A Pediatric Structural MRI Analysis of Healthy Brain Development From Newborns to Young Adults

Jacob Levman,[1,2,3,4]* Patrick MacDonald,[1] Ashley Ruyan Lim,[1] Cynthia Forgeron,[4] and Emi Takahashi[1,2,3]

[1]*Division of Newborn Medicine, Department of Medicine, Boston Children's Hospital, 1 Autumn Street #456, Boston, MA*
[2]*Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, 149 13th Street, Charlestown, MA*
[3]*Department of Pediatrics, Harvard Medical School, Boston, MA*
[4]*Department of Mathematics, Statistics and Computer Science, St. Francis Xavier University, Antigonish, Nova Scotia B2G 2W5, Canada*

Abstract: Assessment of healthy brain maturation can be useful toward better understanding natural patterns of brain growth and toward the characterization of a variety of neurodevelopmental disorders as deviations from normal growth trajectories. Structural magnetic resonance imaging (MRI) provides excellent soft-tissue contrast, which allows for the assessment of gray and white matter in the developing brain. We performed a large-scale retrospective analysis of 993 pediatric structural brain MRI examinations of healthy subjects ($n = 988$, aged 0–32 years) imaged clinically at 3 T, and extracted a wide variety of measurements such as white matter volumes, cortical thickness, and gyral curvature localized to subregions of the brain. All extracted structural biomarkers were tested for their correlation with subject age at time of imaging, providing measurements that may assist in the assessment of neurological maturation. Additional analyses were also performed to assess gender-based differences in the brain at a variety of developmental stages, and to assess hemispheric asymmetries. Results add to the literature by analyzing a realistic distribution of healthy participants imaged clinically, a useful cohort toward the investigation and creation of diagnostic tests for a variety of pathologies as aberrations from healthy growth trajectories. The next generation of diagnostic tests will be responsible for identifying pathological conditions from populations of healthy clinically imaged individuals. *Hum Brain Mapp 00:000–000, 2017.* **© 2017 Wiley Periodicals, Inc.**

Key words: magnetic resonance imaging; hemispheric asymmetry; gender; clinical; maturation

Additional Supporting Information may be found in the online version of this article.

Contract grant sponsor: National Institutes of Health; Contract grant numbers: R01HD078561 and R03NS091587; Contract grant sponsor: Natural Science and Engineering Research Council of Canada's Canada Research Chair; Contract grant number: 231266; Contract grant sponsor: Canada Foundation for Innovation; Contract grant number: R0176004; Contract grant sponsor: Nova Scotia Research and Innovation Trust; Contract grant number: R0176004; Contract grant sponsor: St. Francis Xavier University; Contract grant number: R0168020.

*Correspondence to: Dr Jacob Levman; Division of Newborn Medicine, Department of Medicine, Boston Children's Hospital, Harvard Medical School, 1 Autumn Street #456, Boston, MA 02115, USA. E-mail: jacob.levman@childrens.harvard.edu

Received for publication 18 April 2017; Revised 3 August 2017; Accepted 24 August 2017.

DOI: 10.1002/hbm.23799

Published online 00 Month 2017 in Wiley Online Library (wileyonlinelibrary.com).

INTRODUCTION

Magnetic resonance imaging (MRI) provides a wide variety of physiological/anatomical measurements distributed across a subject's brain, information that may assist in both clinical applications and basic research. In the brain, structural MRI provides the ability to differentiate between gray matter, white matter, and cerebrospinal fluid, which forms the basis for the extraction of a variety of biomarkers such as distributed white matter volume measurements, cortical (gray matter) thickness measurements, cortical folding/gyration-based measurements, cortical surface area measurements, and so on [Fischl, 2012].

The analysis of structural brain MR images from healthy subjects has been performed using automated quantification [Fischl, 2012]. However, existing studies are typically focused on prospective analyses, investigating populations that would not have received medical imaging had they have not been enrolled in the study. This biases the resultant population in favor of subjects with no clinical reason to be referred for medical imaging. A major interest in investigating healthy populations is to determine a baseline from which pathological conditions can be characterized and diagnosed as deviations from expected growth trajectories. In practice, a new diagnostic technology would have to be evaluated in a routine clinical environment to become an integral part of a patient's standard of care. As such, the creation of new diagnostic technologies might benefit from a more clinically realistic baseline of what is considered healthy as compared to those previously established prospectively in the literature. It is possible that prospective trials may yield a healthy population with differing distributions of biomarker growth trajectories than that exhibited in a clinically realistic healthy population and so research toward the creation of new diagnostic tests may benefit from retrospectively assembled datasets of clinically realistic healthy populations.

In addition to multiple relevant review articles [Giedd et al., 2015; White et al., 2010], there have been several studies focused on assessing brain development among healthy participants using structural MRI examinations with automated biomarker extraction technology [Fischl, 2012]. Literature studies have included assessment of variations in brain volumes with age [Ostby et al., 2009], assessing morphological correlations of regional brain volumes [Fan et al., 2011], assessing biomarkers that vary with pubertal stage [Goddings et al., 2014], assessing the effect of birth weight cortical surface area and volumetric measurements [Walhovd et al., 2012], longitudinal brain development [Tamnes et al., 2013; Tamnes et al., 2010; Vijayakumar et al., 2016], frontopolar cortical thickness [O'Donnell et al., 2005], gender differences [Koolschijn and Crone, 2013], brain age estimation [Franke et al., 2012], brain maturity assessment [Erus et al., 2015; Khundrakpam et al., 2015], and asymmetry analyses [Good et al., 2001; Watkins et al., 2001].

It is particularly challenging to assess imaging features in a pediatric population because of the structural changes between children and adults [Bunge et al., 2002; Casey et al., 1997; Fair et al., 2009; Gogtay et al., 2004; Reiss et al., 1996; Supekar et al., 2009; Thomas et al., 2001]. Important information regarding brain function is encoded in distributed patterns of brain activity and structure [Fox et al., 2005; McIntosh et al., 1996; Mesulam, 1981; Vaadia et al., 1995], and identifying these patterns is particularly challenging in a preadult population because of a rapidly changing physiology, a high degree of brain plasticity, small brain sizes, participant motion, and an incomplete understanding of brain development. This article presents an analysis of a large healthy population covering the full range of ages from newborns to young adults. The examinations were acquired as part of clinical imaging and thus represent the type of population that the next generation of clinical diagnostic tests will need to be able to discern from pathological conditions.

Abbreviations

BA	Brodmann's area (typically followed by a number indicating subregion)
BCH	Boston Children's Hospital
CI	curvature index—a biomarker measurement reflecting cortical gyrification
EEG	electroencephalography
FI	folding index—a biomarker measurement reflecting cortical gyrification
fMRI	functional magnetic resonance imaging
GMV	gray matter volume
LH	left hemisphere
MPRAGE	magnetization prepared rapid gradient echo sampling
MR	magnetic resonance
MRI	magnetic resonance imaging
MT	middle temporal visual area
NV	number of vertices on cortical surface
RF	radiofrequency
ROIs	regions-of-interest
S&G	sulcus and gyrus
SD	standard deviation
SI	signal intensity
T1	T1-weighted MRI
WH	whole hemisphere
WM	white matter

METHODS

Participant Population

This retrospective analysis was approved by the Institutional Review Board at *BCH*. The clinical imaging database at *BCH* was retrospectively reviewed for this analysis from 01 January 2008 until 24 February 2016. All MRI examinations that included brain imaging of participants aged 0–32 years at the time of imaging were included for further analysis. Examinations deemed to be of low quality

