Which fetal growth charts should be used? A retrospective observational study in China

Jianxin Zhao¹, Ying Yuan², Jing Tao³, Chunyi Chen¹, Xiaoxia Wu³, Yimei Liao², Linlin Wu³, Qing Zeng², Yin Chen², Ke Wang¹, Xiaohong Li¹, Zheng Liu¹, Jiayuan Zhou¹, Yangwen Zhou¹, Shengli Li², Jun Zhu¹,4,5

¹National Office for Maternal and Child Health Surveillance of China, West China Second University Hospital, Sichuan University, Chengdu, Sichuan 610041, China; ²Department of Obstetrics, Shenzhen Maternity and Child Healthcare Hospital, Southern Medical University, Shenzhen, Guangdong 518028, China; ³Department of Ultrasound, Shenzhen Maternity and Child Healthcare Hospital, Southern Medical University, Shenzhen, Guangdong 518028, China; 4Key Laboratory of Birth Defects and Related Diseases of Women and Children (Sichuan University), Ministry of Education, West China Second University Hospital, Sichuan University, Chengdu, Sichuan 610041, China; 5Sichuan Birth Defects Clinical Research Center, West China Second University Hospital, Sichuan University, Chengdu, Sichuan 610041, China.

Abstract

Background: The fetal growth charts in widest use in China were published by Hadlock >35 years ago and were established on data from several hundred of American pregnant women. After that, >100 fetal growth charts were published around the world. We attempted to assess the impact of applying the long-standing Hadlock charts and other charts in a Chinese population and to compare their ability to predict newborn small for gestational age (SGA).

Methods: For this retrospective observational study, we reviewed all pregnant women (n = 106,455) who booked prenatal care with ultrasound measurements for fetal biometry at the Shenzhen Maternity and Child Healthcare Hospital between 2012 and 2019. A fractional polynomial regression model was applied to generate Shenzhen fetal growth chart ranges for head circumference (HC), biparietal diameter (BPD), abdominal circumference (AC), and femur length (FL). The differences between Shenzhen charts and published charts were quantified by calculating the Z-score. The impact of applying these published charts was quantified by calculating the proportions of fetuses with biometric measurements below the 3rd centile of these charts. The sensitivity and area under the receiver operating characteristic curves of published charts to predict neonatal SGA (birthweight <10th centile) were assessed.

Results: Following selection, 169,980 scans of fetal biometry contributed by 41,032 pregnancies with reliable gestational age were analyzed. When using Hadlock references (<3rd centile), the proportions of small heads and short femurs were as high as 8.9% and 6.6% in late gestation, respectively. The INTERGROWTH-21st standards matched those of our observed curves better than other charts, in particular for fat-free biometry (HC and FL). When using AC<10th centile, all of these references were poor at predicting neonatal SGA.

Conclusions: Applying long-standing Hadlock references could misclassify a large proportion of fetuses as SGA. INTERGROWTH-21st standard appears to be a safe option in China. For fat-based biometry, AC, a reference based on the Chinese population is needed. In addition, when applying published charts, particular care should be taken due to the discrepancy of measurement methods.

Keywords: Infant; newborn; Pregnancy; Growth chart; Gestational age; Birth weight; Prenatal care; Biometry; Fetal growth reference; Ultrasound measurement; Hadlock charts; INTERGROWTH-21st charts; China

Introduction

Measurement of fetal biometry has become an indispensable part of routine ultrasound examination during the second and third trimesters of gestation. Using fetal growth chart ranges for head circumference (HC), biparietal diameter (BPD), abdominal circumference (AC), and femur length (FL), clinicians can identify fetuses of small for gestational age (SGA), appropriate gestational age (AGA), or large gestational age (LGA) in a timely manner. For example, fetal HC below the 3rd
centile might indicate microcephaly, and AC below the 10th centile may indicate fetal growth restriction (FGR), which is closely related to mortality and morbidity.\(^1,2\) More than 100 fetal growth charts have been established over the last four decades; however, large variation is seen in study design and statistical modeling methods, as well as the cutoff points.\(^3,5,6\) For example, the 10th centile for AC at 36 gestational weeks ranged from 276 mm to 292 mm, even among the high-quality studies.\(^3,5,6\) As a result, it is not uncommon for a baby to be diagnosed with SGA by one doctor and later found to be normal when using a different reference chart.

For a considerable number of practitioners, the choice of reference chart is mainly determined by the default chart loaded by the ultrasound manufacturer.\(^5\) This is particularly noticeable in developing countries or in Asia, such as China,\(^7\) India,\(^8\) Indonesia,\(^9\) Brazil,\(^10\) Pakistan,\(^11\) and Korea\(^12\) (representing at least 45% of the world’s population), where local references are rarely applied. The most commonly used chart equipped with ultrasound machines was published by Hadlock in the 1980s and was based on several hundred white, predominantly middle-class women from a single center in the USA.\(^13-16\) However, a recent systematic review has shown that this reference was developed with important methodological flaws.\(^3\) In addition, evidence has shown that fetal growth is largely influenced by ethnicity or nationality;\(^17-23\) therefore, it would be risky if a reference or standard was directly taken into clinical practice without strict evaluation.

This study aimed to assess the impact of applying the long-standing Hadlock references and other fetal growth references in a Chinese population and to compare their ability to predict newborn SGA.

