



Published in final edited form as:

*AJNR Am J Neuroradiol.* 2015 July ; 36(7): 1369–1374. doi:10.3174/ajnr.A4357.

## Assessment of MRI-Based Automated Fetal Cerebral Cortical Folding Measures in Prediction of Gestational Age in the Third Trimester

J. Wu, S.P. Awate, D.J. Licht, C. Clouchoux, A.J. du Plessis, B.B. Avants, A. Vossough, J.C. Gee, and C. Limperopoulos

Department of Radiology (J.W., B.B.A., A.V., J.C.G.), Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania; Department of Computer Science and Engineering (S.P.A.), Indian Institute of Technology Bombay, Mumbai, India; Neurovascular Imaging Lab (D.J.L.), Children's Hospital of Philadelphia, Philadelphia, Pennsylvania; Advanced Pediatric Brain Imaging Research Laboratory (C.C., C.L.), Children's National Medical Center, Washington, DC; and Departments of Neurology, Radiology, and Pediatrics (C.C., A.J.d.P., C.L.), George Washington University School of Medicine and Health Sciences, Washington, DC

### Abstract

**BACKGROUND AND PURPOSE**—Traditional methods of dating a pregnancy based on history or sonographic assessment have a large variation in the third trimester. We aimed to assess the ability of various quantitative measures of brain cortical folding on MR imaging in determining fetal gestational age in the third trimester.

**MATERIALS AND METHODS**—We evaluated 8 different quantitative cortical folding measures to predict gestational age in 33 healthy fetuses by using T2-weighted fetal MR imaging. We compared the accuracy of the prediction of gestational age by these cortical folding measures with the accuracy of prediction by brain volume measurement and by a previously reported semiquantitative visual scale of brain maturity. Regression models were constructed, and measurement biases and variances were determined via a cross-validation procedure.

**RESULTS**—The cortical folding measures are accurate in the estimation and prediction of gestational age (mean of the absolute error,  $0.43 \pm 0.45$  weeks) and perform better than ( $P = .024$ ) brain volume (mean of the absolute error,  $0.72 \pm 0.61$  weeks) or sonography measures (SDs approximately 1.5 weeks, as reported in literature). Prediction accuracy is comparable with that of the semiquantitative visual assessment score (mean,  $0.57 \pm 0.41$  weeks).

**CONCLUSIONS**—Quantitative cortical folding measures such as global average curvedness can be an accurate and reliable estimator of gestational age and brain maturity for healthy fetuses

---

Please address correspondence to Jue Wu, PhD, University of Pennsylvania, 3600 Market St, Suite 370, Philadelphia, PA 19104; wojohn911@gmail.com.

J.C. Gee and C. Limperopoulos are co-senior authors and contributed equally to this article.

Disclosures: James C. Gee—*RELATED: Grant:* National Institutes of Health, \* University of Pennsylvania\*; *UNRELATED: Consultancy:* University of Texas Southwestern Medical Center. *Comments:* consultant on National Institutes of Health–funded project; *Grants/Grants Pending:* National Institutes of Health\*; *Stock/Stock Options:* part of retirement planning; *Travel/Accommodations/Meeting Expenses Unrelated to Activities Listed:* advisory board meetings, international task force meetings, academic lectures, grant review panels. \*Money paid to the institution.

in the third trimester and have the potential to be an indicator of brain-growth delays for at-risk fetuses and preterm neonates.

---

Gestational age (GA) estimation for unborn babies is crucial to any assessment of pregnancy, fetal development, and neonatal care. Obstetricians routinely use the last menstrual period to estimate the beginning date of gestation and calculate the estimated date of delivery. However, it has been reported that 20%–40% of women could not determine their last menstrual period with certainty due to various reasons such as bleeding in the first trimester and pregnancy following the use of oral contraceptives.<sup>1</sup> Fetal growth measurements from sonography biometry are used to estimate the gestation length, if one assumes a normative growth trajectory at the time of the examination.<sup>2</sup> The most widely used fetal biometrics assessed in the first 2 trimesters<sup>3</sup> have been shown to be more accurate than last menstrual period–based estimates.<sup>4,5</sup> Consequently, sonographic measurements are recommended in place of the last menstrual period if there is a large discrepancy between the last menstrual period and sonography-based estimates before 20 weeks of gestation.<sup>6</sup>

