Supplemental Information

Table of Contents

Abbreviations: Total Cerebral Measures (TCMs): Total Brain Volume (TBV), Total Mean Cortical Thickness (MCT), or Total Surface Area (TSA). Grey Matter Volume (GMV). White Matter Volume (WMV).

Abbreviations: Total Cerebral Measures (TCMs): Total Brain Volume (TBV), Total Mean   2
Cortical Thickness (MCT), or Total Surface Area (TSA). Grey Matter Volume (GMV).
White Matter Volume (WMV).

1 – Total Brain Volume (TBV) Choice

60   Both TBV and Total Intracranial Volume (TIV) have been used to adjust for brain allometry

61   to avoid a bias toward isometry when estimating allometric scaling coefficients of large regions

62   (e.g., Jong et al., 2017; Peyre et al., 2020; Reardon et al., 2016; Williams et al., 2020). However,

63   since segmentations yielding the highest number of regions were favored and because we are

64   interested in examining how much a particular volume contributes to the total brain volume at the

65   time of the study (and not at maximal lifetime total brain volume; Herbert et al., 2003; O'Brien et

66   al., 2006), TBV was chosen over TIV. The correlation between TBV and TIV was 88.93%.

67   We did not use the TBV provided by ASEG (26515) due to a mismatch for some participants

68   between the provided TBV measure (26515) and the one calculated by summing the global

69   measures provided by ASEG: Total GMV Volume (26518) + Cerebellum WMV (Left: 26556,

70   Right: 26587) + Cerebral WMV (Left: 26553, Right: 26584). This mismatch yielded a correlation

71   of 99.73% between TBV measures.

72   Instead, we used the TBV calculated by summing the global measures provided by ASEG as

73   our measure of TBV since it correlated at 99.82% with the TBV we calculated from summing the

74   regions of interest in the present study.

75   The ASEG global sum measure was favored over the sum of regions we analyzed in our study

76   to include a maximum number of participants. (We would otherwise, for instance, have to exclude

77   all individuals that do not have the FAST cerebellum segmentations even though they have all the

78   other segmentations).

79

Abbreviations: Total Cerebral Measures (TCMs): Total Brain Volume (TBV), Total Mean   3
Cortical Thickness (MCT), or Total Surface Area (TSA). Grey Matter Volume (GMV).
White Matter Volume (WMV).

80

81
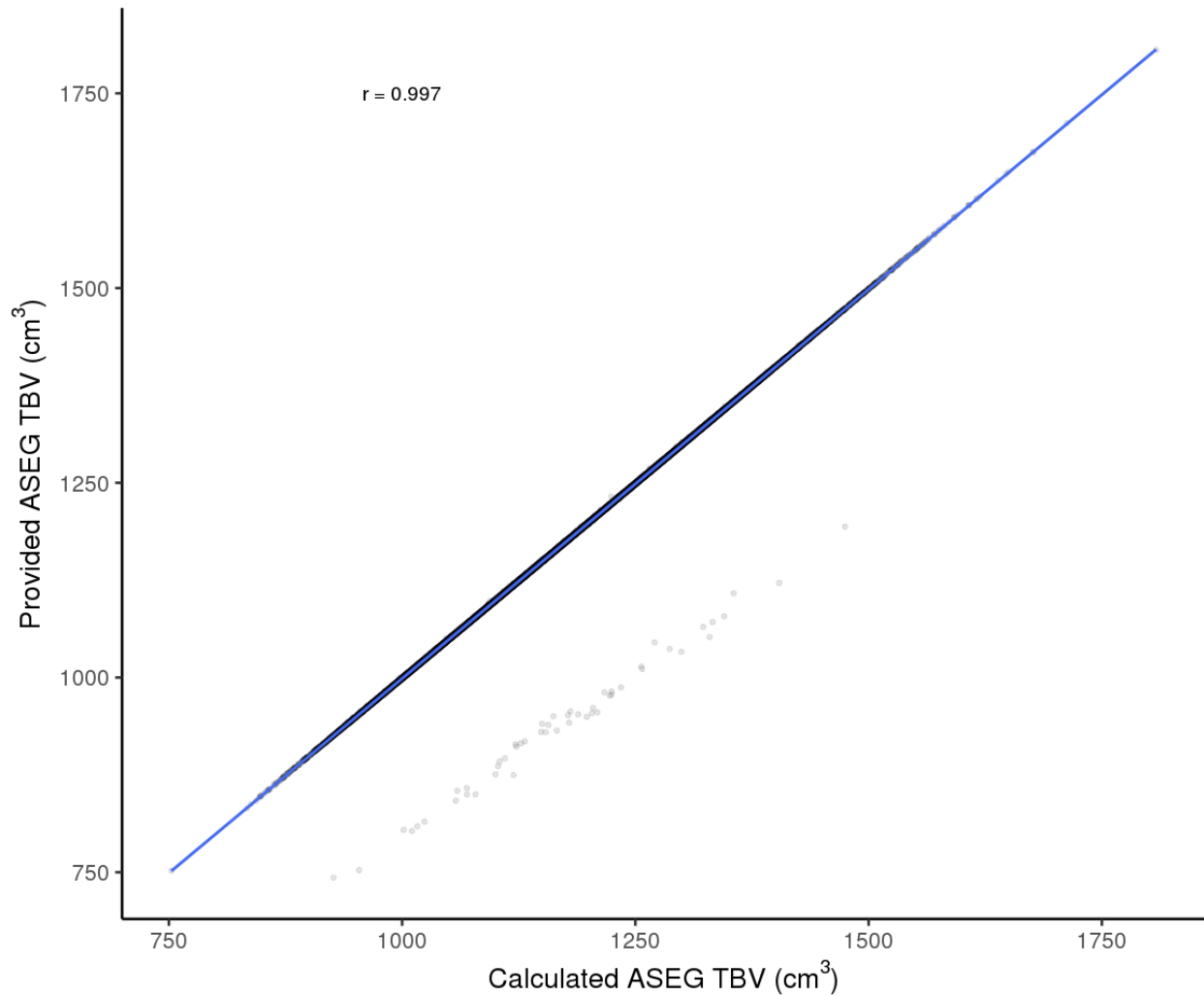
Figure S1. **Correlation between the Provided ASEG Total Brain Volume (TBV) Measure and the Calculated Sum of Global Measures from the Freesurfer ASEG Segmentations.**
The provided TBV corresponds to data-field 26515 and the Calculated ASED TBV to the sum of Total Grey Matter Volume (data-field 26518), Cerebellum White Matter Volume (Left: data-field 26556, Right: data-field 26587), Cerebral White Matter Volume (Left: data-field 26553, Right: data-field 26584). r, rho = 99.73%.

88

r = 0.998

(Y-axis: Calculated TBV (cm^3); X-axis: Calculated ASEG TBV (cm^3))

89
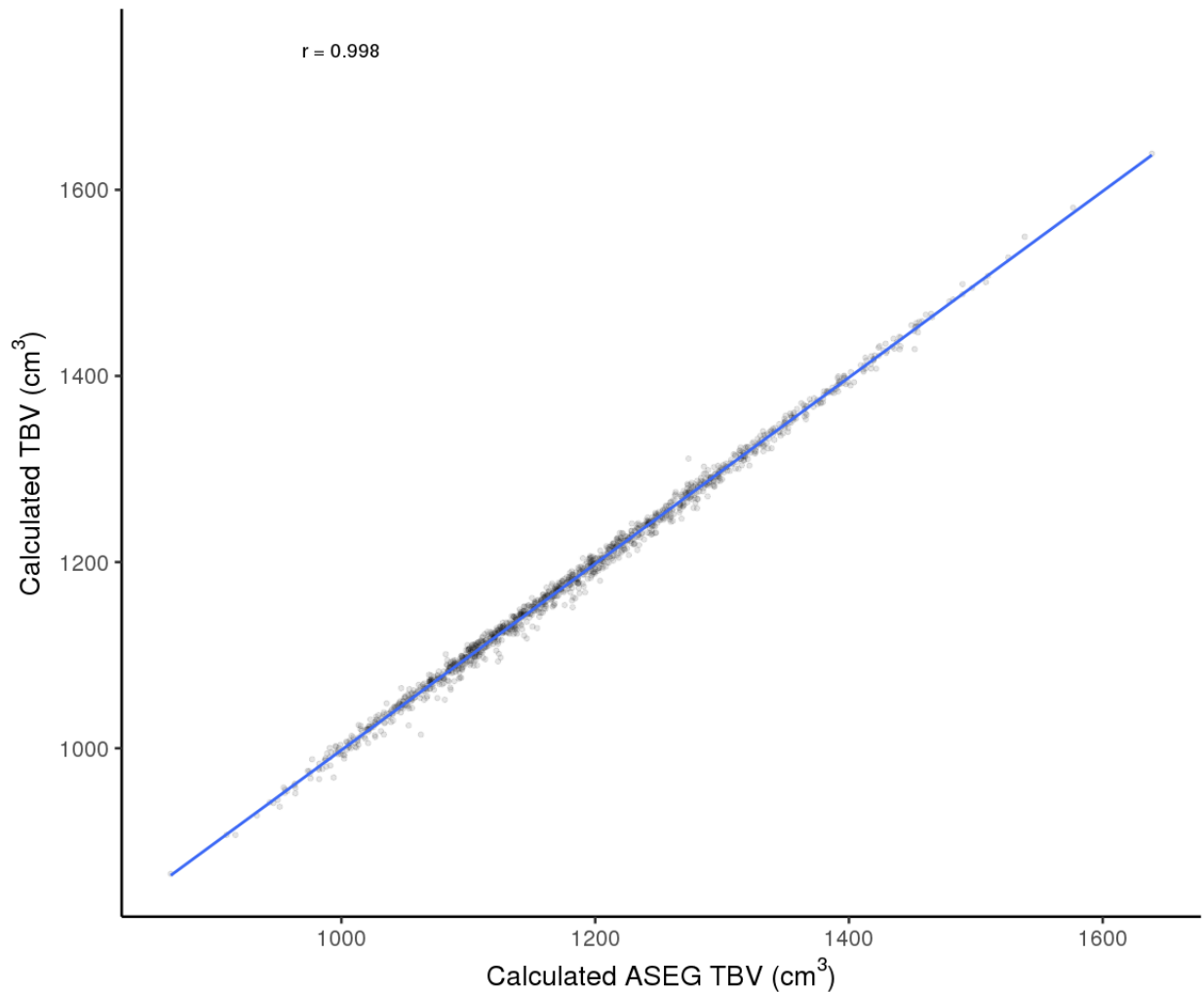
Figure S2. **Correlation between the Calculated ASEG Total Brain Volume (TBV) Measure**

and the Calculated TBV from the Regions examined in this study.

The Calculated ASEG TBV corresponds to the sum of Total Grey Matter Volume (data-field

26518), Cerebellum White Matter Volume (Left: data-field 26556, Right: data-field 26587),

Cerebral White Matter Volume (Left: data-field 26553, Right: data-field 26584). r, Pearson r.

95

96

Abbreviations: Total Cerebral Measures (TCMs): Total Brain Volume (TBV), Total Mean   5
Cortical Thickness (MCT), or Total Surface Area (TSA). Grey Matter Volume (GMV).
White Matter Volume (WMV).

97 *2 - Scanner Site*

98 The number of participants differed across sites ($\chi2$ (2, N = 40 028) = 16 107, p < 2.2e-16) and

99 between sexes ($\chi2$ (2, N = 40 028) = 27.1, p =1.314e-06), with more female than male participants

100 across sites (Site 11025: Male N = 12033 and Female N = 12938, Site 11026: Male N = 2318 and

101 Female N = 2752, and Site 11027: Male N = 4535 and Female N = 5452). Age differed between

102 Site 11025 and Site 11026 (b = 2.27 years, t (40025) = 19.63, p < 2.2e-16), between Site 11025

103 and Site 11027 (b = 1.40 years, t (40025) = 15.78, p < 2.2e-16), and between Site 11026 and Site

104 11027 (b = -0.87 years, t (40025) = -6.70, p = 2.1e-11).

105

106 *3 - Correlation between the calculated global measures and those provided by the Freesurfer*

107 *ASEG Segmentation.*

108 Global measures were calculated as the sum of the regions examined in the present study

109 instead of using predefined measures when available. Since we used regions from different

110 segmentations to perform the most fine-grain analysis possible, we examined the correlation

111 between the summed global measures and the global measures provided by ASEG.