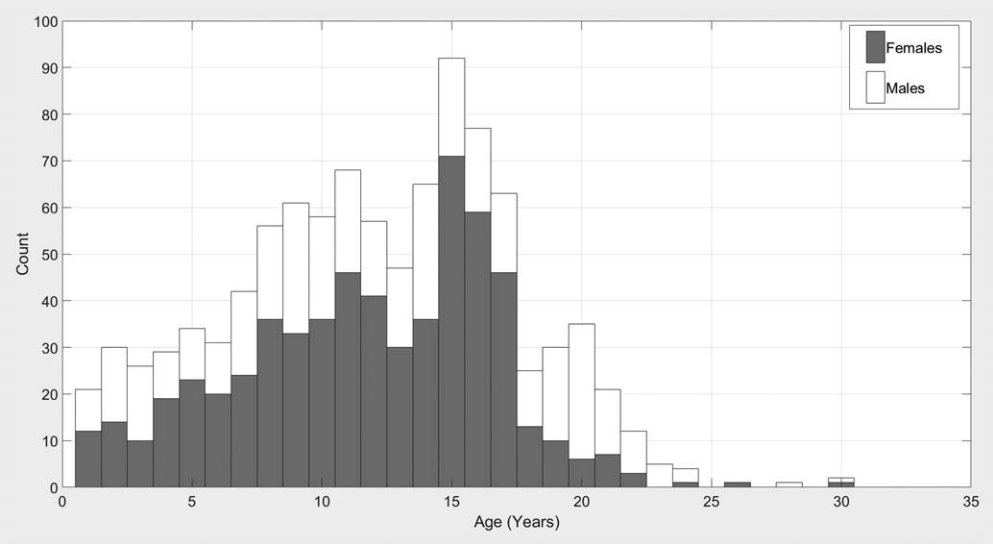

Figure 1.
Histograms showing age distributions for the male and female subjects included in this experiment.

(because of excessive participant motion, large metal artefact from a subject's dental hardware, lack of a T1 structural imaging volume providing diagnostically useful axial, sagittal, and coronal oriented images, etc.) were excluded from the study. Examinations and medical records that were inaccessible because of technical reasons were excluded. Healthy subjects were retrospectively identified from routine clinical imaging by including subjects with a normal MRI examination as assessed by a *BCH* radiologist and whose medical records provide no indication of any substantive neurological problems (subjects with any known neurodevelopmental disorder were excluded such as autism, cerebral palsy, attention deficit hyperactivity disorder, epilepsy, neurofibromatosis, developmental delay, tuberous sclerosis complex, hemiplegia, hallucinations, any brain tumor, bipolar disorder, obsessive compulsive disorder, abnormal psychological factors, meningitis, encephalopathy, postconcussion syndrome, learning disabilities, abnormal EEG examination, paresthesia, Bardet–Biedl syndrome, Waardenburg syndrome, cerebral venous thrombosis, demyelination, etc.). Participants with any type of cancer (including outside the central nervous system) were excluded to avoid neurological data with altered growth trajectories caused by common treatments such as chemotherapy. This yielded 993 examinations (395 male, 598 female) from 988 participants. A joint histogram demonstrating the age distributions for both the male and female healthy participants are provided in Figure 1.

MRI Data Acquisition and Preprocessing

All subjects were imaged with clinical 3 T MRI scanners (Skyra, Siemens Medical Systems, Erlangen, Germany) at *BCH* yielding T1 structural volumetric images for use in this study. Owing to the clinical and retrospective nature of this study, there is variability in the pulse sequences employed to acquire these volumetric T1 examinations including several types of MPRAGE acquisitions and a few traditional T1 structural sequences and volumetric spoiled gradient recalled sequences. Spatial resolution also varied but was, on average, approximately 1 mm. Strengths and limitations of the large-scale varying MR protocol approach taken in this study are addressed in the discussion. Motion correction was not performed, but examinations with substantive motion artefacts were excluded based on visual assessment. T1 structural examinations were processed with automated biomarker extraction technology [Fischl, 2012]. Measurement types extracted (with labels included in the Supporting Information tables provided here in brackets) include cortical thickness average (ThickAvg, the mean thickness of the regional gray matter), cortical thickness standard deviation (ThickStd, the variability in thickness of the regional gray matter), the number of vertices on the cortical surface (NumVert, NVertices, a measure of morphological complexity), vertex separation (also linked with morphological complexity), number of voxels (NVoxels, a volumetric measurement in voxel counts), signal intensity measurements (pertaining to the amount of MR signal observed regionally), signal-to-noise ratios (SNR, the amount of MR signal relative to the observed noise in a region), surface area (SurfArea, mm^2 measurement of cortical surface area), surface curvature (GausCurv, MeanCurv, FoldInd, CurvInd, pertaining to the extent of curvature on the cortical surface), surface integrals (gyral and sulcal biomarkers), and basic volumetric measurements (Volume_mm^3, GrayVol) by region.

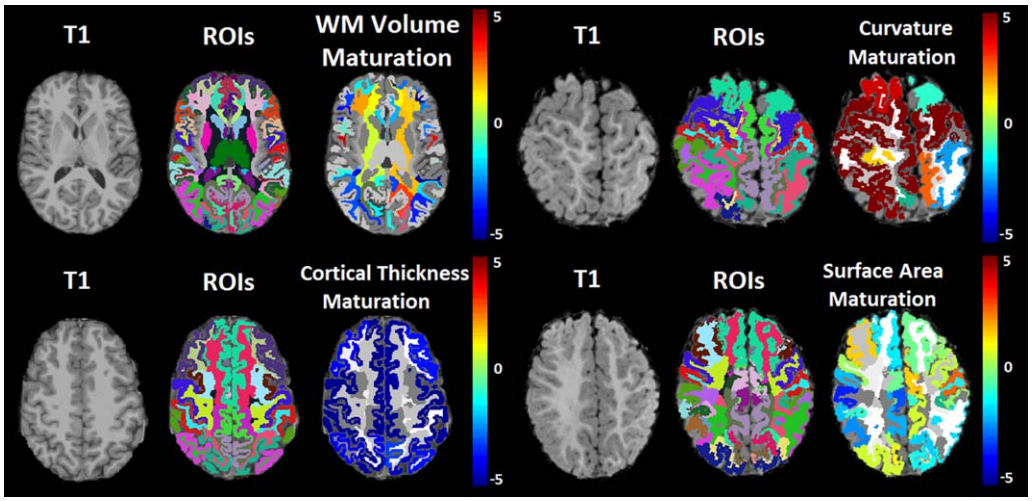

Figure 2.

Structural MRI examinations (labeled T1) of 18-year-old (left half) and 11-month-old (right half) participants, with FreeSurfer segmentations (labeled ROIs) and color maps assessing maturation relative to age expected values for a variety of subregions in the brain (color maps provided, units are in years to be interpreted as a relative measurement not absolute). Red regions indicate precocious development given participant age and blue regions indicate underdevelopment. [Color figure can be viewed at wileyonlinelibrary.com]

Statistical Analysis

An analysis was performed to determine what extracted physiological biomarker measurements are most correlated with subject age. This is provided to help identify those biomarker measurements with the most potential for the assessment of neurological maturation. Pearson's correlation coefficient was computed comparing subject age with each of the $m = 4{,}788$ measurements extracted. A Bonferroni correction was performed to identify those biomarkers statistically significantly correlated with participant age ($P < 0.05/\mathrm{m} = 1.04\mathrm{e}^{-5}$).

