Ultrasound Obstet Gynecol 2025; **65**: 703–711
Published online 5 May 2025 in Wiley Online Library (wileyonlinelibrary.com). **DOI:** 10.1002/uog.29227.

Transvaginal ultrasound assessment of corpus callosal length in the fetus: multicenter cross-sectional study

D. PALADINI[1], S. PARODI[2], H. XIE[3], F. VIÑALS[4], K. HARATZ[5], R. BIRNBAUM[5], G. AZUMENDI[6], L. POMAR[7], E. MONTAGUTI[8], P. ACHARYA[9], P. VOLPE[10], M. PÉREZ-CRUZ[11], K. KARL[12], R. CHAOUI[13], R. POOH[14] and Collaborators†

[1]*Fetal Medicine & Surgery Unit, IRCCS Istituto Giannina Gaslini, Genoa, Italy;* [2]*Unit of Epidemiology and Biostatistics, Scientific Directorate, IRCCS Istituto Giannina Gaslini, Genoa, Italy;* [3]*Department of Ultrasonic Medicine, The First Affiliated Hospital of Sun Yat-sen University, Guangzhou, Guangdong, China;* [4]*Ultrasound Department, Sanatorio Aleman Clinic, Concepción, Chile, and Obstetrics Department, Faculty of Medicine, University of Concepción, Concepción, Chile;* [5]*OB-GYN Ultrasound Unit, Lis Maternity Hospital, Tel Aviv Sourasky Medical Center, Tel Aviv, Israel, and Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel;* [6]*Ultrasound and Fetal Medicine Unit, Centro Gutenberg, Málaga, Spain;* [7]*Ultrasound and Fetal Medicine Unit, Department Woman-Mother-Child, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland, and School of Health Sciences (HESAV), HES-SO University of Applied Sciences and Arts Western Switzerland, Lausanne, Switzerland;* [8]*Obstetric Unit, IRCCS Azienda Ospedaliero-Universitaria di Bologna, Bologna, Italy;* [9]*Paras Advanced Centre for Fetal Medicine, Ahmedabad, India;* [10]*Fetal Medicine Unit, Di Venere Hospital, Department of Human Reproduction, Bari, Italy;* [11]*BCNatal Fetal Medicine Research Center, Hospital Clínic and Hospital Sant Joan de Déu, University of Barcelona, Barcelona, Spain, and Institut de Recerca Sant Joan de Déu, Esplugues de Llobregat, Barcelona, Spain, and Primary Care Interventions to Prevent Maternal and Child Chronic Diseases of Perinatal and Developmental Origin Network (RICORS), Instituto de Salud Carlos III, Madrid, Spain;* [12]*Center for Prenatal Diagnosis Munich, Munich, Germany;* [13]*Feindiagnostik, Berlin, Germany;* [14]*Fetal Brain Center, CRIFM Prenatal Medical Clinic, Osaka, Japan*

KEYWORDS: biometry; corpus callosum; fetus; neurosonography; three-dimensional ultrasound

ABSTRACT

Objective *To produce reference ranges and Z-scores for corpus callosal (CC) length in the fetus, based on transvaginal three-dimensional (3D) ultrasound imaging.*

Methods *This was a cross-sectional multicenter retrospective study based on 3D volume dataset acquisitions of the fetal CC between the 15th and 37th weeks of gestation. Only volume datasets acquired transvaginally through the anterior fontanelle were selected. After plane alignment on multiplanar imaging, the length of the CC was measured edge-to-edge on magnified images. Intra- and interobserver variability were assessed and the related intraclass correlation coefficients (ICC) calculated. Biometric charts to assess the reference values for fetal CC were obtained using the method proposed by Altman in 1993.*

Results *The 13 participating centers provided valid data for 2131 patients. Excellent agreement was observed for both intra- and interobserver analysis, with an ICC range of 0.98–1.00. A quadratic model was used for construction of the reference charts, modified with the insertion of cubic spline coefficients with a single knot at 18 gestational weeks, to recover an apparent lack of fit at lower gestational ages. Centile reference values and the corresponding Z-scores were produced for CC length between 15 and 37 gestational weeks.*

Conclusions *This multicenter study presents growth charts for the fetal CC, addressing the critical methodological weaknesses of several previous studies. An even distribution of cases across all gestational weeks, robust statistical methodology and a standardized, high-resolution transvaginal neurosonographic technique represent key factors supporting the reliability of the biometric curves presented here. © 2025 The Author(s). Ultrasound in Obstetrics & Gynecology published by John Wiley & Sons Ltd on behalf of International Society of Ultrasound in Obstetrics and Gynecology.*

INTRODUCTION

The corpus callosum (CC), which represents the largest interhemispheric commissure of the human brain, is divided into four anatomical components: rostrum, genu, body and splenium. Its development is thought to begin at about 13 gestational weeks, when some guiding axons cross the midline at the level of the commissural plate of

Correspondence: Prof. D. Paladini, Via Zara 24B, 16145, Genoa, Italy (e-mail: dpaladini49@gmail.com)
†Collaborators are listed at the end of the article.

Accepted: 12 March 2025

the lamina terminalis[1]. The CC is anatomically complete by 20 weeks and thereafter grows regularly throughout pregnancy.

The classification of CC abnormalities is controversial, both in the pediatric population and in the fetus, with few certainties and several inconsistencies. In particular, in the fetus, the only unambiguous definitions are complete and partial agenesis of the CC, in which, respectively, all or some of the four segments of the CC are absent. Regrettably, the definition of other rather questionable 'pathologies' of the fetal CC (including thick CC, short CC, hypoplastic (thin) CC) is based necessarily on morphometric assessment. The issue is exacerbated by the considerable heterogeneity among the various charts published on CC biometry, as highlighted in a recent meta-analysis[2]. In particular, it was emphasized that most of them involve several methodological weaknesses. Thus, although we strongly believe that the definition of CC pathological entities ideally should be based on an anatomical approach, we aimed to produce robust reference ranges for CC length based on high-quality images and measurements.