112 The Total Subcortical (Grey Matter, GM) Volumes measure was calculated as the sum of the

113 amygdala, hippocampus, and thalamus volumes from the Freesurfer subsegmentations ([191]) and

114 the caudate, accumbens, pallidum, and putamen of the Freesurfer ASEG segmentations ([190]). This

115 calculated measure correlated at 99.25% with the Total Subcortical Volume measure from the

116 ASEG segmentation ([26517]).

117 The Total Cortical GM, which corresponds to the sum of the a2009s Destrieux volumes

118 correlated at 99.97% with the Total Cortical GM provided by the ASEG segmentation.

119 The calculated total Cerebellum GM, which was obtained by summing the grey matter volumes

120 from the FAST segmentations, correlated with the sum of the left and right cerebellum GM from

121 the ASEG Freesurfer segmentations at 75.83%.

122

123

Abbreviations: Total Cerebral Measures (TCMs): Total Brain Volume (TBV), Total Mean   6
Cortical Thickness (MCT), or Total Surface Area (TSA). Grey Matter Volume (GMV).
White Matter Volume (WMV).

124 *4 - Exploratory Analyses: Methods.*

125 These analyses were not preregistered. Asymmetries in scaling with global measures, and sex and

126 age effects were examined by computing the correlation of the scaling coefficient or the

127 standardized betas of the left and right regions.

128 4.1. Sex Differences in Variance.

129 Differences in the variance of cerebral measures between sexes were assessed with a

130 Levene's test (F-test) performed on the person-level neuroanatomical markers. Variance ratios

131 were calculated as the Female SD / Male SD. We report the correlation between the variance ratio

132 and standardized beta of each region to test if the regions showing the largest sex differences also

133 showed the greatest sex differences in variance. We additionally examined differences in the

134 variance ratio of the global allometric marker between sexes.

135 4.2. Comparing Results of Different TBV Adjustment Techniques.

136 We investigated whether the significance of the TBV, sex, and age (linear and quadratic) effects

137 and interactions varied with the type of adjustment for TBV. We compared results from the

138 allometric approach to results from the proportion and the covariate approach, separately.

139 For the proportion method, we obtained an adjusted region measure by dividing a region

140 by its global measure and then applying equation 4 without the main effect of TBV. For the

141 covariate approach, we applied equation 4 to the raw cerebral measures. The main analyses were

142 initially run without controlling for Scanner Site and since the influence of scanner site was limited

143 to a change of 0.01 in beta coefficients of a few regions in the main analyses, Scanner Site was not

144 included in the analyses performed to compare result consistency across three TBV adjustment

145 analyses.

146 We examined whether the same-sex differences in variance and mean effects and

147 interactions were significant when considering and omitting brain allometry.

148 To identify if results from allometric models diverged more from those obtained with linear

149 models in regions that were more allometric, we examined the correlation between a region's

150 deviance from isometry (| 1 – scaling coefficient|) and the difference in the standardized beta of an

151 effect and interaction across models with differing TCM adjustments (e.g., |Covariate Sex Std Beta

152 – Allometric Sex Std Beta|). We chose to compare the betas from the effects and interactions that

153 were significant in at least one region in our primary analyses.

154

155 ## 4.3. Replication of Previous Studies Examining Brain Allometry.

156 We attempted to replicate age and sex cerebral differences reported by studies that considered

157 brain allometry. When available we used the same segmentations as these studies for a more

158 accurate replication, although our sample was older and much larger. Specifically, we tried to

159 replicate the following findings:

160

161 1. Right and left lobar (frontal, occipital, limbic, parietal, temporal, and cerebellum) sex

162 differences reported by Peyre and colleagues (2020; Figure 1. Right and left cerebellum grey

163 matter: boys greater than girls; right (d = -0.2) and left frontal grey matter: girls greater than boys

164 (d = 0.08-0.1)).

165

166 2. A greater pallidal volume in males and relative caudate head expansion and ventral striatum

167 contraction in females were reported by Reardon and colleagues (2016).

168

169 3. Relative expansion of total cerebellum, flocculus, and Crus II-lobule VIIIB volumes in males

170 and X-specific contraction of Crus II-lobule VIIIB by Mankiw and colleagues, (2017; Figure 6).

171

172 4. Sex differences reported by Sanchis-Segura and colleagues (2019) in Figure 2 when adjusting

173 for allometry with the power-corrected proportion method proposed by Liu and colleagues (2014)

174 were the following:

175   - Left, female greater than male: Frontal superior, superior orbital, mid, med orbital,

176   occipital med, postcentral, parietal inf, cerebellum X; male greater than female cerebellum

177   crus I, pallidum, putamen, amygdala.

178   - Right, female greater than male: Frontal superior orbital, frontal med orbital occipital

179   superior, mid, and parietal inferior; male greater than female: temporal pole mid, pallidum,

180   putamen, amygdala, and lingual.

181

182 ## 4.4. Replication of Ritchie and colleagues (2018) Study and Comparing Results with the Covariate

183 ## and with the Allometric TBV Adjustment.

184   Finally, we attempted to replicate the mean and variance ratio (VR) sex differences

185 reported by Ritchie and colleagues' (2018) while adjusting for age, ethnicity, and linear TBV with

Abbreviations: Total Cerebral Measures (TCMs): Total Brain Volume (TBV), Total Mean   8
Cortical Thickness (MCT), or Total Surface Area (TSA). Grey Matter Volume (GMV).
White Matter Volume (WMV).

186   the Desikan-Killiany atlas segmentations. We additionally controlled for scanner site and used the

187   the Cheadle Site - the site from which Ritchie and colleagues (2018) obtained their data - as the

188   reference. We used the same FIRST subcortical volumes segmentations available in the UK

189   Biobank as Ritchie and colleagues (2018). However, we used the Desikan-Killiany Cortical Atlas

190   segmentations from the UK Biobank while Ritchie and colleagues (2018) performed their own

191   Desikan-Killiany Cortical segmentations.   To examine if brain allometry influenced reported

192   results with Ritchie and colleagues (2018) model and choice of segmentation, we additionally ran

193   Ritchie and colleagues' (2018) model while adjusting for brain allometry (i.e., log transforming

194   cerebral measures) and compared our results from our replication analyses to those obtained when

195   considering brain allometry. As done by Ritchie and colleagues (2018), we report the correlation

196   between the VR and Cohen's D of each region to examine if mean sex differences show sex

197   differences in variance.

198

199   ## 4.5. Height

200   To assess if global differences were due to differences in height, we also ran equation 2

201   with log transformed height instead of the log transformed TBV and extended equation 2 by adding

202   height and TBV effects and interactions with other predictors, as shown in equation 5:

203   Equation 5: $\text{Log10}(\textit{Global Measure}) = \text{Intercept} + \beta1*\text{Log10}(\textit{TBV}) + \beta2*\textit{Age} + \beta3*\textit{Sex} +$

204   $\beta4*\textit{Age}^2 +$

205   $\beta5*\text{Log10}(\textit{Height}) +$

206   $\beta6*\text{Log10}(\textit{TBV}) \times \textit{Age} +$

207   $\beta7 *\text{Log10}(\textit{TBV}) \times \textit{Age}^2 +$

208   $\beta8*\textit{Age} \times \textit{Sex} +$

209   $\beta9*\textit{Age}^2 \times \textit{Sex} +$

210   $\beta10*\text{Log10}(\textit{TBV}) \times \textit{Sex} +$

211   $\beta6*\text{Log10}(\textit{Height}) \times \textit{Age} +$

212   $\beta11 *\text{Log10}(\textit{Height}) \times \textit{Age}^2 +$

213   $\beta12*\text{Log10}(\textit{Height}) \times \textit{Sex} +$

214   $\beta13*\text{Log10}(\textit{Height}) \times \text{Log10}(\textit{TBV}) +$

215   $\beta14*\text{Log10}(\textit{Height}) \times \text{Log10}(\textit{TBV}) \times \textit{Sex} +$

216   $\beta15*\text{Log10}(\textit{Height}) \times \text{Log10}(\textit{TBV}) \times \textit{Age} +$

Abbreviations: Total Cerebral Measures (TCMs): Total Brain Volume (TBV), Total Mean   9
Cortical Thickness (MCT), or Total Surface Area (TSA). Grey Matter Volume (GMV).
White Matter Volume (WMV).

217 β16*Log10(*Height*) *x* Log10(*TBV*) *x Age^2* +

218 β17*Log10(*TBV*) *x Age x Sex* +

219 β18*Log10(*TBV*) *x Age^{2}x Sex* +

220 β19*Log10(*Height*) *x Age x Sex* +

221 β20*Log10(*Height*) *x Age^{2}x Sex* +

222 β21*Log10(*Height*) *x* Log10(*TBV*) *x Age x Sex* +

223 β22*Log10(*Height*) *x* Log10(*TBV*) *x Age^{2}x Sex* +

224 β23**Scanner Site* + Error

225

Abbreviations: Total Cerebral Measures (TCMs): Total Brain Volume (TBV), Total Mean 10
Cortical Thickness (MCT), or Total Surface Area (TSA). Grey Matter Volume (GMV).
White Matter Volume (WMV).

226 ***5 - Supplementary Results and Discussion***

227 Since data for the 5th ventricle was only available for 2% of UK Biobank subjects (830/40 028),

228 all 5th ventricle results may not be reliable.

229

230 5.1. Age by Total Cerebral Measure (TCM) and Sex by TCM Interactions

231 **5.1.1. TCM by Sex.**

232 *Global Measures.* The cerebellum GMV increased more with TBV (i.e., had a larger allometry

233 coefficient α) in males (α = 0.42) compared to females (α = 0.39, β = -0.06).

234 *Regional Measures.* Only 11 regions out of 620 (0.2%) had significant interactions between total

235 cerebral measure and sex and the interactions were the largest for the following regions. As

236 indicated in the main text, the right fimbria volume increases less with TBV in females (α = 0.29)

237 compared to males (α = 0.50, β = -0.12, SE = 0.02), while the left surface area of the pericallosal

238 sulcus increases more with TSA in females (α = 0.71) compared to males (α = 0.59, β = 0.09, SE

239 = 0.01; Figure S3).

240

Abbreviations: Total Cerebral Measures (TCMs): Total Brain Volume (TBV), Total Mean   11
Cortical Thickness (MCT), or Total Surface Area (TSA). Grey Matter Volume (GMV).
White Matter Volume (WMV).

241

242   A)
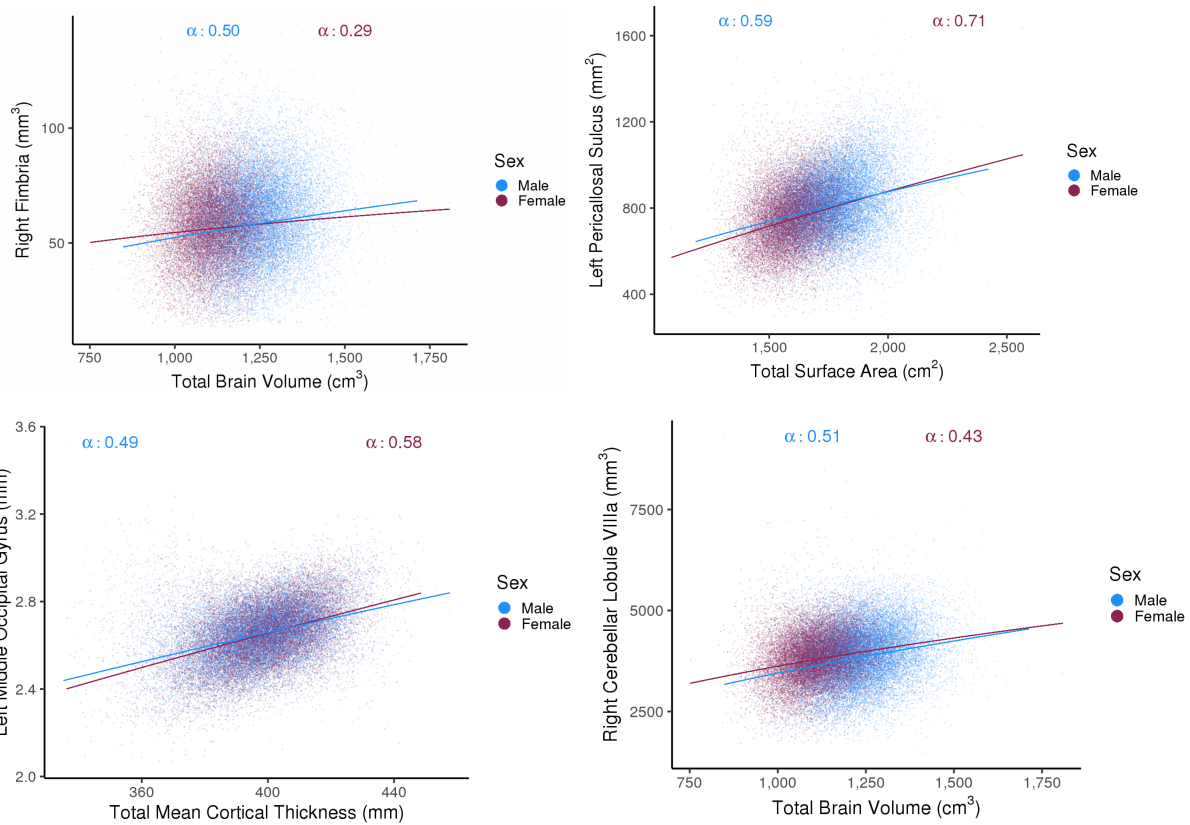
243

244

245

246   *Figure S3.* **Largest Total Cerebral Measure by Sex Interaction**. α is the scaling coefficient of

247   the region with its total cerebral measure. Blue and purple lines correspond to the predicted values

248   of y from the model with TBV, sex, age and TBV, sex, age^2 interactions.