An age-dependent gender-segregated analysis was performed on the 4,788 extracted measurements per imaging examination [Fischl, 2012]. This experiment also involved repeating age-dependent gender segregated analyses from $\alpha - 2$ to $\alpha + 2$ years old (with α varied as an integer from 2 to 18 yielding age-dependent analyses of subjects 0–4, 1–5, 2–6, …, 16–20 years old). This yielded $m = 81{,}396$ groupwise comparative operations. Our dataset only included few samples >20 years old and so this range was not included in these age-dependent analyses because of sample size considerations; however, this age range was included in all scatter plots provided to allow visual assessment. A statistical analysis was performed comparing male and female participants across the outlined age groups using the Bonferroni corrected statistical significance threshold of $P < 0.05/\mathrm{m} = 6.14\mathrm{e}^{-7}$, which provides a much more stringent standard for reaching statistical significance when a large quantity of statistical comparisons are being performed. Statistical testing in the form of P values were computed in age-dependent groupwise comparisons using the standard t test [Student, 1908] for two groups of samples.

An asymmetry analysis was performed by creating an asymmetry index computed as left/right for each biomarker for which both left and right hemispheric measurements are available [Fischl, 2012]. This yielded 2,340 asymmetry measurements for further analysis. The average asymmetry measurement across all participants was computed. A gender segregated analysis that compares male and female participants for differences in asymmetry measurements was repeated for the age ranges described above yielding a Bonferroni corrected threshold for statistical significance of $P < 0.05/\mathrm{m} = 1.26\mathrm{e}^{-6}$. The correlation analysis described above was also repeated for our set of asymmetry measurements.

Structural MRI examinations were color coded to provide a demonstration of how deviations from expected neural maturation growth trajectories can be visually interpreted by a clinician. A second-order polynomial was fitted to the data distribution (Matlab, Natick, MA) allowing estimation of an individual biomarker's maturation by evaluating the expected participant age (that falls on the fitted polynomial) given the extracted biomarker value. A maturation index was computed as the difference between the estimated age (determined by the fitted polynomial) and the actual participant age. Thus, this demonstrative index is measured in years, however, variability in these measurements is such that the values should only be

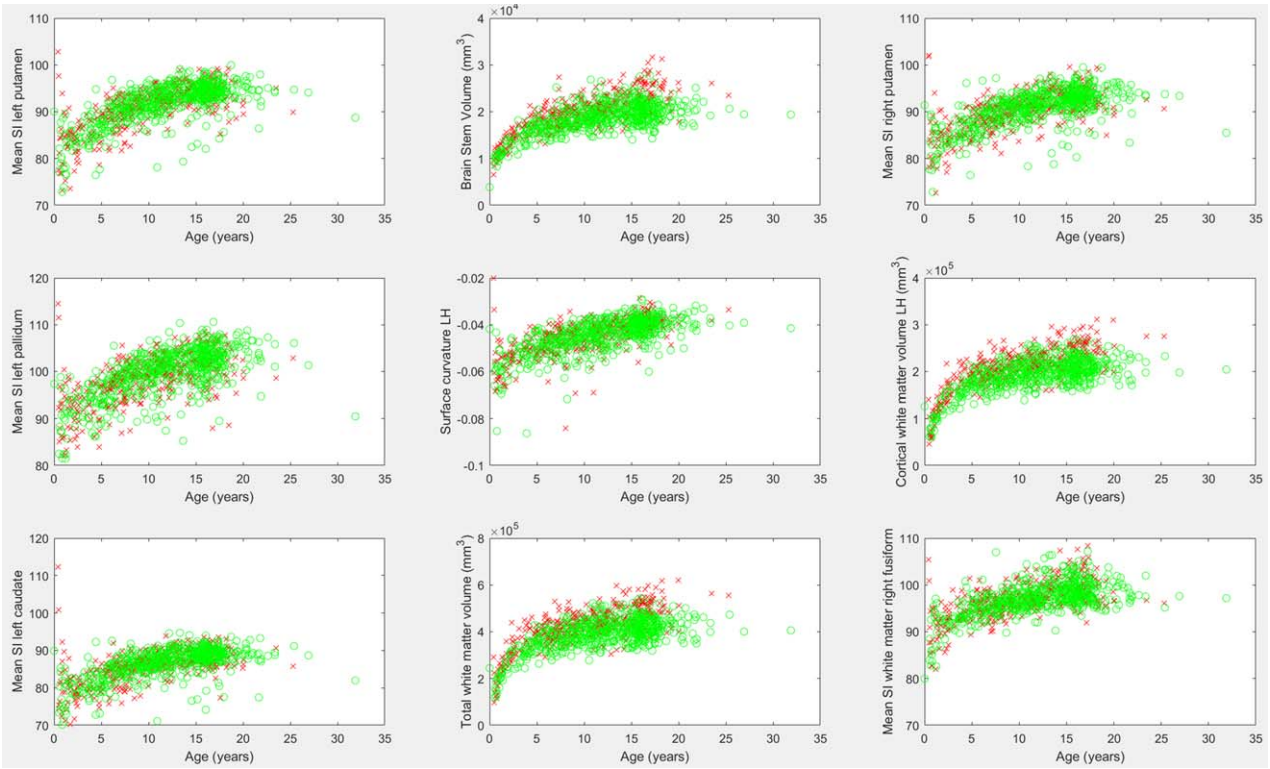

Figure 3.

Biomarker measurements identified as having the highest correlation with participant age. Males are represented by a red "x," females by a green "o." SI, signal intensity; LH, left hemisphere. [Color figure can be viewed at wileyonlinelibrary.com]

interpreted as a relative measure (i.e., color coding helps identify regions that are under- or overdeveloped).

RESULTS

The correlation analysis assessing the relationship between biomarkers and participant age yielded 2,955 measurements exceeding the Bonferroni corrected threshold for statistical significance ($P < 1.04e^{-5}$). All biomarkers are presented in the accompanying Supporting Information spreadsheet (Sheet 1: Age Correlation). The Supporting Information spreadsheet provides the FreeSurfer atlas and biomarker label, the type of measurement, the measurement's correlation coefficient when compared with participant age and the associated P value. Readers may benefit from referring to FreeSurfer biomarker label definitions [Destrieux et al., 2010] and online support (https://mail.nmr.mgh.harvard.edu/pipermail//freesurfer/) for more detailed information on the label listings in the Supporting Information spreadsheet. These results demonstrate that a wide variety of biomarkers have age-dependent profiles resulting in statistically significant correlations between measured biomarkers and participant age in the majority of measurements acquired (2,955/4,788 = 62% of biomarkers exceed the Bonferroni correction). Cortical thickness measurements exhibit negative correlations with participant age and volumetric and surface area measurements exhibit positive correlations with participant age.

Figure 2 presents the examples of segmented subregions of the brain that were used as ROIs for our analyses, and the degree of maturation based on white matter volumetric measurements, cortical thickness measurements, surface area measurements, and gyral curvature measurements. This demonstrates the feasibility of regionally assessing brain maturation by interpreting color-coded images that highlight differences from expected values given the participant's age. Scatter plots demonstrating age-dependent distributions of biomarkers extracted from male and female participants are provided in Figure 3. These biomarkers are most correlated with participant age and thus have potential toward helping assess maturation in the developing brain.

