Sex differences in the adult human brain: Evidence from 5,216 UK Biobank participants

Stuart J. Ritchie\textsuperscript{1,2}, Simon R. Cox\textsuperscript{1,2}, Xueyi Shen\textsuperscript{3}, Michael V. Lombardo\textsuperscript{4,5}, Lianne M. Reus\textsuperscript{6}, Clara Alloza\textsuperscript{7}, Matthew A. Harris\textsuperscript{2,3}, Helen L. Alderson\textsuperscript{7}, Stuart Hunter\textsuperscript{8}, Emma Neilson\textsuperscript{3}, David C. M. Liewald\textsuperscript{1,2}, Bonnie Auyeung\textsuperscript{1}, Heather C. Whalley\textsuperscript{3}, Stephen M. Lawrie\textsuperscript{3}, Catharine R. Gale\textsuperscript{2,9}, Mark E. Bastin\textsuperscript{2,10,11}, Andrew M. McIntosh\textsuperscript{2,3}, Ian J. Deary\textsuperscript{1,2}

\textsuperscript{1}Department of Psychology, The University of Edinburgh, Edinburgh, UK
\textsuperscript{2}Centre for Cognitive Ageing and Cognitive Epidemiology, The University of Edinburgh, Edinburgh, UK
\textsuperscript{3}Division of Psychiatry, The University of Edinburgh, Edinburgh, UK
\textsuperscript{4}Department of Psychology and Center for Applied Neuroscience, University of Cyprus, Nicosia, Cyprus
\textsuperscript{5}Autism Research Centre, Department of Psychiatry, University of Cambridge, Cambridge, UK
\textsuperscript{6}Department of Neurology and Alzheimer Centre, VU University Medical Centre, Amsterdam, The Netherlands
\textsuperscript{7}Department of Psychiatry, Queen Margaret Hospital, Dunfermline, UK
\textsuperscript{8}NHS Lothian, Edinburgh, UK
\textsuperscript{9}MRC Lifecourse Epidemiology Unit, University of Southampton, Southampton, UK
\textsuperscript{10}Brain Research Imaging Centre, The University of Edinburgh, Edinburgh, UK
\textsuperscript{11}Centre for Clinical Brain Sciences, The University of Edinburgh, Edinburgh, UK

Correspondence to: Stuart J. Ritchie, Department of Psychology, The University of Edinburgh, 7 George Square, Edinburgh, EH8 9JZ, UK. Email: stuart.ritchie@ed.ac.uk
Summary

Sex differences in human brain structure and function are of substantial scientific interest because of sex-differential susceptibility to psychiatric disorders [1,2,3] and because of the potential to explain sex differences in psychological traits [4]. Males are known to have larger brain volumes, though the patterns of differences across brain subregions have typically only been examined in small, inconsistent studies [5]. In addition, despite common findings of greater male variability in traits like intelligence [6], personality [7], and physical performance [8], variance differences in the brain have received little attention. Here we report the largest single-sample study of structural and functional sex differences in the human brain to date (2,750 female and 2,466 male participants aged 44-77 years). Males had higher cortical and sub-cortical volumes, cortical surface areas, and white matter diffusion directionality; females had thicker cortices and higher white matter tract complexity. Considerable overlap between the distributions for males and females was common, and subregional differences were smaller after accounting for global differences. There was generally greater male variance across structural measures. The modestly higher male score on two cognitive tests was partly mediated via structural differences. Functional connectome organization showed stronger connectivity for males in unimodal sensorimotor cortices, and stronger connectivity for females in the default mode network. This large-scale characterisation of neurobiological sex differences provides a foundation for attempts to understand the causes of sex differences in brain structure and function, and their associated psychological and psychiatric consequences.
Sex differences have been of enduring biological interest [9]. Nevertheless, in some research fields, such as neuroscience, the potential effects of sex are not always considered [10-12]. A full understanding of how the human sexes differ in their morphological and functional characteristics might assist in explaining their numerous mean-level behavioural differences [4], and might also provide insight into why the prevalence of some psychiatric disorders differs substantially by sex. For example, studies from across the world indicate that rates of Alzheimer’s Disease are higher in females than males, prompting a recent call for the prioritisation of biomedical research into sex differences [13]. In this context, a particularly important target for the investigation of sex differences is the brain.

Broad sex differences in human brain structure are well-known: for instance, males have on average higher total brain volume than females. However, more specific details of where and how the brain differs by sex are unclear, due in part to methodological heterogeneity, small effect sizes scattered throughout the brain [14], and small sample sizes in prior studies [15]. For instance, in the most recent meta-analysis of sex differences in brain structure [5], studies that examined sex differences in specific regions of interest—rather than in broad, overall measures—had a mean sample size of 130 participants (range = 28-465). Here, we set out to characterise sex differences in brain structure and function using data from UK Biobank.

