Archives of Gynecology and Obstetrics (2019) 300:1211–1219
https://doi.org/10.1007/s00404-019-05234-5

MATERNAL-FETAL MEDICINE

Ultrasound diagnosis of microcephaly: a comparison of three reference curves and postnatal diagnosis

Lucas Augusto Monteiro de Castro Doin Trigo[1,3] · Joao Renato Benini-Junior[1] · Luiz Gustavo Oliveira Brito[1] · Sergio Tadeu Martins Marba[2] · Eliana Amaral[1]

Received: 30 December 2018 / Accepted: 27 June 2019 / Published online: 6 September 2019

Abstract

Purpose To evaluate which reference curve (RC)—Snijders, Intergrowth 21st (IG21) and World Health Organization (WHO)—is more accurate for microcephaly diagnosis.

Methods Retrospective cohort study with more than 30,000 exams in more than 11,000 women. Microcephaly was confirmed by a neonatologist at birth and positive predictive values (PPVs) and misdiagnosis were assessed.

Results A total of 71 cases were confirmed as microcephaly at birth. IG21 and Snijders PPVs showed to be more significant over WHO's ($p < 0.001$), without difference between them ($p = 0.39$). All RC were superimposed and did not show significant difference. When evaluated in different fragments, three trends were observed (until 30 weeks, between 30 and 36 and after 36 weeks of gestational age), with the latter interval showing a significant difference between IG21 and WHO ($p = 0.0079$). Conversely, WHO exhibited only one misdiagnosis, a much lower rate than Snijders, who missed eight cases and IG21, nine.

Conclusion WHO's RC appears to misdiagnose fewer cases, which could be useful for a population screening, while IG21's RC presented a more significant PPV, being more useful for a more precise final diagnosis in reference centers.

Keywords Microcephaly · Reference values · Fetal development · Growth curve · Diagnosis

Introduction

Fetal head circumference (HC) measurement by ultrasonography is important to evaluate intrauterine fetal growth and allows the diagnosis of abnormalities such as microcephaly [1, 2]. It also can be related to other fetal malformations, whether in the central nervous system or in other body parts [3–5].

Recently, Zika virus was the most cited cause of microcephaly, but other infections, such as toxoplasmosis and cytomegalovirus may cause similar findings [6–9]. Non-infectious causes like genetic syndromes are also related to microcephaly, including trisomies (Down, Patau and Edwards Syndromes) and maternal use of drugs and alcohol [10–14]. Finally, there are the small for gestational age fetuses due to hereditary factors, nutritional deficiency, or placental disorders [15, 16]. Although the final diagnosis requires other exams, prenatal suspicion relies upon ultrasound evaluation.

Curves of references are designed to define a normality pattern and to help in this evaluation. In 1994, Snijders and Nicolaides (Snijders) published one of the most commonly used reference curves (RC) created from the analysis of more than a thousand normal singleton pregnancies between 1987 and 1993 in the UK [17]. Meanwhile, in a publication on children's growth, the World Health Organization (WHO) suggested that it would be necessary to develop another

✉ Lucas Augusto Monteiro de Castro Doin Trigo
trigoxlv@gmail.com

Joao Renato Benini-Junior
bennini@gmail.com

Luiz Gustavo Oliveira Brito
lgobrito@gmail.com

Sergio Tadeu Martins Marba
sergiomarba@yahoo.com.br

Eliana Amaral
amaraleli@gmail.com

1 Department of Obstetrics and Gynaecology, School of Medical Sciences, University of Campinas, 101 - Cidade Universitária, Campinas, SP 13083-881, Brazil

2 Department of Paediatrics, School of Medical Sciences, University of Campinas, Campinas, Brazil

3 Calle Valencia, 54, Piso 2, Puerta 3, 08015 Barcelona, Spain

reference interval with a more heterogeneous population [18]. Nevertheless, several studies have shown a more accurate result in microcephaly diagnosis when a local developed RC was used [19–21].

In 2013, Intergrowth-21st study (IG21) started to be published and evaluated pregnant women in eight countries from four different continents [22, 23]. Leibovitz et al. compared three references curves, from USA, Israel and the IG21, showing similar positive predictive values (PPV) for the diagnosis of microcephaly, around 60% [24]. Likewise, Stirnemann et al. showed no disparity between IG21 and the French for a low-risk pregnancy group [25]. Later on, but using similar study design as IG21, WHO published a multicentric study in ten different countries from four continents [26, 27].

Brazil was included in both WHO and IG21 studies, but no comparative analysis regarding the regional use of these curves were made. This paper intended to compare both RC and the traditional RC of Snijders, currently used in the regional reference service at a tertiary hospital in Campinas, Brazil. We expected to clarify the real usefulness of each RC and which one would be more appropriate to use in a reference center for a more precise diagnosis.

Methods

Design

Retrospective study conducted at the Women's Hospital (CAISM), University of Campinas, tertiary referral center for an estimated 3,000,000 people in Brazil. Institutional Review Board approved the study under the registration #69181317.3.0000.5404 and STROBE statement was followed.