**Methods**

**Ethics approval**

The study was approved by the ethics committee of Shenzhen Maternity and Child Healthcare Hospital (No. LLYJ2020-017-006) and the need for informed consent was waived.

**Data source and study population**

For this retrospective observational study, we reviewed all examinations of women of singleton pregnancy booked for prenatal care at Shenzhen Maternity and Child Healthcare Hospital between 2012 and 2019, in which >20,000 fetuses were delivered per year. Our center mainly provides routine ultrasound examinations to both low-risk and high-risk pregnant women who lived in Shenzhen (about 200,000 births per year).\(^24\) The Prenatal Diagnosis Center at our hospital was designated a national prenatal diagnosis training center in 2004 and serves as a national referral hospital for patients from throughout China.

A search of the database was performed to identify the scans performed on singleton pregnancies (June 2012 and December 2019), in which each woman contributed 1 to 9 ultrasound examinations. All these data were collected and checked by two investigators. In our center, after the dating scan of crown-rump length (CRL), women were offered routine scans at 18 to 20, 23 to 28, 29 to 32, and 36 to 37 weeks, unless they were suspected of having risk factors for fetal anomalies or having a significant medical condition. The following information was collected: parents information including age, weight, height, body mass index (BMI), education background, income, family history; pregnancy information including parity, gravidity; fetal birth information including sex, date of birth, birth weight, birth-length, Apgar score. Inclusion criteria were: (1) both parents are Chinese; (2) singleton pregnancy; (3) definite last menstrual period (LMP) and regular cycle; (4) difference in gestational age (GA) according to LMP and CRL measurement in the first trimester ≤7 days (based on Hadlock formula\(^25\)); (5) scans at GA between 14\(^{th}\) and 40\(^{th}\) weeks. Pregnancies without first-trimester dating based on CRL and fetuses with major congenital malformations were excluded. We also excluded measurements below or above −5 and +5 standard deviation (SD), respectively.

**Fetal ultrasonic measurements**

In our center, all sonographers were required to undergo extensive training for at least 6 months in ultrasound scanning. All scans of fetal biometry were performed transabdominally by mainstream ultrasound machines (Siemens ACUSON Sequoia 512, Antares S2000, S3000, and SC2000 [Siemens Medical Solutions, Mountain View, CA, USA]; GE Voluson E8 [GE Healthcare, Zipf, Austria]; and Samsung A30 and Samsung WS80A [Samsung Medison Co Ltd, Seoul, Korea]) equipped with the high-resolution curvilinear transabdominal probe. All biometric measurements (HC, BPD, AC, and FL) were made as to the following techniques. Fetal head measurements were made in the axial plane at the level where the continuous midline echo was broken by the cavum septi pellucidi in the anterior third and both thalami could be seen symmetrically. Care was taken to ensure that the calvaria appeared smooth and symmetrical bilaterally. Fetal skull BPD was measured from the outer edge of the proximal calvarial wall to the inner edge of the distal calvarial wall (outer–inner; BPDoi). Fetal HC was measured by fitting a computer-generated ellipse to include the outer edges of the calvarial margins of the fetal skull.\(^7,14,15,26\) AC was measured on a transverse circular plane of the fetal abdomen at the level where the spine, descending aorta, anterior third of the umbilical vein, and stomach bubble could be seen in the same plane.\(^17,16,26\) The FL was measured from the greater trochanter to the lateral condyle, with both ends clearly visible and at a horizontal angle <45\(^\circ\). During the third trimester, particular care was taken not to include the epiphysis.\(^17\) Of note, differing from most previous studies, the greater trochanter was included in our measurement and as a result, this may lengthen the FL when compared to directly measuring two ends of the diaphysis. Strict quality control was performed regularly. Five studies were selected for our comparison. More details are reported in the Supplementary Materials, http://links.lww.com/CMJ/B150.
Statistical analyses

Various statistical methods for constructing fetal growth charts have been suggested, including models of polynomial regression\(^{[27]}\) (usually quadratic or cubic), fractional polynomial regression\(^{[28]}\), and quantile regression\(^{[29]}\). The methods of model curves are presented in the Supplementary Materials, \url{http://links.lww.com/CM9/B150}. After describing the patterns of growth in our population, we quantitatively compared them with other populations from five charts (Intergrowth [IG]-21st, National Institute of Child Health and Human Development [NICHD], World Health Organization [WHO], Hadlock, and Hong Kong) using the Z-score method, as recommended by Salomon et al.\(^{[30]}\) Z-score for each fetal parameter was calculated using the formula: 

\[
Z\-score = \frac{X_{GA} - M_{GA}}{SD_{GA}},
\]