The age-dependent gender segregated analysis was performed on 4,788 FreeSurfer-derived biomarkers emphasizing those that exhibited Bonferroni corrected statistically significant differences between genders among some of

TABLE I. Biomarkers exhibiting hemispheric asymmetry arranged by measurement type

Type	Biomarker region (with mean asymmetry index: left/right provided)
Surface area	BA MT (1.83), BA4a (1.76), BA44 (1.64), perirhinal (1.57), transverse temporal (1.39), primary auditory cortex/anterior transverse temporal gyri (1.36), inferior occipital gyrus and sulcus (1.30), transverse frontopolar sulci and gyri (0.64), occipital pole (0.68)
Number of vertices	BA MT (1.83), BA4a (1.69), BA44 (1.66), perirhinal (1.55), transverse temporal (1.35), primary auditory cortex/anterior transverse temporal gyri (1.31), Superior circular sulcus of the insula (1.31), transverse frontopolar sulci and gyri (0.63), occipital pole (0.69)
Gray matter Volume	BA4a (1.68), BA MT (1.67), BA44 (1.56), transverse temporal (1.33), primary auditory cortex/anterior transverse temporal gyri (1.33), BA3b (1.31), Superior circular sulcus of the insula (1.31), occipital pole (0.67), transverse frontopolar sulci and gyri (0.70)
White matter volume	Transverse temporal (1.44)

BA, Brodmann's Area; MT, middle temporal visual area.

the age ranges analyzed. The Supporting Information spreadsheet provides a detailed listing of mean and standard deviations of all biomarkers for each gender across ages with comparative *P* values (Sheet 2: Gender). These results provide a baseline from which future studies can compare their results toward assessing gender effects in a clinically imaged population. These results also provide a baseline from which developmental disorders which present differently by gender can be characterized as departures from gender-specific healthy growth trajectories. Results demonstrate volumetric and surface area measurements exhibiting statistically significant shifts toward increased values among males relative to females for which groupwise differences tend to increase with age (highlighted in yellow). Results also demonstrate surface integral measurements that exhibit increased values among females relative to males (highlighted in green). Figure 3 shows the examples of summary results of the correlation analysis above, which demonstrates gender differences exhibited in the brain stem volume (center column, top row), the left hemisphere's cortical white matter volume (right column, middle row), and the total white matter volume differences (center column, bottom row).

Table I provides a listing of those biomarkers exhibiting the largest hemispheric asymmetry and includes measurements with a 30% or greater difference between the left and right hemispheres. Minimum signal intensity measurements and curvature measurements were excluded from the table due to the propensity of such biomarkers to have a value of zero among many participants, thus potentially making asymmetry measurements unreliable (results are provided for further inspection in the Supporting Information, sheet 3). Figure 4 provides the scatter plots of asymmetry biomarkers exhibiting leftward asymmetry. Several of Brodmann's Areas presented strong leftward asymmetries (BA44, BA4a, and BA MT—middle temporal visual area). Figure 5 provides scatter plots of asymmetry biomarkers exhibiting rightward asymmetry. The Bonferroni corrected correlation analysis of asymmetry biomarkers is presented in Table II, and demonstrates that some measurements exhibit a mild correlation with participant age which is illustrated in Figure 6. The results demonstrate that the asymmetry biomarkers exhibiting the strongest correlation with participant age do not exhibit strong asymmetries; however, there is a general tendency for hemispheric asymmetry to increase with age among these biomarkers. The Supporting Information spreadsheet includes a full listing of asymmetry measurements along with age correlations and associated *P* values (Sheet 3: Asymmetry with Age Correlation). Biomarkers exceeding the Bonferroni corrected threshold for statistical significance are highlighted in yellow. The gender segregated analysis was repeated for our asymmetry measures, none of which exceeded the Bonferroni corrected threshold for statistical significance, implying that gender differences in left-right asymmetry are not nearly as pronounced as gender differences observed in FreeSurfer's native biomarkers. A full listing of the mean, standard deviation, and comparative *P* values by gender for all age groups is provided in the Supporting Information spreadsheet (Sheet 4: Gender Asymmetry).

DISCUSSION

This study presented a retrospective analysis of brain MRI acquired from 988 healthy participants. Our analysis demonstrates that 62% of FreeSurfer derived biomarkers exceed the Bonferroni corrected threshold for statistical significance in age correlation analyses, implying that the majority of FreeSurfer derived biomarkers are age-dependent. As such, the assessment of pediatric FreeSurfer biomarkers in literature studies should account for participant age. The many strong correlations between participant age and measured biomarkers also provides a foundation on which future research can assess a participant's regionally distributed brain maturation (example demonstration in Fig. 2), which has been shown to be correlated with cognitive performance for both developmentally delayed individuals and those with cognitive precocity [Erus et al., 2015]. This demonstration (Fig. 2) is consistent with studies that have indicated that maturation

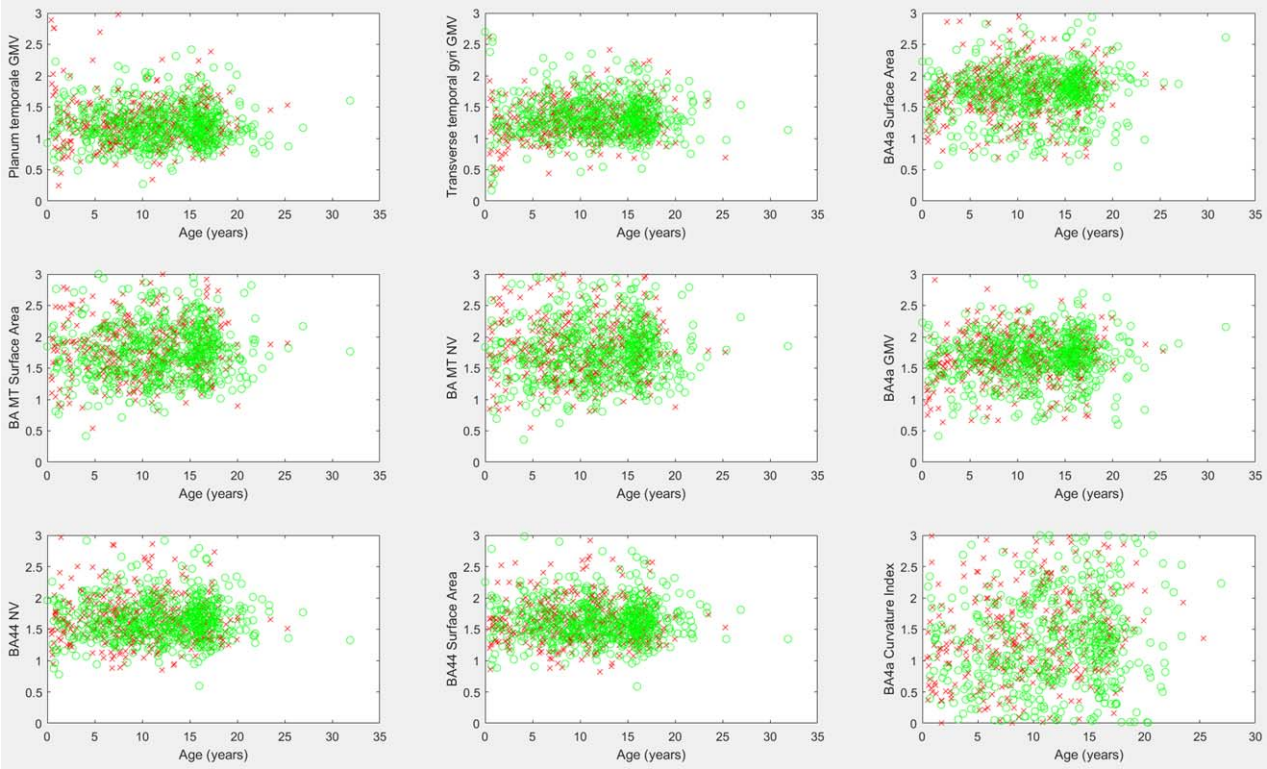

Figure 4.