Study population

Records of all obstetrics ultrasounds (US) between 1 January 2011 and 22 November 2016 were collected from the database available at ASTRAIA® (Astraia software GMBH®, Munich, Germany), comprising 32,181 exams among 11,269 women. Women included in this study had a singleton pregnancy, between 24 and 40 weeks + 6 days, with a living fetus absent of malformations, delivered at CAISM with a last available US assessed no longer than 14 days before birthdate suggesting microcephaly. Identification of these cases was performed through a search on internal database.

Definitions

GA was calculated according to last menstrual period (LMP) and/or the first-trimester ultrasound. If the crown-rump length (CRL) was compatible with LMP or with a variation of 7 days from it, LMP was adopted as reference. If unaware of their LMP or the difference was bigger than 7 days from CRL, we adopted the first-trimester US as reference [28, 29].

Techniques

Exams were performed in the Ultrasound Unit of the hospital with Toshiba Xario® (Toshiba Medical System Corporation, Otawara, Tochigi, Japan), Voluson 730 Expert® or Voluson 8® (GE Healthcare® Ultrasound, Milwaukee, WI, USA) equipment. Data were saved in ASTRAIA® database. Fetal biometry was assessed by ultrasound residents and reviewed and signed by expert ultrasound specialists, senior members of the Imaging Section team.

Biparietal diameter (BPD) was measured from the outer edge of parietal fetal bone to the inner edge of the deeper parietal fetal bone (outer–inner) at the level of the cavum septum pellucidum, thalamus, and choroid plexus in the atrium of the lateral ventricles. Occipitofrontal diameter (OFD) was measured in this same level from the outer edge of frontal fetal bone to the inner edge of occipital fetal bone (outer–inner) in the middle of the bone echoes. HC was calculated from these two previous measures using the formula $3.14 \times (BPD + OFD)/2$, according to Snijders and Nicolaides' method [17].

These respective values were recorded in the software and HC measure was later distributed separately into Snijders, IG21 and WHO's RC for HC. Microcephaly cases were divided according to these RC. Cases without complete information at delivery or delivered elsewhere were excluded.

In order to create a standard to evaluate the PPV of microcephaly diagnosis, this paper used percentile 2.5, also defined as − 1.96 SD below the mean, as the cut-off for HC fetal measurement. This decision was made because it corresponds to the least available percentile provided by WHO's study. Snijders and Nicolaides did not report this percentile, but both provided access to formulas to achieve it (please see statistical analysis). This formula was not available in WHO's publication, not allowing us to standardize a lower percentile for the study [17, 23].

Microcephaly gold standard was defined by the postnatal cephalic measure assessed by an expert neonatologist. It was given using a flexible measuring tape from the occipital prominence at the back to the front above the eyebrows, passing above the ears, immediately post-partum. This measure was distributed in the RC for postnatal HC, through an online tool [30, 31]. Postnatal microcephaly was diagnosed if this value corresponded to more than 2SD below the mean expected for the postnatal GA and gender. This data was calculated by using date of delivery, postnatal HC and gender, obtained directly from the women's medical record at the Hospital.

Statistical analysis

Linear regression was used to identify cases prenatally considered microcephaly for each RC. Snijders disposed a normal range for HC according to GA per week from 14 to 39 weeks with the median and percentile 5 and 95. Percentile 2.5 was calculated through the logarithmic formula as $\log10(CC+1)=1.3369692+0.0596493\times$ weeks $-\ 0.0007494\times \text{weeks}^2$, presented in the original manuscript. With that, it was possible to input fractioned weeks in order to obtain the reference by days and assess the HC expected for the 40th week of GA.

Similarly, the IG21 study presented the normal range for median HC per week of GA from 14 to 40 weeks, starting from the third percentile. The logarithmic formula to assess the median HC equivalent to the 2.5 percentile for each GA was $-28.2849+1.69267\times GA2-0.397485\times GA2\times \log(GA)$. Despite the need of calculate the median of percentiles 2.5 of all days along each week for Snijders and IG21, WHO's publication offers a full growth chart table with HC percentile by GA median per week from 14 to 40 weeks, including the percentile 2.5 used for this study.

The gold standard to detecting microcephaly at birth was measuring 2 SD below the mean expected for the baby's GA and gender by the Intergrowth 21st postnatal criteria, assessed right after birth [32]. To evaluate the diagnostic congruence between Snijders and Nicolaides, IG21 and WHO, the Cohen kappa coefficient was assessed for each two pairs at a time. All three RC were also compared using the linear regression model.

Accuracy of the HC for diagnosis of microcephaly at birth was assessed by the positive predictive value (PPV), which was calculated as the number of cases with microcephaly at birth divided by the number of all cases considered abnormal by HC measure (with or without the postnatal diagnostic). PPVs assessed for each of the optimal HC cut-offs determined by Snijders, IG21 and WHO were compared using Leisenring, Alonzo and Pepe general score test for paired designs [33]. PPV distributions between genders for the three parameters were compared by X^2 test. All statistical calculations were performed with R software (3.4.0 version) and a p value <0.05 was considered statistically significant for all analysis. To define the cut-off for the significance, Bonferroni's correction was used, assuming significance when the p value is below $\alpha/3$, equal to $0.05/3=0.0167$ [34].

