Fetal cerebellar growth and Sylvian fissure maturation: international standards from Fetal Growth Longitudinal Study of INTERGROWTH-21st Project

M. J. RODRIGUEZ-SIBAJA1,2,3, J. VILLAR1,2, E. O. OHUMA1,2,4, R. NAPOLITANO1, S. HEYL1, M. CARVALHO5, Y. A. JAFFER6, J. A. NOBLE7, M. OBERTO8, M. PURWAR9, R. PANG10, L. CHEIKH ISMAIL11,1, A. LAMBERT1,2, M. G. GRAVETT12, L. J. SALOMON13, L. DRUKKER1,2, F. C. BARROS14, S. H. KENNEDY1,2, Z. A. BHUTTA15 and A. T. PAPAGEORGHIOU1,2

1Nuffield Department of Women’s & Reproductive Health, University of Oxford, Oxford, UK; 2Oxford Maternal & Perinatal Health Institute, Green Templeton College, University of Oxford, Oxford, UK; 3Maternal-Fetal Medicine Department, National Institute of Perinatology, Mexico City, Mexico; 4Centre for Statistics in Medicine, Botnar Research Centre, University of Oxford, Oxford, UK; 5Faculty of Health Sciences, Aga Khan University, Nairobi, Kenya; 6Department of Family & Community Health, Ministry of Health, Muscat, Sultanate of Oman; 7Department of Engineering Science, Institute of Biomedical Engineering, University of Oxford, Oxford, UK; 8S.C. Ostitetrica 2U, Città della Salute e della Scienza di Torino, Turin, Italy; 9Nagpur INTERGROWTH-21st Research Centre, Kastur Hospital, Nagpur, India; 10School of Public Health, Peking University, Beijing, China; 11Clinical Nutrition and Dietetics Department, University of Sharjah, Sharjah, United Arab Emirates; 12Departments of Obstetrics & Gynecology and of Public Health, University of Washington, Seattle, WA, USA; 13Department of Obstetrics and Fetal Medicine, Hôpital Necker Enfants Malades, Université Paris Descartes, Paris, France; 14Programa de Pós-Graduação em Saúde e Comportamento, Universidade Católica de Pelotas, Pelotas, Brazil; 15Center for Global Child Health, Hospital for Sick Children, Toronto, Canada

KEYWORDS: cerebellum; cerebral cortex; cortical development; cortical maturation; fetal brain; fissures; gyration; neurodevelopment; neurosonography; operculization; ultrasound

CONTRIBUTION

What are the novel findings of this work?
When constraints on human growth are minimal, the growth and developmental patterns of the fetal cerebellum and Sylvian fissures are similar across diverse populations and there are no sex-related differences.

What are the clinical implications of this work?
We present international standards for fetal cerebellar growth and Sylvian fissure maturation using three-dimensional ultrasound volumes obtained in healthy, well nourished pregnant women from five diverse urban areas worldwide who were enrolled in the Fetal Growth Longitudinal Study of the INTERGROWTH-21st Project. Their babies had low morbidity and adequate growth and development at 2 years of age.

ABSTRACT

Objective To construct international ultrasound-based standards for fetal cerebellar growth and Sylvian fissure maturation.

Methods Healthy, well nourished pregnant women, enrolled at <14 weeks’ gestation in the Fetal Growth Longitudinal Study (FGLS) of INTERGROWTH-21st, an international multicenter, population-based project, underwent serial three-dimensional (3D) fetal ultrasound scans every 5 ± 1 weeks until delivery in study sites located in Brazil, India, Italy, Kenya and the UK. In the present analysis, only those fetuses that underwent developmental assessment at 2 years of age were included. We measured the transcerebellar diameter and assessed Sylvian fissure maturation using two-dimensional ultrasound images extracted from available 3D fetal head volumes. The appropriateness of pooling data from the five sites was assessed using variance component analysis and standardized site differences. For each Sylvian fissure maturation score (left or right side), mean gestational age and 95% CI were calculated. Transcerebellar diameter was modeled using fractional polynomial regression, and goodness of fit was assessed.

Results Of those children in the original FGLS cohort who had developmental assessment at 2 years of age, 1130 also had an available 3D ultrasound fetal

Correspondence to: Prof. A. T. Papageorghiou, Nuffield Department of Women’s & Reproductive Health, University of Oxford, The Women’s Centre, John Radcliffe Hospital, Oxford, OX3 9DU, UK (e-mail: aris.papageorghiou@wrh.ox.ac.uk)
Accepted: 7 March 2020

© 2020 The Authors. Ultrasound in Obstetrics & Gynecology published by John Wiley & Sons Ltd on behalf of International Society of Ultrasound in Obstetrics and Gynecology.
head volume. The sociodemographic characteristics and pregnancy/perinatal outcomes of the study sample confirmed the health and low-risk status of the population studied. In addition, the fetuses had low morbidity and adequate growth and development at 2 years of age. In total, 3016 and 2339 individual volumes were available for transcerebellar-diameter and Sylvian-fissure analysis, respectively. Variance component analysis and standardized site differences showed that the five study populations were sufficiently similar on the basis of predefined criteria for the data to be pooled to produce international standards. A second-degree fractional polynomial provided the best fit for modeling transcerebellar diameter; we then estimated gestational-age-specific 3rd, 50th and 97th smoothed centiles. Goodness-of-fit analysis comparing empirical centiles with smoothed centile curves showed good agreement. The Sylvian fissure increased in maturation with advancing gestation, with complete overlap of the mean gestational age and 95% CIs between the sexes for each development score. No differences in Sylvian fissure maturation between the right and left hemispheres were observed.

Conclusion We present, for the first time, international standards for fetal cerebellar growth and Sylvian fissure maturation throughout pregnancy based on a healthy fetal population that exhibited adequate growth and development at 2 years of age. © 2020 The Authors. Ultrasound in Obstetrics & Gynecology published by John Wiley & Sons Ltd on behalf of International Society of Ultrasound in Obstetrics and Gynecology.

INTRODUCTION

The fetal central nervous system undergoes extraordinary transformation throughout pregnancy. The cerebellum and brain cortex are major landmarks of this complex yet highly organized neurodevelopmental process.

