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

María J. RODRIGUEZ-SIBAJA[1,2,3], José VILLAR[1,2], Eric O. OHUMA[1,2,4], Raffaele NAPOLITANO[1], Sabrina HEYL[1], Maria CARVALHO[5], Yasmin A. JAFFER[6], J. Alison NOBLE[7], Manuela OBERTO[8], Manorama PURWAR[9], Ruyan PANG[10], Leila CHEIKH ISMAIL[11,1], Ann LAMBERT[1,2], Michael G. GRAVETT[12], Laurent J. SALOMON[13], Lior DRUKKER[1,2] Fernando C BARROS[14], Stephen H. KENNEDY[1,2], Zulfiqar A. BHUTTA[15], Aris T. PAPAGEORGHIOU*[1,2]

## Affiliations

1. Nuffield Department of Women's & Reproductive Health, University of Oxford, UK
2. Oxford Maternal & Perinatal Health Institute, Green Templeton College, University of Oxford, Oxford, UK
3. Maternal-Fetal Medicine Department, National Institute of Perinatology, Mexico City, Mexico.
4. Centre for Statistics in Medicine, Botnar Research Centre, University of Oxford, Oxford, UK
5. Faculty of Health Sciences, Aga Khan University, Nairobi, Kenya
6. Department of Family & Community Health, Ministry of Health, Muscat, Sultanate of Oman
7. Department of Engineering Science, Institute of Biomedical Engineering, University of Oxford, Oxford, UK
8. S.C. Ostetricia 2U, Città della Salute e della Scienza di Torino, Italy
9. Nagpur INTERGROWTH-21st Research Centre, Ketkar Hospital, Nagpur, India
10. School of Public Health, Peking University, Beijing, China
11. Clinical Nutrition and Dietetics Department, University of Sharjah, Sharjah, United Arab Emirates
12. Departments of Obstetrics & Gynecology and of Public Health, University of Washington, Seattle 98195 WA USA
13. Department of Obstetrics and Fetal Medicine, Hôpital Necker Enfants Malades, Université Paris Descartes, Paris, France

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/uog.22017

14. Programa de Pós-Graduação em Saúde e Comportamento, Universidade Católica de Pelotas, Pelotas, Brazil
15. Center for Global Child Health, Hospital for Sick Children, Toronto, Canada

* **Corresponding author**
Aris T. Papageorghiou, Nuffield Department of Women's & Reproductive Health, University of Oxford, The Women's Centre, John Radcliffe Hospital, Oxford, OX3 9DU, UK.
Email: aris.papageorghiou@wrh.ox.ac.uk

**Short Title:** International fetal trans-cerebellar diameter and Sylvian fissure maturation standards

**Keywords:** Fetal brain, cerebellum, cerebral cortex, cortical maturation, cortical development, gyration, operculization, fissures, neurodevelopment, neurosonography, ultrasound

## Contribution

### A. What are the novel findings of this work?

When constraints on human growth are minimal, the growth and developmental patterns of the fetal brain structures studied are similar across diverse populations and there are no sex-related differences.

### B. What are the clinical implications of this work?

We present international standards for fetal cerebellar growth and Sylvian fissure maturation using 3D ultrasound volumes obtained during the pregnancies of healthy, well-nourished women from five diverse urban areas worldwide who were enrolled in the Fetal Growth Longitudinal Study of the INTERGROWTH-21$^{st}$ 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 cerebellar growth and Sylvian fissure maturation.

**Methods.** Healthy, well-nourished pregnant women, enrolled at <14 weeks of gestation in the Fetal Growth Longitudinal Study (FGLS) of INTERGROWTH-21st, an international multicenter, population-based project, underwent 3D serial fetal ultrasound scans every 5±1 weeks until delivery in study sites located in Brazil, India, Italy, Kenya and the UK. We measured the trans-cerebellar diameter and assessed Sylvian fissure maturation using 2D images extracted from the available 3D fetal head volumes acquired. For each Sylvian fissure maturation score (left and right), we calculated the mean gestational age and 95% CI. We assessed goodness of fit of the resultant model and the possibility of pooling data from the five sites using variance component analysis and standardized site differences. We modelled trans-cerebellar diameter and Sylvian fissure maturation using fractional polynomial regression and fitted centiles.

**Results.** Of those children in the original FGLS cohort who had a developmental assessment at 2 years of age, 1,130 also had an available 3D fetal head volume. The socio-demographic characteristics and pregnancy/perinatal outcomes of the study sample confirmed the health and low-risk status of the population studied. In

total, 3,016 and 2,359 individual volumes were available for trans-cerebellar 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 modelling trans-cerebellar diameter; we then estimated gestational age-specific 3$^{rd}$, 50$^{th}$ and 97$^{th}$ smoothed centiles. Goodness of fit comparing empirical centiles to 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% CI between the sexes for each development score. No differences in maturation between the right and left hemispheres were observed.

**Conclusion**. We present here, 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.

## 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.[1,2]

The cerebellum, which is associated later in life with sensorimotor, cognitive and affective regulation,[1,3] can be identified during fetal life by ultrasonography as early as 12 weeks of gestation.[4] In clinical practice, ultrasonography is used to evaluate the anatomical integrity of the fetal cerebellum and linear growth of the trans-cerebellar diameter (TCD).[5,1,3] Measuring the TCD also enables estimation of gestational age in cases of uncertain dates in the third trimester,[6,7] and may be more suitable when there are fetal growth disturbances[6] because cerebellar growth is less affected by placental insufficiency due to the 'brain-sparing' phenomenon.[6,8]

The fetal brain cortex displays remarkable gestational age specific maturation.[9,10] A leading marker of these processes, which follow a predictable timetable,[9,11] is the

development (operculization) of the Sylvian fissure (SF) on the lateral convexities of the cerebral hemispheres.