Adjusting a regression logarithmic model with time in weeks as an independent variable as opposed to percentile 2.5, which is a dependent variable, for neonatal HC, we obtained the following formula: $p\ 2.5=\alpha+\beta\times\log(\text{week}) = -\ 474.106+214.897\times\log(\text{week})$.

When analyzed, total sample was stratified into three groups by GA according to the RC's behavior (Fig. 1). Between 24–30 weeks, there were only 63 US diagnosed as microcephaly; 161 cases between 30–36 weeks, and 138 cases with more than 36 weeks. Comparison of PPVs was only possible for cases between 30–36 weeks and more than 36 weeks due to the already expected reduced number of cases from 24–30 weeks. For a proper evaluation of the significance of analysis, Bonferroni's correction was applied to perform multiple comparisons and once more the p value considered statistically significant was $0.05/3=0.0167$ [34].

Results

A total of 32,181 exams were analyzed from 11,859 different gestations. A first exclusion due to missing data, multiple pregnancy and GA over than 40 weeks and 6 days was made, remaining 11,269 exams. From them, 792 cases considered microcephaly according to Snijders, IG21, and/or WHO's criteria. After excluding new missing data, malformed fetus and fetal death, those who were not delivered at CAISM, and a period higher than 14 days since the last US, the final count was 362 microcephaly cases, out of which 251 were diagnosed by Snijders, 256 diagnosed by IG21, and 349 diagnosed by WHO, as illustrated in Fig. 2.

Snijders and WHO simultaneously considered 550 cases as microcephaly, while 51 were considered microcephaly only according to Snijders, and 194 only according to WHO, giving a Cohen's Kappa of 0.80 and a p value <0.001. In IG21 and WHO analysis, 568 cases were considered microcephaly according to both, 21 were considered microcephaly only according to IG21 and 173 only according to WHO, with a Cohen's Kappa of 0.84 and p value <0.001. The third analysis compared Snijders and IG21, with 558 cases considered microcephaly according to both, 43 considered microcephaly only according to Snijders and 31 only according to IG21, with a Cohen's Kappa of 0.93 and a p value <0.001.

A total of 71 cases (19.6%) were confirmed as postnatal microcephaly. Taking into consideration the definition by Snijders, 63 out of the 251 cases (25.1%) were confirmed as postnatal microcephaly. In comparison, IG21 had 62 cases confirmed out of 256 diagnosed (24.2%) whereas WHO confirmed as postnatal microcephaly 70 out of 346 cases (20.2%) Table 1. Snijders and IG21 showed a better PPV in comparison to WHO (binary diagnostic paired test, p value <0.001). However, there was not a significant difference between Snijders and IG21 (p value $=0.39$). While Snijders failed in diagnose eight cases and IG21 failed in nine cases, WHO failed in only one. However, WHO diagnosed 276 false positive cases, while Snijders 188 and IG21 194.

Analyzing the results by gender, according to Snijders, 45 out of the 165 female cases were confirmed (27.3%) as