The cerebellum, which is associated later in life with sensorimotor, cognitive and affective regulation, can be identified during fetal life by ultrasonography as early as 12 weeks’ gestation. In clinical practice, ultrasonography is used to evaluate the anatomical integrity of the fetal cerebellum and linear growth of the transcerebellar diameter (TCD). Measuring the TCD also enables estimation of gestational age in cases of uncertain dates in the third trimester, and may be more suitable when there are fetal growth disturbances, because cerebellar growth is less affected by placental insufficiency due to the ‘brain-sparing’ phenomenon.

The fetal brain cortex displays remarkable gestational-age-specific maturation. A leading marker of these processes, which follow a predictable timetable, is the development (operculization) of the Sylvian fissure (SF) on the lateral convexities of the cerebral hemispheres.

There are several reference ranges for TCD measurements and SF maturation. Many of the studies providing these data have a high risk of methodological bias owing to sample selection, study design, data analysis and lack of ultrasound quality control. These limitations most probably explain the reported variation in the range of values across pregnancy, which makes clinical interpretation difficult. In addition, and importantly, none of the fetal studies continued to assess neurodevelopment into early childhood, which would seem a logical requirement for any tool proposed to evaluate fetal normality.

The Fetal Growth Longitudinal Study (FGLS) of the INTERGROWTH-21st Project has produced international standards, based on World Health Organization (WHO) recommendations, for early and late pregnancy dating, fetal growth and estimated fetal weight, as well as other aspects of pregnancy care. To complement these clinical tools, we aimed to produce international standards for longitudinal TCD growth and SF maturation from the same population of healthy pregnant women contributing data to the FGLS whose babies had adequate growth and development from early fetal life to 2 years of age.

METHODS

Study population

INTERGROWTH-21st is an international, multicenter, population-based project (www.intergrowth21.org.uk). Phase I of the INTERGROWTH-21st Project, conducted between 2009 and 2016, consisted of nine complementary studies designed to describe optimal human growth and development, based conceptually on the WHO prescriptive approach. The FGLS, one of the main studies of the INTERGROWTH-21st Project, enrolled, before 14 weeks’ gestation, a large cohort of healthy, well nourished women with a naturally conceived singleton pregnancy who met rigorous individual inclusion criteria, and whose babies were monitored until 2 years of age, in order to generate international standards.

The INTERGROWTH-21st Project methodology has been described elsewhere in detail. Briefly, participants were first selected at the population level and then at the individual level. At the population level, urban areas (a complete city or county, or part of a city with clear political or geographical limits) in which most deliveries occurred in healthcare facilities were identified. The areas had to be located at an altitude of < 1600 m, with a low risk of fetal growth disturbances, as well as have an absence or low level of major, known, non-microbiological contamination, such as pollution, domestic smoke, radiation or any other toxic abuse. Within each area, all institutions providing pregnancy and neonatal care in which more than 80% of births occurred were selected. From these populations, women were selected at an individual level if they had no clinically relevant obstetric or medical history, initiated antenatal care before 14 weeks and met the entry criteria of optimal health, nutrition, education and socioeconomic status.
The women underwent serial fetal ultrasound scans every 5 weeks (±1 week) until 41 + 6 weeks in eight urban areas worldwide that were geographically delimited to ensure that the study was population based. At each visit, a set of two-dimensional (2D) images and three-dimensional (3D) volumes of fetal biometric parts was obtained and stored digitally. Gestational age was based on the last menstrual period (LMP) provided that its date was certain, the woman had regular 24–32 day cycles, she had not been using hormonal contraception or breastfeeding in the preceding 2 months, and any discrepancy between gestational age based on LMP and gestational age based on crown–rump length, measured by ultrasound at 9 + 0 to 13 + 6 weeks after the LMP, was 7 days or less.

All ultrasound scans were performed by sonographers who were trained in a standardized way and audited regularly according to FGLS requirements. The same ultrasound equipment was used at all study sites (Philips HD-9; Philips Ultrasound, USA), with curvilinear abdominal 2D transducers C5-2, C6-3 and one curvilinear abdominal 3D transducer V7-3, and was specially adapted to ensure that measurement values were not visible on the screen in order to reduce ‘expected value’ bias.

For the present analysis, we included only FGLS participants whose children had neurodevelopmental, nutritional and morbidity assessments at 2 years ±2 months of age, in five of the original eight study sites (Pelotas (Brazil); Turin (Italy); Oxford (UK); the central area of Nagpur (India); and the Parklands suburb of Nairobi (Kenya)). Three sites, in China, Oman and the USA, did not participate in the early childhood development assessments for logistical and administrative reasons pertaining to the timing of the start of the study and/or staff availability.

The INTERGROWTH-21st Project was approved by the Oxfordshire Research Ethics Committee “C” (ref: 08/H0606/139), the research ethics committees of the individual participating institutions, and the corresponding regional health authorities in which the project was implemented. Participants provided written consent to be involved in the project.

Volume acquisition, offline analysis and measurement methodology

TCD measurement and SF assessment were performed using still images, extracted from the available 3D fetal head volumes acquired at the five study sites. Head volumes were acquired at the level of the axial, transthalamic plane. Six predefined quality-control criteria of the 2D images had to be satisfied to acquire the volume: oval shape, symmetrical plane, thalami and cavum septi pellucidi (CSP) visible, cerebellum not visible and head occupying at least 30% of the image.

The acquisition was undertaken with the volume data box and angle of sweep (usually 70°) adjusted to include the entire skull, during fetal quiescence, with the mother holding her breath and with the transducer held steady. The real-time image was observed during acquisition to confirm that the sweep included the entire skull with no maternal or fetal movement during the sweep; otherwise, the volume was discarded and the acquisition repeated. All data were then transferred electronically to the Ultrasound Coordinating Unit in Oxford. Further details of the methodology for volume acquisition are available at https://intergrowth21.org.uk (follow the link to ‘Study Protocol’ to download the ultrasound manual).

Offline image analysis for plane reconstruction and measurements was carried out using the open-source image analysis software program MITK (Medical Imaging Interaction Toolkit MITK, version 0.12.2; German Cancer Research Center, Division of Medical Image Computing, www.mitk.org). All measurements were undertaken by one experienced fetal medicine specialist at the Coordinating Unit in Oxford, who was standardized in ultrasound volume manipulation and who was blinded to the clinical details and gestational age.