Sonography-based GA estimation in the third trimester (26 weeks to birth) should be interpreted with great caution due to remarkably increased variability in organ size in the later stages of pregnancy.<sup>3,7</sup> For example, biparietal diameter varies by approximately 7 days, 14 days, and 21–28 days when measured at 14–20 weeks, 21–30 weeks, and after 30 weeks GA, respectively.<sup>7</sup> Consequently, it has been recommended that menstrual dates be used to establish GA if it is within the error range of these biometric markers during the third trimester.<sup>3</sup>

This limitation in sonographic assessment stems mostly from the inherent variability of organ size and the intrinsic signal properties of ultrasonography. In this article, we attempt to address this important problem from a different perspective—brain cortical folding. During the second half of gestation, the cortical folding of the neocortex initiates at around 20 weeks and drastically changes the brain shape throughout the third trimester, from a largely smooth surface to a complex convoluted one.<sup>8</sup> The process appears to be genetically controlled and largely consistent across the healthy population.<sup>9,10</sup> Primary, secondary, and tertiary gyri sulci emerge in order, and newly evolved sulci related to higher functions, such as auditory, visual, and linguistic functions, appear later than phylogenetically more primitive allocortical folding, such as the hippocampus and olfactory sulcus.<sup>8</sup> Several MR imaging studies on healthy fetuses have found a linear relation between folding measures and GA.<sup>11,12</sup> There was evidence that female and male preterm neonates did not differ in folding while differing in size.<sup>13</sup> Data also showed that the sulcal depth of a folding measure was approximately 5% of the mean of the measure in healthy preterm neonates at around 32 weeks' GA.<sup>14</sup> In comparison, the sonography-based measure, biparietal diameter, has a sulcal depth (1.5 mm) equal to 16% of the average biparietal diameter (9 mm) at 30–36 weeks' GA.<sup>3</sup> The lower variance of folding when controlled for GA potentially leads to a better estimator of GA than sonography in late gestation.

## MATERIALS AND METHODS

### Subjects and Data Acquisition

The study cohort included healthy pregnant women who were recruited in a previously published study with local institutional review board approval.<sup>15</sup> Exclusion criteria included GA of >36 completed weeks, multiple gestations, congenital infection, gestational diabetes, or any maternal contraindication to MR imaging. We also excluded fetuses with sonographic findings of dysmorphic features, dysgenic brain lesions, or anomalies of other organ systems.<sup>15</sup> Written informed consent from mothers was obtained according to a protocol approved by Boston Children's Hospital Committee for Clinical Investigation, and the study was compliant with the Health Insurance Portability and Accountability Act. A cohort of 33 healthy fetuses with high-resolution reconstructed MRI was included in this study, of which 10 subjects were female and 23 were male. The cohort has been previously reported in studies of normative brain development<sup>16</sup> but has not been used to estimate gestational age.

The mean GA at the time of MR imaging was  $29.1 \pm 2.8$  weeks (range, 25.2–35.4 weeks). Estimated gestational age was based on maternal dates and first-trimester sonography measurements, if available, by the pregnant mother's referring obstetrician. The working gestational age as determined by the referring obstetrician at the time of MR imaging was used. All fetal MR imaging studies were performed on a 1.5T scanner (Achieva; Philips Healthcare, Best, the Netherlands) and a 5-channel phased array cardiac coil. Multiplanar single-shot turbo spin-echo imaging was performed (TE = 120 ms, TR = 12,500 ms, 0.625 signal averages, FOV = 330 mm, section thickness = 2-mm, no intersection gap, acquisition matrix =  $256 \times 204$ , acquisition time = 25–55 seconds). All fetal and postnatal MR imaging findings were normal. The Vineland Adaptive Behavior Scale<sup>17</sup> used to assess functional performance in communication, daily living, socialization, and motor skills was age-appropriate in the subjects between 18 and 24 months of age.

### Motion Correction and Manual Brain Segmentation

Spontaneous fetal motion during scanning poses a challenge for the computation of 3D measures. The imaging data used in this study were reconstructed to isotropic volumes ( $1 \text{ mm}^3$ ) on the basis of a superresolution method.<sup>18</sup> Other methods for correcting motion in fetal MR imaging exist,<sup>19,20</sup> but we did not compare the differences among them in this article.