249 ***5.1.2. TCM by Age.***

250 *Global Measures*. The total subcortical volumes and TSA increased more with TBV in individuals

251 64 years old (median) and under compared to individuals more than 64 years old (β = -0.01 for

252 both). However, Cerebellum GMV and Total MCT increased more with TBV in older adults

253 compared to younger adults (β = 0.03 and β = 0.04, respectively). Total MCT and cerebellar GMV

254 additionally decreased more with age in individuals with a smaller TBV (β = -0.37, β = -0.31,

255 respectively) compared to individuals with a larger TBV.

256 *Regional Measures*. Only 67 regions out of 620 (11%) had significant interactions between the

257 total cerebral measure and age and were the largest for the following regions. As indicated in the

258 main text, the right fimbria volume increases less with TBV in younger (α = 0.74) compared to

259 older individuals (α = 1.13, β = 0.03, SE = 0.01), while the mean thickness of the left transverse

260 temporal sulcus increases more with the TBV in younger (α = 1.49) compared to older individuals

261 (α = 1.22, β = -0.04, SE = 0.00; Figure S4).

262
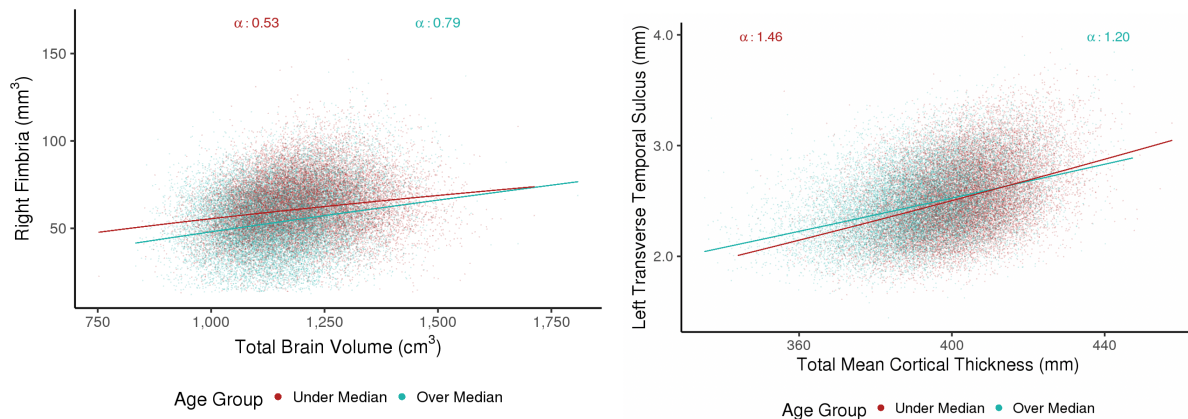
263

264

265 ***Figure S4***. **Largest Negative and Positive Total Cerebral Measure by Age Effects.** Colored

266 lines correspond to the predicted values of y from the model with TBV, sex, age and TBV, sex,

267 age^2 interactions. α is the scaling coefficient of the region with its total cerebral measure.

5.2. Sex Effects

5.2.1 Ventricles and CSF Volumes (ASEG).

All 9 out of 10 ventricles and CSF segmentations were greater in males ranging from -0.40 (left lateral ventricle) to -0.67 (right choroid plexus), except for the 5th ventricle.

5.2.2 Other Volumes (ASEG).

All 14 remaining ASEG volumes exhibited sex differences. The optic chiasm (β = -0.41), cerebellar GMV (Right β = -0.38, Left β = -0.33), and ventral diencephalon (Right β = -0.15, Left β = -0.15) were larger in males, while cerebellum WMV (Right β = 0.27, Left β = 0.26), cerebral WMV (Right β = 0.05, Left β = 0.07), and corpus callosum (β = 0.27 to 0.40) volumes were larger in females.

5.2.3 Subcortical Freesurfer Subsegmentations.

Of the 80 regions out of 116 subcortical regions (69%) exhibiting sex differences, 46 (40%) were larger in males and 34 (29%) were larger in females. The following subcortical sex differences are described by region.

Sex differences across the brainstem volume subsegmentations ranged from -0.49 (medulla) to -0.10 (pons).

The left amygdala was greater in males than females (β = -0.12), while the right amygdala did not show sex differences. However, sex differences were present in the right amygdala subregions, such as the cortical (β = 0.13) and paralaminar (β = -0.13) nuclei. Overall, 8 amygdala subsegmentations out of 20 were greater in males ranging from -0.18 (left paralaminar nucleus) to -0.06 (left accessory basal nucleus). The 3 regions that were larger in females were the left medial nucleus (β = 0.09) and left and right cortical nuclei (β = 0.09, β = 0.13).

The right and left hippocampus were greater in females than males (Right β = 0.07, Left β = 0.06), although some subsegmentations of the hippocampus were larger in males. Specifically, 5 out of 40 (12.5%) hippocampal volumes were larger in males, including the left and right hippocampal fissure (β = -0.29, β = -0.25), left and right parasubiculum (β = -0.11, β = -0.08), and left CA4 head (β = -0.10). Females showed greater volumes in 17 regions (42.5%) ranging from 0.06 (right subiculum body) to 0.19 (left molecular and granule cell layers of the dentate gyrus).

The right and left thalamus were greater in males (Right β = -0.15, Left β = -0.08), with 42 (81%) regions exhibiting sex differences. Male volumes were larger in 28 out of the 52 thalamic

segmentations (54%), ranging from -0.34 (right anterior pulvinar nucleus) to -0.06 (left ventral posterior nucleus), and 14 (27%) regions were greater in females, ranging from 0.07 (left lateral posterior) to 0.25 (right nucleus reuniens of the ventral midline thalamus).

5.3 Age Effects

5.3.1 *ASEG Volumes.*

A total of 28 out of 32 (87.5%) volumes showed linear age effects and 14 out of 32 (43.75%) volumes showed quadratic age effects.

All ventricle and CSF volumes increased with age ranging from 0.18 (4th ventricle volume) to 0.47 (3rd ventricle volume) and were positively associated with age^2 ranging from 0.03 (left choroid plexus) to 0.10 (right inferior lateral ventricle). The 5th ventricle did not show linear or quadratic age effects.

There was no linear or quadratic age effect on cerebellum GM, although cerebellum WMV decreased linearly with age (Left β = -0.13, Right β = -0.11). Volumes of the ventral diencephalon (Left β = -0.08, Right β = -0.09) additionally decreased with linear age, while cerebral WMV (β = 0.04) and the optic chiasm (0.23) increased with linear age.

Interestingly, age effects varied across segments of the corpus callosum. The corpus callosum posterior volume increased linearly with age (β = 0.10), while the corpus callosum mid-posterior (β = -0.14), central (β = -0.20), and mid-anterior (β = -0.22) volumes decreased with age and the anterior volume showed no age effect (β = 0.00).

Finally, putamen and accumbens area volumes decreased with linear age (β = -0.33 to -0.09), while pallidum and caudate volumes increased with linear age (β = 0.03 to 0.15). All volumes showed quadratic age effects ranging from -0.04 (left accumbens area) to 0.06 (left caudate), except for the left putamen and pallidum volumes.

5.3.2 *Subcortical Freesurfer Volumes.*

Linear age was a significant predictor of 108 out of 116 (93%) subsegmentations and quadratic age was a significant predictor of 94 regions (81%).

The amygdala decreased with age (Left β = -0.16, Right β = -0.14), with linear age effects ranging from -0.23 (left cortical nucleus) to -0.05 (left paralaminar nucleus), except for the right paralaminar nucleus which did not show linear age effects.

Hippocampal regions decreased with age (Left β = -0.23, Right β = -0.19), with 37 out of 40 (92.5 %) regions showing linear age effects ranging from -0.29 (right fimbria) to -0.06 (right parasubiculum). In contrast, the left and right hippocampal fissure increased with age (Left β = 0.28, Right β = 0.34).

The linear age effect on the left and right thalamus segmentations was not significant, although linear age effects were reported in thalamic subsegmentations ranging from -0.30 (right lateral geniculate nucleus) to 0.03 (right pulvinar anterior nucleus) and from 0.03 (right ventral lateral posterior nucleus) to 0.33 (right limitans suprageniculate nucleus).

Of the 108 regions with a linear age effect, 94 had a quadratic age effect ranging from -0.09 (right presubiculum head) to 0.11 (left centromedian nucleus). The left and right lateral posterior nuclei (both, $\beta = 0.06$), the right paralaminar nucleus ($\beta = -0.02$), and right pulvinar inferior ($\beta = 0.07$) showed a quadratic but not a linear age effect.

The amygdala (Left $\beta = -0.04$, Right $\beta = -0.05$) and its subsegmentations as well as the hippocampus (Left $\beta = -0.23$, Right $\beta = -0.20$) and its subsegmentations generally had a negative quadratic age effect. The whole thalamus left and right had a positive quadratic age effect (Left $\beta = 0.07$, Right $\beta = 0.06$), although the right lateral geniculate, right reuniens of the ventral midline thalamus, and right medial mediodorsal nuclei increased with quadratic age.

5.3.3 *Cortical Volumes.*

Of the 91 regions with a linear age effect, 5 also had a significant positive quadratic age effect. A quadratic age effect in the absence of a linear effect was also reported for the left parahippocampal gyrus ($\beta = -0.03$), the left superior part of the precentral sulcus ($\beta = 0.03$), and the left anterior part of cingulate gyrus and sulcus ($\beta = 0.03$).

5.3.4 *Cortical Mean Thicknesses.*

Of the 129 regions with an age effect, 8 regions had a positive and 13 had a negative quadratic age effect. The right straight gyrus (gyrus rectus) was the only area with a quadratic age but no linear age effect ($\beta = 0.02$).

5.3.5 *Cortical Surface Areas.*

Of the 85 regions with an age effect, 4 regions had a positive and 3 had a negative quadratic age effect. The left temporal pole surface area was the only area with a quadratic age but no linear age effect ($\beta = -0.02$).

5.4.1 Results

Global Measures. Total MCT (β = 0.10), cerebellum GMV (β = 0.05) and WMV (β = 0.04) decreased more rapidly with age in males compared to females and TSA increased more rapidly with age in males (β = -0.04; Figure 1A).