The TCD was measured in the standard transcerebellar plane, while SF maturation was assessed in the transthalamic plane. First, each stored volume of the fetal head was uploaded onto the multiplanar mode facility. Second, starting from this plane, rotation or scrolling of the volume in orthogonal planes was undertaken with the fulcrum of rotation primarily in the middle of the CSP. As the transthalamic plane was the plane of volume acquisition, it required minimal manipulation to extract the 2D plane. The transcerebellar 2D plane was extracted at the level of the transthalamic plane with a slight posterior tilting and visualization of the frontal horns of the lateral ventricles, CSP, thalami, cerebellum and cisterna magna. Once the appropriate transcerebellar 2D plane had been extracted, the TCD was measured perpendicular to the midline echo (falx cerebri), with the calipers placed ‘outer to outer’ between the distal margins of the hemispheres at the largest transverse diameter of the cerebellum.

Assessment of the SF was performed in the brain hemisphere distal to the probe to prevent shadowing from the fetal skull bones, with a focus on the angle changes between the insula and the temporal lobe. We used a simple, unweighted, scoring system, ranging from Grade 0 (no development) to Grade 5 (maximum development), employed previously for magnetic resonance imaging and ultrasound analysis of cortical maturation. The hemisphere in which the SF was measured (right or left) was identified by combining fetal presentation (cephalic or breech) with head direction at the time of measurement; when the presentation was transverse or oblique, the hemisphere was not determined.

To ensure that the best possible images were obtained for each extracted plane, we used a scoring system to evaluate and grade image quality (1, impossible to assess accurately; 2, possible to assess accurately; 3, good; 4, almost perfect). Only 3D volumes that scored 2–4 were included in the analysis.
Reproducibility

Formal assessment of interobserver reproducibility for plane reconstruction and TCD acquisition was undertaken following the INTERGROWTH-21st quality control strategy\(^{35,46}\) in a randomly selected subset of 132 fetuses (12%). From this subset, a single head volume of each fetus was selected randomly and assessed by two fetal medicine specialists. Both observers independently uploaded each volume, extracted the transcerebellar plane and measured the TCD blinded to all measurements obtained (including their own).

Statistical analysis

Outliers, defined as measurements > 5 SD above the mean at each gestational age, were excluded. To assess the possibility of pooling data across sites, we used two complementary analytical strategies\(^{36}\). Firstly, variance component analysis was used to calculate the percentage of total variance due to between-site variance, as well as an estimation of the percentage of total variance for individuals within each site. Secondly, for each site, at five specific gestational-age windows, we calculated the difference between each site’s mean and the mean of all sites together. Each difference was then expressed as a proportion of all the sites’ SD, i.e. the SD of the data pooled across all sites, at each corresponding gestational age, to give the standardized site difference (SSD). The SSD is similar to the Z-score and is expressed in units of all the sites’ SD (i.e. 1.0 standardized difference = 1.0 of all the sites’ SD). The SSD allows for direct comparisons of biometric measurements in populations across pregnancy, standardized by the corresponding pooled SD. A pattern of SSD values < 0.5 was prespecified in the FGLS protocol, in keeping with WHO recommendations, as an adequate cut-off value for combining data from all sites\(^{36}\).

The distribution of fetal cerebellar measurements was assessed for normality, conditional on gestational age. We then modeled TCD as a function of gestational age using fractional polynomial regression and obtained the fitted centiles\(^{47}\). For the SF maturation analysis, we calculated the mean gestational age and 95% CI for each development score. Goodness of fit of the resultant models was assessed as described previously by Ohuma & Altman for the INTERGROWTH-21st data\(^{48}\), i.e. a plot of the residuals against fitted values, distribution of fitted quantile–quantile (Q–Q) plots of the residuals, plots of residuals against gestational age.

In addition, model fit was assessed visually using quantile–quantile (Q–Q) plots of the residuals, plots of residuals against fitted values, distribution of fitted Z-scores against gestational age and a comparison of the estimated proportions of observations falling below the 3rd centile or above the 97th centile to the expected proportions of 3%.

Measurement reproducibility was assessed using Bland–Altman analysis of the interobserver differences and their 95% limits of agreement\(^{49,50}\). We did not consider the intraclass correlation coefficient (ICC) to be appropriate for the reproducibility study since it depends on the range of the measurement values. Consequently, rather than having fixed values, ICCs vary according to the range of gestational ages being studied\(^{51}\). Therefore, we instead used the method proposed by Bland and Altman, which has been shown to be more appropriate for assessing the repeatability of two measurements\(^{49}\). Analysis was performed using Stata version 15 (StataCorp., College Station, TX, USA).

RESULTS

Of the children in the original FGLS cohort\(^{31}\) who had a developmental assessment at 2 years of age (\(n = 1339\))\(^{38}\), 1130 (84%) also had an available 3D ultrasound fetal head volume. The proportional contribution of cases from the study sites was 26% for India (\(n = 291\)), 25% for Kenya (\(n = 280\)), 23% for Italy (\(n = 262\)), 14% for Brazil (\(n = 156\)) and 12% for the UK (\(n = 141\)).

The sociodemographic characteristics and pregnancy/perinatal outcomes of the study sample are presented in Table 1. They are similar to those of the total FGLS cohort\(^{31}\), and confirm the health and low-risk status of the population studied. In addition, we provide evidence that the fetuses whose brain development is the subject of this analysis had low morbidity and adequate growth and development score. Goodness of fit of the resultant models was assessed visually using Bland–Altman analysis of the interobserver differences and their 95% limits of agreement\(^{49,50}\). We did not consider the intraclass correlation coefficient (ICC) to be appropriate for the reproducibility study since it depends on the range of the measurement values. Consequently, rather than having fixed values, ICCs vary according to the range of gestational ages being studied\(^{51}\). Therefore, we instead used the method proposed by Bland and Altman, which has been shown to be more appropriate for assessing the repeatability of two measurements\(^{49}\). Analysis was performed using Stata version 15 (StataCorp., College Station, TX, USA).