There are several reference ranges for TCD measures [4,6,9,13–22] and SF maturation.[9,22–26] Many of the studies have a high risk of methodological bias due to sample selection, study design, data analysis and lack of ultrasound quality control. These limitations most likely explain the reported variation in the range of values across pregnancy, which makes clinical interpretation difficult. In addition, and very importantly, none of the fetal studies have continued to assess neurodevelopment into early childhood, which would seem a logical requirement for any tool proposed to evaluate fetal normality.[27]

The Fetal Growth Longitudinal Study (FGLS) of the INTERGROWTH-21st Project has produced international standards, based on World Health Organization (WHO) recommendations,[28] for early [29] and late pregnancy dating,[30] fetal growth[31] and estimated weight [32] as well as other aspects of pregnancy care.[33,34] 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 FGLS whose babies had adequate growth and development from early fetal life to 2 years of age.[35]

## Methods

### Study population

INTERGROWTH-21st is an international, multicenter, population-based project (www.intergrowth21.org.uk).[28] 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. FGLS, one of the main studies of the INTERGROWTH-21st Project,[31] enrolled at less than 14 weeks of 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, so as to generate international standards.

The INTERGROWTH-21st Project methodology has been described elsewhere in detail.[31] Briefly, participants were first selected at population and then at individual level. At population level, urban areas (a complete city or county, or part of a city with clear political or geographical limits) where most deliveries occurred in health care facilities were identified. The areas had to be located at an altitude <1600 m, with a low risk of fetal growth disturbances, as well as an absence or low levels of major, known, non-microbiological contamination such as pollution, domestic smoke, radiation or any other toxic abuses. Within each area, all institutions

providing pregnancy and neonatal care where 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 of gestation, and met the entry criteria of optimal health, nutrition, education and socioeconomic status.[31]

Women underwent serial fetal ultrasound scans every five weeks (± 1 week) until 41+6 weeks of gestation in eight urban areas worldwide that were geographically delimited to ensure the study was population-based. At each visit, a set of two-dimensional (2D) images and three-dimensional (3D) volumes of fetal biometric parts were obtained and stored digitally.[31] Gestational age was based on last menstrual period (LMP) provided that: the date of the LMP 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 ages based on LMP and crown-rump length (CRL), measured by ultrasound at 9+0 to 13+6 weeks from the LMP, was 7 days or less.[31,36]

All ultrasound scans were performed by sonographers trained, standardized and regularly audited according to FGLS requirements. Identical 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) that was specially adapted to ensure measurement values were not visible on screen in order to reduce 'expected value' bias.[31,37]

For the present analysis, we included only FGLS participants whose children had neurodevelopmental, nutritional and morbidity assessments at the time of their second birthday ± 2 months, in five of the original eight study sites (the cities of Pelotas (Brazil); Turin (Italy); Oxford (UK); the central area of Nagpur (India) and the Parklands suburb of Nairobi (Kenya).[35,38,39] Three sites in China, Oman and the USA did not participate in the early child 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 where the project was implemented. Participants provided written consent to be involved in the project.[28]

**Volume acquisition, offline analysis and measurement methodology**

TCD measurement and SF assessment were performed using still images, extracted from the available 3D head volumes acquired in the five study sites. Head volumes were acquired at the level of the axial, trans-thalamic plane. Six predefined quality control criteria of the 2D image 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.[40]

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.[41] All data were then transferred electronically to the Ultrasound Coordinating Unit in Oxford. Further details of the methodology for volume acquisition are available at www.intergrowth21.org.uk (follow the link to 'Study Protocol' to download the ultrasound manual).[41]

The 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 [37] and blinded to the clinical details and gestational age.

The TCD was measured in the standard trans-cerebellar plane,[5] whereas SF maturation was assessed in the trans-thalamic plane.[40] First, each stored volume of the fetal head was uploaded onto the multiplanar mode facility.[42] Secondly, 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.[42] As the trans-thalamic plane was the plane of volume acquisition, it required minimal manipulation to extract the 2D plane.[40] The trans-cerebellar 2D plane was extracted at the level of the trans-thalamic plane with a slight posterior tilting, and visualization of the frontal horns of the lateral ventricles, CSP, thalami, cerebellum and cisterna magna.[5]

Once the appropriate trans-cerebellar 2D plane was 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.[5]

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.[22] We used a simple, unweighted, scoring

system, ranging from Grade 0 (no development) to Grade 5 (maximum development), previously employed for magnetic resonance imaging [43] and ultrasound analysis of cortical maturation (Supplementary Figure S1).[9,22,44] 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 the best possible images were obtained for each extracted plane, we used a scoring system to evaluate and grade image quality [45] (1 = impossible to assess accurately , 2 = possible, 3 = good, 4= almost perfect). Only 3D volumes that scored 2-4 were included in the analysis.

**Reproducibility**

A formal assessment of inter-observer reproducibility for plane reconstruction and TCD acquisition was undertaken in a randomly selected subset of 132 fetuses (12%). From this subset, a single head volume of each fetus was randomly selected and assessed by two fetal medicine specialists. Both independently uploaded each volume, extracted the trans-cerebellar plane and measured the TCD, blinded to all measures obtained including their own. Measurement reproducibility was assessed

using Bland-Altman plots of the inter-observer differences and their 95% limits of agreement.[46,47]

**Statistical analysis**

Outliers, defined as measures >5SD 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 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' standard deviation (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).

This SSD allows for direct comparisons of biometric measures in populations across pregnancy, all 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 for combining data from all sites.[36]

Fetal cerebellar measures were assessed for normality of distribution conditional on gestational age. We then modelled TCD as a function of gestational age using fractional polynomial regression and obtained the fitted centiles.[48] For the SF maturation analysis, we calculated the mean gestational age and 95% confidence interval (CI) for each development score. Goodness of fit of the resultant models was assessed as previously described for the INTERGROWTH-21st data by Ohuma & Altman,[49] i.e., a plot of the residuals (observed values minus fitted values) according to 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%.

For the reproducibility study, we did not consider intraclass correlation coefficients (ICCs) to be appropriate since these depend on the range of the measurement values. Consequently, rather than having fixed values, ICCs vary according to the range of gestational ages being studied.[50] Therefore, we used the method proposed by Bland & Altman instead, which has been shown to be more appropriate for assessing the repeatability of two measurements.[46] All analyses were performed using STATA version 15 (StataCorp, College Station, Texas, USA).