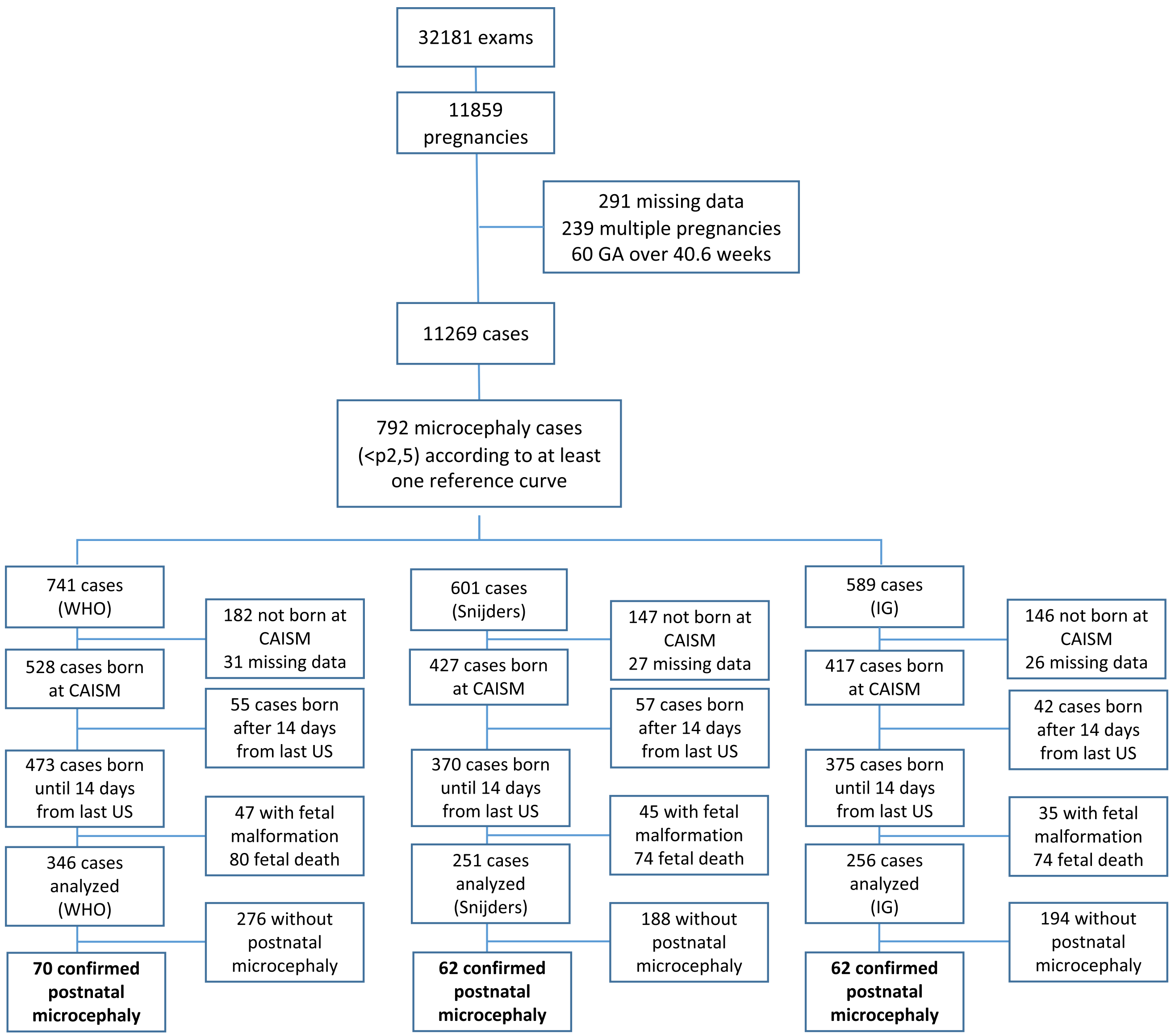

Fig. 1 HC per week curves for Snijders and Nicolaides, WHO and IG21 by gender

postnatal microcephaly, and only 18 out of 86 male cases were confirmed (20.9%). For IG21, 42 out of 170 female cases were confirmed (24.7%) and 20 out of 86 male cases (23.3%). For WHO's cases, 47 out of 222 female cases were confirmed (21.2%) as postnatal microcephaly, and 23 out of 124 male cases (18.5%). In all three references, there was a notable difference between male and female microcephaly diagnosis according to prenatal US, with a proportion of almost two females for each one male. Nevertheless, there was no significant difference between PPV values by gender in any RC (x^2 homogeneity test, Snijders p value = 0.34, IG21 p value = 0.91, and WHO p value = 0.65) (Table 1). As shown in Fig. 1, despite the disparity in microcephaly between male and female fetuses diagnosed by US, there was no difference in microcephaly after birth (green). Females (arrows) and males (triangles) were similar, with a few more female cases in the bottom values.

A graph with the three parameters for HC percentile 2.5 and GA was drawn (Fig. 3), showing that they are very similar until approximately 36–37 weeks, when WHO continues to increase and Snijders and IG21 stopped in a plateau. When compared by pair, Snijders and IG21 showed a p value of 0.679, and WHO and IG21 showed a p value of 0.628. Consequently, Snijders and WHO did not exhibit a significant difference in between them either. The three curves have no statistical difference, meaning that the regression logarithmic model presented was not sufficient to explain the relation between percentile 2.5 and logarithm of weeks.