Table 1 Maternal and perinatal characteristics of pregnancies from INTERGROWTH-21st Fetal Growth Longitudinal Study (FGLS) included in present analysis and full FGLS cohort\(^{31}\)

| Characteristic                           | Current study (\(n = 1130\)) | FGLS (\(n = 4321\)) |
|------------------------------------------|------------------------------|---------------------|
| Maternal age (years)                     | 28.8 ± 3.9                   | 28.4 ± 3.9          |
| Maternal height (cm)                     | 162.1 ± 5.8                  | 162.2 ± 5.8         |
| Maternal weight (kg)                     | 61.2 ± 9.3                   | 61.3 ± 9.1          |
| Paternal height (cm)                     | 172.5 ± 7.9                  | 174.4 ± 7.3         |
| Maternal body mass index (kg/m²)         | 23.3 ± 3.0                   | 23.3 ± 3.0          |
| Gestational age at first visit (weeks)   | 11.6 ± 1.3                   | 11.8 ± 1.4          |
| Time in formal education (years)         | 15.2 ± 2.9                   | 15.0 ± 2.8          |
| Hemoglobin level at < 15 weeks (g/L)†    | 123.6 ± 9.8                  | 125 ± 11            |
| Nulliparous                              | 657 (58.1)                   | 2955 (68.4)         |
| Pre-eclampsia                            | 11 (1.0)                     | 31 (0.7)            |
| Pylonephritis                            | 6 (0.5)                      | 16 (0.4)            |
| Any sexually transmitted infection       | 0 (0.0)                      | 3 (0.1)             |
| Spontaneous initiation of labor          | 754 (66.7)                   | 2868 (66.4)         |
| PPROM (< 37 weeks)                       | 71 (6.3)                     | 80 (1.9)            |
| Cesarean section                         | 275 (24.3)                   | 1541 (35.7)         |
| NICU admission (> 1 day)                 | 29 (2.6)                     | 240 (5.6)           |
| Term low birth weight (< 2500 g)         | 42 (3.7)                     | 128 (3.0)           |
| Neonatal mortality                       | 0 (0.0)                      | 7 (0.2)             |
| Neonatal male sex                        | 553 (48.9)                   | 2149 (49.7)         |
| Exclusive breastfeeding at discharge     | 1048 (92.7)                  | 3786 (87.6)         |
| Mother admitted to intensive care unit   | 2 (0.2)                      | 17 (0.4)            |
| Birth weight (< 37 weeks) (kg)           | 3.2 ± 0.4                    | 3.3 ± 0.4           |
| Birth length (< 37 weeks) (cm)           | 49.2 ± 1.8                   | 49.4 ± 1.9          |
| Birth HC (< 37 weeks) (cm)               | 34.0 ± 1.3                   | 33.9 ± 1.3          |

Data given as mean±SD or n (%). Data missing for < 1% of sample for all variables, unless indicated otherwise. *Data missing for 45% of sample. †Data missing for 32% of sample. HC, head circumference; NICU, neonatal intensive care unit; PPROM, preterm prelabor rupture of membranes.
development at 2 years of age (Figure 1 and Tables 2 and 3). The median number of ultrasound scans per woman was five (range, 1–6); 90% had four or more scans, indicating good adherence to the study protocol. The total number of 3D volumes available was 5746; however, 2730 (48%) volumes could not be assessed. Most of these were obtained after 32 weeks’ gestation when visualization and assessment of brain structures are hampered by acoustic shadowing of the calcified fetal skull, fetal head position in the maternal pelvis and a reduction in amniotic fluid volume. Other reasons for images of limited quality were fetal movement artifact not evident during the original scan, acoustic shadows from proximal structures, reverberation artifacts and unfavorable fetal head orientation. After these exclusions and removal of 11 outliers, 3016 and 2359 volumes from 1130 fetuses were available for TCD and SF analysis, respectively.

Interobserver reproducibility for plane reconstruction and TCD measurement was assessed in a randomized subset of volumes from 132 fetuses (12% of the total sample of 1130 fetuses), across a gestational-age range from 15 + 0 weeks to 36 + 2 weeks. TCD could be measured in 108 (82%) of the 132 volumes. Bland–Altman plots were used to present graphically the interobserver differences and their 95% limits of agreement (Figure 2). The mean interobserver difference was very close to zero (0.07 mm); the inferior and superior limits of agreement (± 2 SD) were −2.40 mm and + 2.55 mm, respectively, suggesting very close agreement. No evidence of consistent bias was seen across the range of measurements.

For both structures assessed, within-site variance was greater than between-site variance. For TCD, within-site variance was 5.55 mm (interobserver) and 2.55 mm (intraobserver). For interobserver and intraobserver reproducibility of plane reconstruction, the mean differences were very close to zero (−0.07 and 0.04 mm, respectively). Bland–Altman plots of interobserver and intraobserver differences for plane reconstruction showed very close agreement (Figure 2).

Table 3 Morbidity at 1 and 2 years of age in children who contributed data to fetal transcerebellar diameter and Sylvian fissure maturation international standards

| Variable | 1 year of age | 2 years of age |
|----------|--------------|---------------|
| Hospitalized at least once | 123 (11.0) | 100 (9.0) |
| Total no. of days hospitalized | 2 (1–4) | 2 (1–3) |
| Any prescription made by healthcare professional | 704 (63.1) | 660 (59.2) |
| Antibiotics (≥ 3 regimens) | 76 (6.8) | 132 (11.8) |
| Iron/tolic acid/vitamin B12/other vitamin supplementation | 452 (40.5) | 183 (16.4) |
| Up-to-date with local vaccination policies | 1062 (95.2) | 1051 (94.3) |
| Otitis media/pneumonia/bronchiolitis | 66 (5.9) | 84 (7.5) |
| Parotitis/diarrhea/vomiting | 42 (3.8) | 42 (3.8) |
| Seizures/cerebral palsy/neurological disorder | 0 (0.0) | 0 (0.0) |
| Exanthema/skin disease | 209 (18.7) | 125 (11.2) |
| UTI/pyelonephritis | 2 (0.2) | 6 (0.5) |
| Fever ≥ 3 days (≥ 3 episodes) | 111 (10.0) | 128 (11.5) |
| Malaria | 0 (0.0) | 0 (0.0) |
| Meningitis | 1 (0.1) | 0 (0.0) |
| Other infection that required antibiotics | 16 (1.4) | 31 (2.8) |
| Hearing problem | 0 (0.0) | 0 (0.0) |
| Asthma | 12 (1.1) | 13 (1.2) |
| Cardiovascular problem | 0 (0.0) | 0 (0.0) |
| Blindness | 0 (0.0) | 0 (0.0) |
| Gastroesophageal reflux | 49 (4.4) | 2 (0.2) |
| Any hemolytic condition | 0 (0.0) | 0 (0.0) |
| Any malignancy | 0 (0.0) | 0 (0.0) |
| Cow's milk protein allergy | NA | 8 (0.7) |
| Food allergy | NA | 13 (1.2) |
| Injury trauma | 13 (1.2) | 26 (2.3) |
| Any condition that required surgery | 14 (1.3) | 8 (0.7) |