## Results

Of the children in the original FGLS cohort [31] who had a developmental assessment at 2 years of age (n = 1,339),[38] 1,130 (84%) also had an available 3D fetal head volume. The proportional contribution from the study sites was 26% India (n = 291), 25% Kenya (n = 280), 23% Italy (n = 262), 14% Brazil (n = 156) and 12% UK (n = 141).

The socio-demographic 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 at 2 years of age [35,38] (Tables 2 and 3, Figure 1).

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 5,746; however, 2,730 (47%) volumes could not be assessed. Most were obtained after 32 weeks of 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.[37,42,51] Other reasons for images of limited quality were fetal

movement artefact not evident during the original scan, acoustic shadows from proximal structures, reverberation artefacts and unfavorable fetal head orientation.[22,37] After these exclusions and removal of eleven outliers, 3,016 and 2,359 volumes from 1,130 fetuses were available for TCD and SF analysis, respectively.

Inter-observer reproducibility for plane reconstruction and TCD measurement was assessed following the INTERGROWTH-21st quality control strategy:[45,52] in a randomized subset of volumes from 132 fetuses (12% of the total sample of 1,130 fetuses) over a gestational age range from 15+0 weeks to 36+2 weeks. The TCD could be measured in 108 (82%) of the 132 volumes. Bland-Altman plots were used graphically to present the inter-observer differences and their 95% limits of agreement [46,47] (Figure 2). The mean inter-observer difference was very close to zero (0.06 mm); the inferior and superior limits of agreement (±2SD) were -2.40 mm and +2.55 mm, respectively, suggesting very close agreement. No evidence of consistent bias was seen across the range of measures.

For both structures assessed, within-site variance was greater than the between-site variance. For TCD, within-site variance was 0.33 (10.9% of the total variance), while the between-site variance was 0.005 (0.02% of the total variance). For SF maturation within-site variance estimate was 0.007 (3.3% of the total variance) while between-site variance was 0.005 (2.2% of the total variance) (Table 4).

The SSD by gestational age for the five sites was expressed as a proportion of the all sites' SD 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 3). The results of these two analyses show that the five study populations are sufficiently similar, according to WHO predefined criteria,[53] for the data to be pooled to produce international standards.

All cerebellar measures were normally distributed 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 4A. Supplementary Table 1 presents the gestational age-specific 3rd, 5th, 10th, 50th, 95th and 97th smoothed centiles for TCD according to gestational age. For clinical purposes, we also present the same fitted centiles of gestational age estimation based on measurement of the TCD (Figure 4B, Supplementary Table 2).

Goodness of fit by gestational age-specific comparisons of empirical centiles to smoothed centile curves showed good agreement and scatter plots 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 Supplementary Table 3, allowing for calculations by readers of any desired centiles according to gestational age in exact weeks. The actual values for these centiles

according to gestational age are presented in Supplementary Table 2. A striking overlap between male and female TCD values was seen (Figure 5).

The SF increased in maturation with advancing gestation, and there was no appreciable difference between study sites with complete overlap of the mean gestational age and 95% CI for each developmental score (Figure 6). Spaghetti plots of SF scores taken longitudinally in the same individuals according to study sites are shown in Figure 7. Similar to TCD, there were no sex differences in SF maturation score (Figure 8). Finally, it was possible to determine which SF was measured (right or left) in 2606 scans; spaghetti plots of the maturation score between right and left SF showed no differences (Supplementary Figure S1).

**Discussion**

We have presented here, 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 have provided follow-up evidence that the fetuses who contributed data to these standards had low morbidity and adequate growth and development at 2 years of age.[35,38]

Variance component analysis showed that only 0.02% and 2.2% 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 FGLS;[36] the 3% variability reported for infant length in the WHO Multicentre Growth Reference Study,[53] 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.[35,38]

Our results refute suggestions that the observed variability in these measures, between unselected samples, is related to genetic differences.[3,17,21,54,55] Similar suggestions have been made in the past 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.[36,56,57]

Our results also do not support the recently published suggestion of *in utero* sexual dimorphism in the development of brain structure and function.[58,59] Specifically, male fetuses have been documented to have larger cerebellar grey matter volume, as well as greater intra-cerebellar functional connectivity.[59] However, we did not find any sex-related differences in the pattern of growth or maturation of the brain structures studied. Finally, 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. 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 two-fold increase in TCD during the second half of pregnancy (Figure 4A). 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 close to term in suspected fetal growth restriction, as the brain-sparing phenomenon may protect cerebellar growth [4,19,21,28] and, in special cases, unusual head shape.[4,7,19,21] In our study the TCD predicted gestational age well, and the prediction interval of +/- 7 days at 20 weeks and +/- 10 days at 32 weeks compares favourably to using Head Circumference. [30] Further work will be needed to assess robustness in pregnancies with abnormal growth.

Longitudinal evaluation of SF maturation shows a characteristic pattern.[44,60,61] 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.[23,44]

We chose to describe and quantify ultrasound patterns of SF maturation, for the following reasons: 1) The SF is the first primary fissure to be sonographically evident at 17-19 weeks of gestation, which means it can be readily assessed at the time of the mid-trimester fetal scan;[9,51] 2) its development follows a predictable timetable during pregnancy, which make it easier to incorporate SF assessment into routine clinical care;[11,51] 3) 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 time,[11,23] and 4) a simple SF scoring system is available.[22] More complex scoring systems or measurements may then be used for detailed neuro-sonographic examination.[11,23,24,62]

A number of studies have reported TCD [4,6,9,13–22] and SF maturation [9,22–24] 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 measures. Not a single study 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 less than 35% of the studies.[27]

We aimed, therefore, to avoid these limitations by producing international standards to complement those from the INTERGROWTH-21st Project that are already being used clinically and for research purposes. [29–34,63,64] The 3D ultrasound volumes were taken specifically for this purpose during the course of conducting FGLS, using rigorous methods, identical ultrasound equipment, blinding of operators and a detailed quality control strategy.[28,31] The images were extracted from the volumes using standardized axial planes recommended in routine clinical practice,[5] and were scored and measured in accordance with a predefined protocol.[40] Our reproducibility study suggested a high level of agreement in fetal imaging. Finally, and crucially, the cohort was followed up to 2 years of age and developmental outcomes confirmed their eligibility for the construction of international standards.[38,39]

It could be said that measures or scores acquired on planes extracted from 3D volumes do not completely accord with 2D measures obtained in real-time.[37,42,65] However, we believe 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 measures obtained in real-time.[22,37,42,51,65]

## Conclusion

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 females. 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-21st grant 49038 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 (BRC). We are grateful to Professor Brenda Eskenazi for her comments regarding the Sylvian fissure side analysis.

