Sex and Anthropomorphic Conserved Relationship of Childhood Brain Growth to Cerebrospinal Fluid

Mallory R. Peterson[1,2,5], Venkateswararao Cherukuri[1,4], Joseph N. Paulson[6], Paddy Ssentongo[2], Abhaya V. Kulkarni[7,8], Benjamin C. Warf[9,10], Vishal Monga[4], Steven J. Schiff[1-3*]

[1]Center for Neural Engineering, Departments of [2]Engineering Science and Mechanics, [3]Neurosurgery, and Physics, and [4]School of Computer Science & Electrical Engineering, The Pennsylvania State University, University Park, Pennsylvania, USA. [5]Medical Scientist Training Program, The Pennsylvania State University College of Medicine, Hershey, Pennsylvania, USA. [6] Department of Biostatistics, Product Development, Genentech Inc., South San Francisco, San Francisco, CA, USA. [7]University of Toronto. [8]The Hospital for Sick Children.[9]Harvard Medical School.[10] Boston Children's Hospital. [*]steven.j.schiff@gmail.com

May 19, 2020, Revised November 30, 2020

Abstract

How brain growth interacts with cerebrospinal fluid (CSF) accumulation has not been well characterized, but is critical in managing diseases of early childhood where brain growth may be affected. We analyzed 1067 magnetic resonance imaging (MRI) scans from 505 healthy pediatric subjects from birth to age 18 to quantify component and regional brain volumes. Brain volume peaked at 10-12 years of age. Males exhibited larger age-adjusted total brain volumes than females, and body size normalization procedures did not eliminate this difference. The ratio of brain to CSF volume, however, revealed a universal age-dependent relationship independent of sex or body size. These findings enable the use of such normative curves in managing a broad range of childhood disease where cognitive development, brain growth, and fluid accumulation are interrelated.

Introduction

The brain growth and CSF accumulation relationship is critical to characterize because it is impacted in numerous conditions of early childhood where brain growth and fluid accumulation are affected such as infection, hemorrhage, hydrocephalus, and a broad range of congenital disorders (*1, 2*). Appropriate management of such diseases requires understanding the normal dynamics of brain growth in relationship to fluid accumulation. However, the study of brain size and growth has a long and contentious history(*3-6*). In the pre-magnetic resonance imaging (MRI) era, post-mortem studies provided insight into brain volume changes over the lifespan, but such methods suffered from inherent inaccuracies(*7*). In the MRI era, the development of advanced computational and statistical algorithms has enabled detailed in-vivo volumetrics including components and regions of the brain(*8-11*).

Studies defining normal brain volume growth patterns in the MRI era have often included small sample sizes, limited algorithm technology, incomplete coverage of the pediatric age range, retrospective cohorts taken from clinical patients, and a disparate array of curve fitting techniques(*8-10, 12, 13*). For instance, previous brain growth studies conducted by Pfefferbaum et al (1994), Giedd et al (1999), and Courchesne et al (2000) included only 30, 243, and 50 subjects in the pediatric age range(*8, 9, 11*). Secondly, none of those studies included the entire pediatric age range from birth to age 18. More recent works such as Coupe et al (2017) and Hedman et al (2012) have incorporated big data or meta-analysis approaches, but by incorporating such large amounts of data, they lose rigor in exclusion criteria, population representation, and calibration between scan site acquisition(*14, 15*). Furthermore, despite leveraging 56 studies and including 2,211 participants, Hedman et al (2012) did not include any subjects under the age of 4, which excluded the period during which the majority of brain growth

takes place. Others such as Knickmeyer et al (2008) have looked closely at the infantile period from birth to age 2, but have not covered the later pediatric ages(*16*). While these studies provide valuable insights into different approaches for characterizing brain volume, the interaction of brain with fluid accumulation during development has not been well characterized.

Results

We quantified brain and CSF volumes from 1067 MRI brain scans from 505 healthy normal subjects (259 female) from birth to age 18. These MRI scans were collected from healthy pediatric participants in the NIH Pediatric MRI Repository (https://nda.nih.gov), which was accessed under an institutional data use agreement between The Pennsylvania State University and the National Institute of Mental Health. This MRI repository was developed based on a scaled-down United States (US) census with rigorous exclusion criteria, with the goal of providing a standard representation of the socioeconomic, male and female, and racial distribution of healthy normal US children(*17*). The cross-sequential cohort contains participants in each year of life ranging from 13-days to 22-years old (Supplemental Figure 1). Uniform MRI scanning protocols were applied at each data collection site, and living phantom calibration was leveraged to ensure volumetric consistency between sites(*17, 18*).

The T2 MRI images of the neonatal and infant population were processed using the Developing Human Connectome Project (dHCP) Pipeline in order to appropriately segment the rapidly growing and incompletely myelinated brains(*19*). This pipeline consists of the Brain Extraction Tool (BET) skull-stripping algorithm, N4 algorithm intensity inhomogeneity correction, and the Developing brain Region Annotation With Expectation-Maximization (Draw-EM) algorithm for segmentation(*20-22*). The T1 images of the older subjects were processed using the Computational Anatomy Toolbox (CAT) within the Statistical Parametric Mapping (SPM12) software. This software utilizes voxel and region-based morphometry upon application of an adaptive probability region-growing (APRG) algorithm for skull-stripping(*23*). Each of the resulting segmentation images (Supplemental Figure 2) were manually inspected to confirm appropriate anatomic assignment.

The brain components and region volumes quantified by the dHCP and CAT12 pipelines were then fit using Smoothing Spline ANOVA (SSANOVA) with random effects (to account for the cross-sequential design of the study) to define time periods of statistically significant differences stratified by sex and hemisphere (Figure 1, 2 , and Supplemental Figure 3)(*24*). In order to define the growth trajectories and population centiles, the volumes were also fit using Generalized Additive Models for Location, Scale, and Shape (GAMLSS) with a Box-Cox power exponential (BCPE) distribution (Supplemental Figures 4 and 5), which is the same platform and distribution leveraged by the World Health Organization to develop their standard growth curves for weight, height, and head circumference(*25, 26*). Finally, the relationship between cognitive scores and brain volume was explored using a linear mixed effects model and age-dependent correlations (Figure 3 and Supplemental Figure 6).

Findings Differentiated by Sex

Males exhibited larger overall brain volumes than females throughout childhood (Figure 1a). The volume for females peaked at 10.7 years, and the volume for males peaked at 11.2 years, followed by a slow, but consistent decrease. This early adolescent peak has been previously described(*12, 14*). CSF increased throughout childhood, with male fluid accumulation significantly larger than female after the third year of life (Figure 1b). Grey matter (Figure 1c) peaked at 7.5 years for males and 7.4 years for females and then began to gradually decrease, while white matter (Figure 1d) continued to progressively increase into early adulthood(*27*). The ratio of grey/white matter (Figure 1e) showed an initial increase, peaking before 2 years of age and followed by a progressive decrease thereafter. Female grey/white matter ratios were significantly larger than male ratios between ages 9 and 11 although the difference was small (Figure 1e).

We performed body size normalization to assess if differences in brain volume persisted between males and females. Normalizing the brain by body size is not a new concept; allometry, or differential growth, of the brain with respect to body size was discussed in depth by D'Arcy Wentworth Thompson in 1917 in *On Growth and Form(28)*. Gould, in *The Mismeasure of Man*, attempted and failed to eliminate sex differences through body size normalization(*3*). Nevertheless, much of the volumetric brain study in the MRI era has not accounted for anthropomorphic normalization. Figures 1f and 1g show brain volume normalized by height-for-age and weight-for-height, respectively, which did not eliminate the sex-based differences in volume. Muscle mass content, greater in males, has been correlated with larger brain volumes although not with higher cognitive capability(*29*). However, the ratio of brain to CSF volume demonstrated no significant sex differences at any age, without requiring anthropomorphic normalization (Figure 1h).

Hemispheric Lateralization Findings

Figure 2a and 2b show that there was no lateralizing difference in size between right and left hemispheres for males and females, nor for cerebellum, frontal, parietal, or occipital lobes (Supplemental Figure 2). Figure 2c and 2d show that the left temporal lobe was significantly larger than the right for both males and females, which has been controversial(*30*). In addition, the hippocampi were significantly larger on the right than the left side for both sexes (Figure 2e and 2f). Previous studies such as Lynch et al (2019) have also found that the right hippocampus is larger than the left(*31*). Age-adjusted hippocampal and temporal lobe volume assessment may be of value in diagnosing and treating medically refractive epilepsy in childhood(*30, 32*).

Correlation Between Cognition and Volume

Cognitive scores showed a small but significant correlation with brain volume in the four years leading up to the peak in volume (Figure 3). The Mental Development Index (MDI) scores for infants from birth to age 3 were not significantly predicted by brain volume (Figure 3b), but the Weschler Abbreviated Scale of Intelligence (WASI) scores for ages 6-18 years were significantly correlated with brain volume z-scores (Figure 3a), as described previously(*33*). The correlation between cognitive score and brain volume was significant for raw brain volume, age-adjusted brain volume z-score, and weight-for-height normalized volumes in the years immediately preceding the brain volume peak (Figure 3c-3e). This correlation was not maintained when separating into smaller cohorts by sex (Supplemental Figure 6).

Growth Curves and Centiles

In 1987, Roche et al created head circumference growth curves from studying 888 healthy US children(*34*), and such normative head circumference curves from US(*35*) and World Health Organization(*25, 36*) cohorts are now in routine clinical practice as indirect metrics of brain growth. Figure 4 illustrates the analogous GAMLSS pediatric brain volume growth curves for males and females (Figure 4a and 4b), with early brain volume growth and CSF volume insets included, as well as the brain/CSF ratio (Figure 4c and 4d) for a more comprehensive presentation of childhood normative brain growth suitable for clinical settings.

Discussion

Brain volume measurement became a field of study of biological determinism pioneered by Samuel Morton in the mid-1800s(*3, 37*). Morton filled over 1000 cranial vaults with mustard seed and lead shot to determine brain volume, which he then compared between races and genders. A century and a half later Gould used Morton as a case study in scientific bias[1]. Decades after the publishing of these analyses in *The Mismeasure of Man*, arguments over biases and flaws continue in the assessment of the volume of the normal human brain(*4-6*).

The sexual dysmorphism found in this present study corroborates previous studies looking at different age groups. Giedd et al (1999), Nopoulos et al (2000), Goldstein et al (2001), Lenroot et al (2007) , Groeschel et al (2010), and Coupe et al (2017) all found similar differences in brain components, with larger total brain volumes in males(*11, 15, 38-41*). We found that normalizing by body measurements did not account for sex differences, we note that Coupe et al (2017) used proportionalization by total intracranial volume (TIV) and failed to find an overall sex difference(*15*), consistent with our discovery of a universal age-dependent ratio between brain and CSF volumes.