Regional Measures. A total of 83 regions out of 620 (13%) had a significant sex by age interaction, ranging from -0.16 (right cerebellar lobule X) to 0.19 (left fimbria). Of the 7 regions (1%) with a significant sex by age^2 interaction, the choroid plexus volumes (β = 0.08) and right CA4 body (β = 0.05) did not show a significant interaction of age by sex. The largest and only negative sex by age^2 effect was reported for the GMV of the cerebellar vermis X (β = -0.10), which also presented an age by sex effect (β = 0.17). The left fimbria and cerebellar vermis volumes decreased more in males than females, while the right cerebellar lobule X decreased more in females than males with age. The left choroid plexus increased more in males than in females with quadratic age (Figure S5).
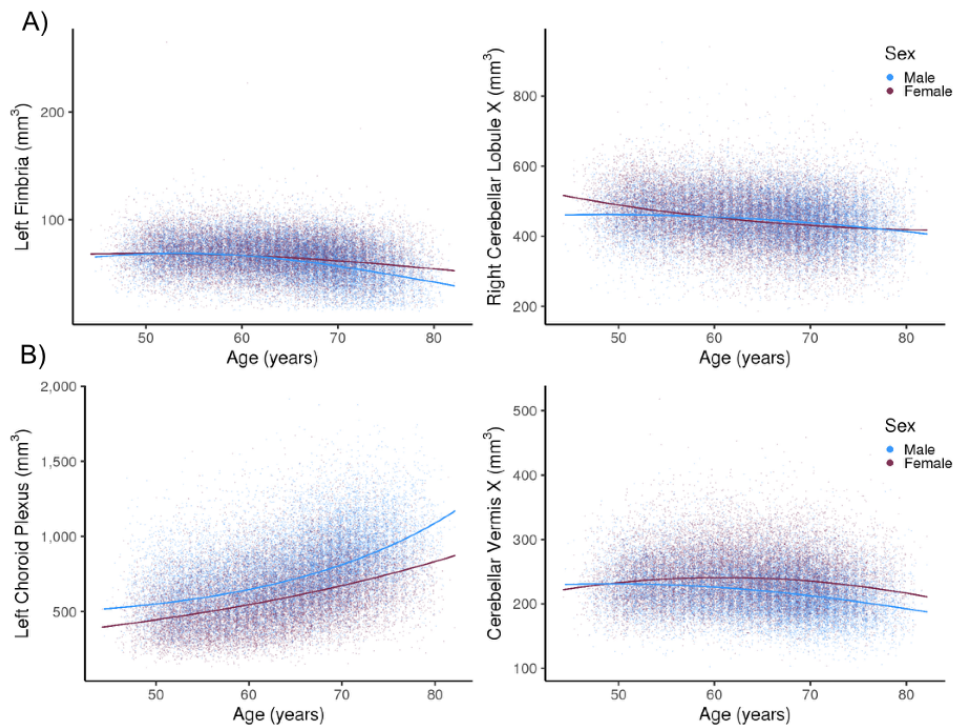
Figure S5. **Largest Sex by Age (A) and Sex by Age^2 (B) Interactions.** Blue and purple lines correspond to the predicted values of y from the model with TBV, sex, age and TBV, sex, age^2 interactions.

5.4.2 Discussion

Sex differences in linear and quadratic age effects were observed for a subset of regions (15%) supporting the presence of sex differences in cerebral aging trajectories (Armstrong et al., 2019; Vinke et al., 2019). Total MCT ($\beta = 0.10$), cerebellum GMV ($\beta = 0.05$) and WMV ($\beta = 0.04$) decreased more rapidly with age in males compared to females and TSA increased more rapidly with age in males ($\beta = -0.04$). Females had a steeper decrease with linear age across occipital temporal mean thicknesses, the thalamus, and its subsegmentations, whereas males showed a more rapid decrease with linear age in frontal and cingulate gyri and sulci mean thicknesses and hippocampal subsegmentations. Females had a steeper decrease with linear age in the cerebellar lobule X, whereas males had a steeper decrease with linear age in the cerebellar lobules iiv, vi, viib, viiia, viiib, ix and vermes viiia, ix, and x. Sex differences in quadratic age effects were only reported in 7 regions including the cerebellum x (flocculus), choroid plexus, and hippocampal CA4 body. Considering that the median absolute effect size for the age by sex interaction was 0.06, we speculate that inconsistent findings with previous studies likely stem from underpowered designs in addition to differences in age modeling (e.g., linear versus spline) and sample characteristics (e.g., age range; Armstrong et al., 2019; Good et al., 2001; Lemaitre et al., 2012; Vinke et al., 2019). Our study further adds to the literature reporting the presence of age-related cerebral changes that vary as a function of sex in late adulthood.

6 - Exploratory Analyses: Results

6.1 Sex Differences in Variance.

Sex differences in variance occurred in 306 out of 629 regions (49%; Supplemental Tables Tables E1-7) across measures types, ranging from 0.82 (for the right cerebellar lobule VIIIa, implying greater male variability) to 1.17 (for the optic chiasm, suggesting greater female variability, Figure S6). A total of 253 (40%) regions exhibited greater male variability and 56 (9%) exhibited greater female variability. A detailed description is made available by segmentation.
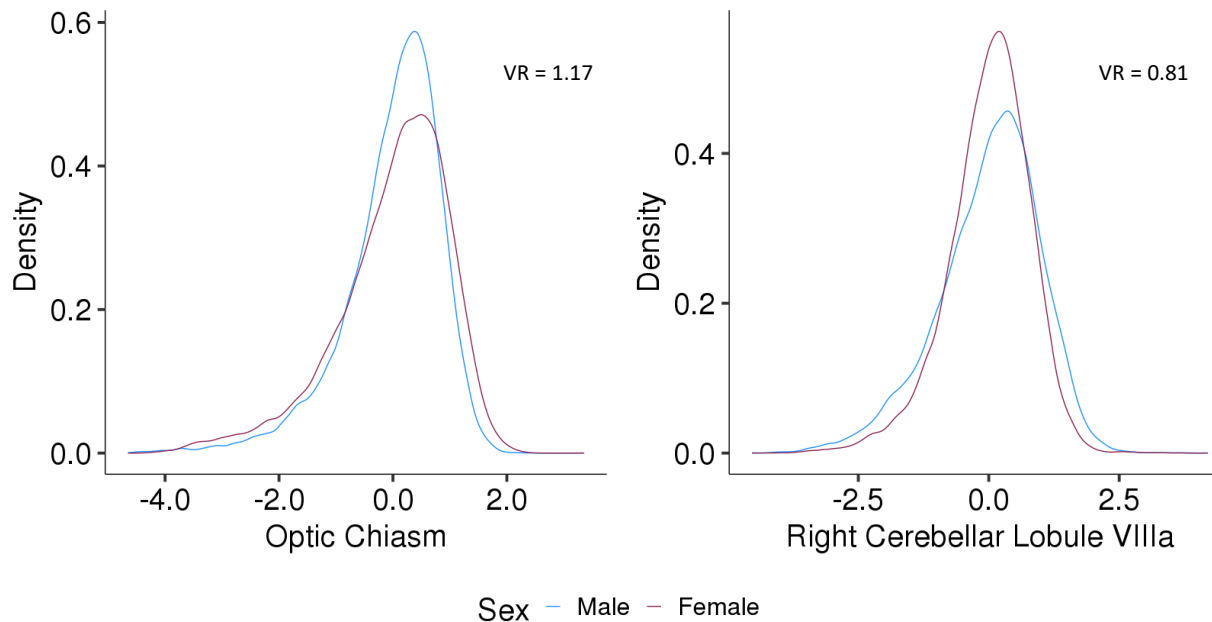
Figure S6. *Largest Sex Differences in Variance Ratio across Brain Regions*. VR: Variance Ratio, Female SD / Male SD. Optic Chiasm was more variable in females and the Right Cerebellar Lobule VIIIa was more variable in males. No sex differences are seen as density is plotted against the residuals of the model with age, total brain volume, and sex as well as age^2, total brain volume, and sex interactions.

6.1.1 Global Measures.

The variance for log of the cerebellum GMV was significantly greater in males compared to females (VR = 0.90, F = 174.04, p = 1.18E-39) as well as the midbrain volume (VR = 0.97, F = 16.06, p = 6.16E-05).

6.1.2 Cortical Volumes.

Sex differences in variance were found for 46% of cortical volumes (68/148) ranging from 0.91 (right parahippocampal gyrus) to 1.14 (right subcallosal gyrus). Of the 68 regions, 5 were more variable in females than males, including the subcallosal gyri, right suborbital sulcus, right pericallosal sulcus, and right horizontal ramus of the anterior segment of the lateral sulcus volumes (Figure S7).

6.1.3 Cortical Mean Thicknesses.

Sex differences in variance were found for 42% of cortical volumes (62/148) ranging from 0.90 (left middle-anterior part of the cingulate gyrus and sulcus) to 1.15 (left anterior transverse collateral sulcus). More than half of the mean thicknesses exhibiting differences in variance between sexes were more variable in females than males (Figure S7).

6.1.4 Cortical Surface Area.

Sex differences in variance were found for 53% of cortical volumes (79/148) ranging from 0.90 (left opercular part of inferior frontal gyrus) to 1.14 (right subcallosal gyrus). Only 8 cortical surface areas were more variable in females than males: the subcallosal gyri, the lateral orbital sulci, the right suborbital sulcus, the left posterior transverse collateral sulcus, the lateral orbital sulcus, and the right inferior temporal sulcus (Figure S7).
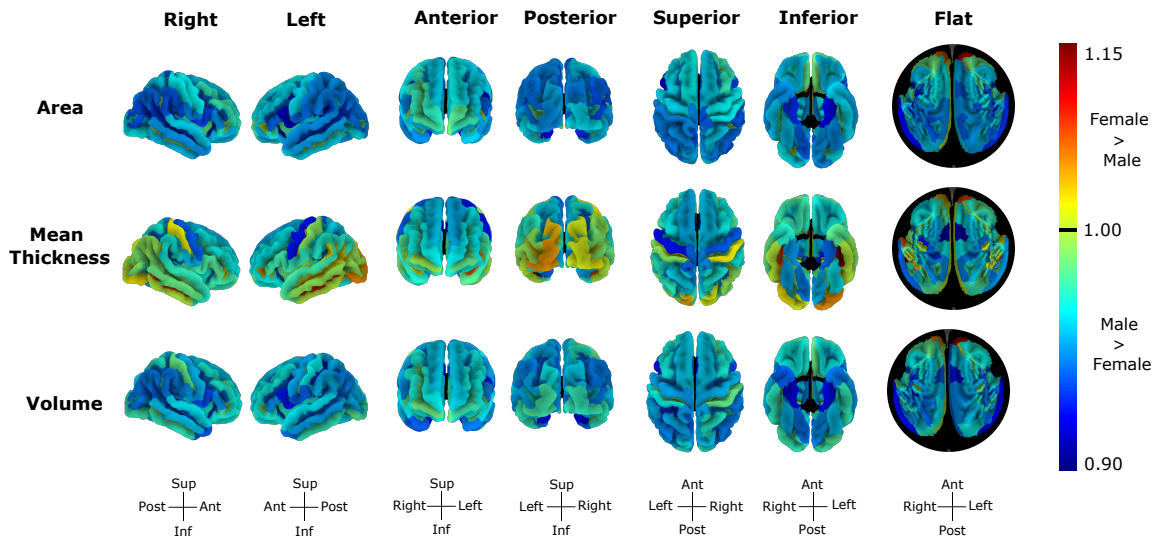
Figure S7. Variance Sex Differences across Cortical Measures. Sex differences in variance (Variance Ratio = SD in females/SD in males) ranged from 0.90 (left surface area of the opercular part of the inferior frontal gyrus) to 1.15 (left mean thickness of the anterior transverse collateral sulcus). The flat representation corresponds to the flatten image of the superior view with the circle middle reflecting regions within the sagittal plane and circle edges reflecting inferior regions.

6.1.5 ASEG Volumes.

A total of 14 out of 32 (44%) regions showed sex differences in variance ranging from 0.93 (posterior corpus callosum) to 1.17 (optic chiasm).

The left and right choroid plexus were more variable in females than males (VR= 1.11, VR = 1.07, respectively), while the 3rd and 4th ventricles were more variable in males than females (VR= 0.95, VR = 0.96, respectively). The left and right white matter volumes did not differ between sexes in terms of variance except for the optic chiasm volume, which was greater in females. Interestingly, the corpus callosum posterior volume variance was greater in males than females (VR= 0.93), while the corpus callosum central and mid-anterior volume variances were greater in females than males (VR= 1.04, VR= 1.02). Finally, the left and right caudate volumes which did not show sex differences in mean were more variable in males than females (VR= 0.95, VR= 0.96).

6.1.6 Subcortical Freesurfer Volumes.

Sex differences in variance were found in 64 out of 116 (55%) of whole subcortical and subsegmentations volumes and were in larger in males. The whole hippocampus (Left VR = 0.96, Right VR= 0.93) was the only whole subcortical volume that differed in terms of variance between sexes. The majority of hippocampal subsegmentations (28/40, 71%) were more variable in males than females, including regions such as the CA1, CA3, and CA4 bodies and heads, which differed in terms of mean sex differences. All hippocampal body volumes of the CA1, CA3, CA4, molecular layer, presubiculum, subiculum, and the granule cell layers of the dentate gyrus were more variable in males. About 60% (31/52) thalamic subsegmentations were more variable in males than females. The anteroventral regions and lateral geniculate nuclei which did not differ between sexes in terms of mean were more variable in males than females. Greater male variability was largely spread across thalamic regions (anterior, posterior, intralaminar, ventral, medial, and lateral regions).