Data given as n (%) or median (interquartile range). NA, not applicable (data were not collected at 1-year follow-up visit); UTI, urinary tract infection.

Table 2 Anthropometric measures at 2 years of age in children who contributed data to fetal transcerebellar diameter and Sylvian fissure maturation international standards, compared with World Health Organization (WHO) Child Growth Standards

| Variable         | Current study | WHO Child Growth Standards |
|------------------|---------------|----------------------------|
|                  | n  | Mean ± SD | Median (IQR) | Mean Z-score ± SD | Median percentile |
| Weight (kg)      | 1095 | 12.3 ± 1.7 | 12.2 (11.2–13.3) | 0.2 ± 1.1 | 58.7 |
| Length (cm)      | 1088 | 86.8 ± 3.3 | 86.7 (84.6–89.1) | −0.2 ± 1.0 | 41.7 |
| Head circumference (cm) | 1098 | 47.9 ± 1.6 | 47.9 (46.8–49.1) | 0.2 ± 1.1 | 55.6 |

Age- and gender-specific Z-scores and centiles. †Mean values were estimated from raw data. IQR, interquartile range.
variance was 0.33 (10.9% of the total variance), while between-site variance was 0.005 (0.2% of the total variance). For SF maturation, the within-site variance estimate was 0.007 (3.2% of the total variance), while between-site variance was 0.005 (2.3% of the total variance) (Table 4).

SSD according to gestational age for the five sites was expressed as a proportion of the SD of all the sites at each gestational-age interval. All SSDs for TCD and SF score were below 0.5 SD for all five fetal gestational-age windows (Figure S2). The results of these analyses show that the five study populations are sufficiently similar, according to WHO predefined criteria53, for the data to be pooled to produce international standards.

All cerebellar measurements were distributed normally, conditional on gestational age. The best fitting powers were provided by a second-degree fractional polynomial. The gestational-age-specific 3rd, 50th and 97th smoothed centiles for TCD are shown in Figure 3a, and the gestational-age-specific 3rd, 5th, 10th, 50th, 90th, 95th and 97th smoothed centiles are presented in Table S1. For clinical purposes, we also present the same fitted centiles of gestational-age estimation based on measurement of TCD (Figure 3b, Table S2).

Assessment of goodness of fit of the TCD model by gestational-age-specific comparisons of empirical centiles to smoothed centile curves showed good agreement, and scatterplots of Z-scores by gestational age did not show any patterns (data not shown). The equations for the mean and SD from the fractional polynomial models for TCD are presented in Table S3, which allow the reader to calculate any desired centiles according to gestational age in exact weeks. The actual values for these centiles according to gestational age are presented in Table S1. Striking overlap of TCD values between male and female fetuses was seen (Figure S3).

The SF increased in maturation with advancing gestation, and there was no appreciable difference between study sites, with complete overlap of the mean gestational
In this study, we present international standards for fetal cerebellar growth and SF maturation based on data from a large, longitudinal sample, obtained under rigorously controlled conditions, from well nourished women living in environments with minimal constraints on fetal growth, across five geographically diverse urban areas worldwide. In addition, and unique to the ultrasound literature, we provide follow-up evidence that the fetuses that contributed data to these standards had low morbidity and adequate growth and development at 2 years of age.

Variance component analysis showed that only 0.2% and 2.3% of the total variability in fetal cerebellar growth and SF maturation, respectively, could be attributed to between-site differences. These results are compatible with the 1.9–3.5% variability between sites reported for fetal skeletal growth and newborn length in the FGLS, the 3% variability reported for infant length in the WHO Multicentre Growth Reference Study, and the 1.3–9.2% variability reported for skeletal growth and neurodevelopmental milestones at 2 years of age in the FGLS follow-up study.

Our results refute suggestions that the observed variability in these measurements, between unselected samples, is related to genetic differences. Similar suggestions have been made previously about human growth. However, there is now consistent evidence that the variability in human skeletal growth within a population is seven times larger than that between populations (genetic variability), which represents less than 10% of the total variance.

Our results also do not support the recently published suggestion of in-utero sexual dimorphism in the development of brain structure and function. Specifically, male fetuses have been documented to have larger cerebellar gray-matter volume, as well as greater intracerebellar functional connectivity. However, we did not find any sex-related differences in the pattern of growth or maturation of the studied brain structures. In addition, we found no differences between the right and left SF in this in-utero population; in the mature brain, the SF follows a steeper trajectory in the right hemisphere, while it extends further posteriorly and is longer (in horizontal length) on the left side. It is possible that these differences are not evident in utero or in the axial ultrasound planes.

The international standards presented here demonstrate a more than 2-fold increase in TCD during the second half of pregnancy (Figure 3). The rapid growth of the cerebellum within a relatively narrow time period suggests that TCD measurement may facilitate gestational age estimation during the second and early third trimesters. This may also apply in suspected fetal growth restriction, as the brain-sparing phenomenon may protect cerebellar growth and close to term when the head may be more difficult to measure. In our study, TCD predicted gestational age well, and the prediction intervals of ±7 days at 20 weeks and ±10 days at 32 weeks compare favorably with those obtained using head circumference. Further work will be needed to assess robustness in pregnancies with abnormal growth.

Longitudinal evaluation of SF maturation shows a characteristic pattern. We estimated the mean gestational age and 95% CI at which each SF maturation categorical score is expected. Hence, assessment of the progress of SF maturation may be a useful adjunct to fetal brain examination, especially if a brain abnormality is suspected.