UK Biobank [16] is a large-scale biomedical study in the United Kingdom. A subset of the full sample of 500,000 participants have contributed neuroimaging data [17], and a portion of these data have been released for analysis while collection is ongoing. For the present study, we had access to brain structural MRI data for 2,750 female (mean age = 61.12 years, SD = 7.42, range = 44.64-77.12) and 2,466 male (mean age = 62.39 years, SD = 7.56, range = 44.23-76.99) participants, all scanned on the same scanner (see Supplementary Experimental Procedures). Having data from such a large sample allows us to compare differences in brain morphology across the sexes with high power and reliability. We tested male-female differences (in mean and in variance; the latter having received little attention in this context) in overall and subcortical brain volumes, mapped the magnitude of sex differences across the cortex in terms of volume, surface area, and cortical thickness, and examined sex differences in white matter microstructure. We tested for potential associations of these structural differences with cognitive variation across the sexes. Finally, we examined sex differences in resting-state functional connectivity.

**Sex differences in overall and subcortical brain volumes**

Raw volumetric sex differences are illustrated in the density plots in Figure 1. The male distributions were further to the right, indicating higher means, and wider, indicating greater variance. There was a substantial degree of overlap between the sexes on all measures.

--Insert Figure 1 here--
SEX DIFFERENCES IN THE HUMAN BRAIN

False Discovery Rate correction, all showing larger volume for males. In what follows, negative effect sizes indicate higher values for males, and positive effect sizes indicate higher values for females. The effect sizes ranged from small to large; for example, Cohen’s $d = -0.39$ and $-0.31$ for the left and right nucleus accumbens volume, respectively; and $-1.41$, $-1.28$, and $-1.49$ for total, grey matter, and white matter volumes respectively. The average difference for the fourteen subcortical volumes was $d = -0.70$.

---Insert Table 1 here---

Given the substantial difference in total brain volume, we tested whether sex differences in the subcortical measures were accounted for by this overall size difference. We regressed each subcortical variable on total brain volume, testing the residuals for sex differences. After this adjustment, there were no longer statistically significant differences in the hippocampus, caudate nucleus, or thalamus (all $p_{adj}$-values > 0.07, absolute $d$-values < 0.03; Table S1). There remained differences in each of the other measures, but with attenuated effect sizes (average $d$ for significant differences after adjustment = 0.17). Females had significantly greater right nucleus accumbens volume after adjustment for total brain volume ($d = 0.10$, $p_{adj} = .003$). Overall, the majority of the sex differences in specific subcortical structures appeared to be linked to the total size of the brain (average attenuation of $d$-values for subcortical structures = 85.0%). We also ran analyses adjusting for height, since overall body size may have influenced these differences (as expected, males were substantially taller: $d = -2.15$). This attenuated all of the effect sizes (average $d$-value attenuation across global and subcortical measures = 71.3%), but males still showed significantly larger volumes for all regions except the nucleus accumbens (see Table S1 for subcortical volumes). For example, post-adjustment $d$-values were $-0.42$ for total brain volume, $-0.31$ for grey matter volume, and $-0.47$ for white matter volume.

A novel aspect of the current study was testing male-female differences in variance. A number of studies have found greater male variance in a variety of cognitive and physical measures (e.g. [8,18]), but to our knowledge no studies have explicitly tested for greater male variance in brain imaging measures. As shown in Table 1, there were statistically significant variance differences in all overall cortical and subcortical brain volumes, with males showing greater variance; the average variance ratio for overall volumes and subcortical volumes was 0.82 (variance ratios < 1.00 indicate greater male variance). After adjusting for total brain volume or height, the variance differences reported in Table 1 remained relatively unchanged (see Table S1).

**Sex differences in brain cortical subregions**

Using Freesurfer to parcellate cortical regions according to the Desikan-Killiany neuroanatomical atlas ([19]; regions shown in Figure S2), we tested for sex differences in volume, surface area, and cortical thickness across 68 cortical subregions. As with the analyses above, we adjusted all subregions for age and ethnicity; $p$-values were also adjusted within each measure type using the False Discovery Rate correction. The results are illustrated in Figure 2A (and fully described in Table S2).

---Insert Figure 2 here---

Males showed larger brain volume across all cortical subregions. The sex differences were statistically significant in every subregion, ranging in size from small ($d = -0.24$ in the right
SEX DIFFERENCES IN THE HUMAN BRAIN

temporal pole) to large \((d = -1.03\) in the right insula); the mean \(d\)-value across all subregions was \(-0.67\) \((p_{adj}\)-values < \(9.00\times10^{-13}\)). Even larger differences, all favouring males, were observed for surface area; these ranged from moderate \((d = -0.43\) in the left caudal anterior cingulate) to large \((d = -1.20\) in the left superior frontal region). The mean \(d\)-value across all subregions was \(-0.83\) \((all p_{adj}\)-values < \(2.00\times10^{-36}\)).