We created a manual mask of the intracranial region for each subject. The mask excluded all the maternal tissue and nonbrain tissue of the fetus. In addition, we also defined a cerebral mask for each subject, which was constructed manually by tracing the interface between developing white matter and cortical gray matter. The mask region excluded the cerebellum and brain stem. The boundary of the mask was used as a representation of the cortex. This mask was also used to calculate the brain/cerebral volume of the subject. Masks for the cortical lobes (frontal, temporal, parietal, occipital, and insular) and left and right hemispheres were generated by registering subject images with a manually labeled template. This template was based on a publicly available atlas, brain-development.org ([www.brain-development.org](http://www.brain-development.org)),<sup>21</sup> which was an average of preterm infants at a postmenstrual age of 29

weeks. Manual tracings of lobes and hemispheres were performed and visually inspected in the free software ITK-SNAP ([www.itksnap.org](http://www.itksnap.org)).<sup>22</sup>

### Cortical Folding Measures

We implemented and evaluated 8 different quantitative measures of cortical folding that have been used in the literature.<sup>23–26</sup> These were average curvedness (AC), Gaussian curvature ( $L^2$ ) norm, mean curvature  $L^2$  norm, intrinsic curvature index, extrinsic curvature index, convexity ratio, isoperimetric index, and average sulcal depth. The convexity ratio and isoperimetric index are globally defined, while the other measures are an average over each point on the cortex. Gaussian curvature  $L^2$  norm, intrinsic curvature index, mean curvature  $L^2$  norm, convexity ratio, isoperimetric index, and AC are all invariant to translation, rotation, and scaling of the cortical surface.<sup>25,27</sup> The extrinsic curvature index is invariant to translation and scaling, but rotation may cause it to change sign. The convexity ratio is also shown to be area-independent (ie, invariant to the change of surface area in question).<sup>26</sup>

The implementation of the folding measure computations was based on a previously established processing pipeline designed for pediatric data.<sup>27,28</sup> The pipeline took in the cerebral mask and upsampled it to a higher resolution (eg,  $0.45^3\text{mm}^3$ ). Our numeric scheme for computing the quantitative folding measures depended on reliably estimating the cortical surface and sampling it fairly uniformly. The adopted numeric scheme did so by obtaining an implicit surface parameterization of the cortical surface, as a level set of a distance transform on a Cartesian grid. We used the sparse field level set representation, which was numerically efficient, on a grid of isotropic voxels with submillimeter dimension. A previous article<sup>28</sup> validated such a scheme for adult brains as well, which exhibited more complex and more folded cortical surfaces than fetal brains, thus indicating the suitability of the scheme for fetal brains. The flowchart of the processing pipeline is shown in Online Fig 1.

### Statistical Analysis

To examine whether the various cortical folding measures provided more information about GA than brain volume, we constructed a linear model with brain volume and each folding measure as 2 predictors of GA and compared the model with the reduced model with only brain volume as the predictor of GA. More specifically, we tested the null hypothesis by using analysis of variance that the following 2 linear regressions are equal:

1.  $GA = a_0 + a_1 \times (\text{Brain Volume}) + a_2 \times (\text{Folding Measure}),$
2.  $GA = a_0 + a_1 \times (\text{Brain Volume}).$

To assess the validity to estimate GA across various samples, we adopted a leave-k-out analysis, where k out of n subjects were chosen as a test set and the regression was performed on the remaining n–k subjects as the training set. Note that when k = 1, this cross-validation degenerates to the commonly used leave-one-out analysis scheme. A prediction error was calculated by averaging the absolute difference between the predicted GA and the known GA in the test set. For a fixed k > 1, a random test set was generated k × n times. The differences in the mean and variance of the absolute error of predicted GA were assessed by t test, F-test, and the Levene test.<sup>29</sup>

To further examine the biases and variances related to the curvedness and volume-based predictions, we performed statistical tests on signed prediction errors in addition to absolute prediction errors. Bias was defined as the mean of the signed errors for a prediction. Variance referred to the variance of the predicted GA. The difference in biases was assessed by *t* test, and the difference in variances was assessed by the *F*-test and Levene test.

## RESULTS

### Relations between Folding and GA

A progressive increment of cortical folding complexity can be observed with increasing GA (Online Fig 2). The Pearson correlation coefficient (*r*) between each individual quantitative cortical folding measure and GA, which is shown in the Table, indicates that these measures have a very strong linear relation with GA ( $P < 10^{-10}$  for all measures). For example, average curvedness and average sulcal depth accounted for 96% and 93% of the variance of GA as calculated by  $r^2$  in the data. Three scatterplots with regression lines are shown in Fig 1. Polynomial fitting suggested that there was some slight degree of nonlinear correlation between the folding measures and age, but linear regression appeared to fit the current data well. We also examined the folding complexity in the left and right hemispheres separately and did not find significant asymmetries in terms of hemispheric folding (see left subfigure in Online Fig 3). When the cerebrum was divided into 5 lobes, we found that the frontal lobe had the largest rate of folding, while the insular region demonstrated the lowest rate of folding (see right subfigure in Online Fig 3).