## Conflict of interest

ATP and JAN are Senior Advisors of Intelligent Ultrasound. All other authors declare no competing interests.

## Source of funding

The authors gratefully acknowledge the contribution of the Bill and Melinda Gates Foundation (Grant ID# 49038) who funded the INTERGROWTH-21st Project. The

funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## References


1. Stoodley CJ, Limperopoulos C. Structure-function relationships in the developing cerebellum: Evidence from early-life cerebellar injury and neurodevelopmental disorders. Semin Fetal Neonatal Med. 2016 Oct;21(5):356–64.
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. Chin Med J (Engl). 2017 Apr;130(8):920–8.
3. Koning I V, Tielemans MJ, Hoebeek FE, Ecury-Goossen GM, Reiss IKM, Steegers-Theunissen RPM, Dudink J. Impacts on prenatal development of the human cerebellum: a systematic review. J Matern Fetal Neonatal Med. 2017 Oct;30(20):2461–8.
4. Goldstein I, Reece EA, Pilu G, Bovicelli L, Hobbins JC. Cerebellar measurements with ultrasonography in the evaluation of fetal growth and development. Am J Obstet Gynecol. 1987 May;156(5):1065–9.
5. Sonographic examination of the fetal central nervous system: guidelines for performing the “basic examination” and the “fetal neurosonogram”. Ultrasound Obstet Gynecol. 2007 Jan;29(1):109–16.
6. Chavez MR, Ananth C V, Smulian JC, Lashley S, Kontopoulos E V, 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 Oct;189(4):1021–5.
7. Reddy RH, Prashanth K, Ajit M. Significance of foetal transcerebellar diameter in foetal biometry: A pilot study. J Clin Diagnostic Res. 2017;11(6):TC01–4.
8. Vinkesteijn AS, Mulder PG, Wladimiroff JW. Fetal transverse cerebellar diameter measurements in normal and reduced fetal growth. Ultrasound Obstet Gynecol. 2000 Jan;15(1):47–51.
9. Quarello E, Stirnemann J, Ville Y, Guibaud L. Assessment of fetal Sylvian

fissure operculization between 22 and 32 weeks: a subjective approach. Ultrasound Obstet Gynecol. 2008 Jul;32(1):44–9.

10. Cohen-Sacher B, Lerman-Sagie T, Lev D, Malinger G. Sonographic developmental milestones of the fetal cerebral cortex: a longitudinal study. Ultrasound Obstet Gynecol. 2006 May;27(5):494–502.
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 Mar;26(3):347–53.
12. Smith PA, Johansson D, Tzannatos C, Campbell S. Prenatal measurement of the fetal cerebellum and cisterna cerebellomedullaris by ultrasound. Prenat Diagn. 1986;6(2):133–41.
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(2):158–61.
14. McLeary RD, Kuhns LR, Barr M. Ultrasonography of the fetal cerebellum. Radiology. 1984;151:439–42.
15. Serhatlioglu S, Kocakoc E, Kiris A, Sapmaz E, Boztosun Y, Bozgeyik Z. Sonographic measurement of the fetal cerebellum, cisterna magna, and cavum septum pellucidum in normal fetuses in the second and third trimesters of pregnancy. J Clin Ultrasound. 2003 May;31(4):194–200.
16. Verburg BO, Steegers EAP, De Ridder M, Snijders RJM, Smith E, Hofman A, Moll HA, Jaddoe VW V, Witteman JCM. New charts for ultrasound dating of pregnancy and assessment of fetal growth: longitudinal data from a population-based cohort study. Ultrasound Obstet Gynecol. 2008 Apr;31(4):388–96.
17. Joshi B. Fetal transcerebellar diameter nomogram in Nepalese population. J Inst Med. 2010;32(1):19–23.
18. Hayata K, Hiramatsu Y, Masuyama H, Etou E, Nobumoto E, Mitsui T. Creation of a cerebellar diameter reference standard and its clinical application to the detection of cerebellar hypoplasia unique to trisomy 18. J

Obstet Gynaecol Res. 2015 Dec;41(12):1899–904.

19. Chang CH, Chang FM, Yu CH, Ko HC, Chen HY. Three-dimensional ultrasound in the assessment of fetal cerebellar transverse and antero-posterior diameters. Ultrasound Med Biol. 2000 Feb;26(2):175–82.

20. Hata K, Hata T, Senoh D, Makihara K, Aoki S, Takamiya O, Kitao M. Ultrasonographic measurement of the fetal transverse cerebellum in utero. Gynecol Obstet Invest. 1989;28(2):111–2.

21. Araujo Junior E, Martins WP, Nardozza LMM, Pires CR, Filho SMZ. Reference range of fetal transverse cerebellar diameter between 18 and 24 weeks of pregnancy in a Brazilian population. J Child Neurol. 2015 Feb;30(2):250–3.

22. Pistorius LR, Stoutenbeek P, Groenendaal F, de Vries L, Manten G, Mulder E, Visser G. Grade and symmetry of normal fetal cortical development: a longitudinal two- and three-dimensional ultrasound study. Ultrasound Obstet Gynecol. 2010 Dec;36(6):700–8.