**DISCUSSION**

In this study, we present international standards for fetal cerebellar growth and SF maturation based on data from a large, longitudinal sample, obtained under rigorously controlled conditions, from well nourished women living in environments with minimal constraints on fetal growth, across five geographically diverse urban areas worldwide. In addition, and unique to the ultrasound literature, we provide follow-up evidence that the fetuses that contributed data to these standards had low morbidity and adequate growth and development at 2 years of age.

Variance component analysis showed that only 0.2% and 2.3% of the total variability in fetal cerebellar growth and SF maturation, respectively, could be attributed to between-site differences. These results are compatible with the 1.9–3.5% variability between sites reported for fetal skeletal growth and newborn length in the FGLS, the 3% variability reported for infant length in the WHO Multicentre Growth Reference Study, and the 1.3–9.2% variability reported for skeletal growth and neurodevelopmental milestones at 2 years of age in the FGLS follow-up study.

Our results refute suggestions that the observed variability in these measurements, between unselected samples, is related to genetic differences. Similar suggestions have been made previously about human growth. However, there is now consistent evidence that the variability in human skeletal growth within a population is seven times larger than that between populations (genetic variability), which represents less than 10% of the total variance.

Our results also do not support the recently published suggestion of in-utero sexual dimorphism in the development of brain structure and function. Specifically, male fetuses have been documented to have larger cerebellar gray-matter volume, as well as greater intracerebellar functional connectivity. However, we did not find any sex-related differences in the pattern of growth or maturation of the studied brain structures. In addition, we found no differences between the right and left SF in this in-utero population; in the mature brain, the SF follows a steeper trajectory in the right hemisphere, while it extends further posteriorly and is longer (in horizontal length) on the left side. It is possible that these differences are not evident in utero or in the axial ultrasound planes.

The international standards presented here demonstrate a more than 2-fold increase in TCD during the second half of pregnancy (Figure 3). The rapid growth of the cerebellum within a relatively narrow time period suggests that TCD measurement may facilitate gestational age estimation during the second and early third trimesters. This may also apply in suspected fetal growth restriction, as the brain-sparing phenomenon may protect cerebellar growth and close to term when the head may be more difficult to measure. In our study, TCD predicted gestational age well, and the prediction intervals of ±7 days at 20 weeks and ±10 days at 32 weeks compare favorably with those obtained using head circumference. Further work will be needed to assess robustness in pregnancies with abnormal growth.

Longitudinal evaluation of SF maturation shows a characteristic pattern. We estimated the mean gestational age and 95% CI at which each SF maturation categorical score is expected. Hence, assessment of the progress of SF maturation may be a useful adjunct to fetal brain examination, especially if a brain abnormality is suspected.
We chose to describe and quantify ultrasound patterns of SF maturation for the following reasons: first, the SF is the first primary fissure to be evident sonographically (at 17–19 weeks’ gestation), which means that it can be assessed readily at the time of the mid-trimester fetal scan9,52; second, its development follows a predictable timetable during pregnancy, which makes it easier to incorporate SF assessment into routine clinical care11,52; third, SF assessment is feasible using the standard 2D ultrasound plane used for routine fetal head biometry, facilitating examination without specialist training in neurosonography or transvaginal scanning, and without increasing the total scanning time11,23; and, fourth, a simple SF scoring system is available22. More complex scoring systems or measurements may then be used for detailed neurosonographic examination1,2,3,4,26.

A number of studies have reported TCD4,6,9,13–22 and SF maturation9,22–24 reference ranges; however, existing charts have several limitations. A recent systematic review found substantial heterogeneity in the methodological quality of studies aimed at developing fetal brain structure charts, which can lead to significant variability in the interpretation of ultrasound measurements. None of the studies had a low risk of bias for sample selection, inclusion/exclusion criteria, quality control, neonatal/infant outcomes or neurological follow-up; in addition, goodness of fit of the proposed model was reported in fewer than 35% of the studies27. We therefore aimed to avoid these limitations by producing international standards to complement those from the INTERGROWTH-21 st Project that are already being used clinically and for research purposes29–34,63,64. The 3D ultrasound volumes were taken specifically for this purpose during the FGLS, using rigorous methods, identical ultrasound equipment, blinding of operators and a detailed quality-control strategy28,31. The images were extracted from the volumes using standardized axial planes recommended for routine clinical practice2, and were measured and scored in accordance with a predefined protocol40. Our reproducibility study suggested a high level of agreement in fetal imaging. Furthermore, and crucially, the cohort was followed up to 2 years of age and developmental outcomes confirmed their eligibility for the construction of international standards38,39.

It could be said that measurements or scores acquired on planes extracted from 3D volumes do not completely accord with 2D measurements obtained in real time37,42,65. However, we believe that this is unlikely because, although volumetry is associated with a degree of variability if not standardized, once rigorous methodology is adopted, 2D assessments from reconstructed planes can be as reproducible as, and consistent with, 2D measurements obtained in real time52,37,42,52,65.

Conclusions

The growth and developmental patterns of the fetal brain structures we studied were similar across diverse geographical regions, and there were no differences between male and female fetuses. Hence, we pooled the data to produce international standards for TCD growth and SF maturation. We suggest that widespread implementation of the standards will enhance the clinical interpretation of fetal brain scans and standardize research findings.

ACKNOWLEDGMENTS

This study was funded by the INTERGROWTH-21 st grant 49 038 from the Bill & Melinda Gates Foundation to the University of Oxford; we gratefully acknowledge their support. A.T.P. is supported by the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre with funding from the NIHR Biomedical Research Centre (BRC) funding scheme. The views expressed herein are those of the authors and not necessarily those of the NHS, the NIHR, the Department of Health or any of the other funders. We are grateful to Professor Brenda Eskenazi for her comments regarding the Sylvian fissure analysis. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Disclosures

A.T.P. and J.A.N. are Senior Advisors for Intelligent Ultrasound. All other authors declare no competing interests.