Cortical thickness displayed a different pattern. Unlike volume and surface area, females had thicker cortex across almost the entire brain. The only area where males showed a statistically significantly thicker cortex was the right (but not left) insula, and the difference was small \((d = 0.14\). In all other areas, there was either no significant thickness difference \((20/68\) areas), or a statistically significant difference favouring females. The mean \(d\)-value in the 47 areas that reached statistical significance after correction was 0.22, ranging from \(d = 0.07\) in the right rostral middle frontal region to \(d = 0.45\) in the left inferior parietal region. Higher female cortical thickness was generally not found in the temporal lobe (except the parahippocampal gyrus) or in the medial orbitofrontal regions. In some regions there appeared to be converse differences: in the motor and somatosensory regions in the parietal lobe, the frontal pole, and the parahippocampal gyrus, females showed relatively higher thickness but males showed relatively higher volume and surface area. In the superior temporal lobe and orbitofrontal regions, males showed relatively higher volume and surface area, but there was no particular sex difference in thickness.

We next adjusted the subregional cortical measures for overall brain volume. As shown in Figure 2B (and Table S2), 14 regions were still significantly larger in volume for males. The most prominent was the right medial orbitofrontal cortex \((d = -0.92)\). However, the majority of regions \((44/68\) no longer showed significant volume differences \((-0.08 < d < 0.08)\). There were also 10 regions where females now had a significantly larger volume relative to the overall size of the brain, the largest being the left isthmus cingulate \((d = 0.22)\). For surface area, males were larger in 31/68 areas (the largest being in the left isthmus cingulate; \(d = -0.22\)), and females were larger in one (the left caudal anterior cingulate, \(d = 0.11\)). For cortical thickness, after correction for total brain volume there were still significant differences favouring females in 46/68 regions \((0.07 < d < 0.41)\), and no regions favouring males.

Variance differences across the three structural measures are illustrated in Figure 2C. For volume and surface area, males showed significantly greater variance than females across almost all brain regions. The volume variance ratio was significant in 64/68 regions, ranging from 0.88 in the right temporal pole to 0.67 in the left isthmus cingulate, with all \(p_{adj}\)-values < .031 after correction. The surface area variance ratio was significant in 66/68 regions, ranging from 0.88 in the left pars orbitalis to 0.65 in the left isthmus cingulate, all \(p_{adj}\)-values < .018 after correction. For cortical thickness (Figure 4C), there were no significant variance differences in any region \((all p_{adj}\)-values > .14) except one: females showed significantly greater variance in the thickness of the left medial orbitofrontal cortex \((VR = 1.19, p_{adj} = .01)\). As can be observed from Figure S2 (full results in Table S3), controlling for overall brain volume made only a negligible difference to the pattern of variance ratios reported above.

To verify whether the pattern of results across the cortical mantle was agnostic to the gyral boundaries of the Desikan-Killiany atlas, we conducted a supplementary analysis, testing sex differences using a vertex-wise approach, the results of which are shown in Figures S5 (for mean differences) and S6 (for variance differences). This very precisely replicated the subregional atlas-based results.
SEX DIFFERENCES IN THE HUMAN BRAIN

Sex differences in white matter microstructure

We tested sex differences in 22 white matter tracts. We focused on two white matter microstructural properties that had previously been shown to demonstrate differences between males and females in the initial release of UK Biobank imaging data [20]. The first was fractional anisotropy (FA), an index of the directionality of water diffusion through the white matter. The second was orientation dispersion (OD), a measure of white matter tract complexity. For FA, there were generally higher values in males, particularly in the cortico-spinal tract ($d = -0.54$) and the acoustic radiation ($d = -0.51$). The average difference across tracts was $d = -0.19$. OD was higher in all tracts for females (average $d = 0.30$). These mean differences are shown in Figure 3, and fully reported in Tables S5 and S6.

--Insert Figure 3 here--

Variance differences are illustrated in Figure S4 (see also Tables S5 and S6). Generally, there was greater male variance in FA (average VR = 0.92); however, there was substantially greater female variance in the cortico-spinal tract in particular (VR = 1.17, $p = .0003$). For OD, the only tract that showed a significant variance difference following FDR correction was the left superior thalamic radiation, where males showed greater variance (VR = 0.79).

Relation of neurostructural differences to cognitive differences

We linked the structural brain differences to scores on two cognitive tests taken at the time of the imaging visit: verbal-numerical reasoning and reaction time (see Supplementary Experimental Procedures). Descriptive statistics for the cognitive tests are shown in Table 1. Note that we coded (reflected) both tests so that higher scores indicated better performance. The test scores correlated positively, but weakly ($r = .12$). Males had a slightly higher mean score than females on verbal-numerical reasoning ($d = -0.18$) and slightly faster mean reaction time ($d = -0.22$); there was no significant variance difference for verbal-numerical reasoning (VR = 0.97, $p = .45$), though males had marginally more variance in reaction time (VR = 0.92, $p = .03$).