where \(X_{GA}\) is the fitted reference value at a known GA from other studies, \(M_{GA}\) is the fitted mean value calculated from our reference equations at this GA, \(SD_{GA}\) is the fitted SD at the same GA from our population. Here is an example. To compare our reference and the Hadlock reference, we calculated Z-score using the formula mentioned above, where \(X_{GA}\) represents the value of the 3rd, 50th, or 97th centile from the Hadlock reference. If the difference between Hadlock curves and ours is very small, the Z-score for the 3rd, 50th, or 97th centile curves should be close to −1.88, 0, and 1.88, respectively. In this study, we only used 3rd centile curves because the assessment of fetal smallness is more widely applied in obstetrics. Z-scores below −1.88 indicate that the lower limiting value of other curves is lower than ours. For example, at 20 weeks, a Z-score of −2.88 is smaller than the 3rd centile of this reference (−1.88) with 1 SD. We also evaluated the impact of applying these published charts in Shenzhen, by calculating the proportions of fetuses below the 3rd centile of these charts. The value of the 3rd centile was measured in days. If a chart identified a proportion below the 3rd centile close to the expected value of 3% in our population, applying this in Shenzhen may be safer. Finally, we compared the ability of predicting SGA by AC<10th centile between our chart and other charts. Neonatal SGA was defined as weight at birth <10th centile in our population at a gestation-specific GA. Sensitivity, specificity, and area under the receiver operating characteristic curves (AUCs) were next performed to measure the discriminatory abilities of these charts for SGA.

If outcome variables or covariates had missing data, the cases were excluded from the model. Statistical analyses were performed using SAS (version 9.4; SAS Institute Inc., Cary, NC, USA). Continuous variables were summarized as mean ± SD or median (Q1–Q3), and categorical variables by number and percentage.

Results

Baseline characteristics

During the period from 2012 to 2019, a search of the database identified 106,455 live-birth singleton pregnancies with a total of 385,282 ultrasound scans (HC, BPD, AC, FL) during the second or third trimester [Figure 1]. A total of 60,629 (56.95%) pregnancies with 189,957 (49.30%) scans were excluded from the analysis due to the absence of first-trimester dating based on CRL length. Among the 45,826 (43.05%) remaining pregnancies, 4794 (10.4%) were excluded for at least one of the following reasons: the absolute value of the difference between GA based on LMP and CRL (Hadlock’s formula\(^{[25]}\)) was >7 days (\(n = 3547\) [7.74%]); GA <14\(^{0}\) weeks or >40\(^{0}\) weeks (\(n = 125\) [0.27%]); fetal biometry measurements were below −5 SD or above 5 SD (\(n = 21\) [0.05%]); and pregnancy without congenital malformations (\(n = 1101\) [2.40%]). Finally, 169,980 (44.11%) scans contributed by 41,032 (38.54%) pregnancies that had a reliable GA that was corroborated by the dating scan in all pregnancies was four (range 1–9), and 39,234 (95.61%) pregnancies had 2 to 6 scans [Supplementary Table 1, \url{http://links.lww.com/CM9/B150}]. Overall, 149,362 (87.87%) scans were performed within the stipulated GA window (23.30% at 22–25 weeks, 23.45% at 29–32 weeks, and 41.12% at 36–39 weeks; Supplementary Figure 1, \url{http://links.lww.com/CM9/B150}).

The demographic characteristics and perinatal events of pregnancies from Shenzhen and other studies are presented in Table 1. The mean maternal age, height, weight, and BMI were 30.8 ± 4.2 years, 159.7 ± 5.8 cm, 53.3 ± 9.8 kg, and 20.9 ± 3.6 kg/m\(^2\), respectively. The mean birth weight and birth length were 3286.2 ± 438.1 g and 50.0 ± 1.4 cm, respectively. In our study population, 4.8% of pregnancies resulted in a preterm birth (defined as <37 weeks of gestation), 1.2% were low birthweight (defined as full-term [≥37 weeks] birthweight <2500 g), and 52.3% were males. Our population structure was similar to Shenzhen city. For
example, 30.6% of immigrants in Shenzhen are from other areas of Guangdong province and this number was 31.3% in our study. [Supplementary Figure 2, http://links.lww.com/CM9/B150].

**Construction of Shenzhen fetal growth curves and comparisons with the curve from other populations**

The equations for the regression models and validations are presented in Supplementary Tables 2 and 3, http://links.lww.com/CM9/B150. The comparisons of the fitted curves between Shenzhen and other populations are presented in Figure 2 and Supplementary Figures 3–5, http://links.lww.com/CM9/B150. For HC, a main fat-free skeletal measure, three centiles of Shenzhen curve performed similar to the curves of the IG-21st and NICHD Asian groups but differed from the Hadlock, Hong Kong, and WHO curves [Figure 2]. For example, the 3rd centile of the Hadlock curve was lower than ours in the early second trimester and clearly higher after 34 weeks. For the other head biometry parameters, the BPD

**Table 1: Comparison of baseline characteristics and perinatal events among populations from Shenzhen (n = 41,032) and other studies.**