6.1.7 Cerebellar GMVs.

About 89% of cerebellar regions exhibited sex differences in variance (25/28) ranging from 0.82 (right cerebellar lobule VIIIa) to 1.04 (left cerebellar lobule X). The left and right cerebellar lobule X were the only regions that were more variable in females than males (VR = 0.96, for both). The cerebellar vermis X, crus I vermis, and vermis VI did not vary differently between sexes.

6.2 Comparing TBV Adjustment Techniques – Allometric Versus Covariate or Proportion Adjustment

6.2.1 Comparing Standardized Beta Coefficients when omitting and considering brain allometry.

We examined result discrepancies between TCM adjustment techniques (allometric vs. linear covariate and allometric vs. proportion) across effects and interactions that were significant in our primary analyses. We did not look at differences in significance between TCM adjustment techniques for the TCM effect, since TCM was always a significant predictor of the brain region (except for the 5th ventricle). See Supplemental Table F1 for the consistency in the number of

significant standardized beta coefficients across global measure adjustment techniques by segmentation.

We found conflicting results in 2.35% of statistical tests between the allometric and the linear covariate approach (102/4340). Discrepancies were reported in 3.44% of volumes (78/2268), 0.97% of surfaces (10/1036), and 1.35% of thicknesses (14/1036). There were more significant effects and interactions across cortical surface areas and mean thicknesses, subcortical subsegmentations, and cerebellar volumes in the models with the allometric adjustment compared to the models with the covariate adjustment. The opposite was true for ASEG and cortical volumes. Inconsistencies between allometric and linear covariate adjustment models were observed for effects and interactions near the threshold of significance (Supplemental Figures 3 https://osf.io/s4qc5/?view_only=7c6e633d701d443e96602a857cf75824).

Meanwhile, we found conflicting results in 14.24% of statistical tests between the allometric and the proportion approach (618/4340). Discrepancies were reported in 18.87 % of volumes (428/2268), 13.51% of surfaces (140/1036), and 4.83% of thicknesses (50/1036). The proportion method models yielded a greater number of significant effects and interactions across all cerebral measures compared to the models with the allometric adjustment.

The standardized beta varied across models, with a difference of up to 0.04 between the allometric and the covariate approach and up to 0.23 between the allometric and the proportion method (Supplemental Figures 3, https://osf.io/s4qc5/?view_only=7c6e633d701d443e96602a857cf75824). For instance, the standardized beta of the interaction of TBV by Sex reported for the right cerebellar VIIIa lobule was at -0.05 in the linear model and -0.09 in the allometric model and the standardized beta of the interaction of Age^2 by Sex reported for the right anterior pulvinar thalamic nucleus was at -0.23 in the linear model and 0.00 in the allometric model.

6.2.2 Comparing the Sex Difference in Variance across TCM Adjustment Techniques.

We found conflicting results in 38.06% of the reported sex differences in variance between the allometric and the linear covariate approach (236/620) and in 29.03% of the reported sex differences in variance between the allometric and the proportion approach (180/620). Discrepancies between the allometric and covariate adjustment models were reported in 46.91% of volumes (152/324), 46.62% of surfaces (69/148), and 10.14% of thicknesses (15/148), while

discrepancies between the allometric and covariate adjustment models were reported in 37.04% of volumes (120/324), 30.41% of surfaces (45/148), and 10.14% of thicknesses (15/148). Moreover, 96 % of the sex differences in variance found in the allometric model were present in the linear model with the covariate adjustment for TBV and 37% for the linear model with the proportion adjustment for TBV. Finally, 57 % of the sex differences in variance found in the linear model with the covariate adjustment for TBV were present in the allometric model and 32% for the linear model with the proportion approach. Sex differences in variance were more common when the variance ratios were calculated from the proportion and linear covariate models than the allometric models across regional and global measures (see Supplemental Tables F2 and F3, respectively and Supplemental Figures 4 https://osf.io/s4qc5/?view_only=7c6e633d701d443e96602a857cf75824).

We also calculated the correlations between the standardized beta ($\beta < 0$: Female < Male) and the VR of the sex difference (VR < 1: Female < Male) to examine whether regions with the largest sex differences were also the most variable (Supplemental Figures 2 https://osf.io/s4qc5/?view_only=7c6e633d701d443e96602a857cf75824). In models with the linear covariate or proportion TCM adjustment, larger regions in females were more variable in females in subcortical ($r_{covariate} = 0.39$, $r_{proportion} = 0.48$), and ventricle, WM, and corpus callosum volumes ($r_{covariate} = 0.64$, $r_{proportion} = 0.71$). In contrast, larger regions in females were more variable in males ($\mu\, r = -0.33$) in the allometric adjustment for TCM and significant across cerebral regions except for the cerebellar and ventricle, WM, and corpus callosum volumes (Supplemental Table F4).

6.2.3 Correlations between a region's deviance from isometry (i.e., |1 – scaling coefficient|) and the difference in standardized beta between the linear (covariate or proportion) and allometric models

We examined the correlations between a region's deviance from isometry (i.e., |1 – scaling coefficient|) and the difference in standardized beta between the linear (covariate or proportion) and allometric models (e.g., |Covariate Standardized Beta – Allometric Standardized Beta|) to investigate whether the difference in standardized beta between the linear and allometric models occurred in regions with a greater degree of allometry.

Across all regions, the correlation between deviance from isometry and the difference in the standardized beta of all effects was 0.16 (p = 1.30e-13) for the allometric and the covariate models and was 0.28 (p = 2.22e-16) for the allometric and the proportion models.

When looking at the correlations by effect or interaction across regions (Supplemental Figures 5 https://osf.io/s4qc5/?view_only=7c6e633d701d443e96602a857cf75824), we found strong significant correlations between the deviance from isometry and the standardized beta difference of the allometric and proportion approach ranging from 0.41 (TBV x Sex) to 0.67 (TBV x Age) and smaller significant correlations for the standardized beta difference of the allometric and linear covariate approach ranging from 0.15 (Sex) to 0.28 (Age^2; Supplemental Table F5A). The strength and direction of the association varied across segmentations (Supplemental Table F5B).

6.3 Replication of Sex Differences in Studies considering Brain Allometry

We attempted to replicate previous sex differences reported by studies taking into account brain allometry and found mixed results (see Supplemental Tables G1- 4).

In short, we replicated the sex differences in the frontal lobes described by Peyre and colleagues (2020) in children who did not adjust for age or age^2, as well as the greater pallidum volume in males reported by Reardon and colleagues' (2016). However, instead of being greater in males (Mankiw et al., 2017), we found that the cerebellum, flocculus, cerebellar lobule VIIb, VIIb, and VIIA, and Crus II volumes were greater in females and that the flocculus volume did not vary between sexes. Finally, Sanchis-Segura and colleagues (2019) adjusted for brain allometry using the power proportion method which does not take into account the intercept of the linear equation (log10(y) = Intercept + slope * log10(x)). We replicated only 10 out of the 25 mean sex differences identified by Sanchis-Segura and colleagues (2019) after an FDR correction for multiple comparison. It is important to note that the effect sizes in our sample and confidence intervals of our study were much smaller. The same significant results were reported when using the logarithmic equation to adjust for brain allometry, although effect sizes were greater and p values were smaller.

6.4 Replication of Ritchie and Colleagues (2018) Study and Comparing Results with the Covariate and with the Allometric TBV Adjustment.

We attempted to replicate the mean and variance sex differences reported by Ritchie and colleagues (2018) when adjusting brain volumes, thicknesses, and surface areas, for age, ethnicity, and total brain volume (for volumes) or total surface area (for areas) or total mean cortical thickness (for mean thicknesses). After providing a summary of the replication analyses and TCM adjustment comparison, we provided a detailed comparison by segmentation where we simultaneously discussed our replication results and compared our results from our replication analyses to those obtained when considering brain allometry. Replication and TCM comparison results are provided in Supplemental Tables G5-G9.

6.4.1 Summary.

The sex differences in mean and variance reported by Ritchie and colleagues (2018) in subcortical volumes from the FIRST segmentations and cortical volumes, mean thicknesses, and surface areas from the Desikan-Killiany Cortical Atlas were generally replicated, with some exceptions. More mean differences in volumes and surface areas and sex differences in variance for mean cortical thickness were reported in our replication compared to Ritchie and colleagues' (2018) study (Table S1).

Mean sex differences were consistent between the allometric and linear covariate TCM adjustments, except for the left pars opercularis and left transverse temporal volumes which no longer significant. When taking into account brain allometry, the sex difference in variance was no longer significant for the left and right pallidum and the number of significant cortical sex differences in variance decreased in volumes and surface areas but increased in mean thicknesses (Table S1).

Finally, we replicated the absence of significant correlations reported by Ritchie and colleagues (2018) between the sex difference's Cohen's d and the VR in cortical and subcortical volumes, cortical mean thicknesses, and cortical surface areas. We found that when taking into account brain allometry, the correlation coefficients between the sex difference's Cohen's d and the VR in cortical and subcortical volumes decreased and became significant for cortical mean thickness (Supplemental Table G10).

Table S1. A Comparison of the Number of Significant Mean Sex Difference reported by Ritchie and colleagues (2018) to those reported in our Linear Covariate and Allometric Replications.

	Total Brain Volume Adjustment	Ritchie and colleagues (2018)	Result Overlap	Present Study	
		Linear Covariate		Linear Covariate	Allometric
Significant Mean Sex Differences	Cortical Volumes	23	19	52	53
	Cortical Surface Areas	25	23	49	50
	Cortical Mean Thicknesses	46	42	52	52
	Subcortical Volumes (N = 14)	6	6	9	11
	Total N	100	90	162	166
	Total %	50	45	81	83
Significant Variance Sex Differences	Cortical Volumes	58	58	62	55
	Cortical Surface Areas	59	59	62	50
	Cortical Mean Thicknesses	2	1	25	32
	Subcortical Volumes	14	14	14	12
	Total N	133	132	163	149
	Total %	66.5	66	81.5	74.5

N.B. Significance based on FDR correction. Result Overlap: Significant results reported both by Ritchie and colleagues (2018) and by our replication of their study with the linear covariate adjustment for Total Brain Volume. Total Cortical Measures from the Desikan-Killiany Atlas, N = 186 and from the Subcortical FIRST segmentations, N= 14.

### 6.4.2 Subcortical Volumes: Mean and Variance Sex Differences.

Mean Sex Differences. We replicated the majority of sex differences reported by Ritchie and colleagues (2018). However, in contrast with Ritchie and colleagues (2018) who did not find sex differences in the thalamic, hippocampal and accumbens volumes, we found that the left hippocampus and accumbens volumes regions were significantly greater in females compared to males in our replication when adjusting for TBV with the covariate approach (Supplemental Figures File 6). With the allometric approach, the left thalamus and the right hippocampus were also greater in females.

Variance Sex Differences. We replicated Ritchie and colleagues' (2018) finding that males show greater variability across subcortical volumes when adjusting for TBV with the covariate approach. However, when adjusting for TBV with the allometric approach sex difference in variability for the left and right pallidum disappeared.

6.4.3 Cortical Volumes: Mean and Variance Sex Differences.

Mean Sex Differences. We replicated 19 of the 23 sex differences reported by Ritchie and colleagues. Specifically, the right paracentral, right superior temporal, left pars opercularis and left transverse temporal volumes did not differ between sexes in our study when adjusting for TBV with the covariate or the allometric approach (*Figures S8-9*). We reported 52/62 significant sex when adjusting for TBV with the covariate approach and 53/62 when adjusting for brain allometry. Covariate and allometric approaches were consistent except for the left precuneus which was only significant when considering brain allometry.