#### Regression correlation between brain measures and GA<sup>a</sup>

Measure	AC	EC	GC	IC	MC	CR	IP	SD	BVol
Linear r	0.98	-0.96	0.97	0.98	0.98	0.95	0.95	0.97	0.95
Poly. r	0.98	-0.97	0.98	0.98	0.98	0.96	0.96	0.97	0.96
R to fTMS	0.98	-0.97	0.98	0.98	0.98	0.95	0.96	0.98	0.94

**Note:**—EC indicates extrinsic curvature index; GC, Gaussian curvature  $L^2$  norm; IC, intrinsic curvature index; MC, mean curvature  $L^2$  norm; CR, convexity ratio; IP, isoperimetric index; SD, sulcal depth; BVol, brain volume; r (R), regression; Poly., polynomial.

<sup>a</sup>The first and second rows are correlation coefficients between each measure and GA in linear and polynomial regressions. The third row is the linear correlation coefficients between folding measures and total maturation score.

The null hypothesis that  $a_2 = 0$  in the linear model (equation 1) was rejected by the *F*-test ( $F > 18.10$ ,  $P < .0002$ ), and the first regression was significantly better than the second regression for any folding measure. This finding indicates that the quantitative cortical folding measures add additional and complementary information to brain volume in the prediction equation for GA. In contrast, when we swapped folding measures with brain volume in the equations, the null hypothesis was not rejected ( $F < 2.34$ ,  $P > .13$ ) for AC, Gaussian curvature  $L^2$  norm, mean curvature  $L^2$  norm, intrinsic curvature index, or sulcal depth. The phenomena implied that some of these folding measures could completely replace brain volume measurements when it came to predicting GA because brain volume provided no extra information about GA compared with AC, Gaussian curvature  $L^2$  norm, mean curvature  $L^2$  norm, intrinsic curvature index, or sulcal depth.

An independent assessment of brain maturity in the same dataset was performed by a pediatric neurologist and a pediatric neuroradiologist (D.J.L. and A.V.), each with 8 years of fetal MR imaging experience. A single fetal total maturation score (fTMS) was calculated by summing up 6 semiquantitative subscores characterizing various visual sulcation observations, the extent and location of myelination, and evolution of the germinal matrix.<sup>16</sup> We found that the folding measures obtained by the proposed framework were in high agreement with the average visual fTMS (Fig 2).

### Prediction of GA

The resultant predictions by quantitative folding measures, brain volume, and the visual fTMS assessment in a leave-one-out scheme are shown in Fig 3. The AC-based prediction appears closer to true GA than brain volume or the fTMS-based prediction. The final errors (averaged across the  $k \times n$  test sets) based on different sizes of the training set are shown in Online Fig 4. It can be observed that using the cortical folding measure to predict GA always resulted in higher prediction accuracy (ie, smaller errors) than using brain volume. Leave-one-out cross-validation showed that the mean of absolute errors is  $0.43 \pm 0.45$  weeks (range, 0.01–1.24 weeks) for AC and  $0.72 \pm 0.61$  weeks (range, 0.02–2.63 weeks) for brain volume. The difference between the 2 prediction errors was statistically significant ( $P = .024$ ). The difference between the 2 variances of the 2 predictions was also significant based on an  $F$ -test ( $P = .002$ ), borderline on a Levene test ( $P = .06$ ) for training size equal to 32, and significant on a Levene test ( $P < 10^{-4}$ ) for any other training size larger than 20. The curvedness predictor also had a slightly lower absolute error than the fTMS predictor (mean,  $0.57 \pm 0.41$  weeks), but the difference was not significant.