The aim of this multicenter study was to produce reference ranges and *Z*-scores for CC length in the fetus, based on transvaginal three-dimensional (3D) ultrasound imaging.

METHODS

Study design and population

This was a cross-sectional retrospective study based on 3D volume dataset acquisitions of the fetal CC between the 15th and 37th weeks of gestation. Thirteen fetal medicine units participated in the study: nine in Europe, three in Asia and one in South America (Table S1). Inclusion and exclusion criteria for the study are shown in Table 1. 3D volume datasets of the CC were acquired for routine assessment or detailed assessment of the fetal brain for research purposes. Only one examination per fetus was included in the study. This was a clinical audit of fully anonymized data; as such, it fits the definition of a quality improvement study and therefore was exempt from ethical evaluation[3]. All neonates were confirmed not to have any major anomaly and had a normal neurological clinical examination on discharge from hospital.

Ultrasound methodology and measurements

Equipment

The following ultrasound equipment was used in the fetal medicine units participating in the study over the study period: Voluson E8, E10 and Expert 22 ultrasound machines (GE Healthcare, Zipf, Austria), equipped with a 5–9-MHz or 6–12-MHz volumetric transvaginal transducer. The 3D volume datasets of the fetal CC were acquired between January 2011 and February 2024. Measurements were taken for the purposes of this study between October 2023 and March 2024.

Neurosonographic technique

3D volume datasets were acquired using transvaginal ultrasound according to the following methodology. The transducer was advanced until it indented the anterior fontanelle or the unossified sagittal suture, as per international guidelines on neurosonography, with the free hand of the operator placed on the maternal abdomen to immobilize the fetal head against the footprint of the transducer[4]. The midsagittal plane of the fetal head was displayed and the acquisition triggered. The woman was asked to hold her breath during the acquisition if significant movement of the fetal head synchronous with maternal breathing was noted. The volume dataset was then evaluated for possible motion artifacts; if present, these would lead to exclusion of the volume from the analysis. Acquisition parameters were: high quality and sweep angle of 50–70°, depending on gestational age (GA).

Measurement methodology

For the purposes of this study all measurements were performed by fetal medicine experts using either a dedicated software platform (4D-View; GE Healthcare) or the ultrasound equipment, as follows. Each 3D volume was opened, the three orthogonal planes aligned and the image rotated so that, in the A window, the frontal part of the fetal head was on the left side of the image. The image was then magnified. The operators were requested to measure the CC length in the multiplanar image, to ensure perfect plane alignment. After appropriate

Table 1 Inclusion and exclusion criteria for study population to produce reference curves for corpus callosal length measurement in the fetus on transvaginal ultrasound imaging

Inclusion criteria
Singleton pregnancy
Gestational age range: 15–37 weeks
Maternal good health
Physiological course of pregnancy
Gestational age confirmed by first-trimester ultrasound
Normal fetal growth
Absence of congenital anomalies at ultrasound and after birth
Normal neurological status at discharge from hospital
Availability of 3D volume datasets acquired transvaginally from midsagittal plane of the fetal head, with no motion/shadowing artifacts
Exclusion criteria
Multiple pregnancy
Gestational age: < 15 or > 37 weeks
Maternal pregestational (e.g. insulin-dependent diabetes) or pregnancy-related (e.g. pre-eclampsia) disease
Absence of dating scan
Growth abnormalities: fetal growth restriction or large-for-gestational age
Diagnosis of congenital anomalies during pregnancy and/or at birth
Abnormal neurological neonatal examination
Transabdominal 3D volume datasets, or transvaginal dataset but with motion/shadowing artifacts

3D, three-dimensional.

magnification, the maximum CC length was then measured twice, edge-to-edge ('on-to-on'): first, on the native two-dimensional (2D) grayscale image (CC-2D), and then adding volume contrast imaging (CC-VCI), with a slice thickness of 1 mm (Figure 1). Each image, with measurements, was then stored for quality control.

Repeatability

Intra- and interobserver variability were assessed as follows. A random sequence of 20 numbers was generated online and used to select randomly 20 cases for repeatability analysis from each participating center. For each case, the first operator measured the CC length, recorded the value and then deleted the measurement. A second operator, blinded to the result of the first measurement, then performed the same measurement, recorded the value and deleted the measurement. The first operator then remeasured the CC length, blinded both to the results of the second operator and to his/her own previous measurement. In this way, intra- and interobserver variability was assessed for each participating center. The procedure was then repeated adding VCI, with a slice thickness of 1 mm.

Data recording

Each center was provided with an Excel spreadsheet (Microsoft Corp., Redmond, WA, USA) to record anonymized data for each case. In addition to CC length with and without VCI (CC-VCI and CC-2D, respectively), the following variables were recorded: GA, biparietal diameter (BPD, in mm) and head circumference (HC, in mm).

Statistical analysis

Absolute frequencies and percentages are reported to describe qualitative variables, while median values and the related interquartile range (IQR) were calculated for quantitative variables. The Wilcoxon rank-sum test was used to compare coupled measurements, i.e. variables measured in the same patient. Spearman's ρ correlation coefficient was applied to assess the association between quantitative variables. Both the intra- and interobserver agreement in CC measurements were assessed by calculating the related intraclass correlation coefficients (ICC), producing Bland–Altman plots and performing Pitman's test[5,6].