Variance Sex Differences. In addition to replicating the reported sex difference in volumetric variance by Ritchie and colleagues (2018), we additionally found greater male variability in the left pars orbitalis, left rostral anterior cingulate, and the left and right pericalcarine volumes. However, after adjusting for brain allometry, the left caudal anterior cingulate, left isthmus cingulate, left and right medial orbitofrontal, left posterior cingulate, left insula, and right lateral orbitofrontal volumes no longer exhibited sex differences in variance.
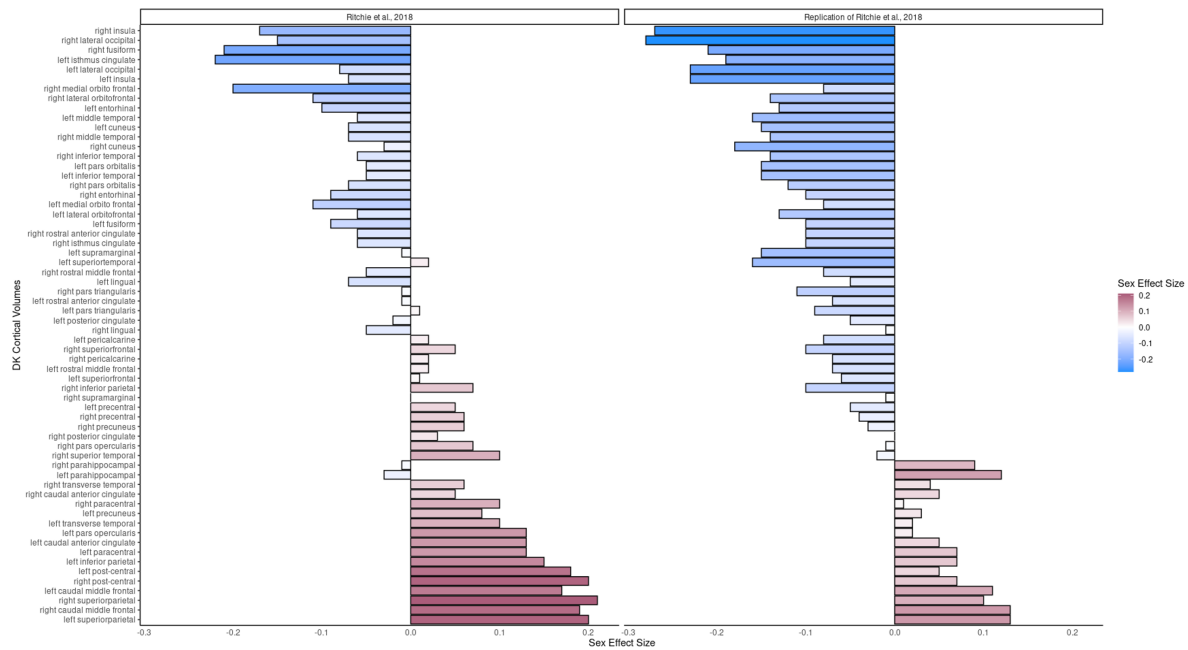
Figure S8. Sex differences in Cortical Volumes reported by Ritchie and colleagues (2018) and in our replication of their study. Effect sizes are Cohen's d. Negative values indicate greater male volumes and positive greater female volumes. Desikan-Killiany-Tourville segmentations were used for Ritchie and colleagues (2018; N = 5,216) and the Replication of Ritchie and colleagues' (2018) analyses (N = 39,363). Sex coefficients from the replication analyses when brain measures were log transformed were similar to those obtained without log transformation.
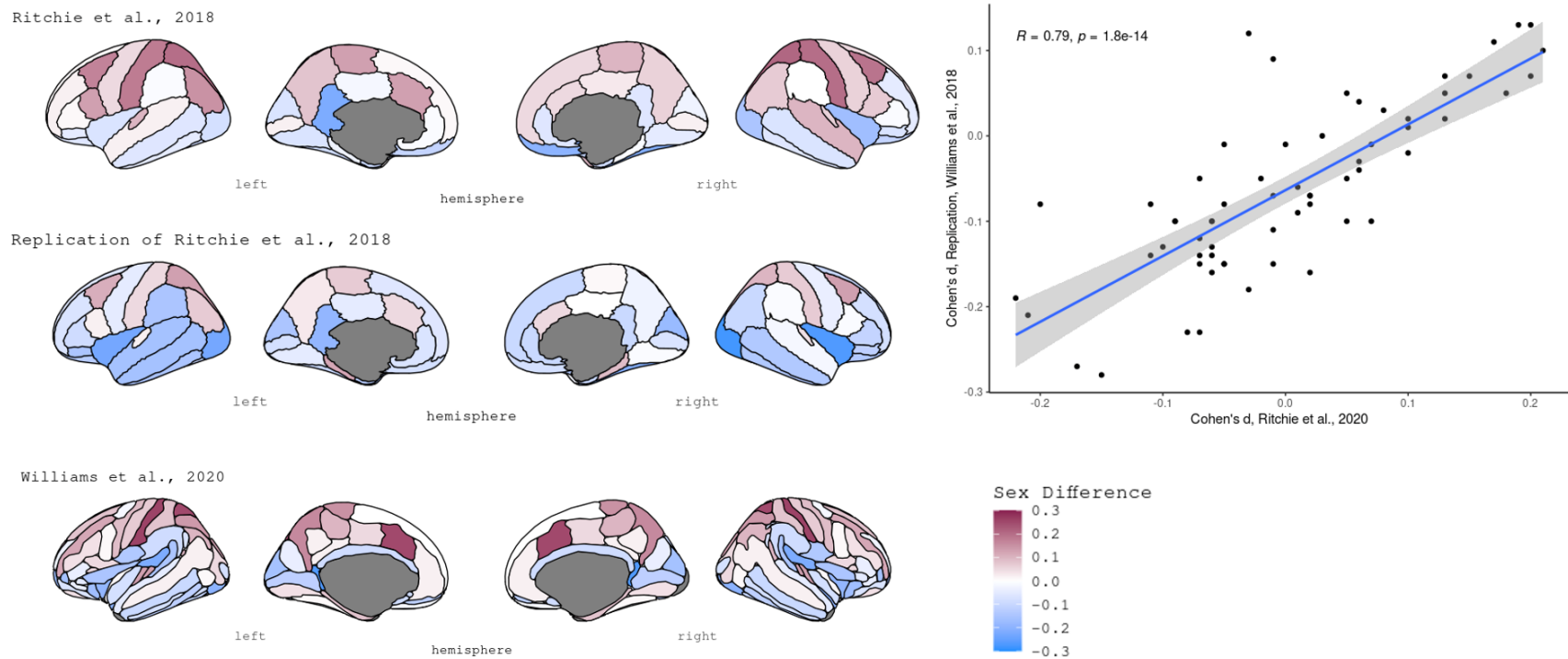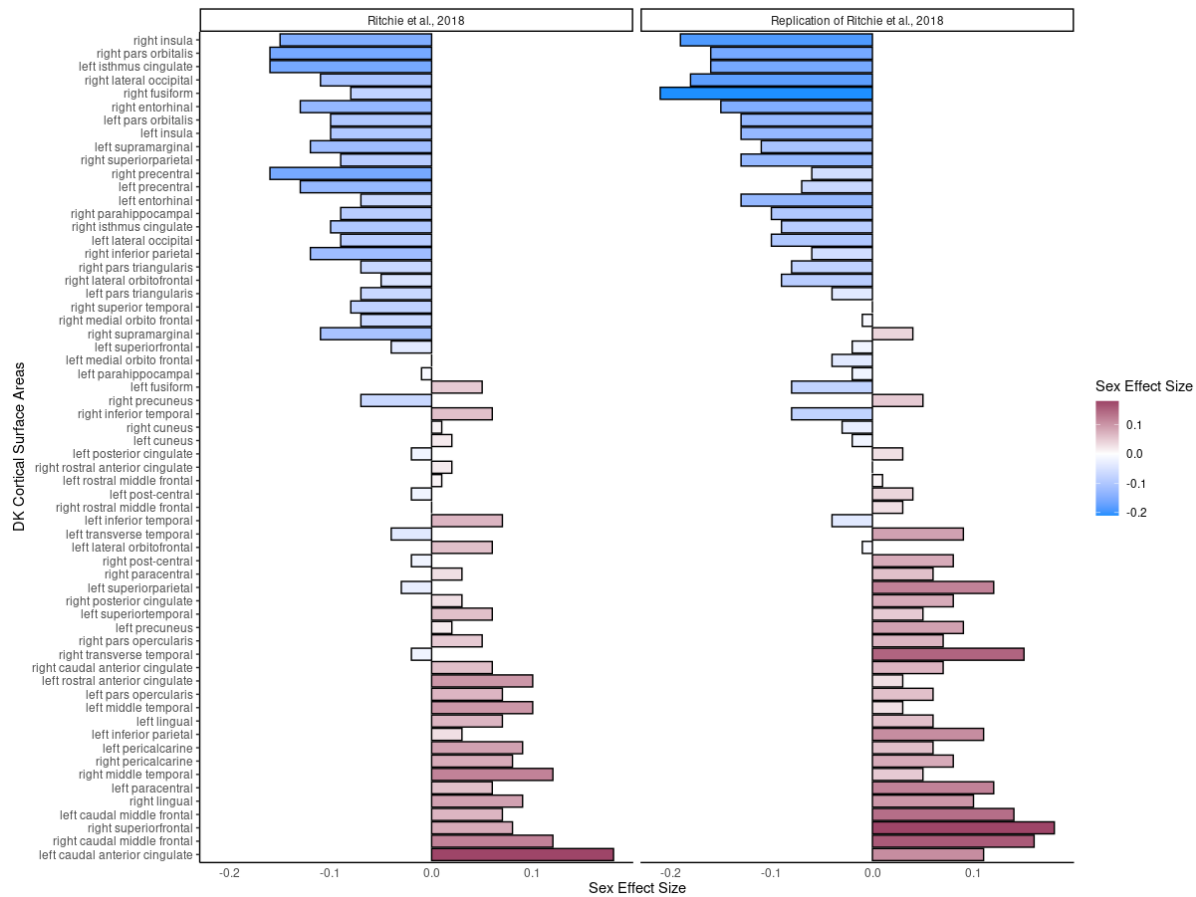
Figure S9. Sex differences in Cortical Volumes by Study. Effect sizes are either standardized betas or Cohen's d. Brain maps from ggseg (Mowinckel & Vidal-Piñeiro, 2019) and ggsegExtra (*Ggseg/GgsegExtra*, 2019/2021) packages, with regions defined using the Desikan-Killiany-Tourville atlas for the Ritchie et al., 2018 study and the replication and the Destrieux a2009s atlas for the Williams et al., 2020 study. Negative values indicate greater male volumes and positive greater female volumes. Sex coefficients from the replication analyses when brain measures were log transformed were similar to those obtained without log transformation. Scatter plot corresponds to the correlation between the Cohen's d of a region. R: pearson's correlation coefficient.

6.4.4 Cortical Surface Areas: Mean and Variance Sex Differences.

Mean Sex Differences. We replicated 23 out of the 25 sex differences in mean reported by Ritchie and colleagues' (2018) when linearly adjusting for TSA, although we found that 49/62 regions differed between sexes (Figure S10-11). Ritchie and colleagues (2018) reported that females had a greater right supramarginal area, while the opposite was true in our analyses. Allometric and covariate approaches were consistent, except for the right rostral middle frontal area which was only significant when considering brain allometry.

Variance Sex Differences. In comparison with Ritchie and colleagues (2018), we replicated all sex differences in variance and additionally reported significant effects in the right and left pericalcarine area and left pars orbitalis area. However, following adjustment for brain allometry, left caudal anterior cingulate, left isthmus cingulate, left lateral medial orbitalfrontal, left medial orbitofrontal, left pericalcarine, left superior frontal, right inferior temporal, right lateral orbitofrontal, right pericalcarine, right precentral, and the right rostral anterior and middle frontal cingulate areas no longer exhibited sex differences in variance.