| Variables                          | From Shenzhen | From the Intergrowth-21st study[28] | From the NICHD study [17] | Form the WHO study [29] | Form the Hong Kong study [26] | Form the Hadlock studies [13-16] |
|-----------------------------------|---------------|-------------------------------------|--------------------------|------------------------|-------------------------------|---------------------------------|
| Maternal age (years)             | 30.8 ± 4.2    | 28.4 ± 3.9                          | 30.5 ± 4.4               | 28 (25–31)             | NS                            | NS                              |
| Maternal height (cm)             | 159.7 ± 5.8   | 162.2 ± 5.8                         | 160.4 ± 6.0              | 163 (157–168)          | NS                            | NS                              |
| Maternal weight (kg)             | 53.3 ± 9.8    | 61.3 ± 9.1                          | 57.0 ± 8.3               | 61 (55–68)             | NS                            | NS                              |
| BMI (kg/m²)                      | 20.9 ± 3.6    | 23.3 ± 3.0                          | 22.1 ± 2.5               | 23.1 (21.0–25.4)       | NS                            | NS                              |
| Maternal birth weight (g)        | 3286 ± 438    | 3300 ± 400                          | NS                       | 3300 (2980–3615)       | NS                            | NS                              |
| Birth length (cm)                | 50.0 ± 1.4    | 49.4 ± 1.9                          | NS                       | NS                     | NS                            | NS                              |
| Preterm (<37 weeks of gestation) | 4.8 (1981/41,032) | 4.5 (195/4321)               | NS                       | 7.5 (99/1312)          | NS                            | NS                              |
| Term LBW <2500 g (≥37 weeks of gestation) | 1.2 (505/41,032) | 3.0 (128/4321)               | NS                       | NS                     | NS                            | NS                              |
| Male infant                      | 52.3 (21,458/41,032) | 49.7 (2149/4321)           | 52.0 (178/342)           | 53.2 (691/1299)        | NS                            | NS                              |

Data are mean ± SD or % (n/N) or median (Q1–Q3). BMI: Body mass index; LBW: Low birth weight; NICHD: National Institute of Child Health and Human Development; NS: Not shown in the original article; WHO: World Health Organization; SD: Standard deviation.

**Figure 2:** Comparison of fitted 3rd (lowest), 50th (middle), and 97th (highest) centile curves of HC between Shenzhen fetal growth curve (blue solid curves; FPs model) and other curves (red dashed curves). FPs: Fractional polynomials; HC: Head circumference; NICHD: National Institute of Child Health and Human Development; WHO: World Health Organization.
curves of Shenzhen population (outer–inner) were similar to those of the NICHD Asian group (outer–inner) and Hong Kong population (outer–inner) but considerably lower than IG-21st (outer–outer) population throughout the middle and late gestation [Supplementary Figure 3, http://links.lww.com/CM9/B150]. The 3rd and 97th centile curves for the BPD of Hadlock were not plotted as data were not available in the original publication. For FL, Shenzhen curves performed very close to the curves of IG-21st before 28 weeks and clearly higher thereafter [Supplementary Figure 4, http://links.lww.com/CM9/B150]. All three centiles of Hadlock were higher than ours. The largest difference was observed in the AC measurements [Supplementary Figure 5, http://links.lww.com/CM9/B150]. Although the 50th centile of the five charts were similar to ours, no chart was matched for the 3rd or 97th centiles.

Z-scores were then used to quantify the differences in fetal biometric measurements between the Shenzhen population and those of five previously published populations across different GAs [Figure 3], with a Z-score of −1.88 representing no difference at that GA in our equation and −2.88 representing that the 3rd centile references in other populations are 1 SD smaller than those in Shenzhen population. The results of this part further support the superiority of the IG-21st. For example, the differences in HC between the 3rd centiles of Shenzhen curve and IG-21st curve were within ±0.5 SD before 38 weeks. Conversely, HC Z-scores of Hadlock tended to progressively increase with advancing GA, with the difference approaching −1 SD at 14 weeks and +1 SD at 40 weeks.

**Impact of adopting published curves in a Chinese population**

To assess the impact of adopting published references in the Shenzhen population, we compared their ability to identify fetal smallness by calculating the proportion of individual biometry measurements below the 3rd centile [Table 2]. NICHD Asian and WHO references were not included, as data specific to gestational day were not available in the original publication. The IG-21st identified a proportion of HCs below the 3rd centile closer to the expected value of 3% than did the other charts. For example, at 14+0 to 17+6 weeks, the proportions of small head fetuses (<3rd centile) were 3.1%, 0.2%, and 0.5% when using IG-21st, Hong Kong, and Hadlock references, respectively. Of note, small HCs were overestimated in late gestation when using the Hadlock reference, reporting 8.9% HCs below the 3rd centile. IG-21st also performed well at detecting the short femur fetuses, especially in the second trimester. A total of 6.6% of short femur fetuses were reported when using Hadlock at 34 to 40 weeks. For BPD and AC, no chart identified a proportion of fetal smallness close to the expected value of 3%.

**Figure 3:** Comparison of our new equations with references from Intergrowth-21st, NICHD, WHO, Hadlock, and Hong Kong curves for HC, BPD, AC, and FL. Z-score = (3rd centile of reference value in other studies − fitted mean value of Shenzhen study)/fitted standard deviation of Shenzhen study. AC: Abdominal circumference; BPD: Biparietal diameter; FL: Femur length; HC: Head circumference; IG-21st: Intergrowth-21st; NICHD: National Institute of Child Health and Human Development; SD: Standard deviation; WHO: World Health Organization. The dotted lines represent the expected Z-scores for the 3rd centiles (horizontal line in the middle) or ±0.5 SD or ±1 SD, calculated from our population; SD: Standard deviation.
Comparisons of the performance of different curves in predicting neonatal SGA