Biometric charts to create the reference values for CC were obtained by the method proposed by Altman in 1993[7]. Briefly, a set of low-degree polynomial regression models were fitted using the two CC length measurements as dependent variables and GA as predictor. The polynomial degree was allowed to vary from 1 (simple linear model) to 4, and the goodness-of-fit of each model was measured by calculating the adjusted R^2 value[7]. The normality of the residuals was evaluated by visual inspection of the related histogram and the normal probability plot. If there was a lack of fit at the extremes of the GA distribution, we planned to add cubic splines to Altman's model, selecting the number and position of the nodes by visual inspection of the related scatterplot. SDs of the obtained measurements, adjusted by GA, were estimated by fitting a quadratic model on the absolute residuals[7], and the reference centile curves for CC (at 10% and 5%) were obtained by exploiting the properties of the normal distribution. Internal validation was done by comparing the fitted values with those obtained by the quantile regression model recently proposed by Muggeo *et al.*[8], which did not assume a normal distribution for the CC length measurements. Furthermore, the proportion of values lying outside the 2.5th to 97.5th centile range was compared with the expected proportion under the normal-distribution assumption. All the analyses were repeated using BPD and HC as dependent variables instead of GA. Finally, fitted values were also compared

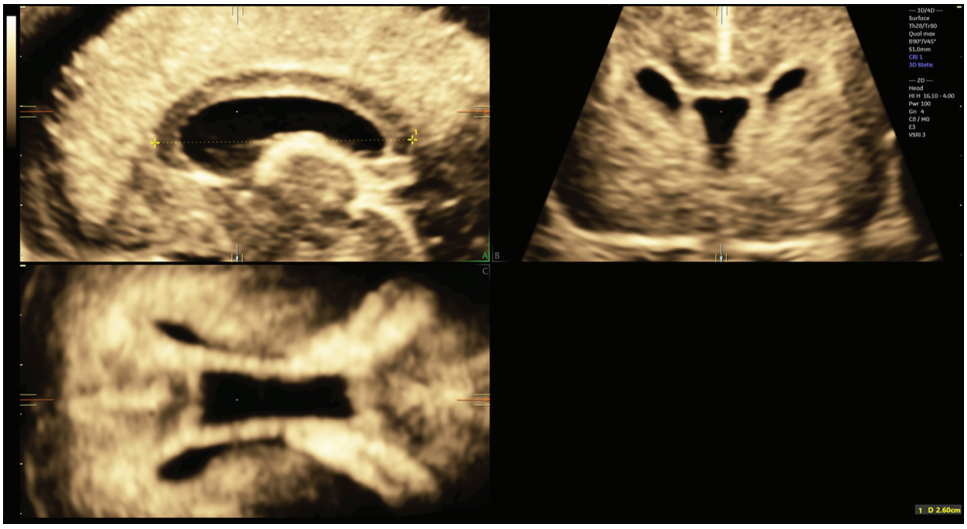

Figure 1 Multiplanar display of fetal corpus callosum (CC) at 22 + 3 weeks' gestation, from a three-dimensional volume dataset acquired transvaginally, with volume contrast imaging (1-mm slice). Maximum CC length was measured in the midsagittal plane (plane A) with calipers positioned on-to-on (edge-to-edge). Plane B is coronal plane and plane C is axial plane.

with those obtained by similar studies[9–20], recently summarized in the review by Corroenne *et al.*[2].

Centile regression was performed by the *quantreg* package developed in *R* (*R* Foundation, Vienna, Austria)[21,22]. All other analyses were performed using Stata for Windows 18.0 statistical software (Stata Corp., College Station, TX, USA). $P < 0.05$ was considered to indicate statistical significance.

RESULTS

The participating centers provided valid data for both measurements of CC length (CC-2D and CC-VCI) from a total of 2131 patients. The number of cases measured per gestational week is shown in Figure 2. Twenty patients selected randomly from all except one center were included in the assessment of reliability; Center 11 provided only 19 valid measurements. Results of the reliability analysis are given in Table 2. Excellent agreement was observed for both the intra- and the interobserver analyses, with ICCs ranging between 0.98 and 1.00 across the centers for the CC-2D measurements and between 0.99 and 1.00 for the CC-VCI measurements. The Bland–Altman analyses of intra- and interobserver reliability for CC-2D are summarized in Figures S1 and S2, respectively, and the corresponding analyses for CC-VCI are shown in Figures S3 and S4, respectively. With one exception, for both CC-2D and CC-VCI measurements there were no violations of the assumption of the homogeneity of variances, assessed by Pitman's test, for interobserver or intraobserver reliability; the exception was for the interobserver analysis of CC-2D in Center 9 ($P = 0.044$, Figure S2i).

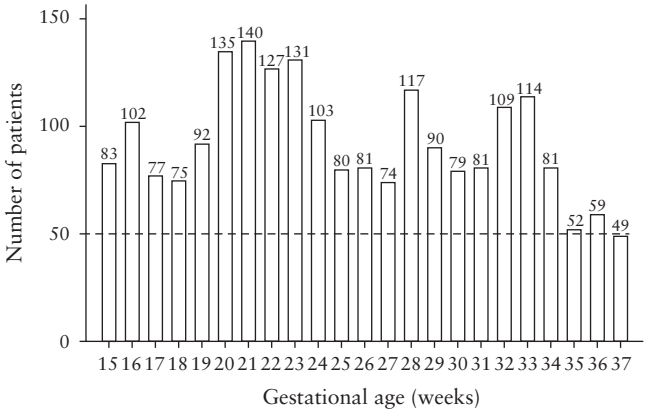

Figure 2 Histogram showing number of patients with transvaginal ultrasound three-dimensional neurosonographic volume datasets used to measure fetal corpus callosal length at each gestational week. Total number of cases: 2131; number of cases per week ranged from 49 to 140.

The CC-2D measurements did not differ significantly from the CC-VCI ones. The corresponding mean values were 27.4 mm and 27.5 mm, respectively, and the median values were 29.9 (IQR, 19.2–37.6) mm and 29.8 (IQR, 19.2–37.5) mm, respectively ($P = 0.344$). The correlation between the two types of measurement was very high (Spearman's $\rho = 0.998$, $P < 0.001$) and the corresponding ICC was 0.999, indicating that CC-2D and CC-VCI measurements were almost identical. Accordingly, for the rest of the results, we present only those for the CC-2D measurements.

Results of the polynomial regression analyses of CC-2D *vs* GA are summarized in Table S2. Regression coefficients were highly statistically significant up to the fourth degree, although the violation of the assumption of the homogeneity of variance (homoscedasticity) prevented the drawing of reliable conclusions about statistical inference for coefficients[6,22,23]. The adjusted R^2 value increased from 0.894 in the linear model to 0.964 in the quadratic model and remained almost unchanged in the cubic and fourth-degree models, suggesting that a quadratic model would be appropriate for construction of the biometric curves for CC length *vs* GA. However, in spite of the excellent overall goodness-of-fit, a lack of fit was evident at low values of GA, which remained unchanged in the cubic regression model (Figures S5 and S6). In order to

Table 2 Intra- and interobserver agreement for measurement of fetal corpus callosal (CC) length, expressed as intraclass correlation coefficient, according to center and measurement method

	CC-2D		*CC-VCI*	
Center	*Intraobserver*	*Interobserver*	*Intraobserver*	*Interobserver*
1	1.00 (1.00–1.00)	1.00 (1.00–1.00)	1.00 (1.00–1.00)	1.00 (0.99–1.00)
2	1.00 (0.99–1.00)	1.00 (0.99–1.00)	1.00 (0.99–1.00)	1.00 (0.99–1.00)
3	0.98 (0.92–0.99)	0.99 (0.97–1.00)	1.00 (0.99–1.00)	0.99 (0.98–1.00)
4	0.99 (0.96–1.00)	0.99 (0.96–0.99)	1.00 (0.99–1.00)	0.99 (0.97–1.00)
5	1.00 (0.99–1.00)	0.99 (0.97–1.00)	1.00 (0.99–1.00)	0.99 (0.98–1.00)
6	1.00 (0.99–1.00)	1.00 (0.99–1.00)	1.00 (0.99–1.00)	1.00 (0.99–1.00)
7	1.00 (0.99–1.00)	1.00 (0.99–1.00)	0.99 (0.98–1.00)	1.00 (0.99–1.00)
8	1.00 (0.99–1.00)	1.00 (0.99–1.00)	1.00 (0.99–1.00)	1.00 (0.99–1.00)
9	1.00 (0.99–1.00)	1.00 (0.99–1.00)	1.00 (0.99–1.00)	0.99 (0.98–1.00)
10	1.00 (0.99–1.00)	1.00 (0.99–1.00)	1.00 (0.99–1.00)	1.00 (0.99–1.00)
11	1.00 (0.99–1.00)	1.00 (0.99–1.00)	1.00 (0.99–1.00)	1.00 (0.99–1.00)
12	0.99 (0.98–1.00)	0.99 (0.97–1.00)	0.99 (0.98–1.00)	0.99 (0.97–0.99)
13	1.00 (0.99–1.00)	1.00 (0.99–1.00)	1.00 (0.99–1.00)	1.00 (0.99–1.00)

Values in parentheses are 95% CI. CC-2D, CC length measured on native two-dimensional image from three-dimensional (3D) multiplanar imaging datasets; CC-VCI, CC length measured after addition of 3D volume contrast imaging.

recover this lack of fit, we modified the Altman approach slightly by adding cubic spline coefficients to the regression models, arbitrarily selecting a single knot at 18 weeks' gestation (Figure S7). Even though only a marginal increase in the overall goodness-of-fit of the regression model was observed (adjusted $R^2 = 0.966$; Table S3), the lack of fit at low GAs was recovered (Figure S8). In addition, the distribution of the residuals showed very good agreement with a normal distribution (Figure S9), indicating that untransformed values of CC-2D were suitable for construction of the corresponding biometric curves[7]. The results in Figures S7 and S8 also indicate that the unexplained variance tends to increase with increasing GA. Accordingly, still following Altman's approach[7], we regressed the absolute values of the residuals *vs* GA using a quadratic regression model, the results of which are shown in Figure S10 and Table S4. SD estimates of the CC-2D measurements by GA values were then obtained by applying Altman's formula to the fitted values of the model. These were used to obtain the centile reference values and the corresponding *Z*-scores, which are shown in Tables 3 and 4, respectively. The corresponding biometric curve is shown in Figure 3. A simple calculator for CC length *Z*-scores according to GA, BPD and HC is available to download online (Appendix S1).

The number of observations outside the 5th to 95th centile range (Figure 3) was 213 (118 above the 95th centile and 95 below the 5th centile curve), corresponding to 10.0% of the whole cohort; similarly, values outside the 2.5th to 97.5th centile range were 61 and 50, respectively, corresponding to 5.2% of the entire cohort. Furthermore, the expected values of CC-2D were very close to those obtained by the non-parametric quantile regression analysis using the model of Muggeo *et al.*[8], which did not assume a normal distribution for the CC length measurements (ICC, 0.997 (95% CI, 0.913–0.999)) (Figures S11 and S12).

Figure 4 shows the fitted values of the cubic spline regression model for CC-2D according to GA and the corresponding estimates reported by the studies included in the systematic review by Corroenne *et al.*[2,9–20]. Our values were quite consistent with those of the other models for GA ≥ 25 weeks, whereas between 16 and 24 weeks, estimates were highly heterogeneous. The largest difference was found between our model and that of Rizzo *et al.*[14], while estimates provided by Malinger and Zakut[9] were very close to our estimates, although their data were available only for GA ≥ 19 weeks.

Analysis of fetal CC length according to BPD and HC provided similar results. Since the same statistical methodology used for regression *vs* GA was employed in the assessment of the relationship of fetal CC length with BPD and HC, we have reported the entire results section related to BPD and HC in Appendix S2, which summarizes the results illustrated in Tables S5–S11 and in Figures S13–S21.

DISCUSSION

In this multicenter study, we provide 3rd to 97th centile reference values and the corresponding *Z*-scores for CC length *vs* GA (Tables 3 and 4), BPD and

Table 3 Centile reference values for fetal corpus callosal (CC) length, measured on native two-dimensional ultrasound images, according to gestational age (GA)

	CC length (mm)						
GA (weeks)	*2.5th centile*	*5th centile*	*10th centile*	*50th centile*	*90th centile*	*95th centile*	*97.5th centile*
15	1.375	1.689	2.051	3.329	4.606	4.969	5.283
16	2.742	3.110	3.533	5.028	6.522	6.946	7.313
17	5.479	5.897	6.380	8.081	9.782	10.264	10.683
18	8.939	9.403	9.939	11.827	13.715	14.250	14.715
19	12.251	12.760	13.346	15.414	17.482	18.068	18.577
20	15.255	15.805	16.439	18.674	20.910	21.543	22.093
21	17.788	18.373	19.047	21.427	23.807	24.481	25.066
22	20.268	20.888	21.604	24.128	26.652	27.367	27.988
23	22.391	23.043	23.794	26.443	29.093	29.844	30.495
24	24.510	25.193	25.980	28.756	31.533	32.320	33.003
25	26.251	26.960	27.776	30.658	33.540	34.357	35.065
26	27.807	28.539	29.382	32.357	35.333	36.176	36.908
27	29.189	29.941	30.808	33.865	36.922	37.789	38.541
28	30.446	31.215	32.102	35.232	38.361	39.248	40.017
29	31.531	32.315	33.218	36.405	39.593	40.496	41.280
30	32.418	33.212	34.128	37.358	40.588	41.504	42.298
31	33.256	34.059	34.984	38.249	41.514	42.439	43.242
32	34.040	34.849	35.781	39.070	42.358	43.291	44.099
33	34.709	35.521	36.456	39.756	43.056	43.992	44.803
34	35.353	36.164	37.100	40.399	43.699	44.635	45.446
35	35.977	36.785	37.717	41.004	44.290	45.222	46.030
36	36.554	37.356	38.281	41.542	44.803	45.728	46.530
37	37.080	37.873	38.788	42.015	45.242	46.157	46.950

Table 4 *Z*-score (SD) reference values for fetal corpus callosal (CC) length, measured on native two-dimensional ultrasound images, according to gestational age (GA)

	CC length Z-score						
GA (weeks)	*−3 SD*	*−2 SD*	*−1 SD*	*0 SD*	*1 SD*	*2 SD*	*3 SD*
15	0.338	1.335	2.332	3.329	4.326	5.323	6.320
16	1.529	2.695	3.862	5.028	6.194	7.360	8.526
17	4.099	5.426	6.753	8.081	9.408	10.736	12.063
18	7.407	8.880	10.353	11.827	13.300	14.774	16.247
19	10.573	12.187	13.800	15.414	17.028	18.641	20.255
20	13.441	15.186	16.930	18.674	20.418	22.163	23.907
21	15.856	17.713	19.570	21.427	23.284	25.141	26.997
22	18.220	20.189	22.158	24.128	26.097	28.066	30.036
23	20.241	22.308	24.376	26.443	28.511	30.578	32.645
24	22.257	24.423	26.590	28.756	30.923	33.090	35.256
25	23.912	26.161	28.410	30.658	32.907	35.155	37.404
26	25.392	27.714	30.036	32.357	34.679	37.001	39.322
27	26.708	29.094	31.479	33.865	36.251	38.636	41.022
28	27.907	30.348	32.790	35.232	37.673	40.115	42.557
29	28.944	31.431	33.918	36.405	38.893	41.380	43.867
30	29.796	32.317	34.838	37.358	39.879	42.399	44.920
31	30.607	33.154	35.702	38.249	40.796	43.344	45.891
32	31.371	33.937	36.504	39.070	41.636	44.202	46.768
33	32.031	34.606	37.181	39.756	42.331	44.906	47.481
34	32.675	35.250	37.825	40.399	42.974	45.549	48.124
35	33.310	35.875	38.439	41.004	43.568	46.132	48.697
36	33.907	36.452	38.997	41.542	44.087	46.632	49.177
37	34.461	36.979	39.497	42.015	44.533	47.051	49.569

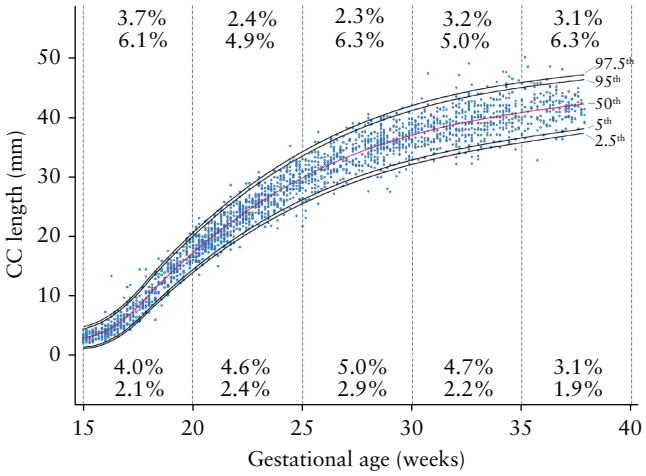

Figure 3 Biometric curve of fetal corpus callosal (CC) length, measured on native two-dimensional ultrasound images from three-dimensional multiplanar imaging datasets, without volume contrast imaging, plotted against gestational age. 2.5th, 5th, 50th, 95th and 97.5th centile curves are shown. Related percentages of cases above 97.5th and 95th centiles and below 5th and 2.5th centiles are shown for five gestational-age classes.

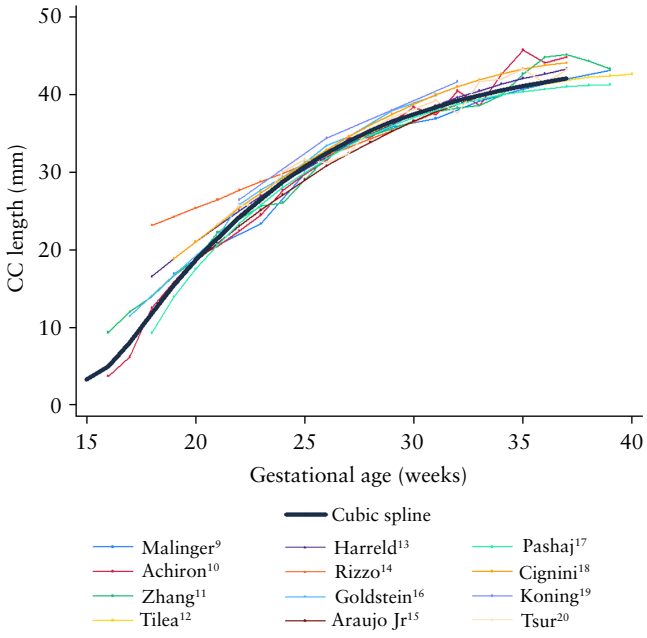

Figure 4 Fitted values of cubic spline regression model for corpus callosal (CC) length, measured on native two-dimensional ultrasound images from three-dimensional multiplanar imaging datasets, without volume contrast imaging, plotted against gestational age, with corresponding estimates of CC length according to gestational age reported by studies included in the systematic review by Corroenne *et al.*[2]. Only first author of each study is given. One implausible value in the data of Malinger and Zakut[9], probably due to a typo, has been corrected.

HC (Tables S8–S11), all developed with robust statistical methodology. We have addressed several of the methodological issues that have led over the years to the significant discrepancies among studies reporting CC biometry, highlighted in the recent review by Corroenne *et al.*[2]. In particular, our study relied on clearly defined

inclusion and exclusion criteria (Table 1), mandatory early-pregnancy dating by ultrasound and a standardized neurosonographic technique. The latter is of the utmost importance, as several studies have used both a transabdominal and a transvaginal approach[14,15,17,18] or a purely transabdominal one[11], which is questionable in terms of resolution, as well as recognition of the anatomical landmarks of the CC and, hence, reproducibility.

Another critical issue highlighted in the review of Corroenne *et al.*[2] is the limitations in methodology, with regard to both the GA distribution of cases and the statistical approach used. In one study, with one of the largest sample populations[18], the number of cases per gestational week ranged between three and 1322, with 2324 of the 2950 (78.8%) cases being confined to gestational weeks 21 and 22. In our multicenter study, the number of cases per gestational week ranged between 49 and 140, and this factor, together with the large and homogeneous sample population, ensured robust results. In fact, the number of observations outside the 5th to 95thand the 2.5th to 97.5th centile ranges (Figure 3) were 10.0% and 5.2% of the entire cohort, respectively, which is in very good agreement with the results expected from the theoretical normal distribution, thus confirming the high accuracy of our model.

An important point to consider is the need to insert a cubic spline coefficient in the regression model, to recover the lack of fit at low GA values, regardless of the excellent overall goodness-of-fit (Figures S5 and S6). The different behavior of the two parts of the curve (below and above 18 weeks) is interesting, because it may relate to a possible differential growth pattern of the CC in the earlier gestational period. In fact, it has recently been proposed that the CC grows initially following a bidirectional vector, with elongation occurring both anteriorly and posteriorly[24]. Hence, it may be speculated that the differential steepness of the CC growth chart before and after 18 weeks (Figures 3 and S7) may reflect, at least in part, a different growth pattern.

With respect to sonographic technique, the methodology described herein is a key factor responsible for the exceedingly high reliability reported (Table 2). In fact, two mandatory inclusion criteria for the sonographic assessment were a transvaginal approach and the use of 3D ultrasound (Table 1). The latter warranted attainment of a perfect midsagittal plane thanks to multiplanar image correlation (Figure 1), which is not possible with 2D imaging, since the other two planes are not available to confirm perfect alignment of the three spatial planes. Hence, we have produced the first chart of fetal CC length based exclusively on robust transvaginal imaging methodology.

Strengths of this study are the large sample population, the strict inclusion and exclusion criteria, the robust statistical methodology, the homogeneous sonographic technique and an even spread of cases across gestation. Three limitations of the study are its retrospective design, i.e. that pregnancies were not enrolled prospectively for the objective of the study, the lack of long-term neurological follow-up of the neonates, which may have prevented diagnosis of late-onset neurodevelopmental delay, and the fact that all 3D volume acquisitions were performed by experts in referral centers. However, this last point, in our opinion, could also be regarded as a strength. In fact, we strongly believe that the CC should be assessed only on targeted neurosonography[4] and not in the screening setting.

At present, most fetal neurology experts believe that the CC should be assessed only subjectively, and that measurement is not advised. Until now, we subscribed to this view, particularly because of the disagreement in biometric studies thus far available[9–20]. However, we have now produced robust data that provide researchers with the opportunity to evaluate the role of biometry in the diagnosis of CC anomalies in future studies. We focused on the length of the CC, as there is no evidence that other antenatal measurements, such as CC thickness, have a clear association with congenital anomalies. The clinical benefit of using these biometry charts is two-fold: first, their use should prevent or at least reduce the chances of a false-positive diagnosis of CC dysplasia and partial agenesis; second, they should enable prospective studies on the neurological development of individuals with a shortened or dysplastic CC, based on a robust and unequivocal biometric assessment. It should be emphasized that these charts were developed for transvaginal neurosonographic assessment of the CC and cannot be used for transabdominal CC measurements. However, the CC should only be measured in targeted neurosonography; the only situation in which transabdominal ultrasound measurement of the CC is indicated is in the context of a neurosonographic assessment for which the transvaginal approach is impossible for some reason.

In conclusion, we have produced reliable reference charts for CC length. Until a clear-cut definition of dysplastic CC is produced and more data are acquired on strong prognostic indicators, we believe that application of these charts will prevent the false-positive diagnosis of dysraphic/dysplastic/short CC, which until now has been based on quantitative criteria using inadequate biometric charts.

ACKNOWLEDGMENTS

For E. Eixarch and M. Perez Cruz, this study was funded by the Instituto de Salud Carlos III (ISCIII) through the project 'PI22/00754' and cofunded by the European Union. Open access funding provided by BIBLIOSAN.

Collaborators

We thank the following researchers for taking part in the reliability study as second operators:
F. Acharya, Paras Advanced Centre for Fetal Medicine, Ahmedabad, India
A. Acharya, Paras Advanced Centre for Fetal Medicine, Ahmedabad, India
I. Alonso, Ultrasound and Fetal Medicine Unit, Centro Gutenberg, Málaga, Spain

D. Baud, Lausanne University Hospital, Lausanne, Switzerland
M. Brusilov, OB-GYN Ultrasound Unit, Lis Maternity Hospital, Tel Aviv Sourasky Medical Center, Tel Aviv, Israel
E. Eixarch, BCNatal Fetal Medicine Research Center, Barcelona Center for Maternal-Fetal and Neonatal Medicine, Hospital Clínic de Barcelona and Hospital Sant Joan de Déu, Barcelona, Spain; Institut Clínic de Ginecología, Obstetrícia i Neonatología, Barcelona, Spain; Fundació de Recerca Clínica Barcelona, Institut d'Investigacions Biomèdiques August Pi I Sunyer, Universitat de Barcelona, Barcelona, Spain; Center for Biomedical Research on Rare Diseases, Institute of Health Carlos III, Madrid, Spain
Y. Jiang, Department of Ultrasonic Medicine, The First Affiliated Hospital of Sun Yat-sen University, Guangzhou, Guangdong, China
M. Megumi, Fetal Brain Center, CRIFM Prenatal Medical Clinic, Osaka, Japan
G. Pilu, Obstetric Unit, IRCCS Azienda Ospedaliero-Universitaria di Bologna, Bologna, Italy
A. O. Rifat, Center for Prenatal Diagnosis Munich, Munich, Germany
G. Quiroz, Sanatorio Aleman Clinic, Concepción, Chile
J. Sichitiu, Lausanne University Hospital, Switzerland
N. Volpe, Fetal Medicine Unit, Di Venere Hospital, Department of Human Reproduction, Bari, Italy

REFERENCES

1. Hochstetter F. Ueber die Nichtexistenz der sogenannten Bogenfurchen an den Gehirnen lebensfrisch konservierter menschlicher Embryonen. *Verhandl d Anat Ges, Anat Anz*. 1904;25:27-34.
2. Corroenne R, Grevent D, Kasprian G, et al. Corpus callosal reference ranges: systematic review of methodology of biometric chart construction and measurements obtained. *Ultrasound Obstet Gynecol*. 2023;62(2):175-184.
3. Casarett D, Karlawish JH, Sugarman J. Determining when quality improvement initiatives should be considered research: proposed criteria and potential implications. *JAMA*. 2000;83(17):2275-2280.
4. Paladini D, Malinger G, Birnbaum R, et al. ISUOG Practice Guidelines (updated): sonographic examination of the fetal central nervous system. Part 2: performance of targeted neurosonography. *Ultrasound Obstet Gynecol*. 2021;57(4):661-671.
5. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1(8476):307-310.
6. Bartlett JW, Frost C. Reliability, repeatability and reproducibility: analysis of measurement errors in continuous variables. *Ultrasound Obstet Gynecol*. 2008;31(4):466-475.
7. Altman D. Construction of age-related reference centiles using absolute residuals. *Stat Med*. 1993;12(10):917-924.
8. Muggeo VMR, Torretta F, Eilers PHC, Sciandra M, Attanasio M. Multiple smoothing parameters selection in additive regression quantiles. *Stat Model*. 2020;21(5):428-448.
9. Malinger G, Zakut H. The corpus callosum: normal fetal development as shown by transvaginal sonography. *AJR Am J Roentgenol*. 1993;161(5):1041-1043.
10. Achiron R, Achiron A. Development of the human fetal corpus callosum: a high-resolution, cross-sectional sonographic study. *Ultrasound Obstet Gynecol*. 2001;18(4):343-347.
11. Zhang H, Yang J, Chen Z, Ma X. Sonographic study of the development of fetal corpus callosum in a Chinese population. *J Clin Ultrasound*. 2009;37(2):75-77.
12. Tilea B, Alberti C, Adamsbaum C, et al. Cerebral biometry in fetal magnetic resonance imaging: new reference data. *Ultrasound Obstet Gynecol*. 2009;33(2):173-181.
13. Harreld JH, Bhore R, Chason DP, Twickler DM. Corpus callosum length by gestational age as evaluated by fetal MR imaging. *AJNR Am J Neuroradiol*. 2011;32(3):490-494.
14. Rizzo G, Pietrolucci ME, Capponi A, Arduini D. Assessment of corpus callosum biometric measurements at 18 to 32 weeks' gestation by 3-dimensional sonography. *J Ultrasound Med*. 2011;30(1):47-53.
15. Araujo Junior E, Visentainer M, Simioni C, Ruano R, Nardozza LM, Moron AF. Reference values for the length and area of the fetal corpus callosum on 3-dimensional sonography using the transfrontal view. *J Ultrasound Med*. 2012;31(2):205-212.
16. Goldstein I, Tamir A, Reece AE, Weiner Z. Corpus callosum growth in normal and growth-restricted fetuses. *Prenat Diagn*. 2011;31(12):1115-1119.
17. Pashaj S, Merz E, Wellek S. Biometry of the fetal corpus callosum by three-dimensional ultrasound. *Ultrasound Obstet Gynecol*. 2013;42(6):691-698.
18. Cignini P, Padula F, Giorlandino M, et al. Reference charts for fetal corpus callosum length: a prospective cross-sectional study of 2950 fetuses. *J Ultrasound Med*. 2014;33(6):1065-1078.
19. Koning IV, Roelants JA, Groenenberg IAL, et al. New ultrasound measurements to bridge the gap between prenatal and neonatal brain growth assessment. *AJNR Am J Neuroradiol*. 2017;38(9):1807-1813.
20. Tsur A, Weisz B, Rosenblat O, et al. Personalized charts for the fetal corpus callosum length. *J Matern Fetal Neonatal Med*. 2019;32(23):3931-3938.
21. Muggeo VMR. Additive Quantile regression with automatic smoothness selection: the R package quantregGrowth. 2021 https://www.researchgate.net/publication/350844895
22. R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing; 2023.
23. Zhou Y, Zhu Y, Wong WK. Statistical tests for homogeneity of variance for clinical trials and recommendations. *Contemp Clin Trials Commun*. 2023;33:101119.
24. Birnbaum R, Barzilay R, Brusilov M, Wolman I, Malinger G. The early pattern of human corpus callosum development: A transvaginal 3D neurosonographic study. *Prenat Diagn*. 2020;40(10):1239-1245.