Another area of great interest in brain volume research is the growth trajectory of components of the brain. The head size increases steadily until at least age 18(*34*). In the MRI era, however, many studies have shown that total brain volume reaches its peak in earlier adolescence. Without including younger children during the early years of rapid brain growth, the location of this volume peak can be somewhat uncertain. Giedd et al (1999) and Hedman et al (2012) showed that brain volume peaks between 12 and 14, but lacked subjects below age four when the majority of brain growth takes place. Lenroot et al (2007) found peaks at age 10.5 years for females and 14.5 for males, but lacked subjects below the age of three years(*40*).

Incorporating more children during the rapid early growth of the first few years of life our analysis placed the peak at 10-12 years of age.

It is interesting to note that Hedman et al (2012) found an early adolescent peak, but also concluded that their meta-analysis highlighted a later resurgence of growth with a secondary growth peak in adulthood(*14*). These findings highlight the need for additional normal growth data throughout the life span. Such lifespan data would gain greater value by incorporating fetal through post-natal volume development, much in the manner of Holland et al (2014)(*42*).

As with total brain volume, we found grey matter to peak in early adolescence and then slowly decrease. Our results showed that grey matter peaked in the seventh year of life for both males and females. Sowell et al (2003) also found that grey matter decreased after the seventh year of life, but they did not include any subjects under the age of seven(*43*). While grey matter peaks in adolescence and starts to decline thereafter, the white matter and CSF components do not reach a peak during the pediatric age range. These findings are also supported by Sowell et al (2003). Sowell and others found that white matter reaches a peak near the age of 40 while CSF accumulation never ceases(*27, 43*). It is appreciated that myelination continues through the second decade of life contributing to white matter growth(*44, 45*). Grey matter, however, undergoes synaptic elimination through pruning, which may account for the peak seen in adolescence(*46*). It is after the grey matter peak at seven years of age that we found a significant correlation between brain volume and cognitive scoring. This correlation continued until the total brain volume peaked near age 12. These findings highlight this critical period of pruning and myelination before overall brain volume begins to decline in late childhood. The small association observed in this study for brain size within the normal range with cognitive

performance within the normal range appears magnified in children with disease early in life that substantially impacts brain growth(*47*).

Our study has limitations. While the growth curves developed in this study provide a standard representation of the United States pediatric population, they can only act as a reference and not as a standard for other geographic regions. Other growth curves specific to particular regions and peoples should be developed using healthy cohorts derived from those regions in order to provide appropriately representative standards representing the diversity and characteristics of non-US populations.

Our findings demonstrate that for a broad spectrum of human disease affecting neurocognitive development and brain growth – ranging from neonatal infection(*48*) to malnutrition(*49*) and hydrocephalus(*1*) – measuring brain growth and CSF with respect to normative values is now feasible. The apparent universal nature of the age-dependent brain/CSF ratio, regardless of sex or body size, suggests that the role of this ratio warrants clinical investigation(*50*). There is also increasing recognition of the genetic underpinnings of children with disordered brain/CSF ratios(*2*). The finding that the brain/CSF ratio is tightly regulated opens up new avenues in the management of diseases affecting the childhood brain.

Methods

Experimental Design

Cohort Characteristics

The MRI scans used in this study were taken from the NIH Pediatric MRI Repository (https://nda.nih.gov) under an institutional data use agreement (214908) between The Pennsylvania State University and the National Institute of Mental Health approved on January 14, 2019. This repository was developing using a scaled down United States census from 2000 (in order to appropriately represent the racial and socioeconomic demographic characteristics of the entire US pediatric population) and included rigorous exclusion criteria to ensure healthy participants with normal brain development. The repository aimed to achieve two-year longitudinal follow-up scans for individual participants, and was able to accomplish this for 378 of the subjects, making this a cross-sequential study(*17*). The cross-sequential format is ideal for the development of growth curves(*9, 51*). The number of subjects in this study was 505 (259 female), with a total of 1067 MRI scans due to the longitudinal nature of the cohort. The minimum age was 13 days, and the maximum age was 22 years, but only scans from subjects up to 18-years-old were included to develop growth curves representative of the pediatric age range. Scans existed for participants in each year of life throughout the pediatric age range, as seen in Supplemental Figure 1.

MRI Acquisition and Quality Control

The Pediatric MRI Repository database was collected at Boston Children's Hospital, Cincinnati Children's Hospital Medical Center, University of Texas Houston Medical School, the Neuropsychiatric Institute and Hospital UCLA, Children's Hospital of Philadelphia, and Washington University in St. Louis. While it is of benefit to use multiple different geographic

locations when cultivating a cohort representative of a particular population that can then be leveraged for growth curves, bias can be introduced through the use of different scanner sites.

MRI conditions such as scanner manufacturer, acquisition protocols, and field strength can all contribute to the underlying compatibility of MRI data gathered at different sites. For the Pediatric MRI Repository, the acquisition protocols were well-defined and reproduced uniformly at each separate site. Dual contrast T2-weighted (Fast/Turbo spin echo-ETL/Turbo factor 8, slice thickness 3 mm, oblique axial acquisition, TR 3300 ms, TEl 83 ms, TE2 165 ms, refocusing pulse 180 degrees, field of view and matrix AP: 256 & LR: 192 mm) scans were acquired for infants and Tl-weighted (3D RF-spoiled gradient recalled echo sequence, slice thickness 1 mm to 1.5 mm, sagittal acquisition, TR 22-25 MS, TE 10-11 ms, excitation pulse 30 degrees, whole head field of view, AP 256 matrix with LR for 1 mm isotropic) scans were acquired for adolescents. More details can be found in Tables 3-5 of Evans (2006) publication discussing the development of the NIH Pediatric MRI Repository(*17*). Furthermore, the field strength, which has been argued to be the most serious concern, was 1.5T at each site.

Lastly, both GE and Siemens machines were used in this study. In an effort to recognize and compensate for any bias due to the use of different machines at the different sites, and in turn compensate for any bias in the derived parameters of the structural MRIs, the Pediatric MRI Repository investigators followed the guidance of Schnack et al (2004) and scanned the same healthy person at each site(*18*). They then calculated the volumes for components such as the Frontal lobe and compared these volumes between sites. When looking at inter-site differences for the left frontal lobe, the investigators found a mean volume of 467 cm^3 for the left frontal lobe with r.m.s. error 5.0 cm^3, inter-site F-ratio of 10.0, and significantly different volumes ($p=0.002$). In order to ameliorate the inter-site differences, and characterize the gradient

distortion field, the investigators applied a geometric distortion correction using a 3D American College of Radiology (ACR) phantom along with spherical harmonic decomposition. Upon application of the geometric distortion correction, the investigators redid the volume calculations and found no significant differences between sites. The right frontal lobe, for instance, was found to have a mean volume of 477 cm^3 with r.m.s error of 6.3 cm^3, and inter-site F-ratio of 1.9 ($p=0.18$). The ACR phantom scanning was completed approximately monthly and the living subject scanning was completed annually to ensure inter-site calibration

In regard to the quality control of the dataset, each MRI scan within the Repository was viewed by a radiologist to ensure no neurologic abnormalities, and the scans were then assessed for quality (in addition, if the first scan taken had too much motion artifact it was redone). If the scan was deemed to be neurologically normal and of acceptable quality, only then was it included in the database. Each scan in the database was then corrected for intensity non-uniformities using the N3-algorithm in tandem with scanner-specific models of B_0 inhomogeneity from each site(*52*).

In our present work with this database, all 1381 available anatomic MRI scans from the database were run through either dHCP or CAT12. Each volume segmentation was then manually audited to ensure reliable results. Only if each segmentation produced by the pipelines was deemed acceptable upon inspection was a processed scan included in our study. We found 1067 of the 1381 scan sets to be successfully processed, leading to our final segmentations used in our analyses.

Segmentation Algorithms

Currently, no single algorithm exists that can reliably segment both young infants and older subjects(*53, 54*). This is particularly due to the myelination changes that do not resolve until approximately two years of age(*55*). These myelination changes lead to difficulty in establishing intensity thresholds between grey matter, white matter, and CSF(*55*). In order to

maximize thresholding intensities, neonatal segmentation techniques most commonly rely on T2 weighted MRI scans, rather than T1 weighted scans that are the predominant scan type used in older cohort segmentation algorithms. Due to these reasons, two different algorithms were used in this study.

The neonates and infants were assessed using the Developing Human Connectome Project (dHCP) pipeline, which required T2 images and was run through a virtual Docker container to access a Linux computer system(*19*). The dHCP pipeline relies on the Brain Extraction Tool (BET) algorithm for skull-stripping, followed by the N4 algorithm for intensity inhomogeneity correction and the Draw-EM algorithm for segmentation (*20-22*). The Draw-EM (Developing brain Region Annotation With Expectation-Maximization) algorithm consists of an adaptive Expectation-Maximization scheme relying on a spatial prior term and intensity model. The spatial prior term is established using labeled atlases that are registered to the image, whereas the intensity model is established using a Gaussian Mixture Model. The labeled atlases were obtained through manual editing and later updated by Markopoulos et al(*20, 56*). The dHCP pipeline is freely available on Github and can be run through code describing the input image location (https://github.com/BioMedIA/dhcp-structural-pipeline).

The older subjects were assessed using the Computational Anatomy Toolbox 12 (CAT12) within the Statistical Parametric Mapping (SPM) platform using Matlab 2019b, which relies on T1 images(*23*). The CAT12 algorithm relies on voxel-based morphometry for segmentation and label-based morphometry for region-of-interest (ROI) classification. Each scan was preprocessed using a modified ICBM Tissue Probabilistic Atlas tissue probability map along with an ICBM space template affine registration. Skull-stripping was accomplished using adaptive probability region-growing (APRG) algorithm followed by surface-based optimization.

The MNI152 Dartel Template was used for Dartel registration, and finally the Hammers atlas was used for ROI analysis(*57*). Each of the resulting scan sets was manually curated to ensure that appropriate skull-stripping and segmentation was accomplished. Upon establishing the volumes determined by each segmentation procedure, the accompanying atlases were used to compile volumes for the desired regions from smaller sections of the brain(*20, 56, 57*).

Statistical Analysis

Smoothing Spline ANOVA

One of the major limitations of the historical fitting efforts is the inability to statistically assess which time points throughout the entire growth process are truly different between groups while also establishing appropriate representative trajectories. Within this study, we split our efforts, using one approach for trajectory modeling and a different approach for statistical analysis. The GAMLSS approach was used to provide smooth trajectory growth curves and centiles for reference in the same manner that the World Health Organization produces their growth curves. These curves, however, do not facilitate an in-depth statistical analysis of differences between groups such as the two different sexes or hemispheres. Therefore, the Smoothing Spline ANOVA technique was also incorporated into this study for the main purpose of examining longitudinal statistical differences between populations.