We tested the extent to which the mean cognitive differences were mediated by any of the overall brain measures (total, grey, and white matter volumes, total surface area, mean cortical thickness, or general factors of FA or OD; Figure S8). We ran a separate model for each brain measure. To assess the replicability of the results and avoid overfitting, these analyses were performed in two randomly-selected halves of the sample. Full results are displayed in Tables S7 and S8 for verbal-numerical reasoning and reaction time, respectively (Tables S9 and S10 contain correlation matrices for each of the variables in this analysis). For verbal-numerical reasoning, the sex difference in test scores was mediated substantially by brain volume measures and by surface area (all mediation percentages >82%). Cortical thickness showed far smaller mediation percentages (7.1% and 5.4% in the two sample halves, respectively). For reaction time, total brain and white matter volumes had mediation percentages >27%, but the other measures all produced smaller percentages (<15.3%), particularly mean cortical thickness (mediating <3% of the variance).

Finally, we correlated performance on the two cognitive tests with the volume, surface area, and thickness of each brain region in males and females separately. These correlations were generally small, with all $r$-values <.17 (Table S11; see [17] for discussion of the associations...
in the full sample). We then compared the size of the correlations across the sexes. After multiple comparisons correction, there were no significant sex differences in these correlations. Thus, there was no evidence for differences in how regional brain structure related to cognitive skills between the sexes.

**Sex differences in resting-state functional connectivity**

In a final set of analyses, we examined sex differences in resting-state functional MRI (rsfMRI) responses within a number of functional networks. We found that 54.7% (811 of 1,485) network connections showed a sex difference (absolute β-values = 0.071 to 0.447 for females; 0.071 to 0.519 for males). A connection map, showing connections between 55 network nodes, is shown in Figure 4A (see also Table S12). Connectivity between sensorimotor, visual, and rostral lateral prefrontal areas was stronger in males than females (see orange cluster of brain regions in Figure 4A), whereas connectivity within the default mode network (DMN; red cluster of regions in Figure 4A) was stronger in females than males.

--Insert Figure 4 here--

To further visualize these sex differences, we calculated the mean strength of all 54 connections to each node, producing a weighted degree statistic. Sex differences in weighted degree are shown in Figure 4B and 4C. A prominent male>female association was apparent in bilateral sensorimotor areas, the visual cortex, and the rostral lateral prefrontal cortex. A female>male association was apparent in cortical areas comprising the DMN: the bilateral posterior cingulate cortex/precuneus, the dorsal anterior cingulate cortex, medial prefrontal cortex, temporo-parietal junction, anterior temporal lobe, medial temporal lobe (e.g. hippocampus and surrounding areas), and some cerebellar regions. A full list of regions is provided in Table S12.
Discussion

In a single-scanner sample of over 5,000 participants from UK Biobank, we mapped the brain regions in which sex differences in volume, surface area, cortical thickness, diffusion parameters, and functional connectivity were found. One main theme of the neurostructural results was that associations with sex were global. Males generally had larger volumes and surface areas, whereas females had thicker cortices. The differences were substantial: in some cases, such as total brain volume, more than a standard deviation. We also found that volume and surface area mediated nearly all of the small sex difference in reasoning ability, but far less of the difference in reaction time. For white matter microstructure, females showed lower directionality (FA) and higher tract complexity (OD); white matter microstructure was a poor mediator of the cognitive sex difference. Resting-state fMRI analyses also revealed a global effect: around 54% of connections showed a sex difference. These differences clustered around specific networks, with stronger connectivity in females in the default mode network and stronger connectivity in males between unimodal sensory and motor cortices as well as high-level cortical areas in the rostral lateral prefrontal cortex. Overall, for every brain region that showed even large sex differences, there was always overlap between males and females, confirming that the human brain cannot—at least for the measures observed here—be described as “sexually dimorphic” (see Figure 1 and [21]).

The principal strengths of the present study are its sample size (providing sensitivity for the identification of small effects with high statistical power), the wide range of MRI modalities, and the consideration of both mean and variance differences. Given the surfeit of small-n studies in neuroscience [15], it is of great importance to test hypotheses in large, well-powered samples, especially given it is possible that many neural sex differences are small [14]. Here, we had excellent statistical power to find even small effects in brain subregions, thus making the current work a robust, definitive, and detailed analysis of sex differences across multiple brain measures. For our subregional analysis, we had a far larger sample size than the most recent meta-analysis [5]. In contrast to that meta-analysis—which found greater volume for females in areas such as the thalamus, the anterior cingulate gyrus, and the lateral occipital cortex—our study found no brain subregions where females had a larger volume than males. The reason for this may be the more restricted age range of the participants in our study (sex may have different effects at different ages), or, more likely, study size: the data for that part of the meta-analysis came from many separate studies, on separate scanners, with small sample sizes (many with \( n < 100 \)), whereas our contrasts were based on a very large, single-scanner study.