Scatter plots of biomarkers exhibiting leftward asymmetries. Males are marked with a red "x," females with a green "o." Asymmetry scores (*y* axes) were acquired by dividing the left hemisphere's biomarker by the corresponding right hemisphere biomarker. Values above 1 indicate a leftward asymmetry. GMV, gray matter volume; BA, Brodmann's area; MT, middle temporal visual area; NV, number of vertices. [Color figure can be viewed at wileyonlinelibrary.com]

can be assessed with gray matter thickness [Khundrakpam et al., 2015; Tamnes et al., 2010] and volumetric measurements [Tamnes et al., 2010]. This study has presented the results of a gender segregated analysis on an age-dependent basis demonstrating volumetric and surface area measurements exhibiting statistically significant shifts toward increased values among males relative to females for which groupwise differences tend to increase with age. Observed gender differences are in agreement with literature findings [Koolschijn and Crone, 2013; Reiss et al., 1996]. This article has also presented an asymmetry analysis adapted for biomarkers extracted with FreeSurfer technology, demonstrating modest correlations with participant age and a lack of substantial asymmetry gender differences. Left dominance in language areas was observed, in agreement with literature findings [Good et al., 2001; Watkins et al., 2001]. We observed asymmetric findings in perirhinal cortical volumes, consistent with literature findings of control subjects [Jutila et al., 2001] and observed increased volumes of BA44 in the left hemisphere, also consistent with literature findings [Amunts et al., 1999; Broca, 1861; Cantalupo and Hopkins, 2001; Foundas et al., 1998; Galaburda, 1984]. We observed hemispheric asymmetries in Brodmann's area 4 which corresponds to the primary motor cortex, findings that confirm known asymmetries [Amunts et al., 1996] which have been linked with handedness. Additionally, we observed hemispheric asymmetries in the middle temporal visual area which is consistent with known asymmetries across regions responsible for visual processing in humans [Amunts et al., 2007] and macaques [Burkhalter et al., 1986]. These results provide a useful baseline for future research toward detecting and characterizing a variety of pathological conditions as departures from expected growth trajectories.

This study focused on a population of healthy participants who were imaged by clinical MRI. This represents both a strength and weakness relative to typical prospective analyses that recruit healthy participants into their studies. Retrospective assessment of clinical imaging induces a bias in favor of populations with symptoms and medical histories that lead to an MRI examination, including many participants with headaches, abnormalities of the ear, eye, pineal gland, pituitary gland, and so on. Theoretically, such a population may exhibit a wider

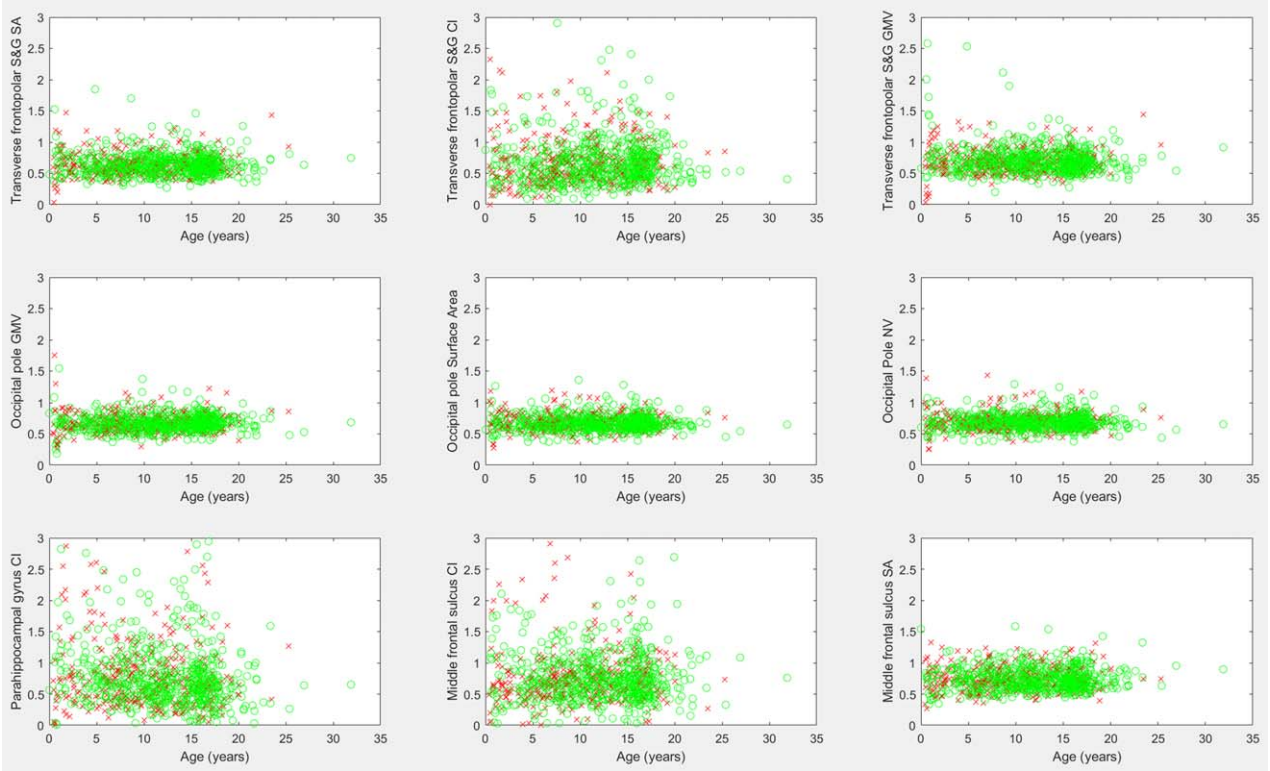

Figure 5.

Scatter plots of biomarkers exhibiting strong rightward asymmetries. Males are marked with a red "x," females with a green "o." Asymmetry scores (*y* axes) were acquired by dividing the left hemisphere's biomarker by the corresponding right hemisphere biomarker. Values below 1 indicate a rightward asymmetry. CI, curvature index; S&G, sulcus and gyrus; SA, surface area; GMV, gray matter volume; NV, number of vertices. [Color figure can be viewed at wileyonlinelibrary.com]

variability in some biomarkers than would be found in a prospectively recruited healthy population that had no medical reason to be referred for an MRI. Thus, although a standard prospective analysis of recruited healthy participants avoids biasing the sample toward participants with pre-existing conditions that lead to imaging, these studies also bias their samples away from a clinically realistic healthy population. In clinical practice, a diagnostic technology responsible for identifying a pathological condition will be responsible for correctly identifying the condition of interest as differing from a pool of neurologically healthy individuals with a wide variety of reasons for

TABLE II. Asymmetry biomarkers (left hemisphere measurement divided by right hemisphere measurement) most highly correlated with participant age with left (L) or right (R) dominance specified after the asymmetry index

Biomarker	Asymmetry index	*P* value	Correlation coefficient
Mean SI medial orbitofrontal WM	0.9949 R	$1.34e^{-26}$	0.3296
Choroid plexus volume (mm^3)	0.9164 R	$1.14e^{-25}$	0.3238
Mean SI putamen	1.0100 L	$1.31e^{-17}$	0.2665
Mean SI lateral orbitofrontal WM	1.0140 L	$1.47e^{-16}$	0.2580
Mean SI cerebellar WM	1.0141 L	$1.04e^{-14}$	0.2421
Thalamus volume (mm^3)	1.1235 L	$1.30e^{-14}$	0.2412
Inferior parietal WM SI standard deviation	1.0253 L	$2.46e^{-13}$	0.2295
Number of vertices on superior temporal surface	1.0500 L	$3.79e^{-13}$	0.2277
Positive surface integral WH (sulcal curvature)	0.9911 R	$5.39e^{-13}$	0.2263
WH cortical white matter volume (mm^3)	0.9862 R	$1.11e^{-11}$	0.2133

SI, signal intensity; WM, white matter; WH, whole hemisphere measurement, not localized to one subregion.