Smoothing Spline analysis of variance (SSANOVA) is a semi-parametric method that models data generated from a smooth function $f(x)$ by assuming that f is a function in a Reproducible Kernel Hilbert Space (RKHS) of the form $\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_1$ (*58*). The set of functions $(\phi_v(x))_{v=1}^{m}$, which spans the finite dimensional subspace $\mathcal{H}_0$ and $\mathcal{H}_1$, is a RKHS induced by a given kernel function k. Therefore, f has a semi-parametric form given by

$$f(x) = \sum_{j=1}^{m} d_j \, \phi_j(x) + g(x)$$

for coefficients d_j, where functions ϕ_j have a parametric form and $g \in \mathcal{H}_1$ which is defined by:

$$g(x) = \sum_{\alpha} g_\alpha \, (x_\alpha) + \sum_{\alpha<\beta} g_{\alpha\beta} \, (x_\alpha, x_\beta) + \cdots$$

where g_α and $g_{\alpha\beta}$ satisfy the standard ANOVA side conditions.

The SS-ANOVA estimate of f given data $(x_i, y_i), i = 1, \ldots, n,$ is given by the solution of a penalized problem,

$$\min_{f \in H} (y_t - f(x))^2 + \lambda J(f(x)) \tag{1}$$

where the first term discourages the lack of fit of f and the second term penalizes the complexity of f with smoothing parameter λ controlling the trade-off between the two.

Following the representor theorem of Kimeldorf et al (1970)(*59*) and the assumption of Gaussian data, the minimizer of the problem in equation (1) has a finite representation of the form:

$$f(x) = \sum_{v=1}^{m} d_v \, \phi_v(x) + \sum_{j=1}^{n} c_j \, k(x_j, x)$$

for coefficients c_i and d_v.

Letting Y be the matrix of observations of size $N \times 1$, where N includes all observations, including repeated measurements for different subjects; S is a $N \times m$ matrix where m represents the number of unpenalized terms in the model; and Q is a $N \times N$ matrix which accounts for all penalized terms in the model; estimation reduces to:

$$\min_{d,c} (Y - Sd - Qc)^T (Y - Sd - Qc) + n\lambda c^T Qc.$$

Here S is the matrix described above with the iv^{th} entry being $\phi_v(x_i)$ and Q is the penalized matrix with the ij^{th} entry being $k(x_i, x_j)$(*60*). We use Generalized Approximate Cross-

Validation (GACV), an approximation to the leave-one-out estimate of the comparative Kullback-Leibler distance between the empirical data $\hat{f}$ and the unknown true f to select regularization parameter λ and θ. This is dependent on the data available.

The smoothing parameters are determined through an iterative cross-validation process, which is then followed by the minimization of the penalized least square functional, the solution of which is a smoothing spline. The details of this process are described in Gu (2014)(*24*).

For this study, the component differences between males and females were explored within the R platform using non-parametric Smoothing Spline ANOVA mixed-effects models with a subject-specific random effects component added to account for the cross-sequential aspect of the data(*24*). The model included age and sex or hemisphere as main factors, as well as an interaction term. We were interested in visualizing and statistically analyzing the longitudinal differences between hemispheres, and therefore incorporated hemisphere as a main factor rather than a nested factor. Furthermore, some subjects only had one data point collected, while others had two or three. This imbalance and the low levels of repetition lead to matrix rank-deficiency in modeling hemisphere as a nested variable, rendering the nested approach mathematically inappropriate for this dataset. Time periods with significant sex or hemispheric differences were defined as regions where there was no overlap between the Bayesian 95% confidence intervals calculated for the sex and hemispheric factors. These regions of significant difference were highlighted on the plotted models, and the time period of significance was documented as well.

Generalized Additive Models for Location, Scale, and Shape (GAMLSS)

The smooth growth curves used to fit the volumes and other growth metrics included in this study were developed using the Generalized Additive Models for Location, Scale, and Shape software implemented in R(*26*). The Box-Cox power exponential (BCPE) distribution, which

was chosen by the World Health Organization (WHO) for their standard growth curves, was used to model the volumes in this study(*61*). This distribution models the median for a nonparametric assessment, and appropriately accounts for kurtosis and skewness within the data. The growth curves were fitted using the default RS algorithm and were smoothed using fractional polynomials of the third order(*26*). A random effects component was added to the curve modeling procedure in order to account for the longitudinal aspect of the data. For the total brain tissue growth curves, we utilized data for subjects between 18-22 years of age to set the 18 year old volume intercepts, and the perinatal volume data from Huppi et al to set the volume intercepts at birth(*62*).

In order to determine the peaks of the brain tissue and grey matter curves, cftool within Matlab 2019b was used to fit differentiable rational polynomial functions (which were applied as the smoothing function in the GAMLSS curves)(*63*). The significant sex and hemispheric differences were found using the Mann-Whitney U-test within Matlab 2019b. A Bonferroni correction was applied so that significance was established with $p<0.000066$.

The weight for height and height for age normalizations were accomplished by fitting a GAMLSS curve to the weight for height and height for age data from the NIH repository for each sex. Based on these fits, percentiles were calculated for each subject, and the 50th percentile was set at 1, with percentiles above and below ranging from 0.5 to 1.5. The corresponding brain volume for each subject was then divided by this percentile value to achieve normalized brain volumes.

Cognitive Score Correlations

Cognitive scoring was included in the NIH Pediatric MRI Repository study, with Wechsler Abbreviated Scale of Intelligence (WASI) Tests undertaken on participants ranging

from 6 years of age to 18 years of age. The infants (from birth to 3 years of age) were assessed using the Bayley Scales of Infant Development, Second Edition (BSID-II) Mental Development Index (MDI). We fit a linear mixed effects model (with subject identification as the random effects component) to the appropriate cognitive score using the brain volume z-score values. While the linear fits showed positive slopes for each metric, only the WASI scores showed a significant fixed effect for the brain volume z-score. Age-dependent correlations were calculated with a subpopulation window-size of 120 and step-size of 20 subjects for raw brain volume, brain volume z-score, and weight-for-height normalized volume z-score.

Data Availability

The volumes used to create the SSANOVA and GAMLSS growth curves presented here are provided in a spread sheet supplied in the online supplemental material (Supplemental Extended Data).

Code Availability

The code needed to reproduce the figures in this study, as well as plot new volumes on the existing growth curves or calculate z-scores, is uploaded on the Github site https://github.com/Schiff-Lab/Brain-Growth .

References

1. A. V. Kulkarni *et al.*, Endoscopic Treatment versus Shunting for Infant Hydrocephalus in Uganda. *New England Journal of Medicine* **377**, 2456-2464 (2017).
2. S. C. Jin *et al.*, Exome sequencing implicates genetic disruption of prenatal neuro-gliogenesis in sporadic congenital hydrocephalus. *Nature Medicine* **26**, 1754-1765 (2020).
3. S. J. Gould, *The mismeasure of man*. (First edition. New York : Norton, [1981] ©1981, 1981).
4. J. E. Lewis *et al.*, The mismeasure of science: Stephen Jay Gould versus Samuel George Morton on skulls and bias. *PLoS Biol* **9**, e1001071 (2011).
5. M. Weisberg, D. B. Paul, Morton, Gould, and Bias: A Comment on "The Mismeasure of Science". *PLoS Biol* **14**, e1002444 (2016).
6. Mismeasure for mismeasure. *Nature* **474**, 419-419 (2011).
7. A. S. Dekaban, Changes in brain weights during the span of human life: relation of brain weights to body heights and body weights. *Ann Neurol* **4**, 345-356 (1978).
8. A. Pfefferbaum *et al.*, A Quantitative Magnetc Resonance Imaging Study of Changes in Brain Morphology From Infancy to Late Adulthood. *JAMA Neurology* **51**, (1994).
9. E. Courchesne *et al.*, Normal brain development and aging: Quantitative analysis at in vivo MR imaging in healthy volunteers. *Radiology* **216**, 672-682 (2000).
10. M. Peterson, B. C. Warf, S. J. Schiff, Normative human brain volume growth. *Journal of Neurosurgery: Pediatrics* **21**, 478-485 (2018).
11. J. N. Giedd *et al.*, Brain development during childhood and adolescence: a longitudinal MRI study. *Nat Neurosci* **2**, 861-863 (1999).
12. J. N. Giedd *et al.*, Quantitative magnetic resonance imaging of human brain development: ages 4-18. *Cereb Cortex* **6**, 551-560 (1996).
13. A. McAllister *et al.*, Quantitative Synthetic MRI in Children: Normative Intracranial Tissue Segmentation Values during Development. *AJNR. American journal of neuroradiology* **38**, 2364-2372 (2017).
14. A. M. Hedman, N. E. van Haren, H. G. Schnack, R. S. Kahn, H. E. Hulshoff Pol, Human brain changes across the life span: a review of 56 longitudinal magnetic resonance imaging studies. *Hum Brain Mapp* **33**, 1987-2002 (2012).
15. P. Coupe, G. Catheline, E. Lanuza, J. V. Manjon, I. Alzheimer's Disease Neuroimaging, Towards a unified analysis of brain maturation and aging across the entire lifespan: A MRI analysis. *Hum Brain Mapp* **38**, 5501-5518 (2017).
16. R. C. Knickmeyer *et al.*, A structural MRI study of human brain development from birth to 2 years. *J Neurosci* **28**, 12176-12182 (2008).
17. A. C. Evans, The NIH MRI study of normal brain development. *NeuroImage* **30**, 184-202 (2006).
18. H. G. Schnack *et al.*, Reliability of brain volumes from multicenter MRI acquisition: a calibration study. *Hum Brain Mapp* **22**, 312-320 (2004).
19. A. Makropoulos *et al.*, The developing human connectome project: A minimal processing pipeline for neonatal cortical surface reconstruction. *Neuroimage* **173**, 88-112 (2018).