23. Alonso I, Borenstein M, Grant G, Narbona I, Azumendi G. Depth of brain fissures in normal fetuses by prenatal ultrasound between 19 and 30 weeks of gestation. Ultrasound Obstet Gynecol. 2010 Dec;36(6):693–9.

24. Alves CMS, Araujo Junior E, Nardozza LMM, Goldman SM, Martinez LH, Martins WP, Oliveira PS, Moron AF. Reference ranges for fetal brain fissure development on 3-dimensional sonography in the multiplanar mode. J Ultrasound Med. 2013 Feb;32(2):269–77.

25. Pooh RK, Machida M, Nakamura T, Uenishi K, Chiyo H, Itoh K, Yoshimatsu J, Ueda H, Ogo K, Chaemsaithong P, Poon LC. Increased Sylvian fissure angle as early sonographic sign of malformation of cortical development. Ultrasound Obstet Gynecol. 2019 Aug;54(2):199–206.

26. Poon LC, Sahota DS, Chaemsaithong P, Nakamura T, Machida M, Naruse K, Wah YM, Leung TY, Pooh RK. Transvaginal three-dimensional ultrasound assessment of Sylvian fissures at 18-30 weeks' gestation. Ultrasound Obstet Gynecol. 2019 Aug;54(2):190–8.

27. Donadono V, Napolitano R, Cavallaro A, Roberts NW, Papageorghiou AT.

OP04.08: Charts of fetal brain structures: a systematic review of methodological quality. Ultrasound Obstet Gynecol. 2017 Sep 1;50:59.

28. Villar J, Altman DG, Purwar M, Noble JA, Knight HE, Ruyan P, Cheikh Ismail L, Barros FC, Lambert A, Papageorghiou AT, Carvalho M, Jaffer YA, Bertino E, Gravett MG, Bhutta ZA, Kennedy SH. The objectives, design and implementation of the INTERGROWTH-21st Project. BJOG. 2013 Sep;120 Suppl:9–26, v.

29. Papageorghiou AT, Kennedy SH, Salomon LJ, Ohuma EO, Cheikh Ismail L, Barros FC, Lambert A, Carvalho M, Jaffer YA, Bertino E, Gravett MG, Altman DG, Purwar M, Noble JA, Pang R, Victora CG, Bhutta ZA, Villar J. International standards for early fetal size and pregnancy dating based on ultrasound measurement of crown-rump length in the first trimester of pregnancy. Ultrasound Obstet Gynecol. 2014 Dec;44(6):641–8.

30. Papageorghiou AT, Kemp B, Stones W, Ohuma EO, Kennedy SH, Purwar M, Salomon LJ, Altman DG, Noble JA, Bertino E, Gravett MG, Pang R, Cheikh Ismail L, Barros FC, Lambert A, Jaffer YA, Victora CG, Bhutta ZA, Villar J. Ultrasound-based gestational-age estimation in late pregnancy. Ultrasound Obstet Gynecol. 2016 Dec;48(6):719–26.

31. Papageorghiou AT, Ohuma EO, Altman DG, Todros T, Cheikh Ismail L, Lambert A, Jaffer YA, Bertino E, Gravett MG, Purwar M, Noble JA, Pang R, Victora CG, Barros FC, Carvalho M, Salomon LJ, Bhutta ZA, Kennedy SH, Villar J. International standards for fetal growth based on serial ultrasound measurements: the Fetal Growth Longitudinal Study of the INTERGROWTH-21st Project. Lancet (London, England). 2014 Sep;384(9946):869–79.

32. Stirnemann J, Villar J, Salomon LJ, Ohuma E, Ruyan P, Altman DG, Nosten F, Craik R, Munim S, Cheikh Ismail L, Barros FC, Lambert A, Norris S, Carvalho M, Jaffer YA, Noble JA, Bertino E, Gravett MG, Purwar M, Victora CG, Uauy R, Bhutta Z, Kennedy S, Papageorghiou AT. International estimated fetal weight standards of the INTERGROWTH-21(st) Project. Ultrasound Obstet Gynecol. 2017 Apr;49(4):478–86.

33. Papageorghiou AT, Ohuma EO, Gravett MG, Hirst J, da Silveira MF, Lambert A, Carvalho M, Jaffer YA, Altman DG, Noble JA, Bertino E, Purwar

M, Pang R, Cheikh Ismail L, Victora C, Bhutta ZA, Kennedy SH, Villar J. International standards for symphysis-fundal height based on serial measurements from the Fetal Growth Longitudinal Study of the INTERGROWTH-21st Project: prospective cohort study in eight countries. BMJ. 2016 Nov;355:i5662.

34. Cheikh Ismail L, Bishop DC, Pang R, Ohuma EO, Kac G, Abrams B, Rasmussen K, Barros FC, Hirst JE, Lambert A, Papageorghiou AT, Stones W, Jaffer YA, Altman DG, Noble JA, Giolito MR, Gravett MG, Purwar M, Kennedy SH, Bhutta ZA, Villar J. Gestational weight gain standards based on women enrolled in the Fetal Growth Longitudinal Study of the INTERGROWTH-21st Project: a prospective longitudinal cohort study. BMJ. 2016 Feb;352:i555.

35. Villar J, Cheikh Ismail L, Staines Urias E, Giuliani F, Ohuma EO, Victora CG, Papageorghiou AT, Altman DG, Garza C, Barros FC, Puglia F, Ochieng R, Jaffer YA, Noble JA, Bertino E, Purwar M, Pang R, Lambert A, Chumlea C, Stein A, Fernandes M, Bhutta ZA, Kennedy SH. The satisfactory growth and development at 2 years of age of the INTERGROWTH-21(st) Fetal Growth Standards cohort support its appropriateness for constructing international standards. Am J Obstet Gynecol. 2018 Feb;218(2S):S841-S854.e2.

36. Villar J, Papageorghiou AT, Pang R, Ohuma EO, Cheikh Ismail L, Barros FC, Lambert A, Carvalho M, Jaffer YA, Bertino E, Gravett MG, Altman DG, Purwar M, Frederick IO, Noble JA, Victora CG, Bhutta ZA, Kennedy SH. The likeness of fetal growth and newborn size across non-isolated populations in the INTERGROWTH-21st Project: the Fetal Growth Longitudinal Study and Newborn Cross-Sectional Study. lancet Diabetes Endocrinol. 2014 Oct;2(10):781–92.

