pregnancy is associated with fewer intrauterine deaths but increased adverse childhood outcomes: an observational study. *BJOG.* 2021;128(01):77–85. Doi: 10.1111/1471-0528.16380
- da Silva AA, Simões VM, Barbieri MA, et al. A protocol to identify non-classical risk factors for preterm births: the Brazilian Ribeirão Preto and São Luís prenatal cohort (BRISA). *Reprod Health.* 2014;11(01):79. Doi: 10.1186/1742-4755-11-79
- Bayley N. 2006 Bayley scales of infant and toddler development: Bayley III (Vol. 7). San Antonio, TX: Harcourt Assessment, Psych. Corporation
- Unterscheider J, Daly S, Geary MP, et al. Optimizing the definition of intrauterine growth restriction: the multicenter prospective PORTO Study. *Am J Obstet Gynecol.* 2013;208(04):290.e1–290.e6. Doi: 10.1016/j.ajog.2013.02.007
- Unterscheider J, O'Donoghue K, Daly S, et al. Fetal growth restriction and the risk of perinatal mortality-case studies from the multicentre PORTO study. *BMC Pregnancy Childbirth.* 2014;14(01):63. Doi: 10.1186/1471-2393-14-63
- Gordijn SJ, Beune IM, Thilaganathan B, et al. Consensus definition of fetal growth restriction: a Delphi procedure. *Ultrasound Obstet Gynecol.* 2016;48(03):333–339. Doi: 10.1002/uog.15884
- Beune IM, Bloomfield FH, Ganzevoort W, et al. Consensus based definition of growth restriction in the newborn. *J Pediatr.* 2018; 196:71–76.e1. Doi: 10.1016/j.jpeds.2017.12.059
- INTERGROWTH-21st [Internet]. Oxford: The University of Oxford; [cited 2019 Nov 10]. Available from <http://intergrowth21.ndog.ox.ac.uk/en/ManualEntry/Compute>
- The Fetal Medicine Foundation [Internet]. London: The Fetal Medicine Foundation; [cited 2019 Nov 10] Available from: <https://fetalmedicine.org/research/assess/bw>
- Villar J, Cheikh Ismail L, Victora CG, et al; International Fetal and Newborn Growth Consortium for the 21st Century (INTERGROWTH-21st) 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):857–868. Doi: 10.1016/S0140-6736(14)60932-6
- de Vries LS, Jongmans MJ. Long-term outcome after neonatal hypoxic-ischaemic encephalopathy. *Arch Dis Child Fetal Neonatal Ed.* 2010;95(03):F220–F224. Doi: 10.1136/adc.2008.148205
- Wallander JL, Bann C, Chomba E, et al. Developmental trajectories of children with birth asphyxia through 36 months of age in low/low-middle income countries. *Early Hum Dev.* 2014;90(07): 343–348. Doi: 10.1016/j.earlhumdev.2014.04.013
- Maitre NL, Slaughter JC, Aschner JL. Early prediction of cerebral palsy after neonatal intensive care using motor development trajectories in infancy. *Early Hum Dev.* 2013;89(10):781–786. *Am J Obstet Gynecol.* 2018;218(02):S692–S699. Doi: 10.1016/j.earlhumdev.2013.06.004
- Monier I, Blondel B, Ego A, Kaminiski M, Goffinet F, Zeitlin J. Poor effectiveness of antenatal detection of fetal growth restriction and consequences for obstetric management and neonatal outcomes: a French national study. *BJOG.* 2015;122(04):518–527. Doi: 10.1111/1471-0528.13148
- Nwabuobi C, Odibo L, Camisasca-Lopina H, Leavitt K, Tuuli M, Odibo AO. Comparing INTERGROWTH-21st Century and Hadlock growth standards to predict small for gestational age and short-term neonatal outcomes. *J Matern Fetal Neonatal Med.* 2020;33(11):1906–1912. Doi: 10.1080/14767058.2018.1533945
- 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 population-based cohort study. *PLoS Med.* 2019; 16(09):e1002902. Doi: 10.1371/journal.pmed.1002902
- Barros AJ, da Silva dos Santos I, Victora CG, et al. The 2004 Pelotas birth cohort: methods and description. *Rev Saude Publica.* 2006; 40(03):402–413. Doi: 10.1590/s0034-89102006000300007

- 32 Passini R Jr, Cecatti JG, Lajos GJ, et al; Brazilian Multicentre Study on Preterm Birth study group. Brazilian multicentre study on preterm birth (EMIP): prevalence and factors associated with spontaneous preterm birth. *PLoS One*. 2014;9(10):e109069. Doi: 10.1371/journal.pone.0109069
- 33 American College of Obstetricians and Gynecologists. Hypertension in pregnancy. Report of the American College of Obstetricians and Gynecologists' task force on hypertension in pregnancy. *Obstet Gynecol*. 2013;122(05):1122–1131. Doi: 10.1097/01.AOG.0000437382.03963.88
- 34 Azagba S, Manzione L, Shan L, King J. Trends in smoking during pregnancy by socioeconomic characteristics in the United States, 2010–2017. *BMC Pregnancy Childbirth*. 2020;20(01):52. Doi: 10.1186/s12884-020-2748-y
- 35 Kajdy A, Modzelewski J, Filipecka-Tyczka D, Pokropek A, Rabi-jewski M. Development of birth weight for gestational age charts and comparison with currently used charts: defining growth in the Polish population. *J Matern Fetal Neonatal Med*. 2019;•••:1–8. Doi: 10.1080/14767058.2019.1676412
- 36 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(04):509.e1–509.e7. Doi: 10.1016/j.ajog.2015.10.931
- 37 Francis A, Hugh O, Gardosi J. Customized vs INTERGROWTH-21<sup>st</sup> standards for the assessment of birthweight and stillbirth risk at term. *Am J Obstet Gynecol*. 2018;218(2S):S692–S699. Doi: 10.1016/j.ajog.2017.12.013
- 38 Duncan J, Leavitt K, Duncan K, Vilchez G. Detection of small for gestational age in preterm prelabor rupture of membranes by Hadlock versus the Fetal Medicine Foundation growth charts. *Obstet Gynecol Sci*. 2021;64(03):248–256. Doi: 10.5468/ogs.20267