REFERENCES

1. Stoodley CJ, Lamperopoulos C. Structure–function relationships in the developing cerebellum: Evidence from early-life cerebellar injury and neurodevelopmental disorders. Semin Fetal Neonatal Med 2016; 21: 356–364.
2. Chen X, Li S-L, Luo G-Y, Norwitz ER, Ouyang S-Y, Wen H-X, Yuan Y, Tian X-X, He J-M. Ultrasonographic Characteristics of Cortical Sulcus Development in the Human Fetus between 18 and 41 Weeks of Gestation. Obh Med J (Engl) 2017; 180: 920–928.
3. Koning IV, Tulemans MJ, Hoebreek SF, Ecart-Goosen GM, Reiss IKM, Steegers-Theunissen RPM, Dukdeik J. Impacts on prenatal development of the human cerebellum: A systematic review. J Matern Fetal Neonatal Med 2017; 30: 2461–2468.
4. Goldstein I, Reece EA, Pilo G, Bovwelli L, Hobbins JC. Cerebellar measurements with ultrasonography in the evaluation of fetal growth and development. Am J Obstet Gynecol 1987; 156: 1063–1069.
5. International Society of Ultrasound in Obstetrics & Gynecology Education Committee. Sonographic examination of the fetal central nervous system: guidelines for performing the ‘basic examination’ and the ‘fetal neurosonogram’. Ultrasound Obstet Gynecol 2007; 29: 109–116.
6. Chavez MR, Ananth CV, Smulian JC, Laskawy S, Kontopoulos EV, Vintzileos AM. Fetal transcerebellar diameter nomogram in singleton gestations with special emphasis in the third trimester: a comparison with previously published nomograms. Am J Obstet Gynecol 2003; 189: 1021–1025.
7. Reddy RH, Prashanth K, Ajit M. Significance of foetal transcerebellar diameter in singleton gestations with special emphasis in the third trimester: a comparison with previously published nomograms. Am J Obstet Gynecol 2003; 189: 1021–1025.
8. Quarello E, Stirnemann J, Ville Y, Guibaud L. Assessment of fetal Sylvian fissure development by multiplanar 3-dimensional ultrasound. J Ultrasound Med 2008; 25: 183–187.
9. Quarello E, Stirnemann J, Ville Y, Guibaud L. Assessment of fetal Sylvian fissure development by multiplanar 3-dimensional ultrasound. J Ultrasound Med 2008; 27: 494–502.
10. Cohen-Sachor R, Lerman-Sagie T, Lev D, Malinger G. Sonographic developmental milestones of the fetal cerebral cortex: a longitudinal study. Ultrasound Obstet Gynecol 2006; 27: 599–603.
11. Mittal P, Goncalves LF, Kusanovic JP, Espinoza J, Lee W, Nien JK, Soto E, Romero R. Objective evaluation of Sylvian fissure development by multiplanar 3-dimensional ultrasonography. J Ultrasound Med 2007; 26: 347–353.
12. Smith PA, Johansson D, Tzannatos C, Campbell S. Prenatal measurement of the fetal cerebellum and cisterna cerebellomedullaris by ultrasound. Chin Med J (Engl) 2001; 114: 1065–1069.
13. Singla M, Jain S, Budhiraja V, Ghal R, Babu CSR, Goel P. Transverse Cerebellar Diameter – A Marker for Estimation of Gestational Age. J Anat Soc India 2013; 59: 158–161.
20. Hata K, Hata T, Senoh D, Makihara K, Aoki S, Takamiya O, Kitao M.

30. Papageorghiou AT, Kemp B, Stones W, Ohuma EO, Kennedy SH, Purwar M, Lambert, A, Bertino E, Papageorghiou V, Gaza, C, Stein, A, Bhatta, Z, Kennedy SH. Neurodevelopmental milestones and associated behaviours are similar among healthy children across diverse geographical locations. Nat Commun 2019; 10.

39. Fernandes M, Stein A, Newton CR, Cheikh-Ismail L, Kihara M, Wulff K, de Leon Quintana E, Arozaitz A, Azeedo J, Buijs, D, Ababakar A, Guiñez, F, Lewis, T, Kennedy, S, Villar J. The INTERGROWTH-21st Project: a novel method for the multi-dimensional assessment of neurodevelopment in pre-school age children. PLoS One 2019; 14(10): e0223198.

48. Ohuma EO, Altman DG; International Fetal and Newborn Growth Consortium for the Evaluation of Agreements between Two Methods of Clinical Measurement. BJOG 2014; 121(Suppl 2): 27–32, v.

49. Papageorghiou AT, Sarris I, Ioannou C, Todros T, Carvalho M, Pilu G, Fortier MV, Meaney MJ, Qiu A. Population differences in brain morphology: how well can offline measurements from three-dimensional ultrasound volumes substitute real-time two-dimensional measurements? Ultrasound Obstet Gynecol 2013; 42: 560–570.

50. Villar J, Fernandes M, Purwar M, Staines-Urivas E, Di Nicola P, Cheikh Ismail L, Ochengo F, Barlows R, Alberness E, Bruno C, Kunnawar N, Temple S, Greenberg J, Salomons D, Carvalho M, Ohuma, E, Jaffer, Y, Aertipi E, Maggy, M, Lambert A, Bertino, E, Papageorghiou V, Gaza, C, Stein, A, Bhatta, Z, Kennedy SH. Clinical accuracy and replicability when constructing international standards of fetal biometry: results from the INTERGROWTH-21st Project. Obstet Gynecol 2018; 132: 506–513.

51. Bland JM, Altman DG. A note on the use of the intraclass correlation coefficient in the evaluation of agreement between two methods of measurement. Comput Biol Med 1990; 20: 377–380.

52. Contro E, Salsi G, Montaguti E, Morganelli G, Pilu G, Rizzo N, Bonasoni P, Ghi T. Morphometric and cognitive assessments in pre-school age children. PLoS One 2013; 8(8): e73280.

53. WHO Multicentre Growth Reference Study Group. Assessment of differences in growth and likeness of fetal growth and newborn size across non-isolated populations in the INTERGROWTH-21st Project. BJOG 2014; 121(Suppl 2): 27–32, v.