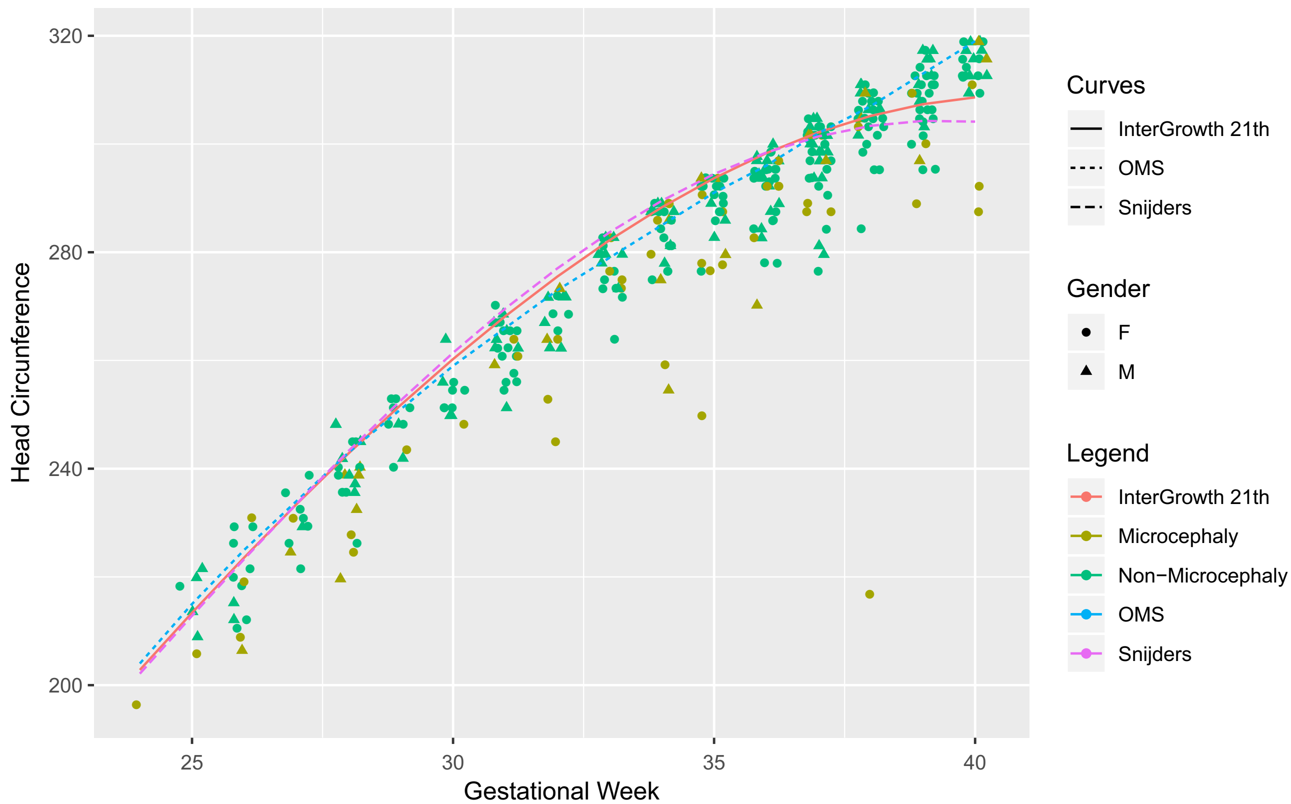

Fig. 2 Fluxogram showing the case selection for each reference (WHO, Snijders and IG21) according to the non-continuity criteria, starting with all the obstetrics ultrasounds analyzed at the beginning of the study

Table 1 PPV values for microcephaly for each US reference analyzed by gender

	Total PPV*	Female PPV	Male PPV
Snijders	25.1	27.27	20.93
WHO	20.2	21.17	18.55
IG21	24.2	24.71	23.26

Snijders and Nicolaides, $p = 0.34$; IG21, $p = 0.91$; and WHO $p = 0.65$

*Binary diagnostic paired test, p value < 0.001

In order to a more profound evaluation, PPVs of each reference were assessed for an interval of 14 days in the gestational week starting at 25 weeks, as shown in Table 2. It was notable that WHO's PPV was higher than the others only between 33 and 34 weeks. The discrepancy became evident at 39–40 weeks, with Snijders' PPV three times higher and IG21's PPV two times higher than WHO's (Table 2). Between 37 and 40 weeks, IG21 and WHO showed a statistically significant difference, with a better result for IG21 (Table 3).

Discussion

This study compared the PPV of three different US references (Snijders, IG21 and WHO) for prenatal microcephaly with the postnatal microcephaly diagnosis as a gold standard. Snijders and IG21 showed a non-statistically significant difference between their PPVs, with a better PPV in comparison to WHO's. However, WHO's reference missed only one case of postnatal microcephaly, while Snijders and IG21 missed eight and nine cases, respectively. This finding indicates a different application for WHO's reference curve when compared to the others, once it can be better used, for instance, in a public health global microcephaly screening, covering a bigger number of cases even with a higher number of false positives, suggesting a tendency in a possible higher sensibility. Unfortunately, we cannot affirm that the sensibility is indeed higher because the NPV could not be assessed, once the cases not diagnosed by US as a microcephaly were not included in the project.

A significant difference among the curves could be expected, considering the disparity between their PPVs. However, this difference was not enough to produce a