20. A. Makropoulos *et al.*, Automatic whole brain MRI segmentation of the developing neonatal brain. *IEEE Trans Med Imaging* **33**, 1818-1831 (2014).
21. S. M. Smith, Fast robust automated brain extraction. *Hum Brain Mapp* **17**, 143-155 (2002).
22. N. J. Tustison *et al.*, N4ITK: improved N3 bias correction. *IEEE Trans Med Imaging* **29**, 1310-1320 (2010).
23. C. Gaser, R. Dahnke, CAT-a computational anatomy toolbox for the analysis of structural MRI data. *HBM* **2016**, 336-348 (2016).
24. C. Gu, *Smoothing Spline ANOVA Models*. (Springer, 2002).
25. E. Borghi *et al.*, Construction of the World Health Organization child growth standards: Selection of methods for attained growth curves. *Statistics in Medicine* **25**, 247-265 (2006).
26. R. A. Rigby, D. M. Stasinopoulos, P. W. Lane, Generalized additive models for location, scale and shape. *Journal of the Royal Statistical Society. Series C: Applied Statistics* **54**, 507-554 (2005).
27. K. Narvacan, S. Treit, R. Camicioli, W. Martin, C. Beaulieu, Evolution of deep gray matter volume across the human lifespan. *Human Brain Mapping* **38**, 3771-3790 (2017).
28. D. A. W. Thompson, *On Growth and Form*. (Cambridge University Press, Cambridge, 1917).
29. A. H. M. Kilgour, O. M. Todd, J. M. Starr, A systematic review of the evidence that brain structure is related to muscle structure and their relationship to brain and muscle function in humans over the lifecourse. *BMC Geriatr* **14**, 85-85 (2014).
30. J. G. Mandell *et al.*, Volumetric brain analysis in neurosurgery: Part 3. Volumetric CT analysis as a predictor of seizure outcome following temporal lobectomy. **15**, 133 (2015).
31. K. M. Lynch *et al.*, Hippocampal Shape Maturation in Childhood and Adolescence. *Cereb Cortex* **29**, 3651-3665 (2019).
32. W. Van Paesschen *et al.*, Quantitative hippocampal MRI and intractable temporal lobe epilepsy. *Neurology* **45**, 2233-2240 (1995).
33. A. L. Reiss, M. T. Abrams, H. S. Singer, J. L. Ross, M. B. Denckla, Brain development, gender and IQ in children: A volumetric imaging study. *Brain* **119**, 1763-1774 (1996).
34. A. F. Roche, D. Mukherjee, S. Guo, W. M. Moore, Head Circumference Reference Data: Birth to 18 Years. *Pediatrics* **79**, 706-712 (1987).
35. C. O. RJ Kuczmarski, LM Grummer-Strawn, *CDC Growth Charts: United States*. Advance data from vital and health statistics (National Center for Health Statistics, Hyattsville, MD, 2000), vol. 314.
36. D. C. M de Onis, C Garza, AW Onyango, WHO Child Growth Standards. *Acta Paediatrica*, (2006).
37. S. G. Morton, *Crania americana: A comparative view of the skulls of various aboriginal nations of North and South America*. (Philadelphia, 1839).
38. P. Nopoulos, M. Flaum, D. O'Leary, N. C. Andreasen, Sexual dimorphism in the human brain: evaluation of tissue volume, tissue composition and surface anatomy using magnetic resonance imaging. *Psychiatry Res* **98**, 1-13 (2000).

39. J. M. Goldstein *et al.*, Normal sexual dimorphism of the adult human brain assessed by in vivo magnetic resonance imaging. *Cereb Cortex* **11**, 490-497 (2001).
40. R. K. Lenroot *et al.*, Sexual dimorphism of brain developmental trajectories during childhood and adolescence. *Neuroimage* **36**, 1065-1073 (2007).
41. S. Groeschel, B. Vollmer, M. D. King, A. Connelly, Developmental changes in cerebral grey and white matter volume from infancy to adulthood. *Int J Dev Neurosci* **28**, 481-489 (2010).
42. D. Holland *et al.*, Structural growth trajectories and rates of change in the first 3 months of infant brain development. *JAMA Neurol* **71**, 1266-1274 (2014).
43. E. R. Sowell *et al.*, Mapping cortical change across the human life span. *Nat Neurosci* **6**, 309-315 (2003).
44. F. M. Benes, M. Turtle, Y. Khan, P. Farol, Myelination of a key relay zone in the hippocampal formation occurs in the human brain during childhood, adolescence, and adulthood. *Arch Gen Psychiatry* **51**, 477-484 (1994).
45. E. R. Sowell, P. M. Thompson, A. W. Toga, Mapping changes in the human cortex throughout the span of life. *Neuroscientist* **10**, 372-392 (2004).
46. W. M. Cowan, J. W. Fawcett, D. D. O'Leary, B. B. Stanfield, Regressive events in neurogenesis. *Science* **225**, 1258-1265 (1984).
47. J. G. Mandell, A. V. Kulkarni, B. C. Warf, S. J. Schiff, Volumetric brain analysis in neurosurgery: Part 2. Brain and CSF volumes discriminate neurocognitive outcomes in hydrocephalus. *Journal of Neurosurgery: Pediatrics* **15**, 125 (2015).
48. A. Grinberg *et al.*, Volumetric MRI Study of the Brain in Fetuses with Intrauterine Cytomegalovirus Infection and Its Correlation to Neurodevelopmental Outcome. *American Journal of Neuroradiology* **40**, 353-358 (2019).
49. K. Tiwari *et al.*, Impact of malnutrition on head size and development quotient. *International Journal of Research in Medical Sciences* **5**, 4 (2017).
50. M. K. Rasmussen, H. Mestre, M. Nedergaard, The glymphatic pathway in neurological disorders. *Lancet Neurol* **17**, 1016-1024 (2018).
51. P. J. Diggle, P. Heagerty, K.-Y. Liang, S. L. Zeger, Analysis of Longitudinal Data. 379-379 (2002).
52. J. G. Sled, A. P. Zijdenbos, A. C. Evans, A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans Med Imaging* **17**, 87-97 (1998).
53. C. N. Devi, A. Chandrasekharan, V. K. Sundararaman, Z. C. Alex, Neonatal brain MRI segmentation: A review. *Computers in Biology and Medicine* **64**, 163-178 (2015).
54. Y. Ding *et al.*, Using Deep Convolutional Neural Networks for Neonatal Brain Image Segmentation. *Frontiers in Neuroscience* **14**, (2020).
55. A. Makropoulos, S. J. Counsell, D. Rueckert, A review on automatic fetal and neonatal brain MRI segmentation. *NeuroImage* **170**, 231-248 (2018).
56. I. S. Gousias *et al.*, Magnetic resonance imaging of the newborn brain: Manual segmentation of labelled atlases in term-born and preterm infants. *NeuroImage* **62**, 1499-1509 (2012).
57. A. Hammers *et al.*, Three-dimensional maximum probability atlas of the human brain, with particular reference to the temporal lobe. *Human Brain Mapping* **19**, 224-247 (2003).

58. P. Craven, G. Wahba, Smoothing noisy data with spline functions. *Numerische Mathematik* **31**, 377-403 (1978).
59. G. S. Kimeldorf, G. Wahba, A Correspondence Between Bayesian Estimation on Stochastic Processes and Smoothing by Splines. *Ann. Math. Statist.* **41**, 495-502 (1970).
60. G. H. Golub, M. Heath, G. Wahba, Generalized Cross-Validation as a Method for Choosing a Good Ridge Parameter. *Technometrics* **21**, 215-223 (1979).
61. R. A. Rigby, D. M. Stasinopoulos, Smooth centile curves for skew and kurtotic data modelled using the Box-Cox power exponential distribution. *Stat Med* **23**, 3053-3076 (2004).
62. P. S. Huppi *et al.*, Quantitative magnetic resonance imaging of brain development in premature and mature newborns. *Ann Neurol* **43**, 224-235 (1998).
63. *Curve fitting toolbox : for use with MATLAB® : user's guide*. (Version 1. Natick, MA : MathWorks, 2001., 2001).

Acknowledgements

We are grateful to T. Sauer and S. Sinnar for helpful discussion, and to Y. Wang and J. Chai for technical help in compiling data. This research was supported by the Penn State and National Science Foundation Center for Healthcare Organization Transformation (CHOT) collaboration, US National Institutes of Health grants R01HD085853 (VC, AK, BCW, VM, SJS), and the NIH Director's Transformative Award R01AI145057 (MRP, JNP, SJS).

Author Contributions

Conception and design: SJS, MRP. Acquisition of data: MRP, VC. Analysis and interpretation of data: MRP, VC, VM, JNP, AK, BCW, SJS. Drafting the article: MRP, SJS. Critically revising the article: all authors. Reviewed and approved submitted version of manuscript: all authors.

Competing Interests

The authors declare no competing interests.

Figure Legends

Figure 1. Brain Component Sex Differences. SSANOVA models with random effects were fit to **a)** total brain volume (cc), **b)** cerebrospinal fluid (CSF) (cc), **c)** grey matter (cc), **d)**white matter (cc), **e)** grey to white matter ratio, **f)** height for age normalized brain volume, **g)** weight for height normalized brain volume, and **h)** brain to CSF ratio. For each plot, the male data is shown in blue and the female data in red. The dashed line represents the fit, with the two solid lines on either side showing the Bayesian 95% confidence intervals. The upper and lower intervals for the right and left sides overlap throughout the entire time frame for the brain/CSF ratio, and therefore there are no statistically significant differences at any age. For the other plots, time periods of significant differences where the intervals do not overlap are shown in yellow.

Figure 2. Brain Region Hemispheric Differences. SSANOVA models with random effects were fit to **a,b)** hemispheres (cc), **c,d)** temporal lobes (cc), and **e,f)** hippocampi (cc). Female regional growth curves from birth to 18-years-old are shown in the left column, and the corresponding male regional curves are shown in the right column. For each plot, the left side is shown in green and the right side in purple. The dashed line represents the fit, with the two solid lines on either side showing the Bayesian 95% confidence intervals. The upper and lower intervals for the right and left sides overlap throughout the entire time frame for the hemispheres, and therefore there are no statistically significant differences. For the temporal lobes and hemispheres, time periods of significant differences where the intervals do not overlap are shown in yellow.

Figure 3. Cognitive Score Correlations. Two cognitive scores were used, the Mental Development Index (MDI) for infants from birth to age 3, and the Wechsler Abbreviated Scale

of Intelligence (WASI) Tests for adolescents aged 6 to 18. Linear mixed effects models were fit for **a)** the relationship between MDI and brain volume z-score, as well as **b)** the relationship between WASI and brain volume z-score. Age-dependent correlations were calculated with a subpopulation window-size of 120 and step-size of 20 subjects to show the Pearson Correlation between cognitive score and **c)** raw brain volume, **d)** brain volume z-score, and **e)** weight-for-height normalized brain volume z-score at different ages.

Figure 4. Standard Brain Volume Growth Curves. To provide a standard tool for researchers and clinicians, normal curves resembling head circumference growth curves were fit using GAMLSS to show the 3rd, 15th, 50th, 85th, and 97th percentiles of normal brain volume growth for **a)** females and **b)** males, as well as the ratio of brain to CSF for **c)** females and **d)** males. The growth curves were modeled using a Box-Cox power exponential distribution and smoothed using fractional polynomials. Both male and female brain volume plots include insets for brain growth from birth to age 2, as well as CSF accumulation from birth to age 18, in order to provide a comprehensive picture of brain growth dynamics in the pediatric age range.

Supplemental Material

Supplemental Figure 1. Cohort Characteristics. Data from each age group from 13 days to 22 years was included in the cohort, for both **a)** males and **b)** females. The cross-sequential study included **c)** 505 total subjects, with most subjects having between 2 and 3 longitudinal visits (separated by 2 years). For this study, only scans from 13 days to 18 years were included in order to focus on the pediatric age range.