Figure 4 shows the ability of an AC cutoff <10th centile to detect SGA at birth (definition of birth weight <10th centile). The NICHD and WHO charts were excluded because reference values of accuracy to the day were not available. The sensitivity and AUC of Shenzhen charts for the prediction of SGA at birth increased as pregnancy developed (sensitivity: 0.18 at 14–18 weeks to 0.46 at 34–40 weeks; AUC: 0.55 at 14–18 weeks to 0.70 at 34–40 weeks). Conversely, IG-21st and Hong Kong curves had performed poorly at predicting SGA at birth, with a sensitivity of only 0.1 after 18 gestational weeks and an AUC not exceeding

Table 2: Proportions of Shenzhen fetus under 3rd centile using prescriptive charts of INTERGROWTH-21st, Hong Kong, and Hadlock.

| Gestational weeks | Intergrowth-21st chart[28] | Hong Kong chart[26] | Hadlock chart[13-16] |
|-------------------|----------------------------|---------------------|---------------------|
| Proportions of HC <3rd centile |                           |                     |                     |
| 14+0–17+6         | 3.1 (108/3485)             | 0.2 (8/3485)        | 0.5 (16/3485)       |
| 18+0–23+6         | 2.0 (351/17467)            | 0.5 (84/17,467)     | 1.2 (207/17,467)    |
| 24+0–27+6         | 3.2 (882/27,319)           | 1.1 (310/27,319)    | 2.4 (648/27,319)    |
| 28+0–33+6         | 1.9 (831/44,104)           | 1.4 (631/44,104)    | 2.5 (1101/44,104)   |
| 34+0–40+0         | 0.8 (599/77,605)           | 1.0 (741/77,605)    | 8.9 (6902/77,605)   |
| Proportions of BPD <3rd centile |                       |                     |                     |
| 14+0–17+6         | 20.4 (770/3780)            | 0.4 (14/3780)       | –                   |
| 18+0–23+6         | 17.1 (3038/17,739)         | 2.4 (419/17,739)    | –                   |
| 24+0–27+6         | 23.2 (6361/27,360)         | 4.6 (1260/27,360)   | –                   |
| 28+0–33+6         | 11.1 (4946/44,584)         | 2.3 (1035/44,584)   | –                   |
| 34+0–40+0         | 3.6 (2181/78,148)          | 1.0 (812/78,148)    | –                   |
| Proportions of AC <3rd centile |                       |                     |                     |
| 14+0–17+6         | 3.6 (136/3778)             | 1.5 (36/3778)       | 0.0 (1/3778)        |
| 18+0–23+6         | 0.8 (14/17,737)            | 0.5 (90/17,737)     | 0.5 (82/17,737)     |
| 24+0–27+6         | 0.8 (22/27,352)            | 0.6 (169/27,352)    | 1.1 (286/27,352)    |
| 28+0–33+6         | 0.4 (164/44,572)           | 0.4 (170/44,572)    | 1.7 (737/44,572)    |
| 34+0–40+0         | 0.2 (160/78,310)           | 0.3 (211/78,310)    | 3.9 (3056/78,310)   |
| Proportions of FL <3rd centile |                       |                     |                     |
| 14+0–17+6         | 3.8 (143/3781)             | 1.1 (43/3781)       | 1.5 (58/3781)       |
| 18+0–23+6         | 1.5 (261/17,734)           | 0.3 (60/17,734)     | 1.8 (324/17,734)    |
| 24+0–27+6         | 2.1 (576/27,355)           | 0.4 (104/27,355)    | 3.9 (1076/27,355)   |
| 28+0–33+6         | 0.9 (407/44,583)           | 0.4 (163/44,583)    | 4.5 (2007/44,583)   |
| 34+0–40+0         | 0.3 (253/78,158)           | 0.4 (287/78,158)    | 6.6 (5164/78,158)   |

Data are shown as % (n/N). The NICHD and WHO charts were excluded because reference values of accuracy to the day were not available in their article. –: Not available; AC: Abdominal circumference; BPD: Biparietal diameter; FL: Femur length; HC: Head circumference; NICHD: National Institute of Child Health and Human Development; WHO: World Health Organization.
0.55. For Hadlock charts, the sensitivity and AUC for the prediction of SGA at birth were lower than Shenzhen chart before 34 weeks (eg, 18–24 weeks, sensitivity 0.12 vs. 0.26; AUC 0.55 vs. 0.58).

Discussion
This large validation study including approximately 170,000 ultrasound scans indicated that the long-standing Hadlock reference curve had a poor ability for identifying fetal smallness, with underdiagnosing SGA in the early second trimester and over-diagnosing in late gestation. For example, extra 3.6% (6.6%–3.0%) and 5.9% (8.9%–3.0%) pregnancies in our database would be misclassified as short femurs and small heads at 34 to 40 weeks, respectively, potentially requiring additional unnecessary investigations and increasing parental anxiety.

With a closer proportion of classified as SGA to the expected value, the IG-21st standard had better overall performance than other references, in particular for the fat-free biometry (HC and FL). However, particular care should be taken to the discrepancy of measurement methods when using a fetal growth chart as we observed a clearly lower BPD in our population compared with the IG-21st population. BPD was measured by the “outer-inner” method in daily practice in China and many other countries and “outer-outer” in the IG-21st. Chitty et al[31] reported that the difference between the two methods ranged from 1.5 to 3.0 mm, and the difference was 2.0 mm in our analysis, indicating that the difference in BPD between IG-21st and Shenzhen curves may be small if measuring consistently. We also observed a higher FL measurement in the third trimester, this could be partly explained by the difference in ultrasound measurement methods: the greater trochanter, which appeared under ultrasound in the late pregnancy, was included in our measure of FL, but not in measure for IG-21st.