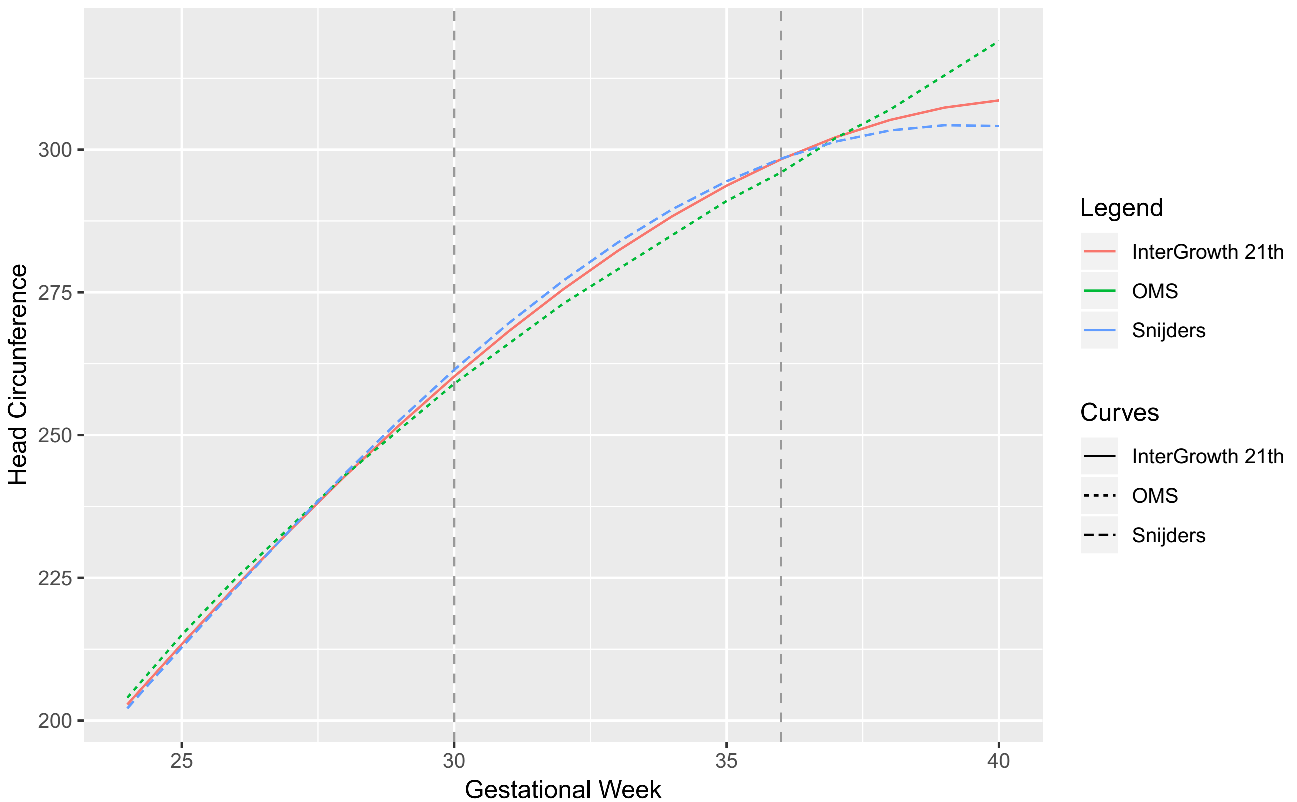

Fig. 3 Curves of percentile 2.5 of fetal head circumference measure per week according to Snijders, WHO and IG21 (IG21 vs WHO, $p=0.628$; IG21 vs Snijders, $p=0.679$)

Table 2 PPV values of microcephaly postnatal diagnosis for each US reference analyzed by a two weeks interval, from 25 to 40 weeks + 6 days

Gestational age	WHO	Snijders	Intergrowth 21st
25+0–26+6	25.00	33.33	33.33
27+0–28+6	28.12	36.00	36.00
29+0–30+6	9.52	11.11	11.11
31+0–32+6	21.62	21.05	21.62
33+0–34+6	30.00	27.66	26.32
35+0–36+6	22.81	24.59	23.64
37+0–38+6	14.10	15.79	18.52
39+0–40+6	15.00	45.45	31.25

Table 3 *p* values PPV values for comparison on microcephaly diagnosis by IG21, WHO, Snijders references according to intervals of gestational weeks

	30–36 weeks	37–40 weeks
IG21 vs WHO	0.1509	**0.0079**
IG21 vs Snijders	0.5883	0.8818
WHO vs Snijders	0.3507	0.0413

The bold meant the only significant *p*-value between groups, according to the Bonferroni correction, in which the $p < 0.017$ was considered

significant divergence between the curves. Furthermore, the total number of confirmed cases of microcephaly was too small when compared to the total amount of US microcephaly cases, not representing a significant change in the curves. This information is expected given the low incidence of microcephaly in general population, resulting in a small number of cases even in a large sample size as ours, which also represents a small variation in the RC. As seen in Fig. 3, the three RC are quite similar, except for the last gestational weeks, which is not the period with higher rate of microcephaly diagnosis. (Fig. 3).

To assess the difference between the RC's behaviors, three intervals were created (24–30 weeks, 30–36 weeks and more than 36 weeks). PPVs for each reference were compared according to their week interval. For a multiple comparison, Bonferroni's method was applied, transforming significant *p* value to 0.0167 [34]. A difference was found between IG21 and WHO RC, with a *p* value of 0.0079. This result was similar to Cheng et al.'s, which showed a difference between IG21 and WHO's curves for fetal estimated weight [35]. He noted that the reference values used by

WHO were always higher than IG21's for a same GA. However, Bonferroni's method represents a conservative adjustment of analysis and its interpretation should be softened. While the comparison between WHO and Snijders showed a p value of 0.0413, not enough to be considered significant according to Bonferroni, the PPV distribution of cases by GA showed that between 39 and 40 weeks, Snijders had a PPV three times bigger than WHO, which should not be overlooked, but considered. It shows that is highly recommended the use of Snijders' RC instead of WHO's for evaluate late term cases of microcephaly. Otherwise, the risk of a false positive increases substantially.

From 36–37 weeks forward, WHO's RC increases in a different pattern compared to Snijders and IG21. However, a regression logarithmic model did not show a statistically significant difference among them. This probably happened because this assessment took into consideration the whole distribution of the curve, not only a single interval, once the incidence of microcephaly is not high enough to allow an interval analysis as previous discussed. Also, Snijders and IG21 RC's behavior are more similar because they are both based on a normal model, while WHO's is based on a quantile regression. The difference in construction added to the small number of cases per week interval did not allow us to provide a comparison to evaluate in which specific point of the curves a statistical significance is detected, which may be seeing as a weakness of this manuscript. However, despite the more than 33,000 US exams with a more than 11,000 different pregnancies in our database, the final number of cases of postnatal microcephaly was only 71. When these cases were spread into two-week GA intervals, the numbers of cases became even smaller. Taking this information in consideration, any statistical comparison would increase considerably the bias of the study.