Supplemental Figure 2. Segmentation Results. Two different algorithms for volume quantification were used, one for the neonates and one for the older cohort. As an example, a **a)** T2 MRI scan from a 16-day-old female is shown. This scan was processed through the dHCP pipeline to produce a **c)** segmentation including grey matter (found by adding the deep grey matter and cortical grey matter), CSF, and white matter. On the right, the **b)** T1 MRI scan from an older patient (a 15-year-old male) is shown. This scan was processed using the CAT12 pipeline within SPM to produce the **d)** segmentation including grey matter, CSF, and white matter.

Supplemental Figure 3. Brain Lobe and Cerebellum Hemispheric Differences. SSANOVA models with random effects were fit to **(a,b)** frontal lobes (cc), **(c,d)** parietal lobes (cc), **(e,f)** occipital lobes (cc), **(g,h)** and cerebella (cc). Female regional growth curves from birth to 18-years-old are shown in the left column, and the corresponding male regional curves are shown in the right column. For each plot, the left side is shown in green and the right side in purple. The dashed line represents the fit, with the two solid lines on either side showing the Bayesian 95% confidence intervals. The upper and lower intervals for the right and left sides overlap throughout the entire time frame for each region, and therefore there are no statistically significant differences.

Supplemental Figure 4. Brain Component Standard Growth Curves. Growth curves from birth to 18-years-old developed with Generalized Additive Models for Location, Shape, and Scale are shown for males in blue and females in red, with the solid median accompanied by dotted one and two standard deviations above and below. Each growth curve was modeled using a Box-Cox power exponential distribution and smoothed using fractional polynomials. The yellow boxes represent Bonferroni corrected significant differences ($p<0.00006$) between males and females for each year, obtained using the Mann-Whitney test. This univariate testing was applied to obtain conservative estimates of the differences in the data for each year of life. The standard curves plotted include **a)** total brain volume (cc), **b)** CSF, **c)** grey matter, **d)**white matter, **e)** the ratio of total brain volume to CSF, **f)** the ratio of grey to white matter, **g)** head circumference (cm), **h)** height normalized brain volume, **i)** weight for height normalized brain volume, and **j)** height for age normalized brain volume.

Supplemental Figure 5. Brain Region Standard Growth Curves. Female regional growth curves from birth to 18-years-old are shown in the left column, and the corresponding male regional curves are shown in the right column, all developed using Generalized Additive Models for Location, Shape, and Scale. Each growth curve was modeled using a Box-Cox power exponential distribution and smoothed using fractional polynomials. Left regions are shown in green, and right regions are shown in purple, with the solid median accompanied by dotted one and two standard deviations above and below. The yellow boxes represent Bonferroni corrected significant differences ($p<0.00006$) between left and right components in that year, obtained using the Mann-Whitney test. The green boxes represent significant differences between the left region for males and females in that year, while the purple boxes represent significant differences between the right region for males and females in that year. The standard curves

plotted include **(a,b)** hemispheres (cc), **(c,d)** frontal lobes (cc), **(e,f)** temporal lobes (cc), **(g,h)** parietal lobes (cc),**(i,j)** occipital lobes (cc), **(k,l)** cerebella (cc), and **(m,n)** hippocampi (cc).

Supplemental Figure 6. Cognitive Score Correlations by Sex. Age-dependent correlations were calculated with a subpopulation window-size of 120 and step-size of 20 subjects to show the Pearson Correlation and confidence intervals between cognitive score and **(a,b)** raw brain volume, **(c,d)** brain volume z-score, and **(e,f)** weight-for-height normalized brain volume z-score at different ages for both females and males.

Expanded Data:

Supplemental Excel File: Supplemental_Master_File.xlsm

Figure 1

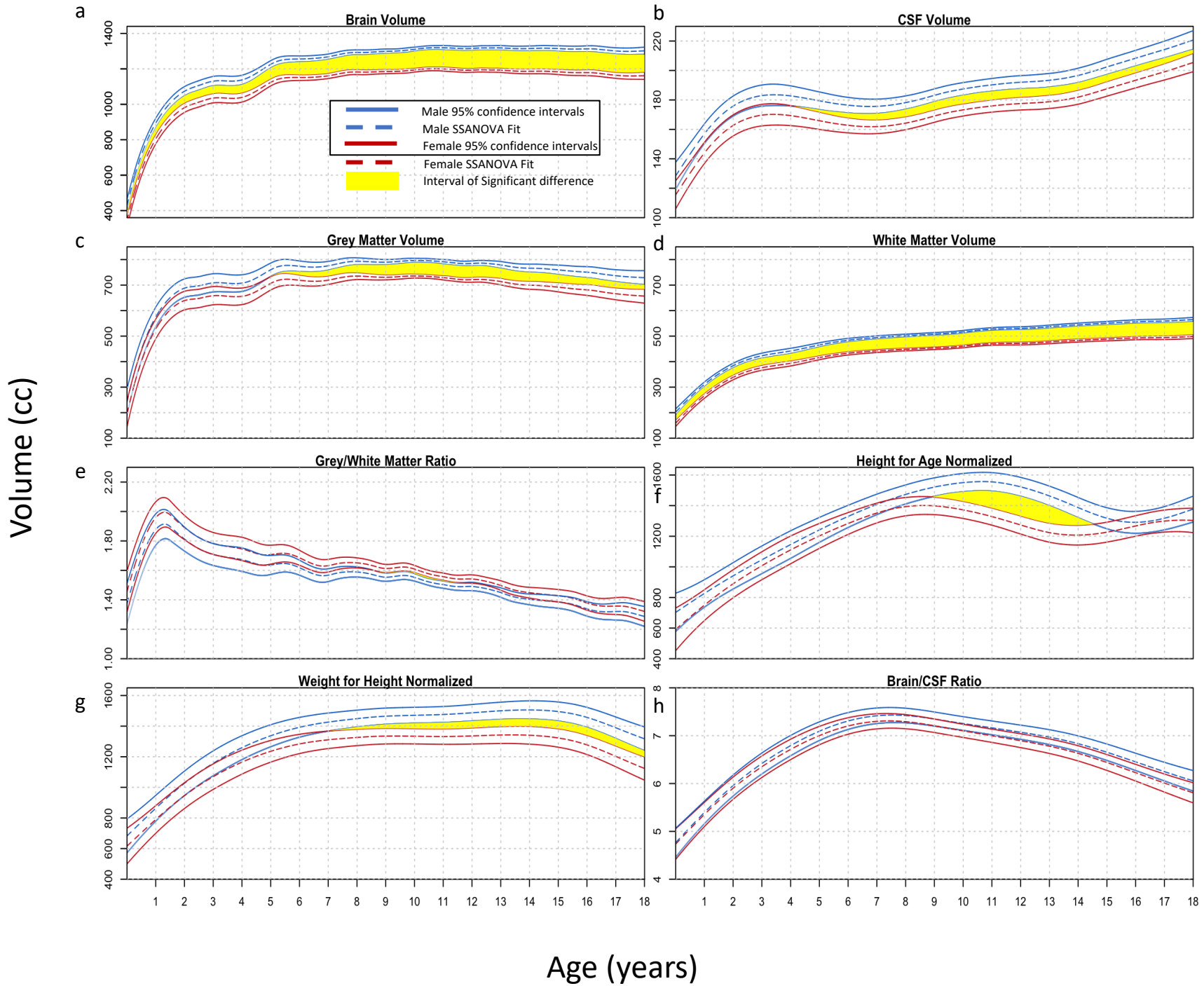

a
Brain Volume
b
CSF Volume
c
Grey Matter Volume
d
White Matter Volume
e
Grey/White Matter Ratio
f
Height for Age Normalized
g
Weight for Height Normalized
h
Brain/CSF Ratio
Male 95% confidence intervals
Male SSANOVA Fit
Female 95% confidence intervals
Female SSANOVA Fit
Interval of Significant difference
Volume (cc)
Age (years)

Figure 2

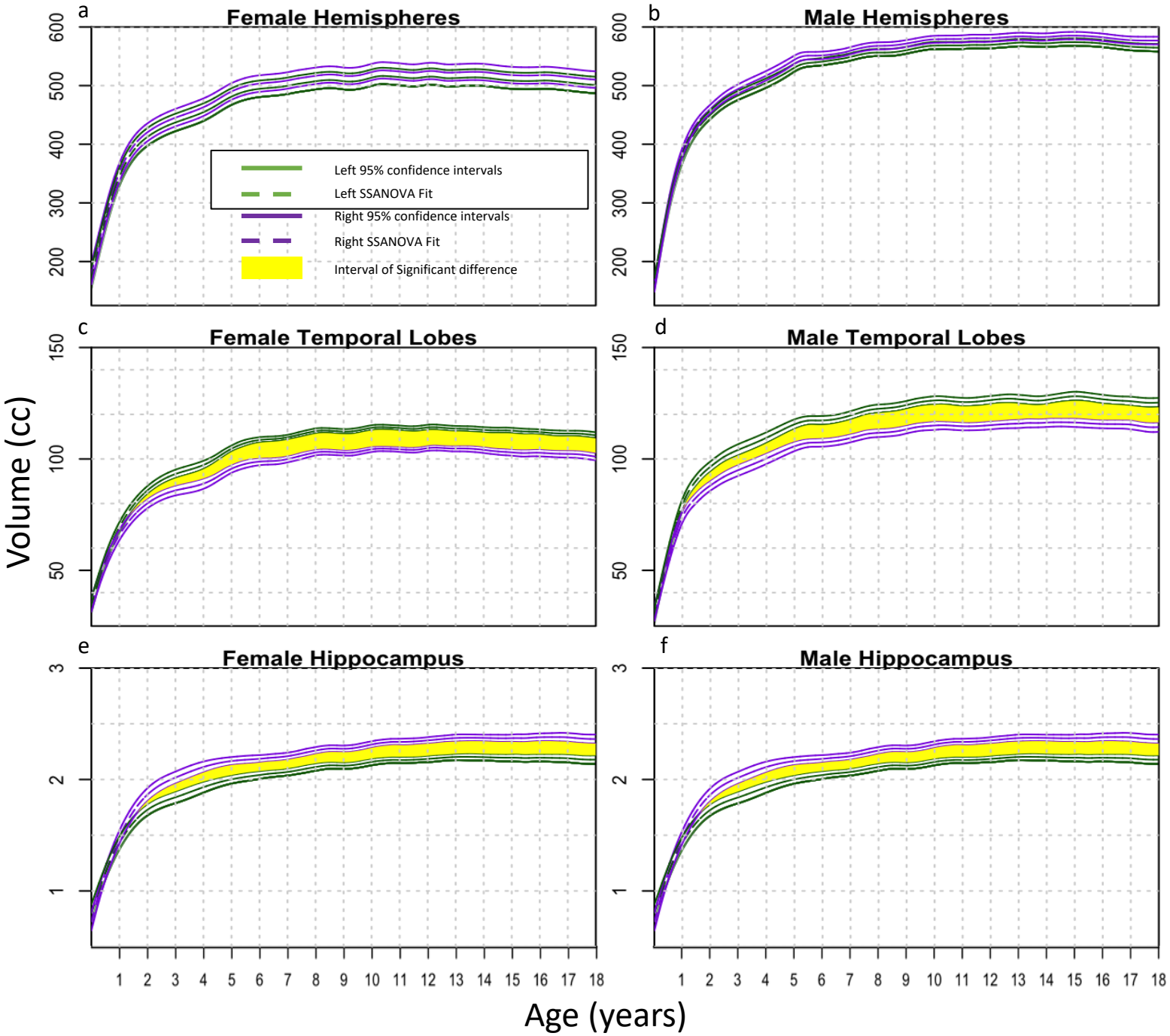

a
Female Hemispheres
b
Male Hemispheres
c
Female Temporal Lobes
d
Male Temporal Lobes
e
Female Hippocampus
f
Male Hippocampus
Left 95% confidence intervals
Left SSANOVA Fit
Right 95% confidence intervals
Right SSANOVA Fit
Interval of Significant difference
Volume (cc)
Age (years)

Figure 3

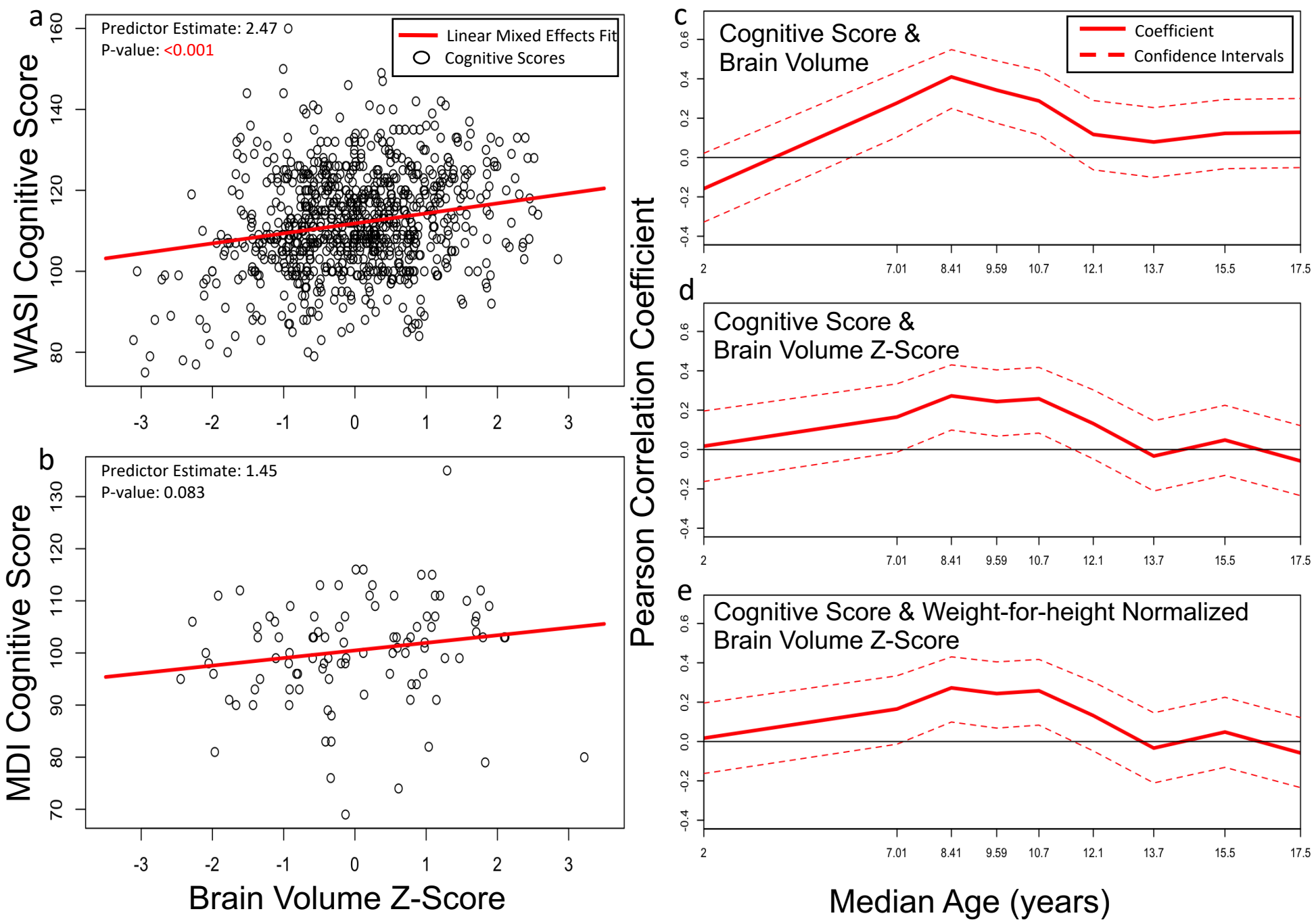
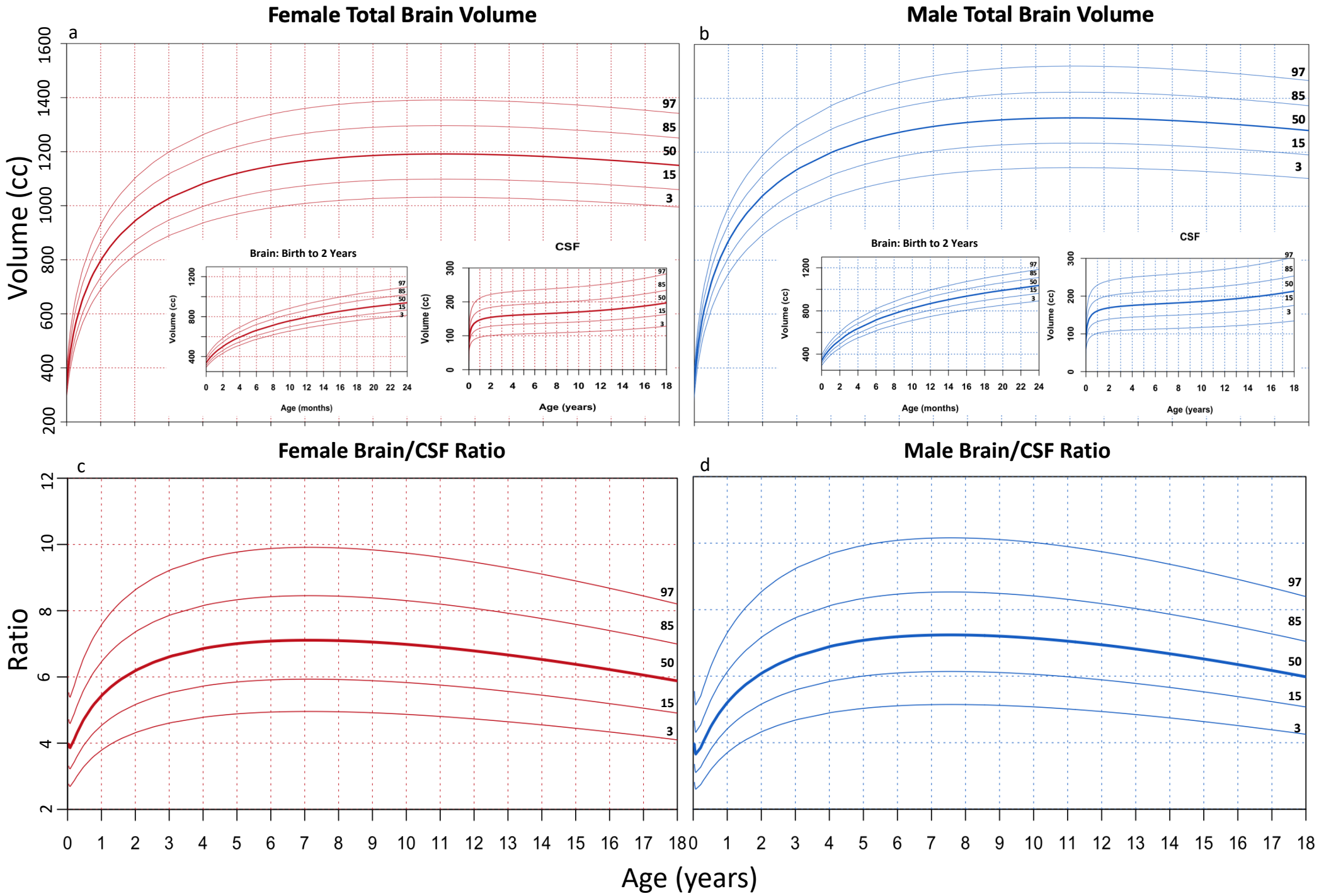

a
Predictor Estimate: 2.47
P-value: <0.001
Linear Mixed Effects Fit
Cognitive Scores
WASI Cognitive Score
b
Predictor Estimate: 1.45
P-value: 0.083
MDI Cognitive Score
Brain Volume Z-Score
c
Cognitive Score &
Brain Volume
Coefficient
Confidence Intervals
d
Cognitive Score &
Brain Volume Z-Score
e
Cognitive Score & Weight-for-height Normalized
Brain Volume Z-Score
Pearson Correlation Coefficient
Median Age (years)

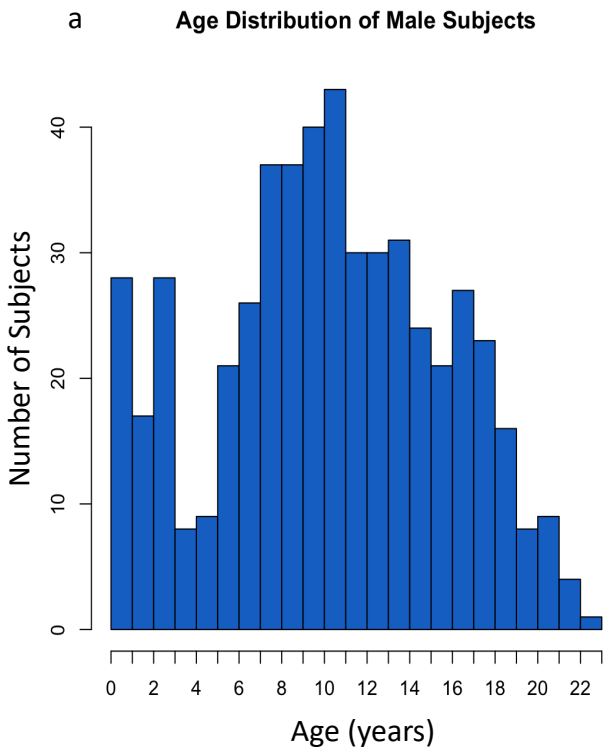
Figure 4


Female Total Brain Volume
a
Volume (cc)
97
85
50
15
3
Brain: Birth to 2 Years
Volume (cc)
Age (months)
CSF
Volume (cc)
Age (years)
Male Total Brain Volume
b
Brain: Birth to 2 Years
Volume (cc)
Age (months)
CSF
Volume (cc)
Age (years)
Female Brain/CSF Ratio
c
Ratio
Male Brain/CSF Ratio
d
Age (years)

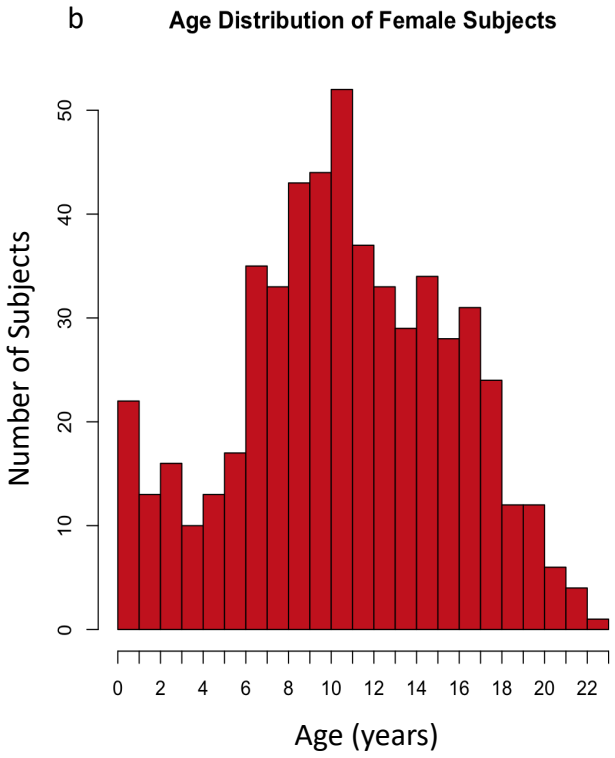
Supplemental Figures

Supplemental Figure 1

a



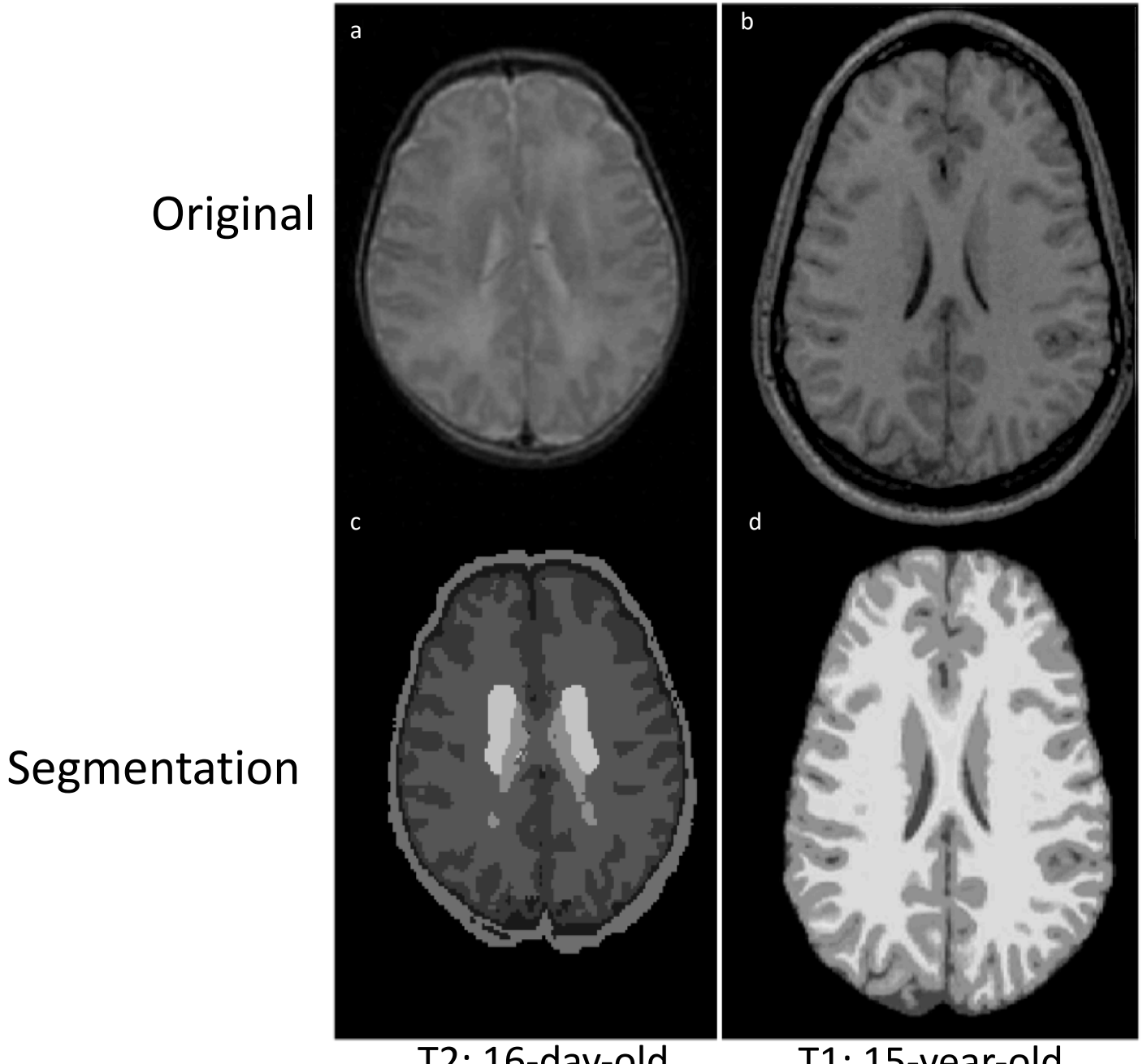
b



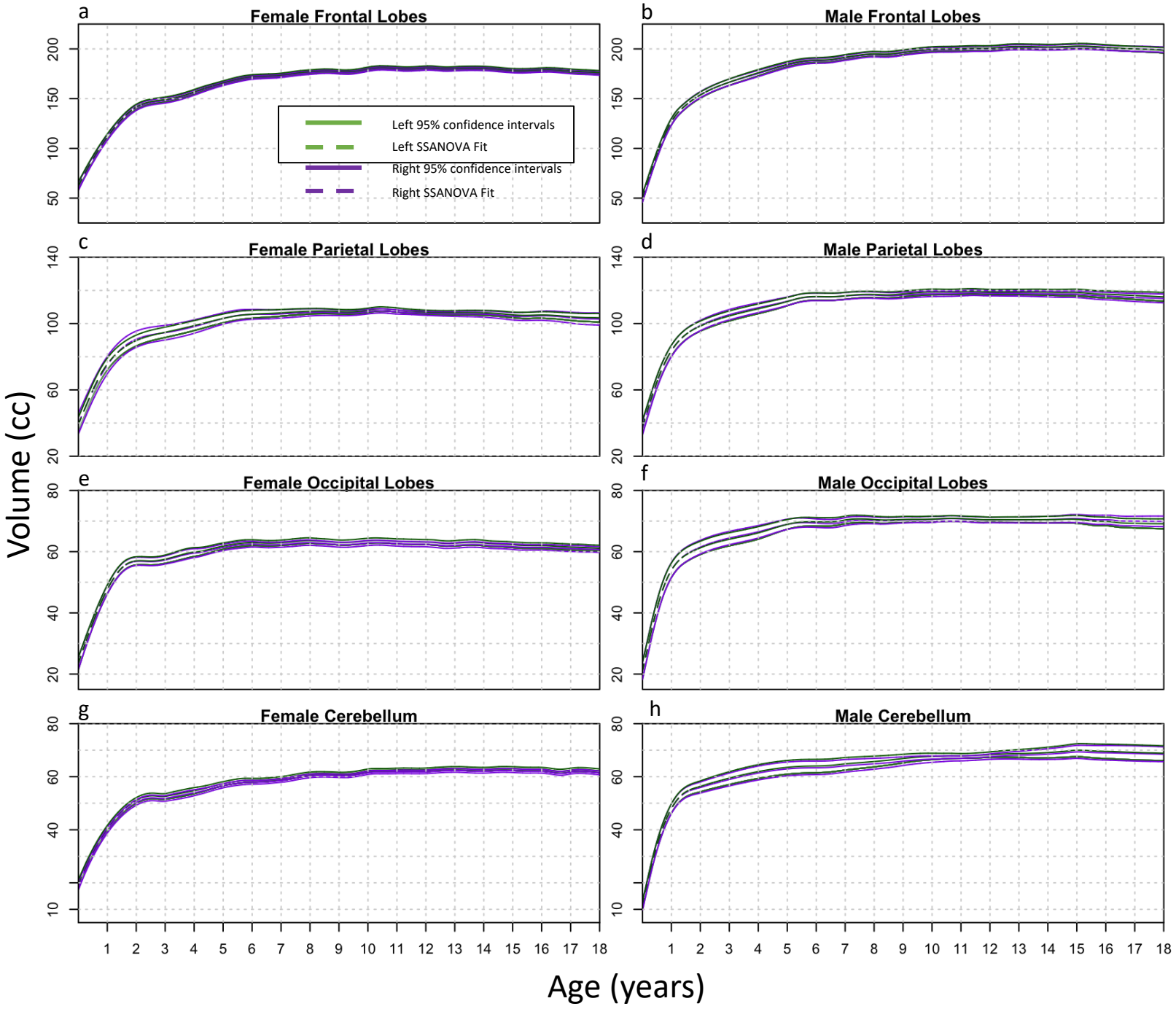
c

	Female	Male	Total
Scans	549	518	1067
Subjects	259	246	505
Min Age (days)	14	13	NA
Max Age (days)	8131	8073	NA