Figure S10. Sex differences in Cortical Surface Areas reported by Ritchie and colleagues (2018) and in our replication of their study. Effect sizes are Cohen's d. Negative values indicate greater male volumes and positive greater female volumes. Desikan-Killiany-Tourville segmentations were used for Ritchie and colleagues (2018; N = 5,216) and the Replication of Ritchie and colleagues' (2018) analyses (N = 39,363). Sex coefficients from the replication analyses when brain measures were log transformed were similar to those obtained without log transformation.
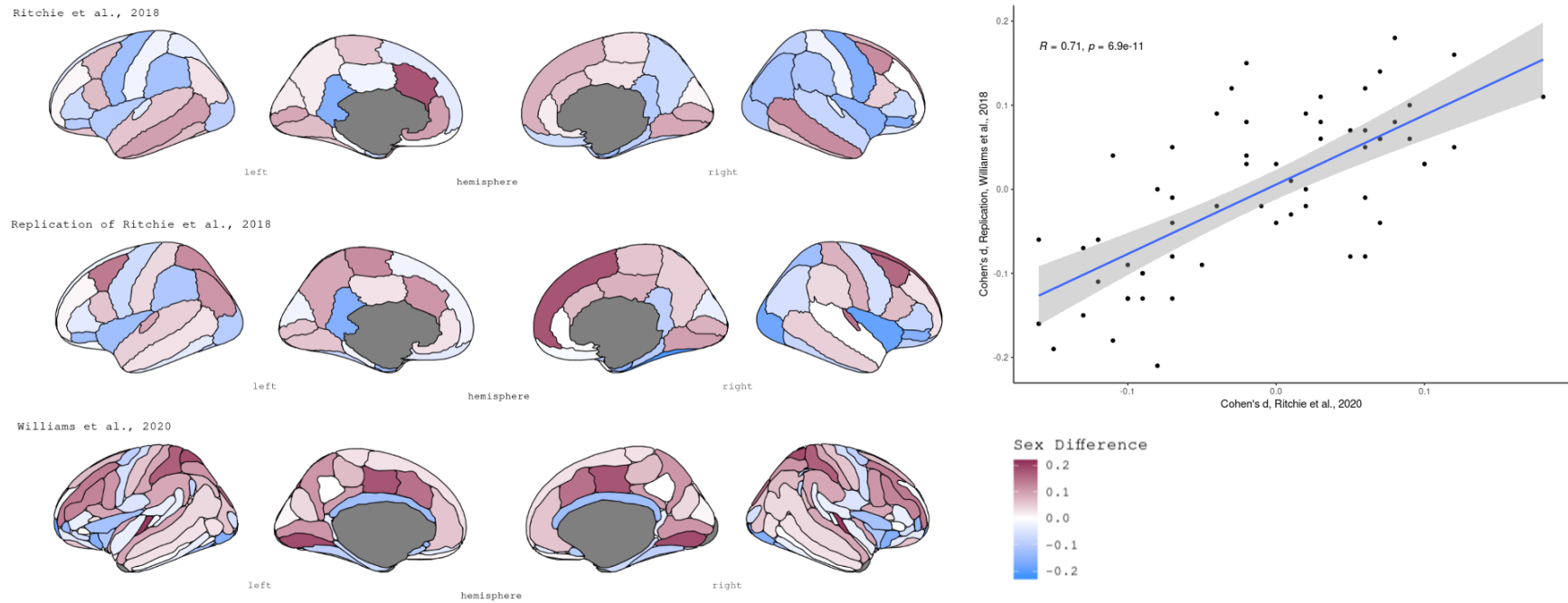
Figure S11. Sex differences in Cortical Surface Areas by Study. Effect sizes are either standardized betas or Cohen's d. Brain maps from ggseg (Mowinckel & Vidal-Piñeiro, 2019) and ggsegExtra (*Ggseg/GgsegExtra*, 2019/2021) packages, with regions defined using the Desikan-Killiany-Tourville atlas for the Ritchie et al., 2018 study and the replication and the Destrieux a2009s atlas for the Williams et al., 2020 study. Negative values indicate greater male volumes and positive greater female volumes. Sex coefficients from the replication analyses when brain measures were log transformed were similar to those obtained without log transformation. Scatter plot corresponds to the correlation between the Cohen's d of a region. R: pearson's correlation coefficient.

6.4.5 Cortical Mean Thicknesses: Mean and Variance Sex Differences.

Mean Sex Differences. We replicated 42 out of the 46 significant sex differences reported by Ritchie and colleagues (2018). We did not replicate the sex differences in the left entorhinal, left pars orbitalis, left transverse temporal, and right fusiform mean thicknesses when linearly adjusting for total mean cortical thickness or when taking into account brain allometry. Moreover, Ritchie and colleagues (2018) reported that the left lateral occipital, right inferior parietal, and were greater in males, while we found them to be greater in females. Finally, Ritchie and colleagues (2018) reported that the right paracentral was greater in males and the left lateral occipital mean thickness was greater in females but the opposite was true in our analyses (Figures S12-13).

We found that 52/62 regions were significantly different between males and females. The left caudal middle frontal, left inferior temporal, left paracentral, left pars triangularis, left posterior cingulate, left rostral anterior cingulate, right caudal middle frontal, right entorhinal, right lateral orbitofrontal, and right rostral anterior cingulate mean thicknesses did not significantly differ between sexes in Ritchie and Colleagues (2018) but did in the present study when considering the linear and allometric relationship of mean thicknesses with total mean cortical thickness.

Variance Sex Differences. In terms of sex differences in mean thickness variability, Ritchie and colleagues (2018) reported greater male variability in the left middle temporal mean thickness and greater female variability in the right inferior parietal mean thickness, while we found greater female variability in the left middle temporal mean thickness and no sex difference in variance in the right inferior parietal mean thickness. We found that 25/62 cortical mean thicknesses differed between sexes in terms of variability when linearly adjusting for total mean cortical thickness and 32/62 when considering brain allometry.
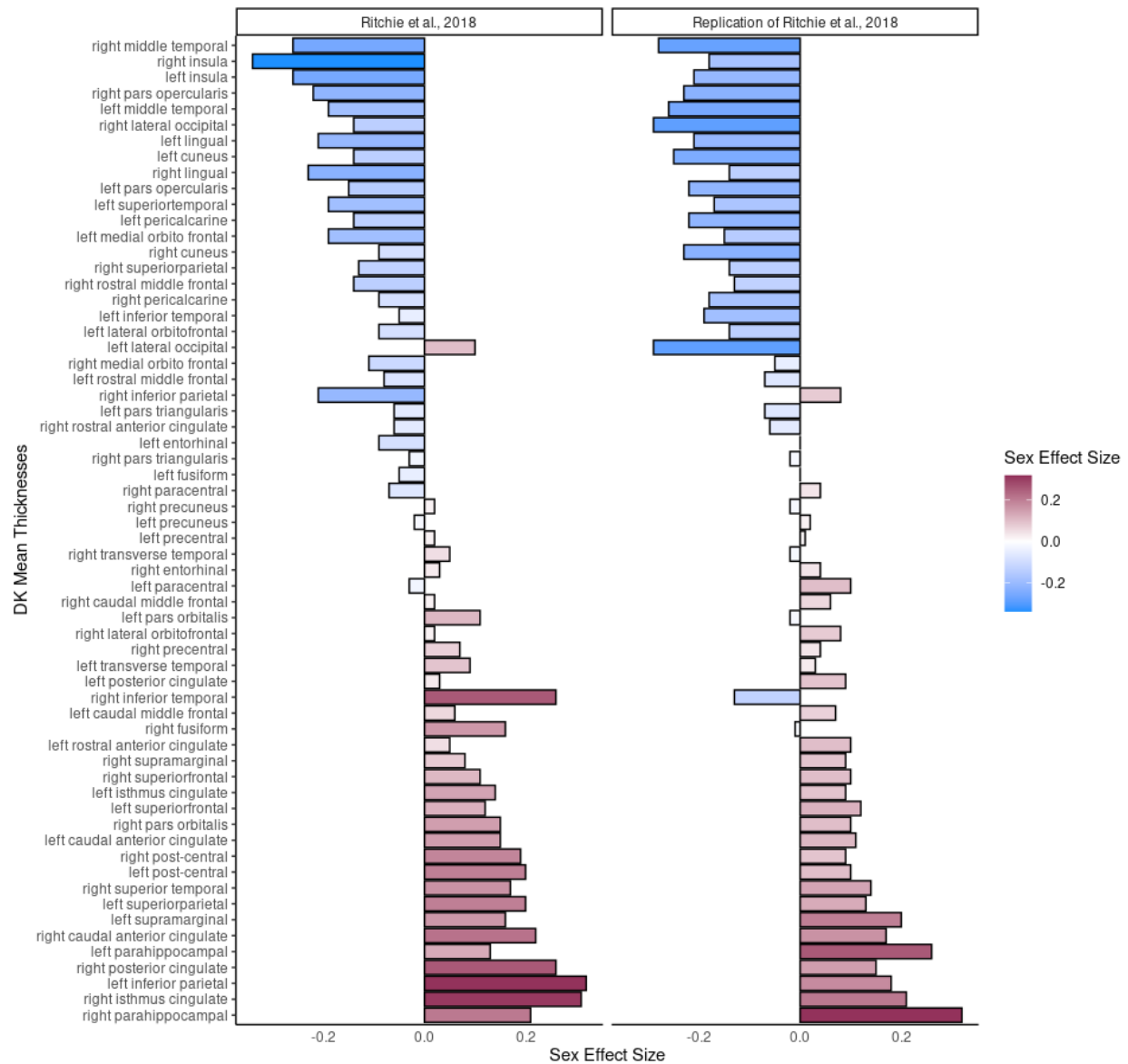
Figure S12. Sex differences in Cortical Mean Thicknesses reported by Ritchie and colleagues (2018) and in our replication of their study. Effect sizes are Cohen's d. Negative values indicate greater male volumes and positive greater female volumes. Desikan-Killiany-Tourville segmentations were used for Ritchie and colleagues (2018; N = 5,216) and the Replication of Ritchie and colleagues' (2018) analyses (N = 39,363). Sex coefficients from the replication analyses when brain measures were log transformed were similar to those obtained without log transformation.
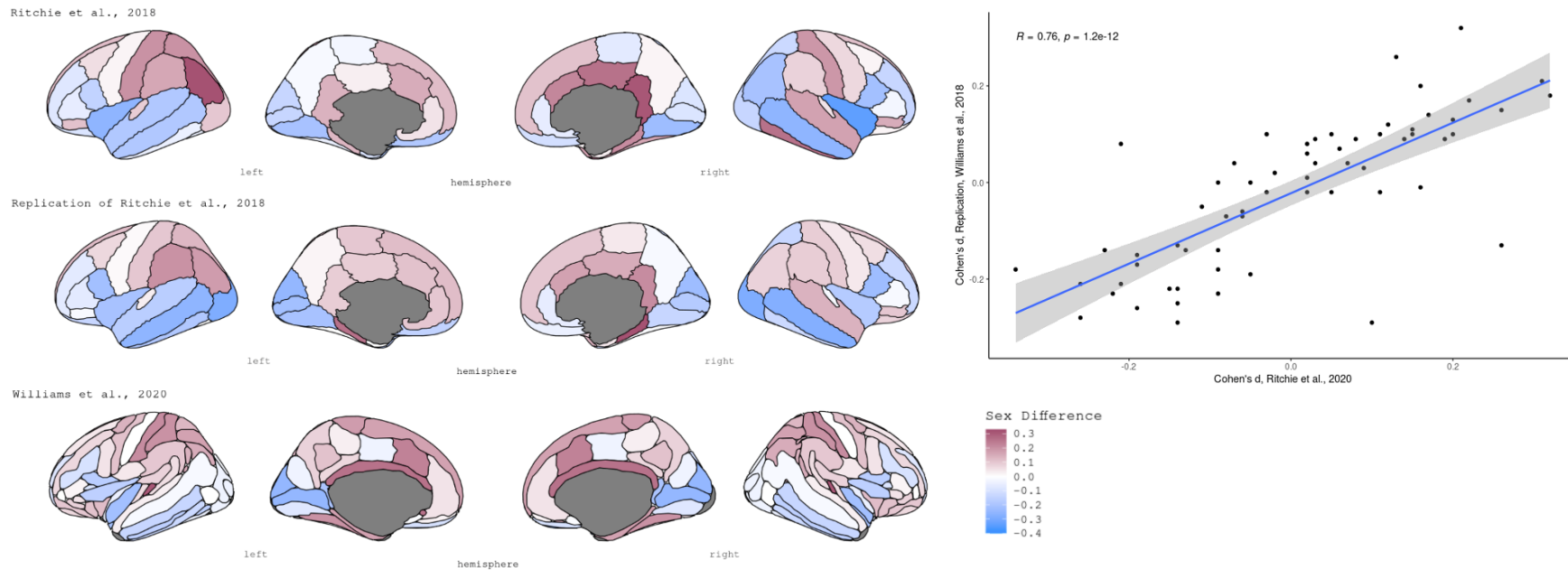
Figure S13. Sex differences in Cortical Mean Thicknesses by Study. Effect sizes are either standardized betas or Cohen's d. Brain maps from ggseg (Mowinckel & Vidal-Piñeiro, 2019) and ggsegExtra (*Ggseg/GgsegExtra*, 2019/2021) packages, with regions defined using the Desikan-Killiany-Tourville atlas for the Ritchie et al., 2018 study and the replication and the Destrieux a2009s atlas for the Williams et al., 2020 study. Negative values indicate greater male volumes and positive greater female volumes. Sex coefficients from the replication analyses when brain measures were log transformed were similar to those obtained without log transformation. Scatter plot corresponds to the correlation between the Cohen's d of a region. R: pearson's correlation coefficient.