The difference between the biases from curvedness and volume-based predictions was insignificant on the basis of a 2-sample  $t$  test when the training set size was reasonably large ( $>20$ ). Regarding variance, the curvedness predictor had consistently smaller variance than the volume predictor (about half) when the training set size was large ( $>20$ ). The 2-sample  $F$ -test showed that the variance difference between curvedness and volume predictors was significant ( $P = .0041$ , for the largest training size 32; for smaller training sizes, the significance was higher with a smaller  $P$  value). The Levene test also showed that the variance difference was significant ( $L = 4.380$ ,  $P = .040$ , for training size 32; for smaller training sizes, the significance was higher). In summary, the biases of the 2 predictors are comparable, and the cortical folding predictor has a smaller variance than the volume predictor, which means the former is more reliable than the latter in the prediction of GA. The quantitative cortical folding predictor and visual fTMS predictor are close in both bias and variance.

## DISCUSSION

To our knowledge, this article is among the first works in the literature to use various MR imaging–based computational cortical folding measures as a means to estimate GA in healthy human fetuses. All the previous studies on fetal MR imaging<sup>11,12,25,30–32</sup> have regarded clinically acquired GA as an independent variable or a known reference point as opposed to a dependent variable to be predicted or estimated. The current work is also

unique in the wide GA range of the cohort, which includes the period close to full term (33–36 weeks), when it is more difficult to date pregnancy.

Many studies have characterized cortical folding in preterm neonates.<sup>13,26,33–35</sup> Although premature neonates are often studied as a surrogate of prenatal development of full-term infants, preterm neonates are exposed to extrauterine stressors and risk factors, which often make their neurodevelopment dissimilar to normal intrauterine development.<sup>14</sup>

On the basis of the data and results, it became clear that linearity was sufficient to account for the relation between GA and folding. More than 90% of variance in gestational age can be explained by the variance in a cortical folding measure. The calculation of folding measures does not use the information of gestational age at all, so it is unlikely that the linearity manifested in the results comes from a bias in the processing that favors a linear model or any other model between GA and folding.

In embryology, GA is defined as postconception time, but in human obstetrics, it is often defined as the time since the first day of the mother's last menstrual period. The latter is approximately 2 weeks longer than the former. In this article, we used the definition of GA involving the last menstrual period, which is the common clinical practice. It is possible that there were errors in the estimation of true GA despite the best clinical practice. The potential inaccuracy of ground truth GA in these data, however, does not greatly diminish the value of the results because the folding process reflects an aspect of brain development that is complementary to size increase, which is the basis of clinical GA estimation. The fact that 2 independently obtained measures of GA (folding-based and size) are so consistent with each other indicates the improbability that either measure is erroneous. The chance that both measures are very imprecise but in high agreement is slim, if one assumes that the errors are random and independent.

The sample size of the current study is relatively small and may reduce the generalizability of the results. When dealing with a small sample size, it is important not to overfit a model by creating too many features/measures or by using a complex type of regression. By using a single whole-brain cortical folding measure and simple linear regression, we sought to avoid the overfitting problem, in which the agreement between the model and data is increased but the generalizability to more data is decreased. Linear regression strikes a good balance between the goodness of fit in the current data and the generalizability to unseen datasets. Although cortical folding may be nonlinear in its entire course, it can be approximated by a linear process within the gestational period of interest (25–36 weeks), as demonstrated in other independent studies.<sup>12,13,30</sup> Another limitation of this study is that there could be some “healthy” fetuses who are later found not healthy because abnormalities were not obvious at short-term follow-ups.

## CONCLUSIONS

This article demonstrated that automated quantitative measurement of fetal cerebral cortical folding can be used to estimate gestational age in the third trimester with high accuracy and reliability on a single-case basis by using clinical fetal MR imaging. The folding measures

accurately predict the gestational age of a fetus in the third trimester (mean error,  $0.43 \pm 0.45$  weeks), which can be a major challenge for sonography-based measures. Improved accuracy and reliability in GA estimation in late gestation can have a positive impact on prenatal care for underserved populations. It is also important to be able to estimate the fetal brain maturity because chronologic GA is no longer a suitable gauge for fetuses with aberrant neurodevelopment later in gestation. Cortical folding measurement offers a potential way to accomplish that as well.

## Acknowledgments

This work was supported by the National Institutes of Health (NS065347, DA022807 and DA014129), the Canadian Institutes of Health Research (MOP-81116), and the SickKids Foundation (XG 06-069).