It is important to understand the considerable difference between prenatal and postnatal diagnosis for microcephaly in our study. Given that as our center is a tertiary referral hospital with a training team who is constantly been renewed, biometry assessment variations could be expected. The fact that we are evaluating fetuses closer to term and with a time difference no more than 14 days from their delivery adds some limitations in the head measurements, perhaps fetal head is fixed on the cervix and thus, changes in the anatomy may occur, fetal bones may be more calcified, increasing artefacts and limiting an ideal evaluation. Furthermore, managing complex cases such as intrauterine grown restriction may have increased the numbers of false positives in terms of small measures. Finally, it is important to clarify that our population has a considerable rate of overweight and obesity, which also may lead to a suboptimum abdominal intrauterine scan.

Some other limitations must be discussed. As this study was performed in a tertiary center with a high volume of exams per day, the Imaging Department expanded rapidly and different machines were used. The need to assess percentile 2.5 for Snijders and IG21 studies may pose some calculation error due to not being the exact value available in the original papers, but estimation. The use of a postnatal reference for microcephaly based on newborn gender was considered another form of bias, granted that the fetal biometry references did not report differences in this variable. The exclusive analysis of cases that were diagnosed as microcephaly through ultrasonography, disregarding all other cases without the ultrasonographic diagnosis, may have led to the loss of other cases of postpartum microcephaly and did not allow us to calculate the NPV nor the accuracy of the RC. However, it is important to take into consideration the great sample size for the project, as well as the time interval need to achieve that number, and the several obstacles to standardize a percentile to study, once each RC presented a different formula and a different presentation of their data. Any comparison between these different constructions would be suggested to a limitation.

Snijders and IG21 showed a better PPV in comparison to WHO. However, WHO's showed a much lower misdiagnosis rate. It is important to highlight that all three curves showed significantly similar behavior until 36–37 weeks regarding the microcephaly diagnosis. From this point forward, WHO's RC increases in a different pattern compared to Snijders and IG21, which highlight that WHO's reference may have a better use for a population-based screening of microcephaly in a large scale and lower GA, until 37 weeks, while Snijders and IG21 may be used for a more accurate final diagnosis on a case-by-case basis due to their better PPVs. Consequently, the use of WHO's reference seems to be more appropriate for the design of health policies, for preterm diagnosis. Once we already know that a screening population comes with a high rate of false positive, the cases should be revaluated according to Snijders or IG21 references for a more definitive diagnosis.

Author contributions LT, EA: Protocol/project development, data collection or management, data analysis, manuscript writing/editing. JRB, STM: Project development, data collection or management. LGB: Data management, data analysis, manuscript writing/editing.

Funding The first author received a Master of Sciences scholarship from Coordination for the Improvement of Higher Education Personnel (CAPES).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This study was performed in accordance with the 1964 Helsinki Declaration and received Institutional Board Review for being performed.

Informed consent Informed consent was not required as it was a retrospective study and all information were de-identified.