However, when using an AC cutoff of <10th centile, all these references had a poor performance for the prediction of SGA at birth (bithweight <10th). For example, only 10.7% of neonatal SGA was detected using IG-21st curve vs. 45.5% using Shenzhen curve during 34 to 40 weeks.

Our study had multiple strengths. First, the structure of our population has great geographical and ethnic diversity. As a pioneering city in China’s reform and opening up, Shenzhen has attracted migrants from all around China, with its population has grown by as much as 40 times from 310,000 in 1979 to 13.43 million in around China, with its population has grown by as much.

We also acknowledge several limitations. First, there was no information on gestational age. However, in our data, 95.7% of pregnancies had 2 to 6 scans, and 87.9% of scans were performed within the stipulated GA window, indicating the vast majority were low-risk pregnancies with routine scans. Additional limitations are the retrospective and single-center nature of the analysis. Finally, we did not perform a multilevel model to create a growth curve, as recommended in IG-21st. However, it was unlikely to affect our findings because we did not aim to create a new Chinese fetal growth chart and this could be confirmed by subsequent validation analysis.

Hadlock charts are methodologically flawed. For example, Hadlock did not consider the variability of measurements with GA (ie, the SD is constant), and as a result, the shape of Hadlock curves is impossibly similar to those of other studies. Indeed, in our analysis, Hadlock curves were clearly lower than ours in the early second trimester and higher in the late third trimester. This is in accordance with findings from many other studies. Merialdi et al[34] compared fetal biometry measurements obtained in a Peruvian population with Hadlock reference and Chitty (United Kingdom) reference.[31,35,36] The results showed that the most pronounced discrepancies were observed in relation to the Hadlock reference, with the averages in the Peruvian population being 1 SD smaller than the medians in the Hadlock population. Similar patterns were observed in a prospective study in Korea: in comparison with the population derived from Hadlock charts, Korean fetuses had higher HC, BPD, AC, and FL in the first half of pregnancy but tended to measure progressively smaller with advancing GA. The trend was more marked for HC, also approaching −1 SD when compared to the Hadlock population in late gestation.[37]

Our analysis shows that the IG-21st standard was more valid than other references for fat-free biometry, in particular HC, but not for fat-based biometry. This is in accordance with the results from a nationwide cross-sectional study that included 4858 pregnancies in France: their observed 50th centile curves of HC closely matched those of IG-21st and the Z-scores were close to 0 across GA.[38] However, IG-21st highly underestimated the proportion of small abdomens, with the proportions of ACs below the 3rd centile of IG were only 1%. This finding was also supported by studies from European countries like France,[21] Greece,[22] the Netherlands,[39] Norway,[40] and Italy,[20] and the United States.[22] It is clear that socioeconomic status is the main determinant of fat-based growth.[31,32] The data from IG-21st show that the fat-free biometry (eg, HC and length at birth), may differ by only 2.5% to 4.0% among populations in eight countries; however, the difference in birthweight may be
as high as 20%. Previous studies have also shown that, based on the 10th centile for birth weight of the IG-21st standards, the proportion of SGA in low-risk pregnancy was reported to be only 2% to 3% in developed countries in contrast to up to 50% in India.

In many developing countries and in particular those with large populations like China, India, Indonesia, Pakistan, Brazil, the choice of fetal growth chart is mainly determined by the default chart loaded by the ultrasound manufacturer for the reason of convenient use. However, we found that applying these long-standing references led to misdiagnosis of SGA, which would have grave consequences and a huge impact. For example, in China, 12 million babies were born in 2020, extra 430,000 (3.6%) may be labeled as short femurs (FL < 3rd centile) according to Hadlock charts at 34 to 40 weeks, potentially requiring unnecessary invasive examination and increasing parental anxiety. As of recently, China has >180,000 midwives, 200,000 obstetricians, and 26,000 midwifery institutes. Our finding alerts these practitioners or institutes to realize the grave consequences of continuing to use old references. We recommend a standard that is more applicable for Chinese babies and if it was to be adopted, overdiagnosis of SGA may be greatly reduced.

In conclusion, our study demonstrated that the Hadlock references had a poor ability to identify SGA in the early second trimester and could lead to serious misdiagnosis for SGA in late gestation, potentially resulting in an unnecessary invasive examination, or an absence of appropriate interventions. Applying the IG-21st standard seems to be a safe option in China, but only for fat-free biometry (HC and FL). For AC, a reference based on the local population is needed. In addition, when applying published charts, particular care should be taken due to the discrepancy of measurement methods.

Acknowledgments

The authors thank the department heads who agreed to share data in their unit. We wish to thank Prof. Ruyan Pang at China Maternal and Child Health Association for her assistance in study design.

Funding

This study was supported by grants from the Shenzhen Science and Technology Project (No. JCYJ20210324130812035) and The National Key Research and Development Program of China (No. 2018YFC1002200).

Conflicts of interest

None.

References

1. Lees CC, Stampalija T, Baschat A, da Silva Costa F, Ferrari 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;56:298–312. doi: 10.1002/uog.22314.