The higher male volume in our study appeared largest in some regions involved in emotion and decision-making, such as the bilateral orbitofrontal cortex, the bilateral insula, and the left isthmus of the cingulate gyrus [22-25], but also areas such as the right fusiform gyrus. For surface area, which showed an even larger difference favouring males, the regions that showed the largest effects were broadly areas involved in the hypothesized intelligence-related circuit in the “P-FIT” model [26]: for example, the bilateral superior frontal gyri, the bilateral precentral gyri, the left supramarginal gyrus, and the bilateral rostral middle frontal areas. However, some of the regions involved in this theorized circuit were also larger, in terms of thickness, for females. For instance, the bilateral inferior parietal regions were the regions with numerically the largest difference favouring females in cortical thickness. Our finding that the cortex was thicker for females is consistent with a number of previous, smaller studies (e.g. [27-29]), though our greater statistical power allowed us to find smaller differences in thickness across the cortex.
SEX DIFFERENCES IN THE HUMAN BRAIN

A recent meta-analysis of sex differences in amygdala volume [32] found that, although a difference favouring males was found in the raw volumes, after correction for total brain volume there was no longer an appreciable sex distinction. However, whereas we found this same pattern for the thalamus, the caudate, and the hippocampus, it was not the case for the amygdala (or the accumbens, pallidum, or putamen): this structure was significantly, but modestly, larger ($d = 0.18$ for both hemispheres) in males even after adjusting for total brain volume. The heterogeneity in the methods of the studies being meta-analysed may have led to the divergent conclusion from our single-sample study. With regard to the hippocampus, however, we found results consistent with another recent meta-analysis [33]: there were no longer significant sex differences after adjustment for total brain volume.

Whereas previous work has found some white matter regions where fractional anisotropy was higher for females [30], we found that males showed higher FA in 18 of the 22 tracts we examined. FA also generally showed greater variance in males. On the other hand, higher orientation dispersion was found for females in all tracts. Unexpectedly, higher OD was found to be related to lower cognitive performance on the two tests examined here. Since OD is a relatively new measure of white matter microstructure [31], further work should aim to clarify its behavioural correlates.

The issue of adjusting for overall brain size in analyses of sex differences (e.g. [34]) was addressed in each of our macrostructural analyses. As can be seen comparing Figures 2 and 3, after this adjustment, the higher male volume and surface area was substantially reduced, often to non-significance. For those latter brain regions, this implies that the sex difference was general: their larger volume or surface area was a by-product of the overall larger male brain. However, for some regions, especially for surface area (particularly in areas such as the left isthmus of the cingulate gyrus and the right precentral gyrus), males still showed a significantly higher measurement, indicating specific sex differences in the proportional configuration of the cortex, holding brain size equal. Most interestingly, for some areas (for example the right insula, the right fusiform gyrus, and the left isthmus of the cingulate gyrus), the difference was reversed, with females showing significantly larger brain volume. We recommend that future studies perform comparisons both before and after adjusting for total volume, since these results pertain to quite different questions.

One question that could not be addressed using the current data regards the underlying biobehavioural causes, ultimate or proximate, for the sex differences that we observed. Sex differences in brain structure are observed early in the life course (e.g. [35]), though this does not imply that the pattern of adult differences we observed is necessarily the same as is found in childhood. The literature on sex differentiation of the brain during development highlights influences of factors such as sex hormones that could not be analysed in the present study (e.g. [36, 37]). Likewise, understanding the potential neurobiological effects of social influences during development [38] was beyond the scope of the present study.

Our analysis also focused on sex differences in variability. The best-studied human phenotype in this context has been cognitive ability: almost universally, studies have found that males show greater variance in this trait ([6,18,39], though see [40]). This has also been found to be the case for academic achievement test results (a potential consequence of intelligence differences [8,41,42]), other psychological characteristics such as personality [7], and a range of physical traits such as athletic performance [43], and both birth and adult weight [8]. Here, for the first time, we directly tested sex differences in the variance of
several brain measures, finding greater male variance across almost the entire brain for volume, surface area, and white matter fractional anisotropy, but only patchy and inconsistent variance differences for cortical thickness and white matter orientation dispersion.