54. International Fetal and Newborn Growth Consortium. International Fetal and Newborn Growth Standards for the 21st Century. Ultrasound Obstet Gynecol 2008; 32(2): 151–159. Available at: https://intergrowth21.tghn.org/articles/category/study-protocols/

55. Papageorghiou AT, Sarris I, Ioannou C, Todros T, Carvalho M, Pilu G, Salomons D, Carvalho M, Ohuma E, Jaffer Y, Bertino E, Papageorghiou V, Gaza C, Stein A, Bhatta Z, Kennedy SH. Ultrasound obstetric measurements of the fetal central nervous system. Ultrasound Obstet Gynecol 2007; 30: 233–241.

56. van der Knaap MS, van Weel-Meijler G, Barth PG, Barkhof F, Ader HJ, Vink JL. Normal gyration and sulcation and term neonates: appearance on MR images. Radiology 1996; 200: 389–396.

57. Toi A, Lister WS, Kong FW. How easily are fetal cerebral sulci visible at prenatal ultrasound and what is the normal pattern of early fetal sulcal development? Ultrasound Obstet Gynecol 2014; 43: 706–715.

58. Villar J., Sarris I., Ioannou C., Bughe M., Mitidieri A., Oberto M., Qinpng W., Shal J., Solonos S., JDziplau W., Hoch L., Altman D.G., Papageorghiou A.T. Standardization of fetal ultrasound measurements: feasibility, agreement, and consistency of measurements. Ultrasound Obstet Gynecol 2011; 38: 681–687.

59. Carvalho A., Azh ST., Napoleon R., Wanyonyi S., Ohuma EO., Mollabolli M., Sande J., Sarris I., Ioannou C., Norris T., Donatondo Y., Carvalho M., Purwar M., Barsos M., Jaffer YA., Bertino E., Pangu, R., Maggy MT., Salomons L., Noble JA., Altman D.G., Papageorghiou AT. Quality control of ultrasound fetal biometry: results from the INTERGROWTH-21st Project. Obstet Gynecol 2014; 123: 332–339.

60. Rosston P., Altman D.G. Regression Using Fractional Polynomials of Continuous Covariates: Parametric and Non-parametric Modelling. Appl Stat 1988; 37: 384–402. [accessed May 28 2019]. Available from: https://www.jstor.org/stable/2986270?origin=crossref

61. Ohuma EO, International Fetal and Newborn Growth Consortium for the Evaluation of Agreements between Two Methods of Clinical Measurement. BMJ 1996; 313: 307–310.

62. Sarris I., Ioannou C., Chamberlain P., Ohuma E., Rosenman F., Hoch L., Altman D.G., Papageorghiou A.T. Intra- and interobserver variability in fetal ultrasound measurements. Ultrasound Obstet Gynecol 2012; 39: 266–273.

63. Bland JM, Altman DG. Are you sure? are you sure? - a guide to constructing confidence intervals for intra-class correlation coefficients in the evaluation of agreement between two methods of measurement. BMJ 1999; 319: 307–310.

64. Contro E., Salis G., Victoria Luigi, M., Morgantin E., Pile G., Roma, N., Bonasoni P., Ghi T. Sequiological analysis of the normal fetal fitsures with three-dimensional ultrasound: a longitudinal study. Prenat Diagn 2015; 35: 493–499.

65. WHO Multicentre Growth Reference Study Group. Assessment of differences in growth patterns among the WHO Multicentre Growth Reference Study. Acta Paediatr Suppl 2006; 540: 56–58.

66. Bai J., Abdul-Rahman M., Reddi-Graban A., Chong Y., Kwok K., Saw S., Godfrey K.M., Gluckman PD, Potter MV, Meaney MJ, Qiu A. Population differences in brain morphology and microstructure among Chinese, Malay, and Indian neonates. PLoS One 2013; 27(4): e74816.

67. Jacquesmyo Y., Svi, Verdonk P. Fetal transverse cerebral diameter in different ethnic groups. iPerinat Med 2009; 8: 1–14.

68. Habicht JP, Martorell R, Yarborough C, Malina RM, Klein RE. Height and weight standards for pre-school children: what are the relevant ethnic differences in growth potential? Lancet 1994; 344: 611–614.

69. Villar J., Papageorghiou AT, Pang R, Solomon LJ, Langer A, Victoria C, Purwar M, Chumlea C, Qinqiang W., Scheron SA, Barsos M., Carvalho M., Altman D.G., Giuliani F., Bertino E., Jaffer YA, Cheikh Ismail L., Ohuma EO, Lambert A, Noble JA, ...
SUPPORTING INFORMATION ON THE INTERNET

The following supporting information may be found in the online version of this article:

Figure S1 Scoring system for Sylvian fissure maturation. (a) Grade 1, smooth indentation; (b) Grade 2, obtuse angular shape; (c) Grade 3, acute angular shape (< 50% operculization); (d) Grade 4, angular closure until most of insula is covered (> 50% operculization); (e) Grade 5, complete closure or operculization.

Figure S2 Standardized site discrepancy (SSD) for fetal transverse cerebellar diameter (TCD) (a) and Sylvian fissure (SF) maturation score (b), according to gestational age. SSD calculated as: (mean of site’s measurements – mean of all sites’ measurements at each gestational age interval)/SD of all sites’ measurements at each gestational age interval. SSD adjusted at median gestational age for all sites at each gestational age interval. ± 0.5 SD is shown (- - - -).

Figure S3 Transverse cerebellar diameter ultrasound measurements according to gestational age and fetal sex (male (green) or female (red)), in fetuses that contributed data to transverse cerebellar diameter international standards. No suggestion of any sex differences was evident.

Figure S4 Spaghetti plot of fetal Sylvian fissure maturation scores, obtained longitudinally in fetuses that contributed data to Sylvian fissure maturation international standards, according to fetal sex. No suggestion of any sex differences was evident.

Figure S5 Spaghetti plot of Sylvian fissure maturation scores, obtained longitudinally in fetuses that contributed data to Sylvian fissure maturation international standards, according to side of the Sylvian fissure. No suggestion of any differences was evident.

Table S1 Smoothed centiles for transverse cerebellar diameter (mm) according to exact gestational age (weeks)

Table S2 Gestational age (weeks) estimation according to transverse cerebellum diameter (mm)

Table S3 Equations for estimation of mean and SD (mm) of transverse cerebellar diameter according to exact gestational age (weeks)