2. Salomon LJ, Alfrevic 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;53:715–723. doi: 10.1002/uog.20722.

3. Ioannou C, Talbot K, Ohuma E, Sarris I, Villar J, Conde-Agudelo A, et al. Systematic review of methodology used in ultrasound studies aimed at creating charts of fetal size. BJOG 2012;119:1435–1439. doi: 10.1111/j.1471-0528.2012.03451.x.

4. Altman DG, Chitty LS. Design and analysis of studies to derive charts of fetal size. Ultrasound Obstet Gynecol 1993;3:378–384. doi: 10.1046/j.1469-0783.1993.03.09.x.

5. Westerway SC, Papageorghiou AT, Hirst J, Costa FD, Hyett J, Walker SP. Intergrowth-21st - time to standardise fetal measurement in Australia. Australas J Ultrasound Med 2015;18:91–95. doi: 10.1016/j.auwm.2015.01.040.

6. Salomon LJ, Bernard JP, Duyme M, Burat I, Ville Y. The impact of choice of reference charts and equations on the assessment of fetal biometry. Ultrasound Obstet Gynecol 2005;25:559–565. doi: 10.1002/uog.1901.

7. Zhang Y, Meng H, Jiang Y, Xu Z, Ouyang Y, Li S, et al. The World Health Organization fetal growth charts: concept, history, findings, interpretation, and application. Am J Obstet Gynecol 2020;32:149–155. doi: 10.1093/aje/kwa130.

8. Aggarwal NH, Sharma GL. Fetal ultrasound parameters: reference values for a local perspective. Indian J Radiol Imaging 2020;30:149–155. doi: 10.4103/ijri.IJR_37.19.

9. Fattah AN, Pratiwi KN, Susilo SA, Bergamas JS, Irwinda R, Wibowo N, et al. Indonesian local fetal-weight standard: a better predictive ability for low Aggar score of SGA neonates. Med J Indonesia 2016;25:228–233. doi: 10.13818/mj.25.4.1301.

10. Araujo Júnior E, Martins Santana EMF, Martins WP, Júnior JE, Ruano R, Pires CR, et al. Reference charts of fetal biometric parameters in 31,476 Brazilian singleton pregnancies. J Ultrasound Med 2020;14:1181–1191. doi: 10.1186/s1371-0118.

11. Munim S, Morris T, Baber N, Ansari Y, Azam SI. Growth charts of fetal biometry: a longitudinal study. J Matern Fetal Neonatal Med 2012;25:692–698. doi: 10.1080/14767058.2011.592878.

12. Lee S, Kwon S, Group KR, Park M, Hwang H, Na S. VP40.18: comparing the new standard chart in Korea to Hadlock fetal growth standard, Tokyo chart, and the INTERGROWTH-21st: multicentre retrospective cohort study. Ultrasound Obstet Gynecol 2020;56:234–235. doi: 10.1002/uog.22956.

13. Hadlock FP, Harrist RB, Deter RL, Park SK. Fetal femur length as a predictor of menstrual age: sonographically measured. AJR Am J Roentgenol 1982;138:875–878. doi: 10.2214/ajr.138.5.875.

14. Hadlock FP, Deter RL, Harrist RB, Park SK. Fetal head circumference: a longitudinal study. J Matern Fetal Neonatal Med 2014;33:1181–1191. doi: 10.1186/s1371-0118.

15. Gratz KL, Hegedur 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;218:5641–5653. doi: 10.1016/j.ajog.2017.11.593.

16. Kissner T, Benach A, Hecher K, Perez RG, Carvalho J, Piaggio G, et al. The World Health Organization fetal growth charts: concept, findings, interpretation, and application. Am J Obstet Gynecol 2018;218:5649–5629. doi: 10.1016/j.ajog.2017.12.010.

17. Vissers GHA, Nicholson WK, Barnea-FR, Ramanauskas D, Nassar AH. FIGO position paper on reference charts for fetal growth and size at birth: which one to use? Int J Gynecol Obstet 2021;152:148–151. doi: 10.1002/ijgo.13500.

18. Stampalija T, Ghi T, Rosolen V, Rizzo G, Ferrari EM, Prefumo F, et al. Current use and performance of the different fetal growth charts in the Italian population. Eur J Obstet Gynecol Reprod Biol 2020;252:323–329. doi: 10.1016/j.ejogrb.2020.06.039.

19. Heide B, Le Guern M, Forhan A, Scherdel P, Kadawathagedara M, Dufour SGN, et al. Are selection criteria for healthy pregnancies responsible for the gap between fetal growth in the French national Elfe birth cohort and the Intergrowth-21st fetal growth standards? Paediatr Perinat Epidemiol 2019;33:47–56. doi: 10.1111/ppe.12526.
22. Hua X, Shen M, Reddy UM, Back Louis G, Souza JP, Guimezoglu AM, et al. Comparison of the INTERGROWTH-21st, International Institute of Child Health and Human Development, and WHO fetal growth standards. Int J Gynaecol Obstet 2018;143:156–163. doi: 10.1002/ijgo.12637.

23. Sottrnadias A, Eletheriades M, Chatziznikolaou F, Hassiakos D, Chrousos GP, Pervanidou P. National curves of foetal growth in singleton foetuses of Greek foetuses. Eur J Clin Invest 2016;46:425–433. doi: 10.1111/eci.12611.