6.5 Height

To investigate whether height explains part of the observed main effects and interactions on brain measures, we ran the models with height as a covariate in models of global brain measures. We found that height, sex, age, and age^2 were generally significant predictors of all global measures (Supplemental Tables B8).

We also ran the analyses of global brain measures with both height and TBV effects and interactions as covariates (Supplemental Tables B9). In general, TBV, sex, age and age2 were generally significant predictors of all global measures, with a few exceptions. Sex was not a significant predictor of Cerebral GM and age^2 was not significant for Cerebral WM, Cerebral GM and Cerebellar WM.

We find that sex differences in brain size cannot be reduced to sex differences in body size. The (within-sex) correlation between height and TBV was only about 0.29. Whereas the sex difference in TBV was d=-1.14, the sex difference in TBV adjusted on height was still large d=-0.7, suggesting that males have a disproportionately large brain compared to females of the same height. The same results were apparent for all other global brain measures (with the exception of MCT), with adjusted effect sizes ranging from -0.09 (Cerebellar GM) to -0.72 (TSA), and even when both height and TBV were adjusted.

7- R packages

The following R packages were used in the study:

- data.table (Dowle et al., 2020)
- ggplot2 (Wickham, Chang, et al., 2020)
- tidyr (Wickham & RStudio, 2020)
- broom (Robinson et al., 2020)
- sjPLot (Lüdecke et al., 2020)
- car (Fox et al., 2020)
- ggpubr (Kassambara, 2020)
- plyr (Wickham, 2020)
- dplyr (Wickham, François, et al., 2020)
- MatchIt (Ho et al., 2011)
- Tables (Murdoch, 2020)
- Scales (Wickham, Seidel, et al., 2020)
- Corrplot (Wei et al., 2017)
- Ggrepel (Slowikowski et al., 2020)
- Ggeseg (Mowinckel & Vidal-Piñeiro, 2019)

References

Armstrong, N. M., An, Y., Beason-Held, L., Doshi, J., Erus, G., Ferrucci, L., Davatzikos, C., &
Resnick, S. M. (2019). Sex differences in brain aging and predictors of
neurodegeneration in cognitively healthy older adults. *Neurobiology of Aging*, *81*, 146–
156. https://doi.org/10.1016/j.neurobiolaging.2019.05.020

Dowle, M., Srinivasan, A., Gorecki, J., Chirico, M., Stetsenko, P., Short, T., Lianoglou, S.,
Antonyan, E., Bonsch, M., Parsonage, H., Ritchie, S., Ren, K., Tan, X., Saporta, R.,
Seiskari, O., Dong, X., Lang, M., Iwasaki, W., Wenchel, S., … Eddelbuettel, D. (2020).
data.table: Extension of "data.frame" (1.13.4) [Computer software]. https://CRAN.R-
project.org/package=data.table

Fox, J., Weisberg, S., Price, B., Adler, D., Bates, D., Baud-Bovy, G., Bolker, B., Ellison, S.,
Firth, D., Friendly, M., Gorjanc, G., Graves, S., Heiberger, R., Krivitsky, P., Laboissiere,
R., Maechler, M., Monette, G., Murdoch, D., Nilsson, H., … R-Core. (2020). *car:
Companion to Applied Regression* (3.0-10) [Computer software]. https://CRAN.R-
project.org/package=car

Ggseg/ggsegExtra. (2021). [R]. ggseg - LCBC - University of Oslo.
https://github.com/ggseg/ggsegExtra (Original work published 2019)

Good, C. D., Johnsrude, I., Ashburner, J., Henson, R. N. A., Friston, K. J., & Frackowiak, R. S.
J. (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*(3), 685–700. https://doi.org/10.1006/nimg.2001.0857

Herbert, M. R., Ziegler, D. A., Deutsch, C. K., O'Brien, L. M., Lange, N., Bakardjiev, A.,
Hodgson, J., Adrien, K. T., Steele, S., Makris, N., Kennedy, D., Harris, G. J., & Caviness,

V. S. (2003). Dissociations of cerebral cortex, subcortical and cerebral white matter volumes in autistic boys. *Brain*, *126*(5), 1182–1192. https://doi.org/10.1093/brain/awg110

Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2011). MatchIt: Nonparametric Preprocessing for Parametric Causal Inference. *Journal of Statistical Software*, *42*(8), 1–28.

Jong, L. W. de, Vidal, J.-S., Forsberg, L. E., Zijdenbos, A. P., Haight, T., Sigurdsson, S., Gudnason, V., Buchem, M. A. van, & Launer, L. J. (2017). Allometric scaling of brain regions to intra-cranial volume: An epidemiological MRI study. *Human Brain Mapping*, *38*(1), 151–164. https://doi.org/10.1002/hbm.23351

Kassambara, A. (2020). *ggpubr: "ggplot2" Based Publication Ready Plots* (0.4.0) [Computer software]. https://CRAN.R-project.org/package=ggpubr

Lemaitre, H., Goldman, A., Sambataro, F., Verchinski, B., Meyer-Lindenberg, A., Weinberger, D., & Mattay, V. (2012). Normal age-related brain morphometric changes: Nonuniformity across cortical thickness, surface area and grey matter volume? *Neurobiology of Aging*, *33*(3), 617.e1-617.e9. https://doi.org/10.1016/j.neurobiolaging.2010.07.013

Liu, D., Johnson, H. J., Long, J. D., Magnotta, V. A., & Paulsen, J. S. (2014). The power-proportion method for intracranial volume correction in volumetric imaging analysis. *Frontiers in Neuroscience*, *8*. https://doi.org/10.3389/fnins.2014.00356

Lüdecke, D., Bartel, A., Schwemmer, C., Powell, C., Djalovski, A., & Titz, J. (2020). *sjPlot: Data Visualization for Statistics in Social Science* (2.8.6) [Computer software]. https://CRAN.R-project.org/package=sjPlot

Mankiw, C., Park, M. T. M., Reardon, P. K., Fish, A. M., Clasen, L. S., Greenstein, D., Giedd, J.
N., Blumenthal, J. D., Lerch, J. P., Chakravarty, M. M., & Raznahan, A. (2017).
Allometric Analysis Detects Brain Size-Independent Effects of Sex and Sex
Chromosome Complement on Human Cerebellar Organization. *The Journal of
Neuroscience*, *37*(21), 5221–5231. https://doi.org/10.1523/JNEUROSCI.2158-16.2017

Mowinckel, A. M., & Vidal-Piñeiro, D. (2019). Visualisation of Brain Statistics with R-packages
ggseg and ggseg3d. *ArXiv:1912.08200 [Stat]*. http://arxiv.org/abs/1912.08200

Murdoch, D. (2020). *tables: Formula-Driven Table Generation* (0.9.6) [Computer software].
https://CRAN.R-project.org/package=tables

O'Brien, L. M., Ziegler, D. A., Deutsch, C. K., Kennedy, D. N., Goldstein, J. M., Seidman, L. J.,
Hodge, S., Makris, N., Caviness, V., Frazier, J. A., & Herbert, M. R. (2006). Adjustment
for whole brain and cranial size in volumetric brain studies: A review of common
adjustment factors and statistical methods. *Harvard Review of Psychiatry*, *14*(3), 141–
151. https://doi.org/10.1080/10673220600784119

Peyre, H., Mohanpuria, N., Jednoróg, K., Heim, S., Grande, M., van Ermingen-Marbach, M.,
Altarelli, I., Monzalvo, K., Williams, C., Germanaud, D., Toro, R., & Ramus, F. (n.d.).
Neuroanatomy of dyslexia: An allometric approach.

Reardon, P. K., Clasen, L., Giedd, J. N., Blumenthal, J., Lerch, J. P., Chakravarty, M. M., &
Raznahan, A. (2016). An Allometric Analysis of Sex and Sex Chromosome Dosage
Effects on Subcortical Anatomy in Humans. *The Journal of Neuroscience: The Official
Journal of the Society for Neuroscience*, *36*(8), 2438–2448.
https://doi.org/10.1523/JNEUROSCI.3195-15.2016

Ritchie, S. J., Cox, S. R., Shen, X., Lombardo, M. V., Reus, L. M., Alloza, C., Harris, M. A.,

      Alderson, H. L., Hunter, S., Neilson, E., Liewald, D. C. M., Auyeung, B., Whalley, H. C.,

      Lawrie, S. M., Gale, C. R., Bastin, M. E., McIntosh, A. M., & Deary, I. J. (2018). Sex

      Differences in the Adult Human Brain: Evidence from 5216 UK Biobank Participants.

      Cerebral Cortex, *28*(8), 2959–2975. https://doi.org/10.1093/cercor/bhy109

Robinson, D., Hayes [aut, A., cre, Couch, S., Patil, I., Chiu, D., Gomez, M., Demeshev, B.,

      Menne, D., Nutter, B., Johnston, L., Bolker, B., Briatte, F., Arnold, J., Gabry, J., Selzer,

      L., Simpson, G., Preussner, J., Hesselberth, J., … Wiernik, B. M. (2020). *broom: Convert*

      Statistical Objects into Tidy Tibbles (0.7.2) [Computer software]. https://CRAN.R-

      project.org/package=broom

Sanchis-Segura, C., Ibañez-Gual, M. V., Adrián-Ventura, J., Aguirre, N., Gómez-Cruz, Á. J.,

      Avila, C., & Forn, C. (2019). Sex differences in gray matter volume: How many and how

      large are they really? *Biology of Sex Differences*, *10*(1), 32.

      https://doi.org/10.1186/s13293-019-0245-7

Slowikowski, K., Schep, A., Hughes, S., Lukauskas, S., Irisson, J.-O., Kamvar, Z. N., Ryan, T.,

      Christophe, D., Hiroaki, Y., & Gramme, P. (2020). *ggrepel: Automatically Position Non-*

      Overlapping Text Labels with "ggplot2" (0.8.2) [Computer software]. https://CRAN.R-

      project.org/package=ggrepel

Vinke, E. J., Huizinga, W., Bergtholdt, M., Adams, H. H., Steketee, R. M. E., Papma, J. M., de

      Jong, F. J., Niessen, W. J., Ikram, M. A., Wenzel, F., & Vernooij, M. W. (2019).

      Normative brain volumetry derived from different reference populations: Impact on

      single-subject diagnostic assessment in dementia. *Neurobiology of Aging*.

      https://doi.org/10.1016/j.neurobiolaging.2019.07.008

Wei, T., Simko, V., Levy, M., Xie, Y., Jin, Y., & Zemla, J. (2017). *corrplot: Visualization of a Correlation Matrix* (0.84) [Computer software]. https://CRAN.R-project.org/package=corrplot

Wickham, H. (2020). *plyr: Tools for Splitting, Applying and Combining Data* (1.8.6) [Computer software]. https://CRAN.R-project.org/package=plyr

Wickham, H., Chang, W., Henry, L., Pedersen, T. L., Takahashi, K., Wilke, C., Woo, K., Yutani, H., Dunnington, D., & RStudio. (2020). *ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics* (3.3.2) [Computer software]. https://CRAN.R-project.org/package=ggplot2

Wickham, H., François, R., Henry, L., Müller, K., & RStudio. (2020). *dplyr: A Grammar of Data Manipulation* (1.0.2) [Computer software]. https://CRAN.R-project.org/package=dplyr

Wickham, H., & RStudio. (2020). *tidyr: Tidy Messy Data* (1.1.2) [Computer software]. https://CRAN.R-project.org/package=tidyr

Wickham, H., Seidel, D., & RStudio. (2020). *scales: Scale Functions for Visualization* (1.1.1) [Computer software]. https://CRAN.R-project.org/package=scales

Williams, C. M., Peyre, H., Toro, R., Beggiato, A., & Ramus, F. (n.d.). Adjusting for allometric scaling in ABIDE I challenges subcortical volume differences in autism spectrum disorder. *Human Brain Mapping*, *n/a*(n/a). https://doi.org/10.1002/hbm.25145