One potential candidate to explain greater male variability across multiple phenotypes is the hypothesized ‘female-protective’ mechanism involving effects of the X chromosome [44,45], or other protective factors that “buffer” females from potential deleterious consequences of rare genetic mutations. For instance, if deleterious genetic variants are found on one X chromosome in (heterozygous) females, they may be buffered by the presence of the opposite allele on the other X chromosome. Since males carry only one X chromosome, this effect cannot occur, increasing the likelihood of the allele being expressed in males, and thus increasing the variation in the phenotype linked to that allele [44,46]. In sex-biased phenotypes like autism (ASD), female protective effects are also frequently discussed. It is known that ASD females typically require a higher burden of rare, deleterious de novo mutations compared to males with ASD [47], and this effect extends into the general population when examining autistic traits in typically-developing individuals [48]. It is possible that higher male variability could be linked to genetic mechanisms that inherently buffer females against deleterious genetic influences, but may have much a more variable and significant effect on average in males. As studies like UK Biobank release even larger amounts of data on individuals who have both neurostructural and genotype data, researchers will be able to perform well-powered tests of these hypotheses.

Using the (limited) data on cognitive abilities available in our sample, we tested whether the data were consistent with any consequences of brain structural differences in terms of ability differences. There were very small correlations between brain variables and the cognitive tests, and these associations did not differ by sex (consistent with a prior meta-analysis on the link between brain volume and intelligence [49]). Mediation modelling suggested that, for verbal-numerical reasoning, a very large portion (up to 99%) of the modest sex difference was mediated by brain volumetric and surface area measures. Smaller fractions (up to 38%) of the modest link between sex and reaction time could be explained by volume or surface area. Perhaps unexpectedly, given evidence and theory linking white matter microstructure to cognitive processing speed [50,51], white matter microstructural measures only mediated a small proportion of the sex difference in reaction time (this may have been due to weaknesses in this cognitive measure; see below). Cortical thickness had trivial mediating effects compared to volume and surface area: no more than 7.1% of the sex-cognitive relation was mediated by thickness in any analysis. Thus, the data are consistent with higher volume and cortical surface area (but not cortical thickness or microstructural characteristics) being of particular relevance to sex differences in reasoning abilities (but not particularly reaction time).

Sex differences in intrinsic functional connectome organization also revealed results that corroborate and extend prior work. Notably, the original study of the 1,000 Functional Connectomes dataset reported sex differences similar to those we identified – that is, Female>Male connectivity within the default mode network and some evidence for a Male>Female effect in sensorimotor and visual cortices [52]. The higher female connectivity within circuits like the DMN may be particularly important, given that DMN regions are typically considered as the core part of the ‘social brain’ [53]. Whether such an effect can help explain higher average female ability in domains like social cognition [54], and whether such functional differences can be integrated with differences in the structural connectome [55], remains to be seen, and are avenues for future work. Finally, it is noteworthy that recent
SEX DIFFERENCES IN THE HUMAN BRAIN

work has shown that intrinsic functional connectome organization can be parsimoniously described as a small number of connectivity gradients [56]. The most prominent connectivity gradient has at one pole the DMN and at the other pole has unimodal sensory and motor cortices. Intriguingly, the observed pattern of sex differences in functional connectome organization observed here recapitulates these two main poles of that principal connectivity gradient [56]; see Figure S10. Thus, one interesting and parsimonious way of describing the biological significance of these functional sex differences is that mechanisms involved in shaping these sex differences (e.g., biological, cultural, developmental) may be influencing this main principal connectivity gradient in the human brain; the end result may be the sex differences we discovered across multiple networks.

Limitations

Whereas the sample was large, it was also selective. It covered only one part of the life course (from approximately age 45 to age 75 years). In addition, UK Biobank had a very low response rate to invitations to participate (5.47% in the full sample of ~500,000; [16]). We would thus expect the individuals studied here would not be fully representative of males and females from the general UK population. This was the case for education: individuals with college or university degrees were over-represented (see Supplementary Experimental Procedures), though the male:female education ratio itself appeared representative.

Caution should be taken in interpreting the results of the analyses involving the cognitive tests. Whereas previous, representative studies (e.g. [6]) have found no mean difference, but a variance difference, in cognitive test performance, the tests examined here showed mean differences but no strong variance differences. This may be due to problems of sample representativeness (see [57]), or due to the tests tapping specific cognitive skills rather than general ability [58]. The cognitive measures were generally psychometrically poor, the verbal-numerical reasoning having only 13 items, and the reaction time test having only 4 trials that counted towards the final score (see [59] for analyses of their reliability). Although the tests—particularly verbal-numerical reasoning—have some external validity [60], the above issues mean that the cognitive analyses reported here should be considered preliminary. Fuller cognitive testing, currently underway in UK Biobank, will allow a more comprehensive exploration. It should also be noted that cross-sectional mediation models of observational data, such as those used here, are inherently limited: they cannot address causal relations between variables. The models were very simple, including only three main variables (sex, the brain measure, and cognitive ability; Figure S3). More complex model specifications, using longitudinal data, and preferably using latent cognitive variables (derived from a battery of tests), will be required fully to understand sex differentiation in both the brain and cognitive abilities.