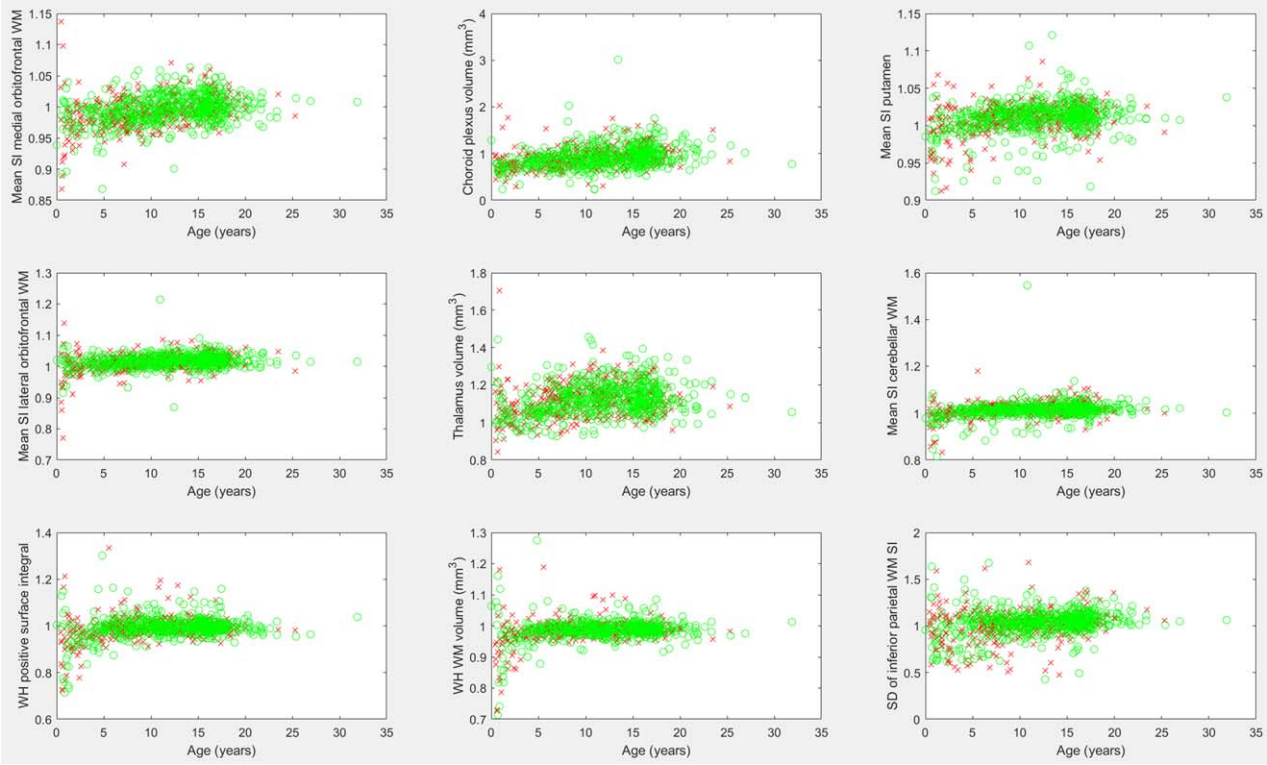

Figure 6.

The leading asymmetry measurements (left hemisphere measurement divided by right hemisphere measurement) based on correlation with participant age. Males are represented with a red "x," females with a green "o." SD, standard deviation; WH, whole hemisphere; SI, signal intensity; WM, white matter. [Color figure can be viewed at wileyonlinelibrary.com]

being referred to imaging. The approach taken in this article is uncommon as most imaging centers employ a standard clinical MRI examination which is liable to be of limited research benefit compared with the sets of MR protocols included in a typical prospective study. However, at *BCH*, advanced MR protocols that might be considered research sequences at other centers are included in many routine clinical imaging examinations such as high-resolution volumetric T1 structural examinations compatible with FreeSurfer technology and diffusion tensor imaging and resting-state functional MRI. As such, the baseline of healthy participants included in this analysis will form the foundation for future research into the neurodevelopment of a wide variety of medical pathologies imaged extensively at *BCH* with many types of MRI including autism, epilepsy, cerebral palsy, tuberous sclerosis complex, neurofibromatosis, migraine with aura, and much more.

The main strength of this research is that it is a large study that provides baseline data for what to expect from a group of healthy subjects that received clinical imaging. Large sample sizes help to overcome limitations caused by natural variability in the biomarker measurements being evaluated. MRI data acquisition involves a certain amount of measurement noise. Additional noise is introduced to our biomarker measurements evaluated in this analysis based on the computational technologies employed by FreeSurfer. Furthermore, there is a natural amount of variability in the healthy population investigated. These factors result in a substantive amount of measurement variability when employing MRI and FreeSurfer to assess healthy subjects. With few samples available, differences between pathological and healthy groups may appear present when in fact the observed effect could be little more than a by-product of high levels of measurement variability. Furthermore, measurement variability can obscure real effects as nonstatistically significant when sample sizes are very low. Including such a large dataset facilitates us breaking down our analysis based on age ranges. This allows us to evaluate possible changes in the brain at various developmental stages. By including all the samples available, we not only help overcome sample size limitations but also provide a complete collection of imaging from clinical scanners, which is ideal for the assessment and development of diagnostic tests which ultimately would be applied to this very same type of population.

An additional strength of this work was obtained by incorporating all measurements produced by FreeSurfer, making this study considerably more thorough than many in the literature which often focus on a single type of FreeSurfer measurement such as volumetric measurements or cortical thickness measurements. Furthermore, this analysis has been extended to include asymmetry biomarkers for every FreeSurfer derived measurement, for which both a left and right hemisphere pair are acquired.

A limitation of this study is that FreeSurfer is not optimized for the youngest subjects in our analysis. As such, its failure rate increases substantially for subjects aged 0 to 8 months and the reliability of the results successfully produced by FreeSurfer on subjects from this age range is uncertain. FreeSurfer's reliability was assessed as reasonable for subjects aged 8 months and later (given that the technology was not validated on this age range) at which point myelination patterns have inverted, so as to match the general pattern exhibited through the rest of life. We are aware of current research aimed at overcoming the problem of FreeSurfer's applicability and reliability in very young populations [de Macedo Rodrigues et al., 2015] and any developments in this venue will be incorporated into future work, which will also involve the extension of this analysis to tractography and functional MRI.