## ABBREVIATIONS

<b>AC</b>	average curvedness
<b>fTMS</b>	fetal total maturation score
<b>GA</b>	gestational age

## References

- Dewhurst CJ, Beazley JM, Campbell S. Assessment of fetal maturity and dysmaturity. *Am J Obstet Gynecol.* 1972; 113:141–49. [PubMed: 5025871]
- Drumm JE, Clinch J, MacKenzie G. The ultrasonic measurement of fetal crown-rump length as a method of assessing gestational age. *Br J Obstet Gynaecol.* 1976; 83:417–21. [PubMed: 1276097]
- MacGregor, SN.; Sabbagha, RE. [Accessed April 12, 2013] Assessment of gestational age by ultrasound. *The Global Library of Women's Medicine.* 2008. [http://www.glowm.com/section\\_view/heading/Assessment%20of%20Gestational%20Age%20by%20Ultrasound/item/206](http://www.glowm.com/section_view/heading/Assessment%20of%20Gestational%20Age%20by%20Ultrasound/item/206)
- Mongelli M, Wilcox M, Gardosi J. Estimating the date of confinement: ultrasonographic biometry versus certain menstrual dates. *Am J Obstet Gynecol.* 1996; 174:278–81. [PubMed: 8572021]
- Bennett KA, Crane JMG, O'shea P, et al. First trimester ultrasound screening is effective in reducing postterm labor induction rates: a randomized controlled trial. *Am J Obstet Gynecol.* 2004; 190:1077–81. [PubMed: 15118645]
- American College of Obstetricians and Gynecologists. ACOG Practice Bulletin No. 101: ultrasonography in pregnancy. *Obstet Gynecol.* 2009; 113(2 pt 1):451–61. [PubMed: 19155920]
- Benson CB, Doubilet PM. Sonographic prediction of gestational age: accuracy of second- and third-trimester fetal measurements. *AJR Am J Roentgenol.* 1991; 157:1275–77. [PubMed: 1950881]
- Glenn O. Normal development of the fetal brain by MRI. *Semin Perinatol.* 2009; 33:208–19. [PubMed: 19631082]
- Rogers J, Kochunov P, Zilles K, et al. On the genetic architecture of cortical folding and brain volume in primates. *Neuroimage.* 2010; 53:1103–08. [PubMed: 20176115]
- Garel C, Chantrel E, Brisse H, et al. Fetal cerebral cortex: normal gestational landmarks identified using prenatal MR imaging. *AJNR Am J Neuroradiol.* 2001; 22:184–89. [PubMed: 11158907]
- Habas P, Scott J, Roosta A, et al. Early folding patterns and asymmetries of the normal human brain detected from in utero MRI. *Cereb Cortex.* 2012; 22:13–25. [PubMed: 21571694]
- Hu HH, Hung CI, Wu YT, et al. Regional quantification of developing human cortical shape with a three-dimensional surface-based magnetic resonance imaging analysis in utero. *Eur J Neurosci.* 2011; 34:1310–19. [PubMed: 21995768]
- Dubois J, Benders M, Cachia A, et al. Mapping the early cortical folding process in the preterm newborn brain. *Cereb Cortex.* 2008; 18:1444–54. [PubMed: 17934189]