Conclusions

The present study is the largest single-sample study of neuroanatomical sex differences to date. We report evidence on the pattern of sex differences in brain volume, surface area, cortical thickness, white matter microstructure, and functional connections between adult males and females in the range between middle- and older-age. As has previously been argued [61], providing a clear characterisation of neurobiological sex differences is a step towards understanding patterns of differential susceptibility to neurodevelopmental disorders such as autism spectrum disorder [1], a variety of psychiatric conditions [2], and neurodegenerative disorders such as Alzheimer’s Disease [13,62].
SEX DIFFERENCES IN THE HUMAN BRAIN

of further MRI scans (to a maximum sample of 100,000 with MRI data) will be available for UK Biobank in the coming months and years, in addition to more complex cognitive testing batteries and genotypic data. Future studies will thus be able to explore in much greater depth the links between sex differences in the brain, their potential causes, and their medical and behavioural consequences.
Acknowledgements

We are grateful to the UK Biobank members for their participation in the study, and to the UK Biobank team, who collected, processed, and made available the data for analysis. This work was carried out under UK Biobank application number 10279. The work was primarily carried out in The University of Edinburgh Centre for Cognitive Ageing and Cognitive Epidemiology (CCACE), is part of the UK Research Councils’ cross-council Lifelong Health and Wellbeing initiative (MR/K028992/1). Additional funding has been received from the Biotechnology and Biological Sciences Research Council (BBSRC), the Medical Research Council (MRC), and Age UK’s Disconnected Mind project. Authors S.J.R., S.R.C., D.C.M.L., and I.J.D. were supported by the MRC award to CCACE (MR/K026992/1). Author S.R.C. was supported by MRC grant MR/M013111/1. Author I.J.D. is additionally supported by the Dementias Platform UK (MR/L015382/1). The work was also supported by a Wellcome Trust Strategic Award, “Stratifying Resilience and Depression Longitudinally” (STRADL; reference 104036/Z/14/Z). Author X.S. receives support from the China Scholarship Council. Author L.M.S. was supported by an Erasmus Traineeship Grant. Author H.C.W. is supported by a JMAS SIM fellowship from the Royal College of Physicians of Edinburgh, and by an ESAT College Fellowship from the University of Edinburgh. Authors A.M.M., H.C.W., and S.M.L. gratefully acknowledge the support of the Sackler Foundation. Author A.M.M. has previously received grant support from Pfizer, Lilly, and Janssen. We are grateful to Anne Scheel for helpful comments on a previous draft.
SEX DIFFERENCES IN THE HUMAN BRAIN

References

system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. NeuroImage 31, 968-980.


Table 1. Descriptive statistics with mean and variance comparisons for overall volumes, subcortical volumes, and cognitive tests.