37. Sarris I, Ohuma E, Ioannou C, Sande J, Altman DG, Papageorghiou AT. Fetal biometry: how well can offline measurements from three-dimensional volumes substitute real-time two-dimensional measurements? Ultrasound Obstet Gynecol. 2013 Nov;42(5):560–70.

38. Villar J, Fernandes M, Purwar M, Staines-Urias E, Di Nicola P, Cheikh Ismail L, Ochieng R, Barros F, Albernaz E, Victora C, Kunnawar N, Temple S, Giuliani F, Sandells T, Carvalho M, Ohuma E, Jaffer Y, Noble A, Gravett M,

Pang R, Lambert A, Bertino E, Papageorghiou A, Garza C, Stein A, Bhutta Z, Kennedy S. Neurodevelopmental milestones and associated behaviours are similar among healthy children across diverse geographical locations. Nat Commun. 2019 Jan;10(1):511.

39. Fernandes M, Stein A, Newton CR, Cheikh-Ismail L, Kihara M, Wulff K, de Leon Quintana E, Aranzeta L, Soria-Frisch A, Acedo J, Ibanez D, Abubakar A, Giuliani F, Lewis T, Kennedy S, Villar J. The INTERGROWTH-21st Project Neurodevelopment Package: a novel method for the multi-dimensional assessment of neurodevelopment in pre-school age children. PLoS One. 2014;9(11):e113360.

40. Papageorghiou AT, Sarris I, Ioannou C, Todros T, Carvalho M, Pilu G, Salomon LJ. Ultrasound methodology used to construct the fetal growth standards in the INTERGROWTH-21st Project. BJOG. 2013 Sep;120 Suppl:27–32, v.

41. The International Fetal and Newborn Growth Consortium. International Fetal and Newborn Growth Standards for the 21 st Century. Ultrasound Operations Manual. INTERGROWTH-21st. 2012. p. 1–59.

42. Pilu G, Ghi T, Carletti A, Segata M, Perolo A, Rizzo N. Three-dimensional ultrasound examination of the fetal central nervous system. Ultrasound Obstet Gynecol. 2007 Aug;30(2):233–45.

43. van der Knaap MS, van Wezel-Meijler G, Barth PG, Barkhof F, Ader HJ, Valk J. Normal gyration and sulcation in preterm and term neonates: appearance on MR images. Radiology. 1996 Aug;200(2):389–96.

44. Toi A, Lister WS, Fong KW. How early are fetal cerebral sulci visible at prenatal ultrasound and what is the normal pattern of early fetal sulcal development? Ultrasound Obstet Gynecol. 2004 Dec;24(7):706–15.

45. Sarris I, Ioannou C, Dighe M, Mitidieri A, Oberto M, Qingqing W, Shah J, Sohoni S, Al Zidjali W, Hoch L, Altman DG, Papageorghiou AT. Standardization of fetal ultrasound biometry measurements: improving the quality and consistency of measurements. Ultrasound Obstet Gynecol. 2011 Dec;38(6):681–7.

46. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet (London, England). 1986 Feb;1(8476):307–10.

47. Sarris I, Ioannou C, Chamberlain P, Ohuma E, Roseman F, Hoch L, Altman DG, Papageorghiou AT. Intra- and interobserver variability in fetal ultrasound measurements. Ultrasound Obstet Gynecol. 2012 Mar;39(3):266–73.

48. Royston P, Altman DG. Regression Using Fractional Polynomials of Continuous Covariates: Parsimonious Parametric Modelling. Appl Stat [Internet]. 1994 [cited 2019 May 28];43(3):429. Available from: https://www.jstor.org/stable/2986270?origin=crossref

49. Ohuma EO, Altman DG. Statistical methodology for constructing gestational age-related charts using cross-sectional and longitudinal data: The INTERGROWTH-21(st) project as a case study. Stat Med. 2018 Nov;

50. 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(5):337–40.

51. Contro E, Salsi G, Montaguti E, Morganelli G, Pilu G, Rizzo N, Bonasoni P, Ghi T. Sequential analysis of the normal fetal fissures with three-dimensional ultrasound: a longitudinal study. Prenat Diagn. 2015 May;35(5):493–9.

52. Cavallaro A, Ash ST, Napolitano R, Wanyonyi S, Ohuma EO, Molloholli M, Sande J, Sarris I, Ioannou C, Norris T, Donadono V, Carvalho M, Purwar M, Barros FC, Jaffer YA, Bertino E, Pang R, Gravett MG, Salomon LJ, Noble JA, Altman DG, Papageorghiou AT. Quality control of ultrasound for fetal biometry: results from the INTERGROWTH-21(st) Project. Ultrasound Obstet Gynecol. 2018 Sep;52(3):332–9.

53. WHO Multicentre Growth Reference Study Group. Assessment of differences in linear growth among populations in the WHO Multicentre Growth Reference Study. Acta Paediatr Suppl. 2006 Apr;450:56–65.

54. Bai J, Abdul-Rahman MF, Rifkin-Graboi A, Chong Y-S, Kwek K, Saw S-M, Godfrey KM, Gluckman PD, Fortier M V, Meaney MJ, Qiu A. Population differences in brain morphology and microstructure among Chinese, Malay,

and Indian neonates. PLoS One. 2012;7(10):e47816.

55. Jacquemyn Y, Sys SU, Verdonk P. Fetal transverse cerebellar diameter in different ethnic groups. J Perinat Med. 2000;28(1):14–9.

56. Habicht JP, Martorell R, Yarbrough C, Malina RM, Klein RE. Height and weight standards for preschool children. How relevant are ethnic differences in growth potential? Lancet (London, England). 1974 Apr;1(7858):611–4.

57. Villar J, Papageorghiou AT, Pang R, Salomon LJ, Langer A, Victora C, Purwar M, Chumlea C, Qingqing W, Scherjon SA, Barros FC, Carvalho M, Altman DG, Giuliani F, Bertino E, Jaffer YA, Cheikh Ismail L, Ohuma EO, Lambert A, Noble JA, Gravett MG, Bhutta ZA, Kennedy SH. Monitoring human growth and development: a continuum from the womb to the classroom. Am J Obstet Gynecol. 2015 Oct;213(4):494–9.

58. Ruigrok AN V, Salimi-Khorshidi G, Lai M-C, Baron-Cohen S, Lombardo M V, Tait RJ, Suckling J. A meta-analysis of sex differences in human brain structure. Neurosci Biobehav Rev. 2014 Feb;39:34–50.

59. Wheelock MD, Hect JL, Hernandez-Andrade E, Hassan SS, Romero R, Eggebrecht AT, Thomason ME. Sex differences in functional connectivity during fetal brain development. Dev Cogn Neurosci. 2019 Mar;36:100632.

60. Monteagudo A, Timor-Tritsch IE. Normal sonographic development of the central nervous system from the second trimester onwards using 2D, 3D and transvaginal sonography. Prenat Diagn. 2009 Apr;29(4):326–39.

61. Garel C, Chantrel E, Brisse H, Elmaleh M, Luton D, Oury JF, Sebag G, Hassan M. Fetal cerebral cortex: normal gestational landmarks identified using prenatal MR imaging. AJNR Am J Neuroradiol. 2001 Jan;22(1):184–9.

62. Napolitano R, Molloholli M, Donadono V, Ohuma EO, Wanyonyi SZ, Kemp B, Yaqub MK, Ash S, Barros FC, Carvalho M, Jaffer YA, Noble JA, Oberto M, Purwar M, Pang R, Cheikh Ismail L, Lambert A, Gravett MG, Salomon LJ, Bhutta ZA, Kennedy SH, Villar J, Papageorghiou AT. International standards for fetal brain structures based on serial ultrasound measurements from the Fetal Growth Longitudinal Study of the INTERGROWTH-21(st) Project. Ultrasound Obstet Gynecol. 2020 Feb;

63. Villar J, Cheikh Ismail L, Victora CG, Ohuma EO, Bertino E, Altman DG, Lambert A, Papageorghiou AT, Carvalho M, Jaffer YA, Gravett MG, Purwar M, Frederick IO, Noble AJ, Pang R, Barros FC, Chumlea C, Bhutta ZA, Kennedy SH. International standards for newborn weight, length, and head circumference by gestational age and sex: the Newborn Cross-Sectional Study of the INTERGROWTH-21st Project. Lancet (London, England). 2014 Sep;384(9946):857–68.

64. Villar J, Giuliani F, Bhutta ZA, Bertino E, Ohuma EO, Ismail LC, Barros FC, Altman DG, Victora C, Noble JA, Gravett MG, Purwar M, Pang R, Lambert A, Papageorghiou AT, Ochieng R, Jaffer YA, Kennedy SH. Postnatal growth standards for preterm infants: the Preterm Postnatal Follow-up Study of the INTERGROWTH-21(st) Project. Lancet Glob Heal. 2015 Nov;3(11):e681-91.

65. Ioannou C, Sarris I, Salomon LJ, Papageorghiou AT. A review of fetal volumetry: the need for standardization and definitions in measurement methodology. Ultrasound Obstet Gynecol. 2011 Dec;38(6):613–9.

66. WHO Motor Development Study: windows of achievement for six gross motor development milestones. Acta Paediatr Suppl. 2006 Apr;450:86–95.

**Figure legends**

**Figure 1.** Median age of achievement (3rd and 97th centiles) of four gross motor development milestones. Data are for children who contributed data to the trans-cerebellar diameter and Sylvian fissure maturation standards. For comparison, the 3rd and 97th centiles of the World Health Organization windows of achievement for the same milestones [66] are presented in *grey* (with the median shown as a *vertical line*).

**Figure 2.** Inter-observer reproducibility for trans-cerebellar diameter measures. Bland-Altman plot showing inter-observer reproducibility for trans-cerebellar diameter measures in 108 randomly selected fetuses. The black horizontal line represents the mean inter-observer difference and the dashed red horizontal lines represent the inferior and superior 95% limits of agreement (± 2SD).

**Figure 3.** Standardized site discrepancy for trans-cerebellar diameter (TCD) **(A)** and Sylvian fissure (SF) maturation scores **(B)**. Standardized site discrepancy (SSD) calculated by: (site mean of trans-cerebellar diameter or Sylvian fissure maturation score minus all sites' mean of trans-cerebellar diameter or Sylvian fissure maturation score at each gestational age interval) divided by all sites' SD of either trans-cerebellar diameter or Sylvian fissure maturation score at each gestational age

interval. SSD adjusted at the median gestational age for all sites at each gestational age interval. The dashed red horizontal line is the 0.5 SD.

**Figure 4.** Cerebellar measures according to gestational age. Fitted 3rd, 50th and 97th smoothed centile curves for trans-cerebellar diameter (TCD) measured by ultrasound according to gestational age (GA) (**A).** Gestational age estimation by trans-cerebellar diameter measures (**B)**.

**Figure 5.** Trans-cerebellar diameter through gestational age by fetal sex. Data are for fetuses who contributed data to the trans-cerebellar diameter standards. For comparison, data are plotted by fetal sex (girls [red crosses], boys [green crosses]. No suggestion of any sex differences was evident.

**Figure 6.** Mean gestational age according to Sylvian fissure maturation score. Mean gestational age and 95% CI according to Sylvian fissure maturation score calculated and plotted by study site.

**Figure 7.** Sylvian fissure maturation through gestational age. Spaghetti plot of Sylvian fissure maturation scores taken longitudinally in the same individuals according to study sites.

**Figure 8.** Sylvian fissure maturation scores by fetal sex. Data are for fetuses who contributed data to the Sylvian fissure maturation standards. For comparison, data are plotted by fetal sex (girls [pink lines], boys [blue lines]. No suggestion of any sex differences was evident.

**Supplementary Figure S1.** Sylvian fissure maturation scores. 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 the insula is covered (>50% operculization); (E) Grade 5, complete closure or operculization.

**Supplementary Figure S2.** Sylvian fissure maturation score by side of the Sylvian fissure. Data are for fetuses who contributed data to the Sylvian fissure maturation standards and where the side of the fissure could be determined (see text). Right - blue lines, Left – red lines. No suggestion of any differences was evident.

**Table 1.** Maternal and perinatal characteristics of the women included in the present analysis compared to the full INTERGROWTH-21st Fetal Growth Longitudinal Study cohort.[31]

| Characteristics | Current study | FGLS* |
|---|---|---|
| | N = 1130 | 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 ($kg/m^2$)** | 23.3 (3.0) | 23.3 (3.0) |
| **Gestational age at first visit (weeks)** | 11.6 (1.3) | 11.8 (1.4) |
| **Years of formal education (years)** | 15.2 (2.9) | 15.0 (2.8) |
| **Hemoglobin level at <15 weeks of gestation (g/L)** + | 123.6 (9.8) | 125 (11) |
| **Nulliparous (%)** | 657 (58%) | 2955 (68%) |
| **Pre-eclampsia** | 11 (<1%) | 31 (<1%) |

| | | |
|---|---|---|
| **Pyelonephritis** | 6 (<1%) | 16 (<1%) |
| **Any sexually transmitted infection** | 0 (0%) | 3 (<1%) |
| **Spontaneous initiation of labor** | 754 (66.7) | 2868 (66%) |
| **PPROM ‡ (<37 weeks of gestation)** | 71 (6.3%) | 80 (2%) |
| **Cesarean section** | 275 (24.3%) | 1541 (36%) |
| **NICU § admission ( >1 day)** | 29 (2.6%) | 240 (6%) |
| **Term low birth weight (<2500 g)** | 42 (3.7%) | 128 (3%) |
| **Neonatal mortality** | 0 (0%) | 7 (<1%) |
| **Male sex** | 553 (48.9) | 2149 (50%) |
| **Exclusive breastfeeding at discharge** | 1048 (92.7) | 3786 (88%) |
| **Mother admitted to intensive care unit** | 2 (<1%) | 17 (<1%) |
| **Birth weight (≥37 weeks of gestation) (kg)** | 3.2 (0.4) | 3.3 (0.4) |
| **Birth length (≥37 weeks of gestation) (cm)** | 49.2 (1.8) | 49.4 (1.9) |

| | | |
|---|---|---|
| **Birth head circumference (≥37 weeks of gestation) (cm)** | 34.0 (1.3) | 33.9 (1.3) |

* FGLS = Fetal Growth Longitudinal Study.

[+] Data missing for 45% (paternal height) and 32% (hemoglobin <15 weeks of gestation) of the sample. All other variables with <1% missing data.

[‡] PPROM = Preterm Prelabor Rupture of Membranes.

[§] NICU = Neonatal Intensive Care Unit.

**Table 2.** Anthropometric measures at 2 years of age in children who contributed data to the fetal trans-cerebellar diameter growth and Sylvian fissure maturation international standards compared with the World Health Organization (WHO) Child Growth Standards.[*53]

| | | Current study | | WHO Child Growth Standards | |
|---|---|---|---|---|---|
| **Variable** | **N** | **Mean[+] ± SD[‡]** | **Median (interquartile range)** | **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 compared with the WHO Child Growth Standards.

[+] Mean values were estimated from raw data

[‡] SD = Standard Deviation.

**Table 3.** Morbidity at 1 and 2 years of age of children who contributed data to the international fetal trans-cerebellar diameter and Sylvian fissure maturation standards in the current study.

| **Medical condition** | **1 year of age** | **2 years of age** |
|---|---|---|
| | N= 1115 (%) | N = 1115 (%) |
| Hospitalized at least once | 123 (11) | 100 (9) |
| Total no. of days hospitalized * | 2 (1, 4) | 2 (1, 3) |
| Any prescription made by a healthcare professional | 704 (63.1) | 660 (59.2) |

| | | |
|---|---|---|
| Antibiotics (≥ 3 regimens) | 76 (6.8) | 132 (11.8) |
| Iron/folic acid/vitamin B12/other vitamins | 452 (40.5) | 183 (16.4) |
| Up-to-date with local vaccination policies | 1062 (95.3) | 1051 (94.3) |
| Otitis media/pneumonia/bronchiolitis | 66 (5.9) | 84 (7.5) |
| Parasitosis/diarrhea/vomiting | 42 (3.8) | 42 (3.8) |
| Seizures/cerebral palsy/neurological disorders | 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) |
| Meningitis | 1 (0.1) | 0 (0.0) |
| Other infections that required antibiotics | 16 (1.4) | 31 (2.8) |
| Hearing problems | 0 (0) | 0 (0.0) |
| Asthma | 12 (1.1) | 13 (1.2) |

| | | |
|---|---|---|
| Cardiovascular problems | 0 (0) | 0 (0.0) |
| Blindness | 0 (0) | 0 (0.0) |
| Gatroesophageal 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 allergies | NA‡ | 13 (1.2) |
| Injury trauma | 13 (1.2) | 26 (2.3) |
| Any condition that required surgery | 14 (1.3) | 8 (0.7) |

*Data are given as median (interquartile range).

+ UTI = Urinary tract infection

‡ NA, not applicable (data were not collected at the 1-year follow-up visit)

**Table 4**. Variance component analysis for trans-cerebellar diameter and Sylvian fissure maturation score.

| | Trans-cerebellar diameter (n=3,016) | | Sylvian fissure score (n=2,359) | |
|---|---|---|---|---|
| | **Variance estimate** | **Proportion of variance** | **Variance estimate** | **Proportion of variance** |
| **Variance between sites** | 0.005 | 0.2% | 0.005 | 2.2% |
| **Variance between individual within a site** | 0.330 | 10.9% | 0.007 | 3.3% |
| **Residual variance** | 2.706 | 89.0% | 0.206 | 94.5% |