24. Shenzhen, China Population. Populationstat, 2022. Available from: https://populationstat.com/china/shenzhen. [Accessed on March 9th, 2022]

25. Hadlock FP, Shah YP, Kanon DJ, Lindsey JV. Fetal crown-rump length: re-evaluation of relation to menstrual age (5-18 weeks) with high-resolution real-time US. Radiology 1992;182:501–505. doi: 10.1148/radiology.182.2.1732970.

26. Leung TN, Pang MW, Daljit SS, Leung TY, Poon CF, Wong SM, et al. Fetal biometry in ethnic Chinese: biparietal diameter, head circumference, abdominal circumference and femur length. Ultrasound Obstet Gynecol 2008;31:321–327. doi: 10.1002/uog.5192.

27. Altman DG, Chitty LS. Charts of fetal size: 1. Methodology. Br J Obstet Gynaecol 1994;101:29–34. doi: 10.1111/j.1471-0528.1994.tb13006.x.

28. 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;384:869–879. doi: 10.1016/S0140-6736(14)61490-2.

29. Kiserud T, Piaggio 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:e1002220. doi: 10.1371/journal.pmed.1002220.

30. Salomon LJ, Duyme M, Croquart J, Brodaty G, Talmant C, Fries N, et al. French fetal biometry: reference equations and comparison with other charts. Ultrasound Obstet Gynecol 2006;28:193–198. doi: 10.1002/uog.2733.

31. Chitty LS, Altman DG, Henderson A, Campbell S. Charts of fetal size: 2. Head measurements. Br J Obstet Gynaecol 1994;101:35–43. doi: 10.1111/j.1471-0528.1994.tb13007.x.

32. Shenzhen Basics. Shenzhen Government Online, 2021. Available from: http://www.sz.gov.cn/en_szgov/aboutsz/profile/content/post_1357629.html. [Accessed on March 15th, 2022]

33. Shenzhen, Wikipedia. 2021. Available from: https://en.wikipedia.ahau.cf/wiki/Shenzhen#cite–23–191. [Accessed on March 15, 2022]

34. Merzadi M, Caulfield LE, Zavaleta N, Figueroa A, Costigan KA, Dominici F, et al. Fetal growth in Peru: comparisons with international fetal size charts and implications for fetal growth assessment. Ultrasound Obstet Gynecol 2005;26:123–128. doi: 10.1002/uog.1954.

35. Chitty LS, Altman DG, Henderson A, Campbell S. Charts of fetal size: 4. Femur length. Br J Obstet Gynaecol 1994;101:132–135. doi: 10.1111/j.1471-0528.1994.tb13078.x.

36. Chitty LS, Altman DG, Henderson A, Campbell S. Charts of fetal size: 3. Abdominal measurements. Br J Obstet Gynaecol 1994;101:125–131. doi: 10.1111/j.1471-0528.1994.tb13077.x.

37. Kwon JY, Park JY, Wae JH, Cho S, Kim CJ, Shin JC. Fetal biometry in the Korean population: reference charts and comparison with charts from other populations. Prenat Diagn 2014;34:927–934. doi: 10.1002/pd.4394.

38. Stinnermann JJ, Fries N, Bessis R, Fontanges M, Mangione R, Salomon LJ. Implementing the INTERGROWTH-21(st) fetal growth standards in France: a ‘flash study’ of the College Français d’Echographie Foetale (CFEF). Ultrasound Obstet Gynecol 2017;49:487–492. doi: 10.1002/uog.2723.

39. van de Kamp K, Pajket E, Zwenderman AH, van der Post JA, Snijders RJM. Validation of reference charts for mid-trimester fetal biometry. Fetal Diagn Ther 2019;45:42–49. doi: 10.1159/000486094.

40. Sletner L, Kiserud T, Vangen S, Nakstad B, Jenum AK. Effects of applying universal fetal growth standards in a Scandinavian multi-ethnic population. Acta Obstet Gynecol Scand 2018;97:168–179. doi: 10.1111/aogs.13269.

41. Bhandari N, Bahl R, Taneja S, de Onis M, Bhan MK. Growth performance of affluent Indian children is similar to that in developed countries. Bull World Health Organ 2002;80:189–195. doi: 10.1590/S0042-96862002000300004.

42. Stevens GA, Finucane MM, Paciorek CJ, Flaxman SR, White RA, Donner AJ, et al. Trends in mild, moderate, and severe stunting and underweight, and progress towards MDG 1 in 141 developing countries: a systematic analysis of population representative data. Lancet 2012;380:824–834. doi: 10.1016/S0140-6736(12)60647-3.

43. 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;218:509.e1–509.e7. doi: 10.1016/j.ajog.2015.10.931.

44. 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;214:509.e1–509.e7. doi: 10.1016/j.ajog.2015.10.931.

45. Number of Births Per Year in China From 2011 to 2021. Statista, 2022. Available from: https://www.statista.com/statistics/250650/number-of-births-in-china/. [Accessed on March 20th,2022]

46. Maternal, Child Health Not Impacted by COVID-19: Health Authorities. China NHCorPdR, 2020. Available from: http://en.nhc.gov.cn/gw2020/2020-07/13/c_81053.htm. [Accessed on March 20th, 2022]