<table>
<thead>
<tr>
<th>Measure type</th>
<th>Measure</th>
<th>Female (n = 2,750)</th>
<th>Male (n = 2,466)</th>
<th>Mean difference test</th>
<th>Variance Ratio test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M (SD)</td>
<td>M (SD)</td>
<td>t</td>
<td>p</td>
<td>d</td>
</tr>
<tr>
<td>Overall volumes (cm³)</td>
<td>Total brain volume</td>
<td>1115.76 (89.68)</td>
<td>1233.58 (98.31)</td>
<td>-48.91 ~0.00</td>
<td>-1.41</td>
</tr>
<tr>
<td></td>
<td>Grey matter volume</td>
<td>597.02 (47.78)</td>
<td>643.45 (52.08)</td>
<td>-38.97 1.75×10⁻²⁸⁷</td>
<td>-1.28</td>
</tr>
<tr>
<td></td>
<td>White matter volume</td>
<td>518.85 (47.89)</td>
<td>589.59 (52.69)</td>
<td>-51.53 ~0.00</td>
<td>-1.49</td>
</tr>
<tr>
<td>Subcortical volumes (cm³)</td>
<td>Left hippocampus</td>
<td>3.73 (0.42)</td>
<td>3.94 (0.46)</td>
<td>-18.91 2.69×10⁻⁷⁶</td>
<td>-0.55</td>
</tr>
<tr>
<td></td>
<td>Right hippocampus</td>
<td>3.82 (0.42)</td>
<td>4.04 (0.48)</td>
<td>-18.43 1.16×10⁻⁷²</td>
<td>-0.54</td>
</tr>
<tr>
<td></td>
<td>Left accumbens</td>
<td>0.49 (0.11)</td>
<td>0.53 (0.12)</td>
<td>-13.42 5.19×10⁻³⁹</td>
<td>-0.39</td>
</tr>
<tr>
<td></td>
<td>Right accumbens</td>
<td>0.40 (0.10)</td>
<td>0.42 (0.11)</td>
<td>-10.64 3.82×10⁻²⁶</td>
<td>-0.31</td>
</tr>
<tr>
<td></td>
<td>Left amygdala</td>
<td>1.21 (0.22)</td>
<td>1.35 (0.25)</td>
<td>-20.04 5.23×10⁻⁸⁵</td>
<td>-0.59</td>
</tr>
<tr>
<td></td>
<td>Right amygdala</td>
<td>1.18 (0.24)</td>
<td>1.31 (0.27)</td>
<td>-17.55 2.16×10⁻⁶⁶</td>
<td>-0.51</td>
</tr>
<tr>
<td></td>
<td>Left caudate</td>
<td>3.28 (0.38)</td>
<td>3.54 (0.41)</td>
<td>-23.00 3.04×10⁻¹⁰</td>
<td>-0.66</td>
</tr>
<tr>
<td></td>
<td>Right caudate</td>
<td>3.45 (0.40)</td>
<td>3.72 (0.44)</td>
<td>-22.67 2.37×10⁻⁷⁰</td>
<td>-0.65</td>
</tr>
<tr>
<td></td>
<td>Left pallidum</td>
<td>1.69 (0.21)</td>
<td>1.85 (0.22)</td>
<td>-26.64 4.87×10⁻¹⁴⁵</td>
<td>-0.77</td>
</tr>
<tr>
<td></td>
<td>Right pallidum</td>
<td>1.74 (0.20)</td>
<td>1.89 (0.22)</td>
<td>-26.96 3.82×10⁻¹³⁸</td>
<td>-0.78</td>
</tr>
<tr>
<td></td>
<td>Left putamen</td>
<td>4.61 (0.50)</td>
<td>5.07 (0.56)</td>
<td>-34.72 1.73×10⁻²³⁴</td>
<td>-1.01</td>
</tr>
<tr>
<td></td>
<td>Right putamen</td>
<td>4.64 (0.49)</td>
<td>5.13 (0.55)</td>
<td>-37.13 4.76×10⁻²⁶⁴</td>
<td>-1.08</td>
</tr>
<tr>
<td></td>
<td>Left thalamus</td>
<td>7.54 (0.64)</td>
<td>8.11 (0.72)</td>
<td>-33.73 7.76×10⁻²²³</td>
<td>-0.98</td>
</tr>
<tr>
<td></td>
<td>Right thalamus</td>
<td>7.34 (0.62)</td>
<td>7.92 (0.69)</td>
<td>-35.76 2.42×10⁻²⁴⁷</td>
<td>-1.03</td>
</tr>
<tr>
<td>Cognitive tests</td>
<td>Verbal-numerical reasoning (max. score 13)</td>
<td>6.80 (2.10)</td>
<td>7.14 (2.13)</td>
<td>-6.21 5.77×10⁻⁷⁰</td>
<td>-0.18</td>
</tr>
<tr>
<td></td>
<td>Reaction time (ms)</td>
<td>590.37 (98.04)</td>
<td>574.71 (100.71)</td>
<td>-7.63 2.71×10⁻³⁴</td>
<td>-0.21</td>
</tr>
</tbody>
</table>

Note: Means and SDs are shown prior to adjustment for age and ethnicity; statistical tests are performed after this adjustment. Reaction Time is shown here in raw millisecond units, but was reverse-scored for analysis so that higher scores indicated better performance. Negative t- and d-values mean higher male means. VR = Variance ratio (values < 1 indicate greater male variance). p-values for brain variables corrected for multiple comparisons using the False Discovery Rate correction. † = sex difference in subcortical region still significant after adjustment for total brain volume (see Table S1).
Figure Legends

Figure 1. Density plots of sex differences in overall brain volumes (left section) and subcortical structures (right section). $d =$ Cohen’s $d$ (mean difference); VR = Variance Ratio (variance difference). All mean differences were statistically significant at $p < 3.0 \times 10^{-25}$, all variance differences were significant at $p < .003$, after correction for multiple comparisons (see Table 1).

Figure 2. Sex differences across the subregions in volume, surface area, and cortical thickness. Shown are A) mean differences, B) mean differences adjusted for total brain volume, and C) variance differences. Variance differences corrected for total brain volume were near-identical to those shown in C); see Figure S4. See Figure S2 for subregional atlas.

Figure 3. Mean sex differences in white matter microstructural measures A) fractional anisotropy and B) orientation dispersion across 22 white matter tracts. For both measures, numerically the largest effect was found in the right cortico-spinal tract. See Figure S3 for tract atlas.

Figure 4. Resting-state fMRI results. and weighted degree of nodes. A) Spatial maps for individual connections. Colours represent the effect sizes of sex on the strength of connections (red = female stronger; blue = male stronger). For the purpose of illustration, only effect sizes larger than $\pm 0.3$ are shown. B) and C) Weighted degrees of nodes with higher values in males and females, respectively. The spatial maps of significant group-ICA nodes were multiplied by the size of the sex correlation.