That this study was performed on a single retrospective dataset is another limitation as the study does not evaluate differences between imaging technologies and imaging centers. Additionally, there is some variability in imaging parameters (spatial resolution, signal-to-noise ratio, etc.) because of variations in the pulse sequences employed; however, imaging was performed with a consistent set of 3 T Siemens MRI scanners all installed at *Boston Children's Hospital* in 2007, just prior to the review period included as part of this study. Ideally, this study would be performed on only a single MR imaging protocol; however, doing so would greatly reduce the number of samples available for inclusion in this analysis. Large sample sizes help to overcome potential bias associated with measurements that exhibit considerable variability. While limiting the analysis to a single imaging protocol would reduce potential bias caused by scan parameter variability, it would increase bias caused by sample size effects. Furthermore, work focused on autism has demonstrated that sample size effects are likely to introduce substantial bias when the number of examinations is small [Levman et al., 2017] and so we expect that the bias introduced by variable scan parameters to be a lesser problem than that introduced by greatly reducing our sample size. Many FreeSurfer-based biomarker measurements from our dataset demonstrate that the discriminating power (between autistic and healthy subjects) of volumetric measurements (in mm^3) is approximately identical to the discriminating power of the voxel counts in those same regions, including ventricular volumes/voxel counts and corpus callosum volumes/voxel counts [Levman et al., 2017]. As voxel counts vary greatly based on spatial resolution variations, we believe the effect of varying spatial resolutions in our dataset to be minimal.

An additional limitation of this study is that the age distributions of available male and female subjects vary considerably (Fig. 1) because of the availability of appropriate subjects that met our inclusion criteria from a large clinical population, resulting in imbalanced pools of subjects for further analysis. Our experiment did not involve participant matching between our female and male groups. This would typically be performed by pairing subjects based on age and gender. Instead, we have opted to perform our statistical analyses in a groupwise manner varying the age range under consideration and to plot our main findings on an age-dependent basis while differentiating between male and female subjects in our scatter plots. This methodology was selected to avoid the reduced sample size that would arise from only including those subjects who have a counterpart with identical age in the opposite gender group. Additionally, this methodology was selected to avoid having our analysis be influenced by the extent of difference between matched pairs of individuals for which a variety of factors beyond age and gender might influence how appropriate it was for the subjects to have been paired (brain volume, substructure volume, co-morbidities, etc.).

Our experiment did not analyze participants based on handedness, as this information was unavailable in electronic medical records. However, no effect of handedness on brain morphology was found in a population of 465 adult subjects [Good et al., 2001]; however, it should be noted that handedness has been linked with hemispheric asymmetries [Amunts et al., 1996] and so future work will investigate the relationship between MRI-assessed biomarkers and participant handedness.

Future work will be devoted to overcoming the limitations associated with including 0–8-month-old subjects by combining this study's approach with ongoing neonatal statistical atlas FreeSurfer research [de Macedo Rodriguez et al., 2015]. Future work will also involve extending this study to the analysis of tractography and functional (fMRI) data which is available clinically on many of the examinations included in this analysis. Future work will entail comparing FreeSurfer measurements with potentially correlated clinical variables (such as subject IQ, etc.). Future work will also investigate the potential use of multivariate analyses (including machine learning) toward the identification of sets of biomarker measurements that best characterize neurological maturation. We will investigate the possibility of assessing the maturation levels of each of the brain's subregions segmented by FreeSurfer individually by combining several biomarker measurements with multivariate techniques in the hope of creating technology that can identify regions of the brain that are over- or underdeveloped. This can help both in healthy populations and in populations with neurodevelopmental

disorders, which may be detected, assessed, and characterized based on deviations from expected growth trajectories in various subregions of the brain. Healthy individuals with moderately underdeveloped regions of the brain will benefit from additional knowledge of their physiological status. Parents of such children may elect to have their offspring train in tasks dependent on underdeveloped regions of the brain to assist in increasing the rate at which that region of the brain develops.

The scientific literature has seen enormous growth in developmental imaging of preadult populations. However, this work has been largely focused in pediatric imaging with considerably less focus on neonatal and fetal imaging. Neonatal imaging is more challenging as brain size is considerably smaller than in pediatrics and it can be difficult to get a neonate to remain still over the course of their imaging examination. Furthermore, FreeSurfer is substantially less reliable in a neonate population. Participant movement induces multiple types of imaging artefacts, which make studies on neonate populations challenging. Fetal imaging is the greatest challenge of the three as the brain sizes are the smallest, movement remains a substantial issue and FreeSurfer is not compatible with fetal imaging which often does not include volumetric exams but only a set of image slices. Furthermore, MRI technology is reliant on the spatial proximity of a radiofrequency (RF) coil (antenna) to the organ being imaged. Normal brain imaging benefits from a specialized head RF coil that is mounted immediately adjacent to the subject's cranium; however, in fetal imaging, this is not possible and so coils located outside the mother's abdomen are used, inherently reducing image quality. Additional challenges exist in fetal imaging because of variations in tissue contrast observed *in utero* in fetal subjects relative to later developmental ages. Regardless of the many challenges inherent in the imaging of fetal, neonatal, and pediatric populations, there is considerable potential for ongoing research toward better understanding of healthy brain development. Healthy baselines for biomarker trajectories derived from clinical MRI scanners may also help toward the creation of the next generation of diagnostic tests responsible for identifying a variety of pathological conditions as departures from expected growth trajectories established among healthy participants.

CONCLUSION

We have performed a large-scale analysis of T1 MRI examinations that included hemispheric asymmetry measures and demonstrated their potential to help elucidate structural neurological differences between healthy male and female participants. We performed a correlation analysis between biomarkers and participant age and also extended it to include asymmetry biomarkers. These approaches can play a useful role in helping to answer questions about organization and development in the brain. Automatic segmentation techniques such as FreeSurfer have considerable potential toward the creation of the next generation of clinical diagnostic tests informed by the large amounts of information acquired with clinical MRI (using structural MRI and combining it with alternative modalities such as MR tractography). Results from this study reveal sets of biomarkers correlated with age, sets exhibiting considerable asymmetry, and gender comparisons. Future work will look at improving our ability to characterize healthy brain development with the help of additional MR imaging modalities and multivariate analysis techniques. It is hoped that these research avenues will assist toward better characterizing healthy brain development and toward diagnosing and detecting a variety of pathological conditions as deviations from expected growth trajectories as defined by healthy populations imaged in a clinical context.

ACKNOWLEDGMENT

The authors report no conflicts of interest.

REFERENCES

Amunts K, Schlaug G, Schleicher A, Steinmetz H, Dabringhaus A, Roland PE, Zilles K (1996): Asymmetry in the human motor cortex and handedness. NeuroImage 4:216–222.

Amunts K, Schleicher A, Bürgel U, Mohlberg H, Uylings HB, Zilles K (1999): Broca's region revisited: Cytoarchitecture and intersubject variability. J Comp Neurol 412:319–341.

Amunts K, Armstrong E, Malikovic A, Homke L, Mohlberg H, Schleicher A, Zilles K (2007): Gender-specific left-right asymmetries in human visual cortex. J Neurosci 27:1356–1364.

Broca P (1861): Remarques sur le siège de la faculté de langage articulé, suivies d'une observation d'aphémie. Bull Soc Anat 2: 330–357.

Bunge SA, Dudukovic NM, Thomason ME, Vaidya CJ, Gabrieli JD (2002): Immature frontal lobe contributions to cognitive control in children: Evidence from fMRI. Neuron 33:301–311.

Burkhalter A, Felleman DJ, Newsome WT, Van Essen DC (1986): Anatomical and physiological asymmetries related to visual areas V3 and VP in macaque extrastriate cortex. Vis Res 26: 63–80.

Cantalupo C, Hopkins WD (2001): Asymmetric Broca's area in great apes. Nature 414:505.