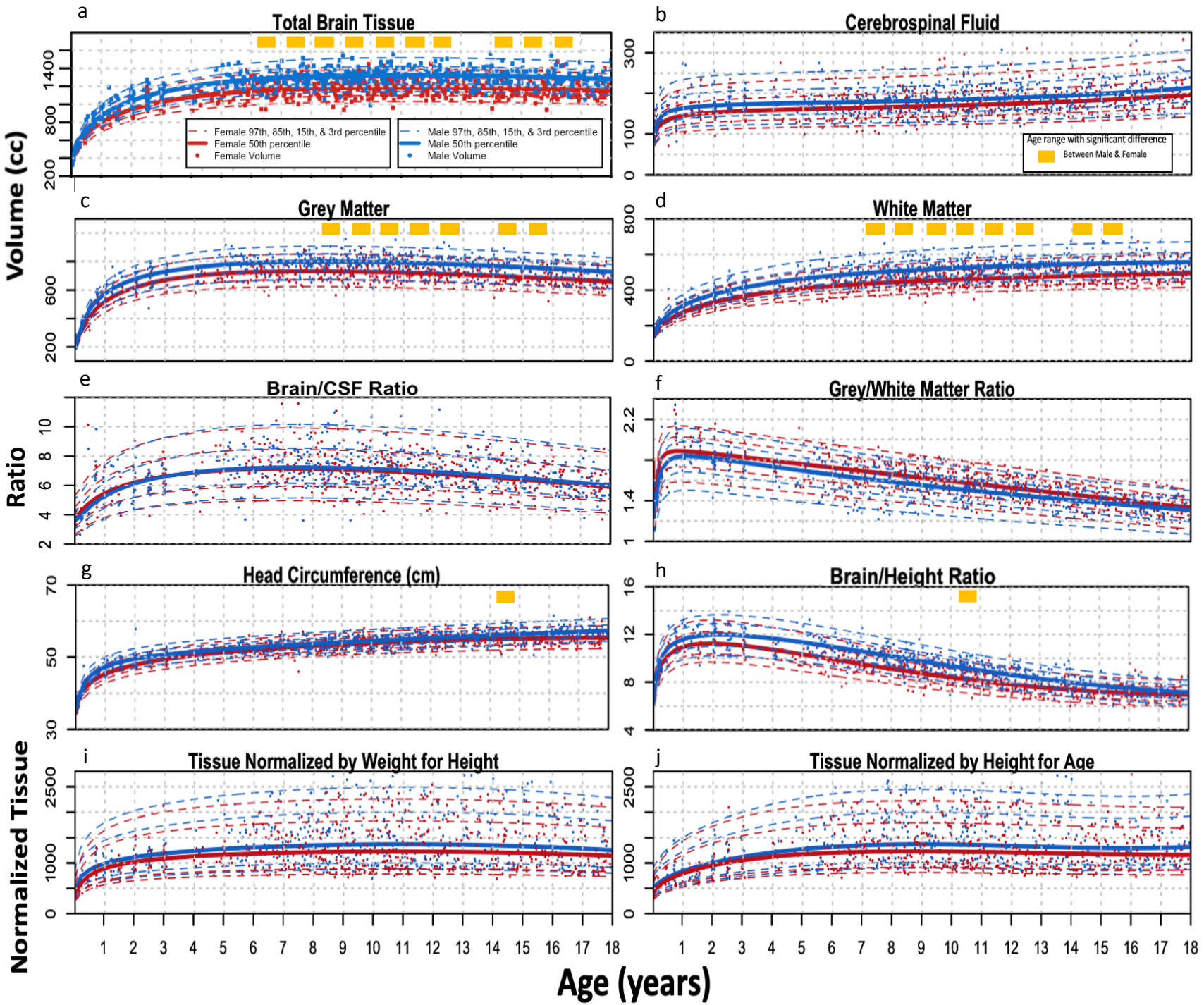
Supplemental Figure 2



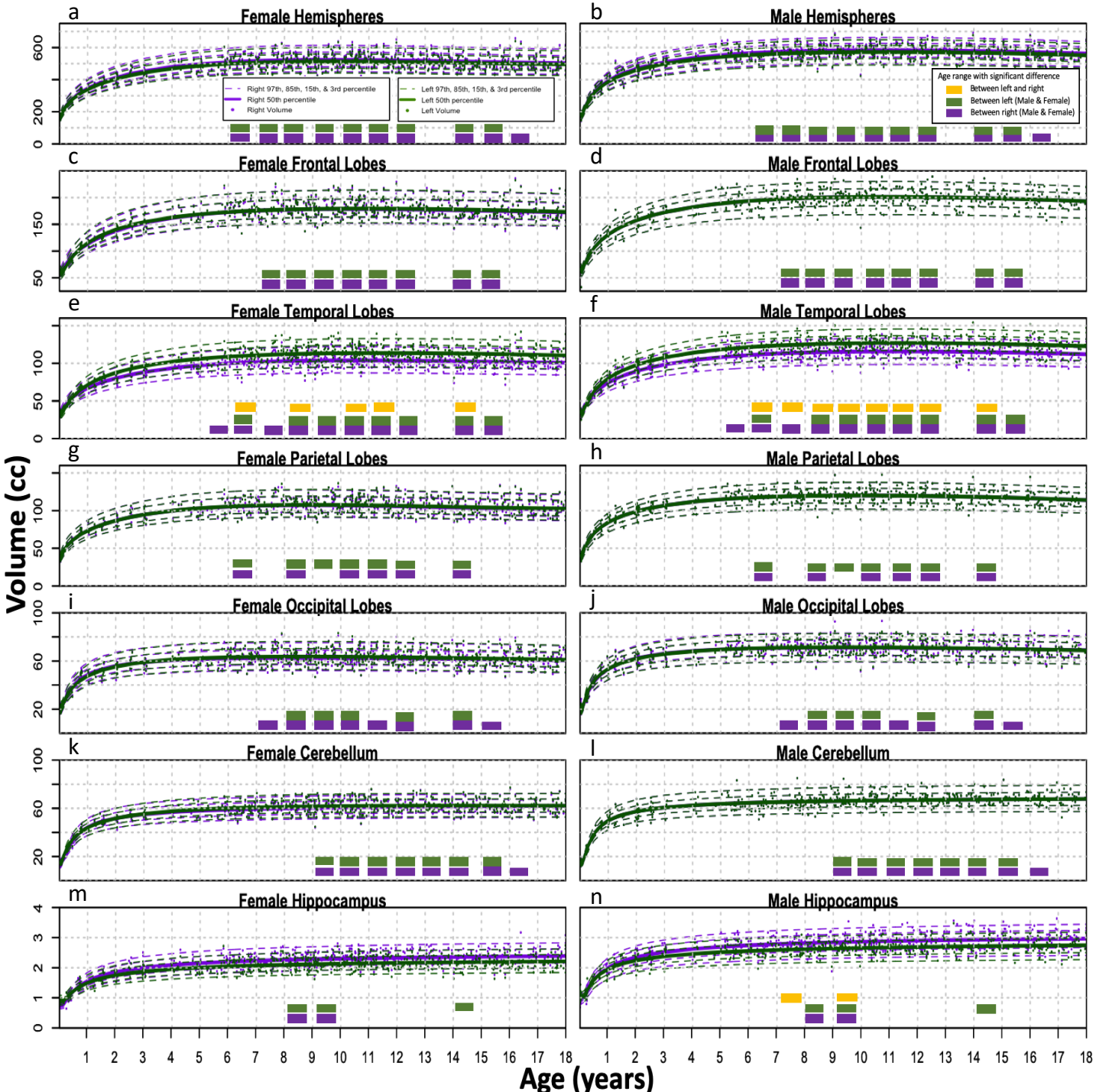
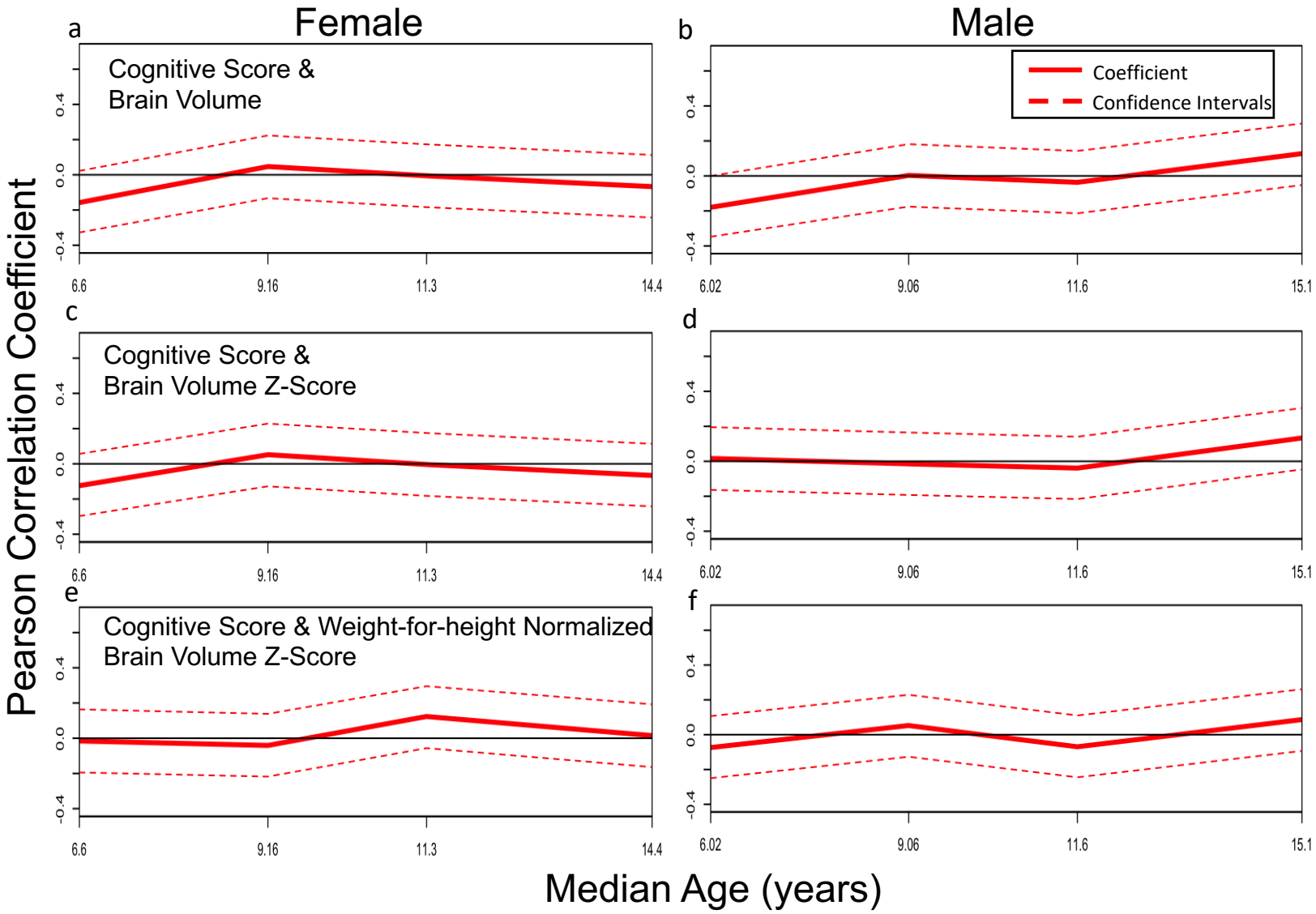
Supplemental Figure 3



Supplemental Figure 4



Supplemental Figure 5



Supplemental Figure 6