References

1. Johnsen SL, Rasmussen S, Wilsgaard T, Sollien R, Kiserud T, Lian Johnsen S (2006) Longitudinal reference ranges for estimated fetal weight. Acta Obstet Gynecol Scand 85(3):286–297
2. Kiserud T, Johnsen SL (2009) Biometric assessment. Best Pract Res Clin Obstet Gynaecol 23(6):819–31
3. Chervenak FA, Jeanty P, Cantraine F et al (1984) The diagnosis of fetal microcephaly. Am J Obstet Gynecol 149(5):512–517
4. Smith PA, Chudleigh P, Campbell S (1984) Prenatal diagnosis. Ultrasound Br J Hosp Med. 31:421–426
5. Pante FR, Madi JM, de Araújo BF, Zatti H, Madi SRC, Rombaldi RL (2011) Malformações congênitas do sistema nervoso central: prevalência e impacto perinatal. AMRIGS 55(4):339–344
6. Sakkas H, Economou V, Papadopoulou C (2016) Zika virus infection: past and present of another emerging vector-borne disease. J Vect Borne Dis 53:305–311
7. Neu N, Duchon J, Zachariah P (2015) TORCH infections. Clin Perinatol 42(1):77–103
8. Graham KA, Fox DJ, Talati A et al (2017) Prevalence and clinical attributes of congenital microcephaly—New York, 2013–2015. MMWR 66(5):125–129
9. Rawlinson WD, Boppana SB, Fowler KB et al (2017) Congenital cytomegalovirus infection in pregnancy and the neonate: consensus recommendations for prevention, diagnosis, and therapy. Lancet Infect Dis 17:e177–e188
10. Delgado Peña YP, Rodríguez Martínez G, Samper Villagrasa MP et al (2012) Características socioculturales, obstétricas y antropométricas de los recién nacidos hijos de madre fumadora. An Pediatr (Barc) 76(1):4–9
11. Ortega-Garcia JA, Gutierrez-Churango JE, Sanchez-Sauco MF et al (2012) Head circumference at birth and exposure to tobacco, alcohol and illegal drugs during early pregnancy. Childs Nerv Syst 28(3):433–439
12. del Campo M (2017) A review of the physical features of the fetal alcohol spectrum disorders. Eur J Med Genet 60(1):55–64
13. Kagan KO, Sonek J, Wagner P, Hoopmann M (2017) Principles of first trimester screening in the age of non-invasive prenatal diagnosis: screening for chromosomal abnormalities. Arch Gynecol Obstet 296(4):645–651
14. Santorum M, Wright D, Syngelaki A, Karagioti N, Nicolaides KH (2017) Accuracy of first-trimester combined test in screening for trisomies 21, 18 and 13. Ultrasound Obstet Gynecol 49(6):714–720
15. Gordijn SJ, Beune IM, Thilaganathan B et al (2016) Consensus definition of fetal growth restriction: a Delphi procedure. Ultrasound Obstet Gynecol 48(3):333–339
16. Figueras F, Gratacos E (2017) An integrated approach to fetal growth restriction. Best Pract Res Clin Obstet Gynaecol 38:48–58
17. Snijders RJM, Nicolaides KH (1994) Fetal biometry at 14–40 weeks' gestation. Ultrasound Obstet Gynecol 4(1):34–48
18. WHO Multicentre Growth Reference Study Group (2006) WHO Child Growth Standards based on length/height, weight and age. Acta Paediatr Suppl 450:76–85
19. Juliusson PB, Roelants M, Hoppenbrouwers K, Hauspie R, Bjerknes R (2011) Growth of Belgian and Norwegian children compared to the WHO growth standards: prevalence below -2 and above +2 SD and the effect of breastfeeding. Arch Dis Child 96(10):916–921
20. Natale V, Rajagopalan A (2014) Worldwide variation in human growth and the World Health Organization growth standards: a systematic review. BMJ Open 4(1):e003735
21. Buck Louis GM, Grewal J, Albert PS et al (2015) Racial/ethnic standards for fetal growth: the NICHD fetal growth studies. Am J Obstet Gynecol 213(4):449.e1–449.e41
22. Altman DG, Ohuma EO, International Fetal and Newborn Growth Consortium for the 21st Century (2013) Statistical considerations for the development of prescriptive fetal and newborn growth standards in the INTERGROWTH-21st Project. BJOG 120(Suppl. 2):71–6
23. Papageorghiou AT, Ohuma EO, Altman DG et al (2014) 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-21 st Project. Lancet 384:869–879
24. Leibovitz Z, Daniel-Spiegel E, Malinger G et al (2016) Prediction of microcephaly at birth using three reference ranges for fetal head circumference: can we improve prenatal diagnosis? Ultrasound Obstet Gynecol 47(5):586–592
25. Stirnemann JJ, Fries N, Bessis R, Fontanges M, Mangione R, Salomon LJ (2017) Implementing the INTERGROWTH-21st fetal growth standards in France: a "flash study" of the College Français d'Echographie Foetale (CFEF). Ultrasound Obstet Gynecol 49(4):487–492
26. Kiserud T, Piaggio G, Carroli G et al (2017) The World Health Organization fetal growth charts: a multinational longitudinal study of ultrasound biometric measurements and estimated fetal weight. PLOS Med 14(1):e1002220
27. Kiserud T, Piaggio G, Carroli G et al (2017) Correction: The World Health Organization fetal growth charts: a multinational longitudinal study of ultrasound biometric measurements and estimated fetal weight. PLOS Med 14(3):e1002284
28. American Institute of Ultrasound in Medicine (2013) AIUM practice parameter for the performance of obstetric ultrasound examinations. Laurel, USA
29. Committee on Obstetric Pratice, the American Institue of Ultrasound in Medicine, and the Society for Maternal-Fetal Medicine (2017) Methods for estimating the due date. Committee opinion no. 700. Obstet Gynecol 129(5):e150–e154
30. Villar J, Cheikh Ismail L, Victora CG, International Fetal and Newborn Growth Consortium for the 21st Century (INTERGROWTH-21st), et al (2014) 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 384(9946):857–868
31. Villar J, Giuliani F, Bhutta ZA, Bertino E, International Fetal and Newborn Growth Consortium for the 21(st) Century (INTERGROWTH-21(st) et al (2015) Postnatal growth standards for preterm infants: the Preterm Postnatal Follow-up Study of the INTERGROWTH-21st Project. Lancet Glob Health 3(11):e681–e691
32. Intergrowth 21st. Intergrowth 21st standards and tools. https://intergrowth21.tghn.org/standards-tools/.

33. Leisenring W, Alonzo T, Pepe MS (2000) Comparisons of predictive values of binary medical diagnostic tests for paired designs. Biometrics 56(2):345–351
34. Bonferroni CE (1936) Teoria statistica delle classi e calcolo delle probabilità. R Ist Super di Sci Econ e Commer di Firenze 8:3–62
35. Cheng YKY, Lu J, Leung TY, Chan YM, Sahota DS (2018) Prospective assessment of the INTERGROWTH-21 and WHO estimated fetal weight reference curve. Ultrasound Obstet Gynecol 51(6):792–798