Casey BJ, Trainor RJ, Orendi JL, Schubert AB, Nystrom LE, Giedd JN, Castellanos FX, Haxby JV, Noll DC, Cohen JD, Forman SD, Dahl RE, Rapoport JL (1997): A developmental functional MRI study of prefrontal activation during performance of a Go-No-Go task. J Cogn Neurosci 9:835–847.

de Macedo Rodrigues K, Ben-Avi E, Sliva DD, Choe M, Drottar M, Wang R, Fischl B, Grant PE, Zollei L (2015): A FreeSurfer-compliant consistent manual segmentation of infant brains spanning the 0–2 year age range. Front Hum Neurosci 9:21.

Destrieux C, Fischl B, Dale A, Halgren E (2010): Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. NeuroImage 53:1–15.

Erus G, Battapady H, Satterthwaite TD, Hakonarson H, Gur RE, Davatzikos C, Gur RC (2015): Imaging patterns of brain

development and their relationship to cognition. Cereb Cortex 25: 1676–1684.

Fair DA, Cohen AL, Power JD, Dosenbach NU, Church JA, Miezin FM, Schlaggar BL, Petersen SE (2009): Functional brain networks develop from a "local to distributed" organization. PLoS Comput Biol 5:e1000381.

Fan Y, Shi F, Smith JK, Lin W, Gilmore JH, Shen D (2011): Brain anatomical networks in early human brain development. NeuroImage 54:1862–1871.

Fischl B (2012): FreeSurfer. NeuroImage 62:774–781.

Foundas AL, Eure KF, Luevano LF, Weinberger DR (1998): MRI asymmetries of Broca's area: The pars triangularis and pars opercularis. Brain Lang 64:282–296.

Fox MD, Snyder AZ, Vincent JL, Corbetta M, Van Essen DC, Raichle ME (2005): The human brain is intrinsically organized into dynamic, anticorrelated functional networks. Proc Natl Acad Sci 102:9673–9678.

Franke K, Luders E, May A, Wilke M, Gaser C (2012): Brain maturation: Predicting individual *BrainAGE* in children and adolescents using structural MRI. NeuroImage 63:1305–1312.

Galaburda, AM (1984): Cerebral Dominance: The Biological Foundations. In: Geschwind N, Galaburda AM, editors. Cambridge, Massachusetts: Harvard Univ. Press. pp 11–25.

Giedd JN, Raznahan A, Alexander-Bloch A, Schmitt E, Gogtay N, Rapoport JL (2015): Child psychiatry branch of the National Institute of Mental Health longitudinal structural magnetic resonance imaging study of human brain development. Neuropsychopharmacol Rev 40:43–49.

Goddings AL, Mills KL, Clasen LS, Giedd JN, Viner RM, Blakemore SJ (2014): The influence of puberty on subcortical brain development. NeuroImage 88:242–251.

Gogtay N, Giedd JN, Lusk L, Hayashi KM, Greenstein D, Vaituzis AC, Nugent TF 3rd, Herman DH, Clasen LS, Toga AW, Rapoport JL, Thompson PM (2004): Dynamic mapping of human cortical development during childhood through early adulthood. Proc Natl Acad Sci 101:8174–8179.

Good CD, Johnsrude I, Ashburner J, Henson RNA, Friston KJ, Frackowiak RSJ (2001): Cerebral asymmetry and the effects of sex and handedness on brain structure: A voxel-based morphometric analysis of 465 normal adult human brains. NeuroImage 14:685–700.

Jutila L, Ylinen A, Partanen K, Alafuzoff I, Mervaala E, Partanen J, Vapalahti M, Vainio P, Pitkanen A (2001): MR volumetry of the entorhinal, perirhinal, and temporopolar cortices in drug-refractory temporal lobe epilepsy. Brain 22:1490–1501.

Khundrakpam BS, Tohka J, Evans AC, Brain Development Cooperative Group (2015): Prediction of brain maturity based on cortical thickness at different spatial resolutions. NeuroImage 111:350–359.

Koolschijn PC, Crone EA (2013): Sex differences and structural brain maturation from childhood to early adulthood. Dev Cogn Neurosci 5:106–118.

Levman J, MacDonald P, Stewart N, Lim A, Galaburda A, Takahashi E (2017): Pediatric autism: Assessing clinical MRI's potential with a large-scale retrospective analysis. Proceedings of the Organization for Human Brain Mapping Conference, Vancouver, Canada.

Mesulam MM (1981): A cortical network for directed attention and unilateral neglect. Ann Neurol 10:309–325.

McIntosh AR, Bookstein FL, Haxby JV, Grady CL (1996): Spatial pattern analysis of functional brain images using partial least squares. NeuroImage 3:143–157.

O'Donnell S, Noseworthy MD, Levine B, Dennis M (2005): Cortical thickness of the frontopolar area in typically developing children and adolescents. NeuroImage 24:948–954.

Ostby Y, Tamnes CK, Fjell AM, Westlye LT, Due-Tonnessen P, Walhovd KB (2009): Heterogeneity in subcortical brain development: A structural magnetic resonance imaging study of brain maturation from 8 to 30 years. J Neurosci 29: 11772–11782.

Reiss AL, Abrams MT, Singer HS, Ross JL, Denckla MB (1996): Brain development, gender and IQ in children. A volumetric imaging study. Brain 119:1763–1774.

Student S (1908): The probable error of a mean. Biometrika 6:1–25.

Supekar K, Musen M, Menon V (2009): Development of large-scale functional brain networks in children. PLoS Biol 7: e1000157.

Tamnes CK, Ostby Y, Fjell AM, Westlye LT, Due-Tonnessen P, Walhovd KB (2010): Brain maturation in adolescence and young adulthood: Regional age-related changes in cortical thickness and white matter volume and microstructure. Cereb Cortex 20:534–548.

Tamnes CK, Walhovd KB, Dale AM, Ostby Y, Grydeland H, Richardson G (2013): Brain development and aging: Overlapping and unique patterns of change. NeuroImage 68:63–74.

Thomas KM, Drevets WC, Whalen PJ, Eccard CH, Dahl RE, Ryan ND, Casey BJ (2001): Amygdala response to facial expressions in children and adults. Biol Psychiatry 49:309–316.

Vaadia E, Haalman I, Abeles M, Bergman H, Prut Y, Slovin H, Aertsen A (1995): Dynamics of neuronal interactions in monkey cortex in relation to behavioural events. Nature 373: 515–518.

Vijayakumar N, Allen NB, Youssef G, Dennison M, Yucel M, Simmons JG, Whittle S (2016): Brain development during adolescence: A mixed-longitudinal investigation of cortical thickness, surface area, and volume. Hum Brain Mapp 37: 2027–2038.

Watkins KE, Paus T, Lerch JP, Zijdenbos A, Collins DL, Neelin P, Taylor J, Worsley KJ, Evans AC (2001): Structural asymmetries in the human brain: A voxel-based statistical analysis of 142 MRI scans. Cereb Cortex 11:868–877.

Walhovd KB, Fjell AM, Brown TT, Kuperman JM, Chung Y, Hagler DJ, Jr, Roddey JC, Erhart M, McCabe C, Akshoomoff N, Amaral DG, Bloss CS, Libiger O, Schork NJ, Darst BF, Chang L, Ernst TM, Frazier J, Gruen JR, Kaufmann WE, Murray SS, van Zijl P, Mostofsky S, Dale AM, Pediatric Imaging, Neurocognition, and Genetics Study (2012): Long-term influence of normal variation in neonatal characteristics on human brain development. Proc Natl Acad Sci 109: 20089–20094.

White T, Su S, Schmidt M, Kao CY, Sapiro G (2010): The development of gyrification in childhood and adolescence. Brain Cogn 72:36–45.