14. Dubois J, Benders M, Borradori-Tolsa C, et al. Primary cortical folding in the human newborn: an early marker of later functional development. *Brain*. 2008; 131:2028–41. [PubMed: 18587151]
15. Limperopoulos C, Tworetzky W, McElhinney D, et al. Brain volume and metabolism in fetuses with congenital heart disease: evaluation with quantitative magnetic resonance imaging and spectroscopy. *Circulation*. 2010; 121:26–33. [PubMed: 20026783]
16. Vossough A, Limperopoulos C, Putt ME, et al. Development and validation of a semiquantitative brain maturation score on fetal MRI: initial results. *Radiology*. 2013; 268:200–07. [PubMed: 23440324]
17. Sparrow, S.; Cicchetti, D.; Balla, D. *Vineland Adaptive Behavioral Scales (Vineland II)*. Circle Pines: AGS Publishing; 2005.
18. Gholipour A, Estroff JA, Warfield SK. Robust super-resolution volume reconstruction from slice acquisitions: application to fetal brain MRI. *IEEE Trans Med Imaging*. 2010; 29:1739–58. [PubMed: 20529730]
19. Jiang S, Xue H, Glover A, et al. MRI of moving subjects using multislice snapshot images with volume reconstruction (SVR): application to fetal, neonatal, and adult brain studies. *IEEE Trans Med Imaging*. 2007; 26:967–80. [PubMed: 17649910]
20. Kim K, Habas P, Rousseau F, et al. Intersection based motion correction of multislice MRI for 3-D in utero fetal brain image formation. *IEEE Trans Med Imaging*. 2010; 29:146–58. [PubMed: 19744911]
21. Kuklisova-Murgasova M, Aljabar P, Srinivasan L, et al. A dynamic 4D probabilistic atlas of the developing brain. *Neuroimage*. 2011; 54:2750–63. [PubMed: 20969966]
22. Yushkevich PA, Piven J, Hazlett HC, et al. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. *Neuroimage*. 2006; 31:1116–28. [PubMed: 16545965]
23. Koenderink JJ, van Doorn AJ. Surface shape and curvature scales. *Image and Vision Computing*. 1992; 10:557–65.
24. Van Essen DV, Drury H. Structural and functional analyses of human cerebral cortex using a surface-based atlas. *J Neurosci*. 1997; 17:7079–102. [PubMed: 9278543]
25. Batchelor P, Castellano-Smith A, Hill D, et al. Measures of folding applied to the development of the human fetal brain. *IEEE Trans Med Imaging*. 2002; 21:953–65. [PubMed: 12472268]
26. Rodriguez-Carranza C, Mukherjee P, Vigneron D, et al. A framework for in vivo quantification of regional brain folding in premature neonates. *Neuroimage*. 2008; 41:462–78. [PubMed: 18400518]
27. Awate S, Win L, Yushkevich P, et al. 3D cerebral cortical morphometry in autism: increased folding in children and adolescents in frontal, parietal, and temporal lobes. *Med Image Comput Assist Interv*. 2008; 11(pt 1):559–67. [PubMed: 18979791]
28. Awate SP, Yushkevich PA, Song Z, et al. Cerebral cortical folding analysis with multivariate modeling and testing: studies on gender differences and neonatal development. *Neuroimage*. 2010; 53:450–59. [PubMed: 20630489]
29. Levene, H. Robust tests for equality of variances. In: Olkin, I.; Ghurye, SG.; Hoefding, W., et al., editors. *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*. Stanford, CA: Stanford University Press; 1960. p. 278-92.
30. Rajagopalan V, Scott J, Habas PA, et al. Local tissue growth patterns underlying normal fetal human brain gyration quantified in utero. *J Neurosci*. 2011; 31:2878–87. [PubMed: 21414909]
31. Hu H, Guo W, Chen H, et al. Morphological regionalization using fetal magnetic resonance images of normal developing brains. *Eur J Neurosci*. 2009; 29:1560–67. [PubMed: 19419421]
32. Clouchoux C, Kudelski D, Gholipour A. Quantitative in vivo MRI measurement of cortical development in the fetus. *Brain Struct Funct*. 2012; 217:127–39. [PubMed: 21562906]
33. Battin M, Maalouf E, Counsell S, et al. Magnetic resonance imaging of the brain in very preterm infants: visualization of the germinal matrix, early myelination, and cortical folding. *Pediatrics*. 1998; 101:957–62. [PubMed: 9606219]
34. Deipolyi A, Mukherjee P, Gill K, et al. Comparing microstructural and macrostructural development of the cerebral cortex in premature newborns: diffusion tensor imaging versus cortical gyration. *Neuroimage*. 2005; 27:579–86. [PubMed: 15921934]

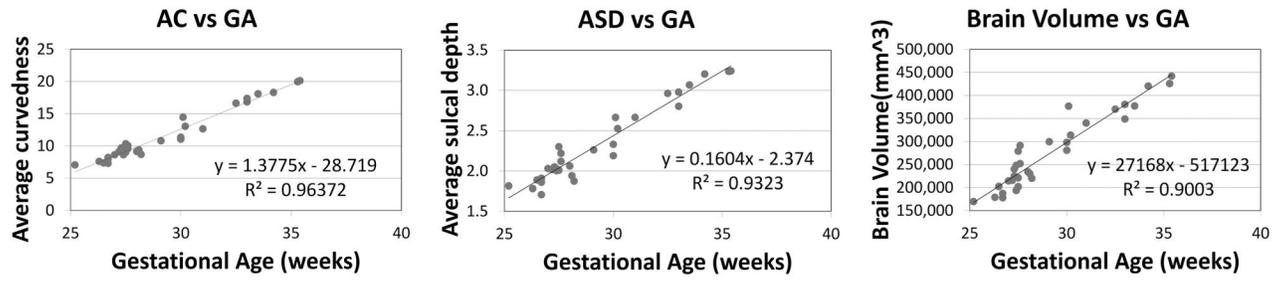
35. Smyser CD, Kidokoro H, Inder TE. Magnetic resonance imaging of the brain at term equivalent age in extremely premature neonates: to scan or not to scan? *J Paediatr Child Health*. 2012; 48:794–800. [PubMed: 22970674]