SUPPORTING INFORMATION ON THE INTERNET

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

Table S1 Transvaginal assessment of corpus callosal length in the fetus: list of participating centers

Tables S2–S4 Regression coefficients and related adjusted R^2 value for low-degree polynomial regression models (Table S2), for cubic spline regression models (Table S3) and for cubic regression models of absolute residuals obtained from cubic spline regression models reported in Table S3 (Table S4) for fetal corpus callosal length measurements on native 2D ultrasound image *vs* gestational age

Tables S5–S7 Regression coefficients and related adjusted R^2 values for low-degree polynomial regression models (Table S5), for cubic spline regression models (Table S6) and for cubic regression models of absolute residuals obtained from cubic spline regression models reported in Table S6 (Table S7) for fetal corpus callosal length measurements on native 2D ultrasound image *vs* biparietal diameter and head circumference

Tables S8–S11 Centile (Tables S8 and S10) and *Z*-score (SD) (Tables S9 and S11) reference values for fetal corpus callosal length measurements on native 2D ultrasound images, according to biparietal diameter (Tables S8 and S9) and head circumference (Tables S10 and S11)

Figures S1–S4 Intraobserver (Figures S1 and S3) and interobserver (Figures S2 and S4) reliability for fetal corpus callosal length measurement on native 2D ultrasound image (Figures S1 and S2) and on 3D ultrasound volume contrast image (Figures S3 and S4), evaluated by Bland–Altman plot.

Figures S5 and S13 Fetal corpus callosal length measurements on native 2D ultrasound image, according to gestational age (Figure S5) and according to biparietal diameter and head circumference (Figure S13), obtained from the fitted values of quadratic and cubic polynomial models described in Table S2 and Table S4, respectively.

Figures S6 and S14 Residuals from cubic regression models for fetal corpus callosal length measurements on native 2D ultrasound image *vs* gestational age (Figure S6) and *vs* biparietal diameter and *vs* head circumference (Figure S14), plotted against corresponding fitted values.

Figures S7 and S15 Cubic spline regression models for fetal corpus callosal length measurements on native 2D ultrasound image *vs* gestational age with a single knot at 18 gestational weeks (Figure S7), and *vs* biparietal diameter with a single knot at 40 mm and *vs* head circumference with a single knot at 140 mm (Figure S15).

Figures S8 and S16 Residuals *vs* fitted values from cubic spline regression models for fetal corpus callosal length measurements on native 2D ultrasound image *vs* gestational age (Figure S8) and *vs* biparietal diameter and *vs* head circumference (Figure S16).

Figures S9 and S17 Histograms and normal quantile plots (Q-Q plots) of residual distribution from cubic spline regression models for fetal corpus callosal length measurements on native 2D ultrasound image *vs* gestational age (Figure S9) and *vs* biparietal diameter and *vs* head circumference (Figure S17).

Figures S10 and S18 Association between absolute residuals from cubic spline regression model for fetal corpus callosal length measurements on native 2D ultrasound image and gestational age (Figure S10) and biparietal diameter and head circumference (Figure S18).

Figures S11 and S20 Association between fitted values of fetal corpus callosal length measurements on native 2D ultrasound image *vs* gestational age (Figure S11) and *vs* biparietal diameter and *vs* head circumference (Figure S20) from cubic spline regression models and from non-parametric quantile regression models of Muggeo *et al.*[8].

Figures S12 and S21 Biometric curves of fetal corpus callosal length measurements on native 2D ultrasound image *vs* gestational age (Figure S12) and *vs* biparietal diameter and *vs* head circumference (Figure S21), obtained from non-parametric quantile regression model of Muggeo *et al.*[8].

Figure S19 Biometric curves of fetal corpus callosal length measurements on native 2D ultrasound image according to biparietal diameter (a) and head circumference (b). 2.5th, 5th, 50th, 95th and 97.5th centile curves are shown.

Appendix S1 Calculator for fetal corpus callosal length

Appendix S2 Results of the statistical analysis of fetal corpus callosal (CC) length *vs* biparietal diameter (BPD) and *vs* head circumference (HC)