Author Manuscript

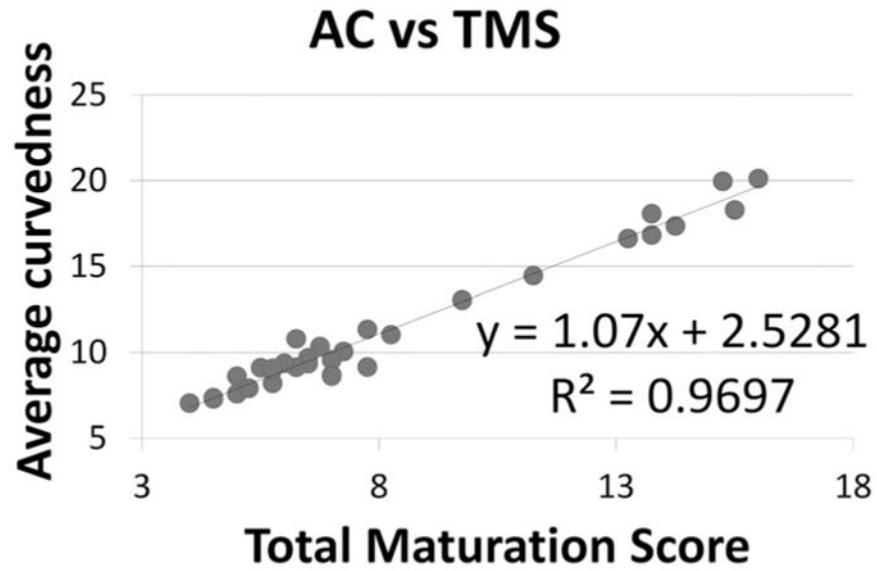
Author Manuscript

Author Manuscript

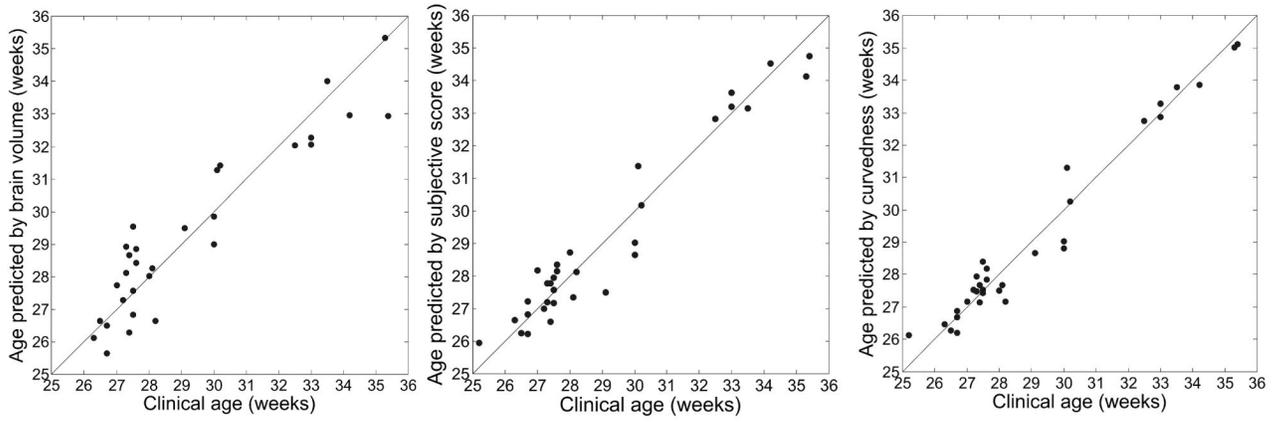
Author Manuscript



**FIG 1.** Bivariate plots of 2 folding measures and brain volume against GA. Squared correlation coefficients  $r^2$ , linear regression lines, and their equations are shown on the plots.



**FIG 2.** The plot of automatically computed average curvedness against the average total maturation score obtained by 1 pediatric neurologist and 1 pediatric neuroradiologist for the 33 healthy fetuses.



**FIG 3.** The left panel is a plot of GA predicted by brain volume (fTMS, AC) versus clinically estimated GA. The line in the plot represents a perfect match between the 2 variables. All the predictions were performed in a leave-one-out scheme.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript