Left-right asymmetry of the human brain: associations with neurodevelopmental disorders and genetic factors

The educational component of the doctoral training was provided by the International Max Planck Research School (IMPRS) for Language Sciences. The graduate school is a joint initiative between the Max Planck Institute for Psycholinguistics and two partner institutes at Radboud University – the Centre for Language Studies, and the Donders Institute for Brain, Cognition and Behaviour. The IMPRS curriculum, which is funded by the Max Planck Society for the Advancement of Science, ensures that each member receives interdisciplinary training in the language sciences and develops a well-rounded skill set in preparation for fulfilling careers in academia and beyond.

More information can be found at www.mpi.nl/imprs.

ISBN: 978-94-92910-29-5

Cover design and layout by Merel Charlotte Postema.
Printed and bound by Ipskamp Drukkers b.v.

The research reported in this thesis was conducted at the Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands.

Left-right asymmetry of the human brain: associations with neurodevelopmental disorders and genetic factors

Proefschrift

ter verkrijging van de graad van doctor

aan de Radboud Universiteit Nijmegen

op gezag van de rector magnificus prof. dr. J.H.H.M. van Krieken,

volgens besluit van het college van de decanen

in het openbaar de verdedigen op woensdag 7 juli 2021

om 12.30 uur precies

door

Merel Charlotte Postema

geboren op 30 oktober 1991

te Soest

Promotor

Prof. dr. S.E. Fisher

Copromotor

Dr. C. Francks (MPI)

Manuscriptcommissie:

Prof. dr. G.S.E. Fernandez

Prof. dr. I.C. McManus (University College London, Verenigd Koninkrijk)

Dr. E. Sprooten

Left-right asymmetry of the human brain: associations with neurodevelopmental disorders and genetic factors

Doctoral Thesis

to obtain the degree of doctor

from Radboud University Nijmegen

on the authority of the Rector Magnificus prof. dr. J.H.H.M. van Krieken,

according to the decision of the Council of Deans

to be defended in public on Wednesday, July 7, 2021

at 12.30 hours

by

Merel Charlotte Postema

born on October 30, 1991

in Soest (The Netherlands)

Supervisor

Prof. dr. S.E. Fisher

Co-supervisor

Dr. C. Francks (MPI)

Doctoral Thesis Committee

Prof. dr. G.S.E. Fernandez

Prof. dr. I.C. McManus (University College London, United Kingdom)

Dr. E. Sprooten

Voor mijn lieve moekie, die altijd in mij geloofde.

Table of contents

CHAPTER

General introduction 1

1.1 Left-right asymmetries of the human brain

1.1.2 Brief history of human brain asymmetry research

Around 160 years ago, one of the first discoveries in the field of human brain asymmetry was made by Paul Broca, a French anatomist, who found that left-sided lesions in the posterior inferior frontal gyrus resulted in problems with speech production (Broca, 1863; Renteria, 2012). Soon thereafter, Karl Wernicke, a German physician, reported that injury of the posterior superior temporal gyrus in the left hemisphere resulted in loss of language comprehension (Renteria, 2012; Wernicke, 1874). This pioneering work gave rise to the idea that the left hemisphere was dominant over the right; for many decades the left hemisphere was termed the 'major hemisphere' and the right was termed the 'minor hemisphere' (Corballis & Haberling, 2017). Early evidence for right-hemispheric dominance for non-verbal functions, such as emotion (Luys, 1881) was reported, but it was not until the 1960's that consensus views on human brain asymmetry changed (Bogen, 1969; Corballis & Haberling, 2017; Gazzaniga, Bogen, & Sperry, 1962; Ornstein, 1972).

Around that time, a patient with epilepsy underwent a split-brain surgery (Gazzaniga et al., 1962), which involved cutting the corpus callosum (the major white matter tract connecting the left and right sides of the brain), to halt the epilepsy from spreading between the two hemispheres (Staub, 2016). The surgery indeed stopped seizure activity, but also resulted in impairments in functions that require interhemispheric communications, such as visual

perception. Yet certain functions remained intact, and for that reason, it was concluded that the two hemispheres were independent in terms of these highly lateralized processes (Staub, 2016). Findings like these supported a shift in view from major and minor hemispheres to 'propositional' and 'appositional' hemispheres, which relate to dichotomous concepts, like *Yin* and *Yan* (Corballis & Haberling, 2017).

This idea still holds in the public mind, where the left hemisphere is typically associated with verbal communication, logical, and analytical thinking, while the right is associated with non-verbal communication, emotion, and creativity (Corballis & Haberling, 2017). While some component processes of these functions are indeed typically lateralized to either left or right hemisphere, a dualism between the two sides of the brain is overstated in popular discourse, and may reflect the human disposition for categorization when faced with complex and continuous features (Corballis & Haberling, 2017). In fact, functional activation is often bilateral to some degree, especially in situations that are cognitively challenging, such as when performing certain working memory tasks (Holler-Wallscheid, Thier, Pomper, & Lindner, 2017). The involvement of the non-dominant, contralateral hemisphere for a given task reflects deep homology between the two hemispheres, while each is fine-tuned to preferentially control different types of function (Holler-Wallscheid et al., 2017).

1.2.3 Structural and functional human brain asymmetries

At the macroscopic level, prominent structural asymmetries of the human brain include the Yakovlevian torque and Petalia, which involve a horizontal skewing of the hemispheres (***Fig.1***) (X.-Z. Kong et al., 2020) such that the right frontal lobe protrudes forwards and the left occipital lobe protrudes backwards (Toga & Thompson, 2003). A similar skew is also evident in the vertical plane (***Fig.1***). Numerous structural asymmetries can also be found in the volumes, shapes, surface areas or thicknesses of homotopic regions across the two hemispheres (X. Z. Kong et al., 2020; Renteria, 2012). For example, the superior temporal sulcus shows a marked asymmetry of depth, typically being deeper on the right side than on the left, which may relate to functional language dominance (Leroy et al., 2015). Nearby regions of auditory cortex

(Heschl's gyrus and *planum temporale*) are also typically asymmetric structures (Le Guen et al., 2018).

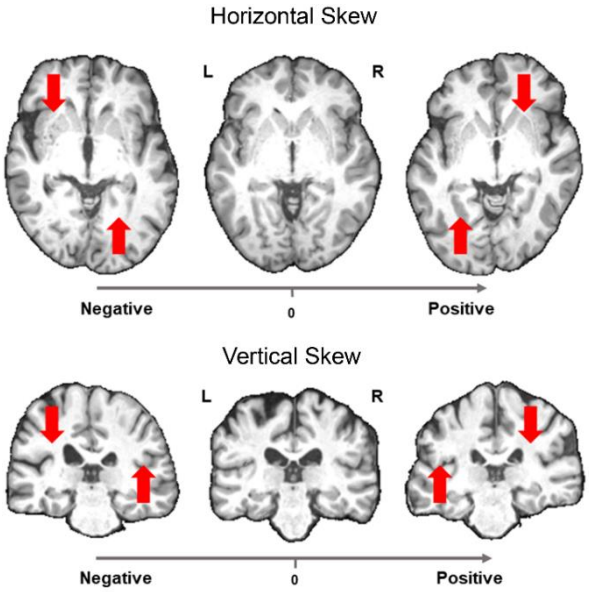


***Fig.1**. Human brain asymmetrical skews. Figure adapted from (X.-Z. Kong et al., 2020).*

Two well-known functional asymmetries include left-hemispheric dominance for language (as mentioned further above) (Josse & Tzourio-Mazoyer, 2004) and handedness (Dassonville, Zhu, Uurbil, Kim, & Ashe, 1997; Peters, Reimers, & Manning, 2006). About 90% of the human population is right-handed (Peters et al., 2006), due to left-hemispheric dominance for hand motor control (Dassonville et al., 1997). Other typically lateralized functions are, to name a few, verbal working memory (left dominant) (Langel, Hakun, Zhu, & Ravizza, 2014), visuospatial attention (right dominant) (Thiebaut de Schotten et al., 2011), and aspects of emotional processing (right dominant) (Wyczesany, Capotosto, Zappasodi, & Prete, 2018). While both hemispheres are needed for optimal functioning, it has been proposed that right hemisphere-dominant processes more often interact with the contralateral hemisphere than left hemisphere-dominant processes, particularly considering language and motor control (Gotts et al., 2013).

Left-right differences in functional activity patterns between the human cerebral hemispheres, e.g. related to language processing, have been revealed using diverse imaging modalities, including hemodynamic-(Holland et al., 2007), electrophysiological- (Youssofzadeh & Babajani-Feremi, 2019) and magnetoencephalographic imaging (Sha et al., 2020), both during task performance (Hugdahl & Westerhausen, 2016), as well as during the resting state (Labache et al., 2019). In the latter case, left-right differences in functional activity are usually measured as a degree of similarity between functional connectivity patterns (Raemaekers, Schellekens, Petridou, & Ramsey, 2018). Other approaches that can be used to determine language dominance are the Wada test (Kundu, Rolston, & Grandhi, 2019) and dichotic listening (Westerhausen & Kompus, 2018). The Wada test is mostly used in studies with epileptic patients, because it is an invasive method that involves alternately injecting a barbiturate drug into the left and right internal carotid arteries to achieve hemisphere-limited inhibition (Kundu et al., 2019). Speech will then temporarily stop whenever the dominant hemisphere is exposed to the drug. In contrast, dichotic listening is a non-invasive method that is widely used in brain asymmetry studies to determine hemispheric dominance for the auditory processing of spoken syllables (Westerhausen & Kompus, 2018), based on the fact that individuals perceive these stimuli primarily with the ear contralateral to the dominant hemisphere.

Brain asymmetries have also been studied at other levels of investigation (Renteria, 2012), including the histological level, where inter-hemispheric asymmetries can be observed in the degree of dendritic or axonal arborization of neurons in some cortical regions (Sa, Ruela, & Madeira, 2007), and the genetic level, where left-right differences in cortical gene expression have been observed, relating especially to synaptic transmission (Karlebach & Francks, 2015). These molecular and cellular asymmetries seem likely to relate to neurophysiological differences between homotopic regions across the two hemispheres, which help to fine tune neural circuitry differently for distinct functional dominances.

1.2 An evolutionary perspective on left-right asymmetries

Left-right asymmetries can be found in all sorts of organisms, including vertebrates and invertebrates. For example, plaice (bottom-dwelling flatfish) have both eyes located on their side facing upwards, male fiddler crabs have one major claw, and different snail species can show either predominantly clockwise or anti-clockwise coiled shells. However, not all left-right asymmetries are outwardly visible. For instance, at first sight, vertebrates appear to be left-right symmetrical. Yet internally there is a strikingly asymmetrical pattern of the visceral organs (heart, liver etc.), which is a conserved characteristic of the vertebrate body plan (Matsui & Bessho, 2012).

In terms of brain and behaviour, examples of laterality include honeybees being better at responding to odors learnt through their right antenna compared to their left antenna (Frasnelli & Vallortigara, 2018; Letzkus et al., 2006), eye dominance which relates to tool construction and use in New Caledonian crows (Martinho, Burns, von Bayern, & Kacelnik, 2014), and left-right biased responses to predators in fish and other species (Vallortigara & Rogers, 2005). Additionally, population-level lateralization for handedness has been observed in certain great apes (albeit to a lesser degree than humans), namely chimpanzees and bonobos (Hopkins, 2006) as well as a number of monkey species (Hopkins, Misiura, Pope, & Latash, 2015). Toads also have lateralized motor programs on the population level (Bisazza, Cantalupo, Robins, Rogers, & Vallortigara, 1997).

In fact, brain and behavioural asymmetries are present in many (if not all) vertebrates, as well as various invertebrates (Frasnelli, 2013), reflecting a highly conserved and hence fundamental feature of brain organization (L. Rogers & Vallortigara, 2015). Brain asymmetries may improve cognitive performance by enhancing efficiency through the availability of parallel processing (Duboc, Dufourcq, Blader, & Roussigne, 2015). Parallel processing allows division of labour which can increase the brain's capacity for multi-tasking (L. J. Rogers, Zucca, & Vallortigara, 2004). For example, a study showed that chicks with lateralization for eye use were better at performing two tasks simultaneously (i.e., finding food and detecting an

overhead predator) than chicks without lateralization for eye use (L. J. Rogers et al., 2004). Lateralization also avoids ineffective duplication of functions in the two hemispheres, such that neural capacity may be increased overall (Vallortigara, 2006). It furthermore reduces the connection distances within a lateralized module, which not only allows processing to be faster (Karolis, Corbetta, & Thiebaut de Schotten, 2019), but also costs less energy (L. Rogers & Vallortigara, 2015). Additionally, hemispheric dominance ensures that the two hemispheres do not issue conflicting instructions for behaviour.

However, advantages of asymmetry at the level of individual organisms cannot necessarily explain average asymmetry at the population level (Frasnelli & Vallortigara, 2018). In an attempt to account for population-level lateralization, one theory states that it is a form of evolutionary stable strategy (ESS), which arises from inter-individual interactions (Frasnelli & Vallortigara, 2018). As a result of such interactions, the individual or species (depending on the ecological context) may benefit from a behavioural asymmetry pattern that is harmonized with those of other organisms with whom it interacts (Frasnelli & Vallortigara, 2018). For example, shoaling of fish decreases the individual risk of being preyed upon, such that a coordinated lateralized flee response (e.g., turning behaviour) can increase the chance of survival for those typically lateralized, by gaining protection from the group. Nevertheless, predators can exploit behavioural biases of prey, which might then favour a minority of prey organisms with atypical lateralization, who have the advantage of surprise (Vallortigara & Rogers, 2005). Thus a stable equilibrium could be reached with a certain percentage showing atypical laterality. Additionally, the theory of ESS can be extended to include intraspecific interactions (i.e., interactions between individuals from the same species), whereby a stable equilibrium of left- and right lateralized individuals can be reached when there is a balance between competitive and cooperative interactions (Ghirlanda, Frasnelli, & Vallortigara, 2009). Of note, the theory of ESS does not hold for all population-level asymmetries, as some of these (e.g. limb usage in toads) appear to be irrelevant to inter- or intra-species interactions (Vallortigara, 2006). It is therefore possible that some asymmetries have arisen as a by-product of others that were selected in the context of an ESS, i.e. simply because the brain already demonstrated laterality for some functionally

relevant domains when new behaviours evolved (Vallortigara, 2006). It is also possible that some population-level asymmetries of brain and behaviour are merely a secondary consequence of fundamental molecular asymmetries acting very early in development, which have effects on a wide variety of later outcomes, without each lateralized function necessarily having been selected for as such (Francks, 2015).

In addition to benefits of brain asymmetries, there may also be costs. As previously mentioned, predators can exploit the behavioural responses of their prey, such that typically lateralized individuals can be easier to predate. Furthermore, lateralized behaviour of the predator may result in missed feeding opportunities (Vallortigara & Rogers, 2005). In a general sense, brain laterality might come at the expense of processes that require relatively high interhemispheric communication, decreasing the efficiency of such processes (Wiper, 2017). As an illustration, topminnow fish that were trained to use the middle door in a row of nine, performed better when they exhibited a low degree of behavioural lateralization (Dadda, Zandona, Agrillo, & Bisazza, 2009). Fish that demonstrated a high degree of lateralization likely performed worse as the task required integration from both left and right visual hemifields (Dadda et al., 2009). In addition, lateralization can result in less efficient responses when an environmental cue from the 'wrong' side is presented (Wiper, 2017). To illustrate, toads were more efficient in capturing prey that appeared on their right side compared to their left (Robins & Rogers, 2006).

1.3 The development of left-right asymmetry of the brain and body

1.3.1 Initial symmetry breaking

Studies in animal models have provided important insights into the core developmental mechanisms contributing to left-right patterning in the vertebrate body (Hamada, Meno, Watanabe, & Saijoh, 2002). In particular, four main steps have been identified: 1) initial symmetry breaking, 2) transfer of asymmetric signals to the lateral plate mesoderm, 3) asymmetric gene

expression in the lateral plate mesoderm, and 4) asymmetric morphogenesis (Hamada et al., 2002). These steps are shared among most, if not all, vertebrates, and involve genetic modulation of *Pitx2* expression by *Nodal*, *Lefty1*, and *Lefty2*, key genes linked to asymmetry (Hamada et al., 2002).

However, the underlying mechanisms that lead to the initial breaking of symmetry seem to have diverged somewhat during evolution (Nakamura & Hamada, 2012). For example, most mammals exhibit motile cilia at the embryonic node (a pitted structure at the ventral surface of the early embryo), which generate a fluid flow specifically towards the left side of the lateral plate mesoderm, accordingly inducing asymmetric expression of key genes (Hamada et al., 2002). The cilia produce a flow predominantly in one direction due their inherent structural chirality which means that they rotate in a clockwise direction (Nonaka, Shiratori, Saijoh, & Hamada, 2002). This is due ultimately to the chirality of amino acid molecules used in life on earth (Inaki, Liu, & Matsuno, 2016). Chirality refers to an object or structure that cannot be superimposed on its mirror image form. However, pigs and chicks do not have nodal cilia, and left-right symmetry breaking in these species seems to be driven by a leftward movement of cells around the node, which themselves express key genes (Gros, Feistel, Viebahn, Blum, & Tabin, 2009). While it is thought that the movement of these cells is driven by a left-sided depolarization of the membrane potential of the primitive streak, the initial symmetry breaking events in pig and chick embryos remain unclear (Gros et al., 2009).

Impairments in ciliary motility can lead to *situs ambiguus*, as observed in *iv* mutant mice, or to *situs inversus*, as seen in *inv* mutant mice (Okada et al., 1999). S*itus ambiguus*/heterotaxy refers to partially reversed asymmetry of the visceral organs, and *situs inversus totalis* refers to a complete mirror reversal on the left-right axis (Fliegauf, Benzing, & Omran, 2007). An artificially induced rightward Nodal flow can also result in *situs inversus* in wild-type mouse embryos (Nonaka et al., 2002). In mice and humans, *situs inversus* is observed in roughly 50% of people with mutations in genes affecting components of primary cilia, such as those encoding axonemal dynein arms (Deng, Xia, & Deng, 2015). This indicates that when the initial,

direction-giving developmental process is lost, asymmetrical development still occurs, but with randomized direction.

There is also increasing evidence from invertebrates (including fruit flies, snails, and worms) that left-right symmetry can be broken by cellular chirality, i.e. chirality of cellular shape or structure, without cilia necessarily being involved (Inaki, Sasamura, & Matsuno, 2018). For example, cell chirality is important in establishing asymmetry of the embryonic hindgut of the fruit fly, where the actin cytoskeleton is involved (Lebreton et al., 2018). In worms and snails, cell chirality is present in the blastomere (Inaki et al., 2018). It is unclear to what extent these types of chirality are present in vertebrate species *in vivo*, but it is known from cell cultures (*in vitro*) that many vertebrate cells may exhibit chirality under certain conditions (Inaki et al., 2018). As I will note further below, a recent large-scale genetic study of adult human brain asymmetry has implicated cytoskeletal genes, and genes expressed relatively highly in the embryonic brain (Sha et al., 2020). This may indicate that brain asymmetry is induced by a mechanism that starts with cellular chirality.

1.3.2 Zebrafish epithalamus asymmetry development

The best studied brain asymmetry in vertebrates at the molecular level is that of the zebrafish epithalamus, which consists of two asymmetrically arranged structures: the habenulae and the parapineal organ (Concha et al., 2009). Just as Nodal signaling is central in the formation of visceral organ asymmetry in vertebrates, it plays an important role in the formation of epithalamus asymmetry in zebrafish. Mutant fish that do not express left-sided Nodal-signaling pathway genes, or express them bilaterally, show a randomization of epithalamus asymmetry (Concha, Burdine, Russell, Schier, & Wilson, 2000). Therefore, although Nodal signaling does not seem to be essential for the establishment of epithalamus asymmetry in zebrafish, proper signaling of Nodal is required for consistent directionality of asymmetry (Concha et al., 2000).

The establishment of epithalamus asymmetry is also mediated by Fgf8, as demonstrated by the loss of brain asymmetry in zebrafish *fgf8* mutants (Regan, Concha, Roussigne, Russell, & Wilson, 2009). Reduced FGF-

signaling can disrupt brain asymmetry independently of nodal-pathway induced asymmetry, by reducing expression of two transcription factors involved in *lefty1* repression (i.e., *six3b* and *six7*), which accordingly induces bilateral expression of *lefty1* in the brain (Neugebauer & Yost, 2014). Nodal pathway induced asymmetry is initiated by asymmetric fluid flow of cilia within the zebrafish node (called Kupffer's vesicle), leading to asymmetric expression of *southpaw* (*spaw;* a *nodal* signaling pathway gene) in the lateral plate mesoderm (Neugebauer & Yost, 2014). Furthermore, both Notch and Wnt/beta-catenin signaling are needed for left-sided expression of *Nodal* (Gunturkun & Ocklenburg, 2017), as Notch signaling controls the length of motile cilia within Kupffer's vesicle (Lopes et al., 2010), and because Wnt/beta-catenin signaling coordinates ciliogenesis via the transcription factor Foxj1a, which is known to inhibit repression of *Nodal* (Husken & Carl, 2013).

1.3.3 The development of human brain asymmetry

While visceral and brain asymmetries are clearly linked through the Nodal signaling pathway in zebrafish, there appears to be a disconnection of at least some types of brain asymmetry from visceral asymmetry in humans (*Box 1.1*), which means that another developmental pathway is likely to be involved in human brain asymmetry.

The human brain starts taking shape from around four weeks gestational age, when an embryonic structure, called the neural tube, initially gives rise to three main structures, namely the prosencephalon, mesencephalon, and rhombencephalon (also referred to as forebrain, midbrain, and hindbrain respectively) (Schmitz, Gunturkun, & Ocklenburg, 2019). The earliest developmental observation of human brain anatomical asymmetry has been made at 11 gestational weeks, involving the choroid plexus which is larger on the left than the right (Abu-Rustum, Ziade, & Abu-Rustum, 2013; Schmitz et al., 2019). This is after the establishment of visceral organ asymmetry in humans is nearly completed (Schmitz et al., 2019). Even earlier than this, subtle differences in global gene expression profiles have been observed between the left and right sides of the embryonic forebrain as early as 7.5 gestational weeks (de Kovel, Lisgo, Fisher, & Francks, 2018). Furthermore,

Box 1.1 The link between brain and body asymmetry

Asymmetry of the zebrafish epithalamus is perhaps the best studied brain asymmetry in vertebrates at the molecular level, which is not only present on the population-level, with more than 95% of zebrafish showing a leftward asymmetry of the epithalamus, but also coupled to asymmetry of the visceral organs (Concha, Signore, & Colombo, 2009). A study in the *frequent-situs-inversus* (*fsi*) line of zebrafish showed that those with *situs inversus* demonstrated a similar reversal of asymmetry in the epithalamus, along with reversals of behaviour, such as approaching a target from the opposite side than seen in wild types (Barth et al., 2005). Although some lateralized behaviours were not reversed in the *situs inversus* fish, such as turning behaviour when startled, the findings nevertheless indicated a link between visceral and brain asymmetry in zebrafish (Barth et al., 2005).

In humans, however, the existence of a link between visceral and brain asymmetry remains unclear. Studies in people with *situs inversus* and primary ciliary dyskinesia (PCD; an autosomal recessive disorder featured by deficits in primary cilia due to mutations in ciliary component proteins) have shown that reversal of visceral organ asymmetry is not associated with reversal of hemispheric language dominance (Kennedy et al., 1999; Vingerhoets et al., 2018), nor a change in right ear advantage during dichotic listening to spoken syllables, (Tanaka, Kanzaki, Yoshibayashi, Kamiya, & Sugishita, 1999), nor an increased rate of left-handedness (McManus, Martin, Stubbings, Chung, & Mitchison, 2004). However, a study that distinguished between *situs inversus* with and without PCD reported increased levels of left-handedness specifically in those without PCD (Vingerhoets et al., 2018). This observation suggested a potential link between body and brain asymmetry in humans that is independent of the ciliary-nodal mechanism of left-right axis formation, a possibility that was followed up in Chapter 4 of this thesis.

gene expression differences have been reported between the left and right embryonic spinal cords and hindbrains at 6-10 weeks of gestational age, which indicated a subtle left-right difference in maturation rates (de Kovel et al., 2017). Other human brain asymmetries, or manifestations thereof, that can be observed already before birth are asymmetries of the lateral ventricles (Hering-Hanit, Achiron, Lipitz, & Achiron, 2001), and lateralization of motor behaviour (Hepper, McCartney, & Shannon, 1998). Indeed, about 85% of human fetuses at 10 gestational weeks have been reported to move their right arms more often their left (Hepper et al., 1998), a percentage roughly comparable to the adult rate of right-handedness.

One study in human fetuses suggested that brain asymmetries might be influenced by epigenetic factors within the spinal cord (Ocklenburg et al., 2017). This study analyzed mRNA expression, microRNA expression, and DNA methylation patterns in fetal cervical and anterior thoracic segments of spinal cord, and found these to be lateralized (Ocklenburg et al., 2017). Epigenetic mechanisms, including DNA methylation and histone modification, provide a mechanism for the regulation of gene expression (Mazzio & Soliman, 2012). These mechanisms impact the accessibility of genomic DNA for transcription, rather than affecting the DNA sequence directly. The epigenome, which refers to the arrangement and modification of DNA-associated proteins as well as DNA methylation marks, can be affected by numerous biological as well as environmental factors. It is known that environmentally induced epigenetic modifications can sometimes be inherited through mitosis and meiosis (Mazzio & Soliman, 2012). However, left-right epigenetic differences observed in fetal spinal cord may simply be downstream manifestations of a genetically controlled program of asymmetrical development.

1.4 Population variation and heritability of brain asymmetry

1.4.1 Variation in brain asymmetries

For a given measure of brain asymmetry (e.g. of a specific cortical regional volume), deviations from the typical, population-average asymmetry can be

indicated in at least four ways: (1) an increase or (2) decrease of the default asymmetry pattern, but where the asymmetry remains in the same direction as the typical form, (3) a reversal of the default pattern, or (4) a loss of the default pattern (i.e., symmetry) (X. Z. Kong et al., 2020). For example, people with schizophrenia have been reported to show a decrease in left > right asymmetry in temporal lobe regions, both in terms of grey matter volume and functional activity (Oertel et al., 2010). I will return to associations with psychiatric disorders in section 6 below, and in Chapters 2 and 3. Other examples of atypical asymmetries are that roughly 1% of the population has rightward hemispheric language dominance, compared to the majority being leftward lateralized in the population (Mazoyer et al., 2014), and that 11% of the population have a larger *planum temporale* (a cerebral cortical region located at the posterior end of the sylvian fissure) on the right hemisphere than the left, in contrast to the majority of the population (Geschwind & Levitsky, 1968).

It is important to mention that one individual can simultaneously exhibit different alterations of asymmetry, depending on the specific brain region or measure investigated. In fact, various aspects of brain asymmetry can vary largely independently in the population, as has been shown from low correlations between structure-structure, structure-function, and/or function-function pairs of asymmetry measures (Tulio Guadalupe, Kong, Akkermans, Fisher, & Francks, 2020; X. Z. Kong et al., 2020; Mazoyer et al., 2014). For example, handedness and functional language dominance vary largely independently of each other in the population, although people with fully reversed, right-hemisphere language dominance do tend to be mostly left-handed (Mazoyer et al., 2014). Brain asymmetry must therefore be considered as a complex and multivariate trait. Indeed, as regards the arrangement of functional lateralities for different cognitive domains, a distinction has been proposed between complete reversal of all functions versus reversal of some but not all functions, whereby the first may be referred to as 'reversed typical' rather than 'atypical', to highlight that functional complementarity is retained in some individuals (Vingerhoets, 2019).

1.4.2 Heritability of brain asymmetry

As summarized above, various brain asymmetries have strong population-level biases which have early developmental origins (e.g. roughly 90% of people are right-handed, only 10% left-handed). This indicates a developmental program that is inherently lateralized and under genetic control. However, this does not necessarily imply that variation in brain and behavioural asymmetry is particularly heritable, i.e. when development does not follow the typical lateralized course as specified by the genome, it may be largely due to non-heritable factors. Heritability is defined as the proportion of variation in a trait that is due to genetic differences between individuals in a population, and is therefore a distinct concept from the notion of a default developmental program being specified genetically. In principle, many factors, including randomness in early development, and possible environmental influences, might affect a complex trait like brain asymmetry (see section 5 below). It might therefore not be surprising that twin- and family-based estimates of heritability for brain asymmetries are only moderate to low, such as those for subcortical volume asymmetries (15-27% heritabilities) (T. Guadalupe et al., 2016), regional asymmetries of the cerebral cortex (thickness or surface area; up to roughly 25% heritable) (Kong et al., 2018), or estimates for asymmetries of white matter fibre integrity (20-37%) (Jahanshad et al., 2010).

Heritability estimates for left handedness range from 2% (based on genome-wide genotype data for common polymorphisms (de Kovel, Carrion-Castillo, & Francks, 2019; de Kovel & Francks, 2019; Neale, 2017)), up to 25% (based on twin and family data (Medland et al., 2009)). A difference between single nucleotide polymorphism (SNP)-based and twin-based heritability estimates is consistently observed for many traits, and is often referred to as 'missing heritability', which can reflect certain differences in methods and assumptions between the two approaches (*Box 2*). Another consideration is that different studies have used different ways of assessing handedness: sometimes by means of hand preference questionnaires over multiple items (e.g. the Edinburgh handedness Inventory (Oldfield, 1971)), hand performance measures (e.g. the pegboard task) (Corey, Hurley, & Foundas, 2001), or by means of self-report according to a simple choice between left, right or mixed handed.

Some researchers have posited single-gene theories in relation to handedness (Annett, 1998; McManus, 1985) (i.e., theories assuming that one single genomic locus accounts for variation in handedness), but these theories were not supported empirically into the modern molecular-genetic age (Armour, Davison, & McManus, 2014). Instead, it is now considered that handedness, and likely other brain asymmetries too, involve complex multifactorial aetiologies, rather than representing monogenic traits (de Kovel et al., 2019; de Kovel & Francks, 2019; Francks, 2015; Wiberg et al., 2019). Notably, a recent GWAS study associated genetic variants at 41 separate genomic loci with left handedness, and seven loci to ambidexterity (Cuellar-Partida et al., 2020). This study involved more than 1 million individuals, with each of the loci influencing handedness by only a tiny amount. Interestingly, pathway enrichment analysis of significantly associated loci revealed significant associations with several molecular pathways, including microtubule organization (Cuellar-Partida et al., 2020).

Genetic variants associated with variation in anatomical brain asymmetries in the population have also been identified recently by large-scale imaging genetics analysis (Carrion-Castillo et al., 2020; Le Guen et al., 2020; Sha et al., 2020). Microtubule-related genes have been particularly implicated, which is consistent with a role of the cytoskeleton in setting up cellular chirality during embryonic development of the left-right axis of other organs in other species (see above). However, a variety of other genes involved in brain development have also been implicated (Sha et al., 2020), and it is unclear why they might be especially involved in left-right asymmetry.

1.5 Environmental factors that may affect brain asymmetry

1.5.1 Hormones

Early in life, fetal brain development can be impacted by *in utero* exposure of the fetus to the mother's hormones, particularly levels of androgens (Duboc et al., 2015). Prenatal testosterone levels have been positively associated with language lateralization in 6 year old children (Lust et al., 2011; Lust et al.,

2010), which indicates that long-term effects of sex hormones are at least partly established during prenatal brain development. Additionally, both prenatal and pubertal testosterone levels have been linked to brain laterality during cognitive performance (Beking et al., 2018). For various brain regions, males demonstrate stronger hemispheric asymmetries than females on average (T. Guadalupe et al., 2016; Hirnstein, Hugdahl, & Hausmann, 2019; Kong et al., 2018). However, these effects are small, with largely overlapping ranges of variation between the sexes.

Contrary to popular belief, average sex differences in brain asymmetry may not be related to average sex differences in performance profiles across different cognitive domains, as the two have been shown to generally occur independently (Hirnstein et al., 2019). The maternal hormonal milieu of a developing fetus may be considered as an environmental influence on the fetus, although circulating hormone levels in the fetus will also be affected by its own genome.

1.5.2 Training

Training can also influence brain/behavioural asymmetries (Esteves, Ganz, Sousa, & Leite-Almeida, 2020). For example, left-handers who are forced to use their right hand can become right-handed, involving plasticity to enable functional adaptations of the brain (Kloppel, Vongerichten, van Eimeren, Frackowiak, & Siebner, 2007). Additionally, musicians with perfect pitch exhibited larger *planum temporale* asymmetry compared to non-musicians, or early beginning musicians without perfect pitch (Keenan, Thangaraj, Halpern, & Schlaug, 2001). This finding suggests that *planum temporale* asymmetry may be influenced by long-term training as a musician (Esteves et al., 2020; Keenan et al., 2001), although gene-environment interactions are also likely to play a role via innate individual differences in musicality, which lead to increased likelihood of becoming a musician (Gingras, Honing, Peretz, Trainor, & Fisher, 2015). Another example of training-induced changes in brain asymmetry comes from children who were raised bilingually. Such children have shown increased bilateral activation for both languages as compared to children who had learnt a second language later in life (Esteves et al., 2020; Hull & Vaid, 2007). Finally, experienced London taxi drivers have shown increased rightward asymmetry of hippocampal gray

matter as compared to London bus drivers (who followed constrained routes) (Maguire, Woollett, & Spiers, 2006), further supporting the idea that training can influence brain asymmetry (Esteves et al., 2020).

1.5.3 Ageing

Ageing (both in health and in neurodegenerative disorders) can impact brain asymmetries, such as functional lateralization in the default mode network (Banks et al., 2018). Indeed, the brains of healthy elderly people can plastically adapt to age-related cognitive decline, by recruiting more brain regions for lateralized functions, including regions located in the contralateral hemisphere (Esteves et al., 2020). Additionally, structural grey matter volume asymmetries can be influenced by age (Kong et al., 2018; Minkova et al., 2017; Roe et al., 2020). These effects are region-specific, as will become clear later in this thesis.

1.5.4 Early life factor

It has been suggested that the *in utero* position of human fetuses in the final trimester can influence their brain lateralization for hand use and language (Duboc et al., 2015). To clarify, most human fetuses take on a head-down orientation on the left side of the mother's womb, probably because the visceral asymmetry of the mother might impose physical constraints otherwise (Duboc et al., 2015). It is suggested that this head-down orientation, which is seen in about 96% of fetuses (Noli et al., 2019), allows better right-arm movement, and that it may even trigger language-lateralization due to greater stimulation of the right ear (Duboc et al., 2015). This would be analagous to the situation in chicks, where the posture of the embryo in the egg can influence asymmetrical development of the visual pathway, as it affects whether the right or left eye is more exposed to light (Duboc et al., 2015).

Other early life factors are also likely to be involved in brain asymmetry. In a recent cross-sectional study of over 400,000 people (de Kovel et al., 2019), the probability of being left-handed was found to associate with the year and location of birth within the United Kingdom, as well as birthweight, being part of a multiple birth, season of birth, breastfeeding, and sex, with each effect remaining significant after accounting for all others. However, all

together, these factors had only a minimal predictive value for individual hand preference, as each effect was very small. Birthweight was itself partly heritable in this study, but showed no evidence for co-heritability with left-handedness.

1.5.5 Lesions

Homotopic regions of the contralateral hemisphere can show increased activity patterns after recovery from lesions, for example caused by stroke (Esteves et al., 2020). This again demonstrates that neural plasticity can influence functional brain asymmetry.

1.5.6 Developmental randomness

As none of the above factors has been shown to have a large effect on population variation in brain asymmetry, and the heritabilities of brain asymmetry measures are also relatively low, then it is likely that a large part of the variance is due to randomness in brain development (Mitchell, 2018), perhaps stemming particularly from early embryonic patterning. See chapter 6 for more discussion of this subject.

1.6 Altered brain asymmetries in disorders

Altered brain asymmetry has been associated with numerous cognitive and psychiatric disorders, including dyslexia (Altarelli et al., 2014), Alzheimer's disease (Thompson et al., 1998), attention-deficit/hyperactivity disorder (ADHD) (Shaw et al., 2009), psychotic disorders (Crow, 1990; M. Yucel et al., 2002; M. Yucel et al., 2003), autism (Eyler, Pierce, & Courchesne, 2012), and mood disorders (K. Yucel et al., 2009). These studies have used a wide range of different functional, structural and behavioural measures of laterality. It is perhaps unsurprising that different aspects of asymmetry can be altered in a wide range of disorders, considering that asymmetry is a feature of healthy brain organization for so many different functions.

In addition, while some atypical asymmetries of brain and/or behaviour, such as left-handedness, are not important clinically, they can actually be more frequent in various disorders than in the general population (Brandler &

Paracchini, 2014). For example, it has been reported that the prevalence of left- and mixed-handedness is increased in autism spectrum disorder (ASD) (Markou, Ahtam, & Papadatou-Pastou, 2017), anxiety disorder (Hardie, Wright, & Clark, 2016), bipolar disorder (Nowakowska et al., 2008), and schizophrenia (Hirnstein & Hugdahl, 2014). These observations suggest that alterations of brain asymmetry might serve as a risk-factor for any of the aforementioned disorders (Logue, Logue, Kaufmann, & Belcher, 2015). However, although the term 'risk-factor' might imply that atypical brain asymmetry is a contributing cause of disorder, it may also be a consequence of the pathology, or else both altered asymmetry and disorder can be consequences of another underlying causal factor (*Box 3*).

It is also important to note that the literature has not been consistent with respect to altered asymmetry in disorders (X. Z. Kong et al., 2020). As well as varying methods, prior studies have been marked by relatively low sample sizes, typically in the range of tens to low hundreds of individuals at most. Limited sample sizes provide low statistical power for detecting subtle associations (i.e., low probability of detecting a true but small effect), as well as increase the chance that positive findings do not reflect true effects (Munafo & Flint, 2010). For that reason, many prior studies may have reported effects that were unrealistically large and/or unreliable. In addition, inconsistent findings could be explained by study-specific differences in clinical features of the disorder, which can arise, for example, due to distinct case recruitment criteria, and/or differences in diagnostic practices. Diagnosis of psychiatric disorders often partly depends on subjective measures (Emser et al., 2018), making it more prone to bias than diagnosis of non-psychiatric disorders, which are more often based on objective measures. Moreover, differences in methodology between studies can account for inconsistent findings between studies, such as the use of distinct hardware, software, and processing pipelines for MRI data (Biberacher et al., 2016).

This thesis includes two large-scale studies, combining data from many different datasets around the world, to study asymmetries of brain anatomy in ASD (chapter 2) and in ADHD (chapter 3) in thousands of individuals. The largest previous studies of brain asymmetry in these disorders comprised, to

Box 1.3. Different models of causality for asymmetry-disorder associations

A paper by D. Bishop described four different causal models (Fig.1.2) that might explain association patterns between genotypes, language and literacy impairment, and brain lateralization (Bishop, 2013). These are the endophenotype model, the pleiotropy model, the additive/interactive risks model, and the neuroplasticity model (Bishop, 2013). Notably, these models are 'simplified abstractions and not mutually exclusive' (Bishop, 2013), and they are not restricted to language and literacy impairment, but instead can apply to any disorder that is associated with altered brain laterality.

Briefly, genetic factors that are causal for the disorder can likewise affect brain laterality, either via a direct causal link (endophenotype model) or without (pleiotropy model). If genetic factors solely affect the disorder, and not brain laterality, then altered laterality itself can exert an independent causal influence on the disorder (additive/interactive risks model), or it can be a consequence of disorder pathology (neuroplasticity model) (Bishop, 2013).

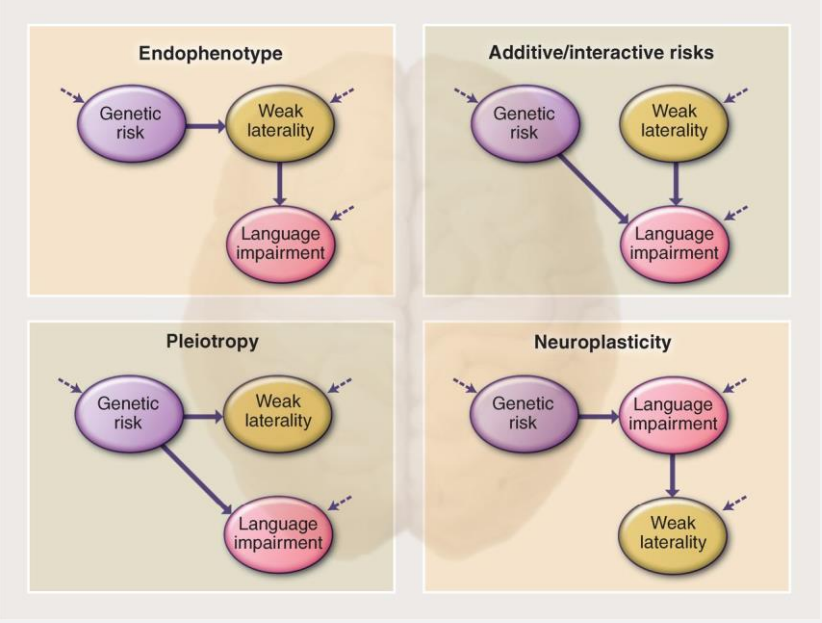


Fig.1.2. Overview of causality models (derived from (Bishop, 2013)).

our knowledge, no more than 128 individuals with ASD (Dougherty, Evans, Katuwal, & Michael, 2016), and 341 individuals with ADHD (Douglas et al., 2018). ASD and ADHD are genetically correlated neurodevelopmental disorders (Ghirardi et al., 2019) that often co-occur (Leitner, 2014). Both disorders have their onsets in childhood, and both can be linked to lateralized functions (see Chapters 2 and 3). To illustrate, ASD often features an impairment in social communication and language performance (American Psychiatric Association, 2013), while language processing is typically performed dominantly by the left hemisphere (Mazoyer et al., 2014). As regards ADHD, which often features symptoms of inattention (American Psychiatric Association, 2013), the right hemisphere is usually dominant for some aspects of attention and arousal (Heilman, Bowers, Valenstein, & Watson, 1986). Previous large-scale structural imaging studies of these disorders, which did not investigate brain asymmetry, have found very small case-control differences bilaterally (i.e., Cohen's *d* effect sizes between -0.21 and 0.06) (Hoogman et al., 2017; Hoogman et al., 2019; van Rooij et al., 2018). I therefore hypothesized that if case-control differences in asymmetry are equivalently small as the bilateral alterations, then larger sample sizes than studied previously will be needed to identify them. Identifying specific alterations of brain asymmetry in ASD and ADHD, however small, would help to better describe their neurobiology and pathogenesis.

1.7 Methods and approaches used in this thesis

The current thesis comprises two studies of altered brain asymmetry in major neurodevelopmental disorders, and two gene mapping studies aimed at discovering new genetic influences on brain asymmetry. Here I will summarize general principles and histories of the key approaches that were used for the studies in this thesis.

1.7.1 Magnetic Resonance Imaging (MRI)

1.7.1.1 Technical aspects underlying MRI

The MRI signal is based on the spinning and magnetic properties of hydrogen protons in the tissue under study (Broadhouse, 2019). Briefly, spinning of

protons in the hydrogen nucleus generates a local magnetic field (Grover et al., 2015). When a strong, external magnetic field (B0) is applied, the nuclei of the hydrogen atoms will align to it. Nuclear spins in a low-energy state will align parallel to the B0 field, while those in a high-energy state will align perpendicular to it. Generally, solids or liquids tend to have an excess of low-energy state spins, which can be excited by applying a second magnetic field (B1) perpendicular to B0. This B1 magnetic field is typically applied in short radiofrequency (RF) pulses, and each time it is switched off, the nuclei will emit the energy they had absorbed, resulting in a voltage that can be captured by a tuned coil of wire. The voltage is accordingly amplified and displayed as a signal. Since many RF pulses are applied (each lasting microseconds), the average is taken, resulting in a time-domain signal that can be mathematically solved using the Fourier transformation, accordingly enabling the creation of an MRI image or a frequency spectrum (Grover et al., 2015). Depending on the type and parameters of the RF pulse used, that is, the MRI sequence (e.g., spin echo, gradient echo, diffusion weighted), images with different tissue contrasts can be obtained (Bitar et al., 2006). Those brain images are typically stored in DICOM format, but can be converted to NIfTI format for subsequent analyses (Li, Morgan, Ashburner, Smith, & Rorden, 2016). After several pre-processing steps (e.g. spatial realignment, normalization, smoothing), individual-specific images can be mapped with respect to an atlas based on previously analyzed brain data, enabling parcellation of the image into a number of distinct regions defined according to neuro-anatomical or functional criteria (Yaakub et al., 2020).

1.7.1.2 Functional MRI

Gradient echo sequences are typically used to reveal the effects of deoxygenated hemoglobin in venous blood (Glover, 2011; Ogawa, Lee, Kay, & Tank, 1990), resulting in an image contrast known as the blood-oxygen-level-dependent (BOLD) contrast (Glover, 2011). The BOLD contrast is based on the principle that deoxygenated hemoglobin is highly paramagnetic, in contrast to oxygenated hemoglobin. Upregulation of certain brain regions, for example when performing a cognitive task, results in increased neural firing and a subsequent drop in blood oxygen levels. Thus, the BOLD contrast is an indirect measure of neural activity. Since neural activity fluctuates over

time, multiple images of the brain are acquired during functional MRI (fMRI), where each voxel reflects a time series that can subsequently be analyzed (Glover, 2011). While task-based fMRI is accompanied by a specific task, no cognitive tasks are used during resting-state fMRI. Hierarchical clustering of resting-state time series data has shown, at the highest level, two anti-correlated clusters; one for intrinsic and one for extrinsic processing (Doucet et al., 2011). Furthermore, the intrinsic networks could be divided into three modules, related to cognitive control and switching activity, among other processes (Doucet et al., 2011), which indicates that resting-state activity can partly reflect functional networks.

1.7.1.3 Deriving asymmetry measures from MRI data

After MRI images have been parcellated, regional anatomical- or functional measures (depending on the MRI sequence used) can be extracted, such as regional cortical thickness and surface area, or the average BOLD-signal per region. Any of these measures can be extracted per hemisphere, and the difference between homotopic regions in the two hemispheres can be calculated to acquire an index of asymmetry. The asymmetry index (AI) is widely used in brain asymmetry studies, and can be calculated as (Left-Right)/((Left+Right)/2) (Kong et al., 2018). Several variations on the formula exist, such as multiplying, instead of dividing the denominator by a factor 2 (Leroy et al., 2015), but these variations generally do not impact the results for group comparisons. A property of the AI is that it does not necessarily scale with the bilateral magnitude of the feature under study (i.e L+R), due to division by the denominator term. This is a useful feature for many applications, including those in this thesis, although it should also be acknowledged that other approaches can also sometimes be useful (e.g. treating L and R as paired, within-subject measures in linear mixed models, or analyzing L-R without adjustment for the denominator L+R).

One important issue is that the application of pre-prepared brain atlases can affect the measurement of asymmetry, to the extent that the brains used for deriving the atlas themselves were asymmetrical on average (which is likely as general population data tend to be used for atlas construction). However, I was primarily interested in relative group differences in asymmetry, not the measurement of absolute values of asymmetry, so that a

classical atlas-based approach was suitable. Indeed, an asymmetrical atlas has the advantage of being a closer match than an artificially symmetrized atlas, when data are themselves asymmetrical on average.

1.7.2 Genome sequencing and genotyping

1.7.2.1 DNA sequencing

Research at the molecular genetic level generally involves techniques for characterizing (parts of) the genetic code of an organism (Petersen, Fredrich, Hoeppner, Ellinghaus, & Franke, 2017). The first technique for DNA sequencing was developed in the 1970s by a British chemist named Frederick Sanger, and was called Sanger sequencing (Petersen et al., 2017; Sanger & Coulson, 1975). Sanger sequencing uses nucleotide-specific chain-terminating inhibitors to determine the genetic code of a specific DNA fragment (Petersen et al., 2017; Sanger & Coulson, 1975). Despite its high accuracy, the method has high costs, as well as a low throughput. That is to say, it can only process small DNA fragments, typically of less than 1000 base pairs (bp) in length, in a one-by-one manner (Petersen et al., 2017). It took more than a decade (1990-2003), and nearly 3 billion dollars, to sequence the first entire human genome using Sanger sequencing, a huge effort of many thousands of scientists across the world, as realized by the Human Genome Project (Lander et al., 2001).

In the mid-2000s, however, new sequencing methods, typically referred to as next generation sequencing (NGS), were introduced, allowing researchers to sequence the entire human genome for merely ~10,000 dollars in a matter of weeks (Goodwin, McPherson, & McCombie, 2016; Petersen et al., 2017). Since then, techniques have become even faster and cheaper, so that a genome can now be sequenced within days for roughly $1,000 (Hayden, 2014). NGS methods enable the sequencing of tens to hundreds of millions of DNA fragments simultaneously (Yegnasubramanian, 2013), rather than sequencing one DNA fragment at a time (Petersen et al., 2017), by means of sequencing by ligation or sequencing by synthesis of different fragments in parallel (Goodwin et al., 2016). However, notable limitations of NGS platforms are the lower accuracy (~0.1-2%) (Liu et al., 2012), and shorter read lengths (35-700bp) compared to Sanger sequencing (Goodwin et al., 2016). Yet the

possibility to sequence the whole genome or exome (i.e., the protein-coding regions specifically) in a relatively short amount of time, and with low costs, far outweighs the limitations of NGS. Accuracy can be improved by repeated sequencing from the same sample and then looking for consensus across the reads that map to a given region. Indeed, whole-genome sequencing (WGS) and whole-exome sequencing (WES) have now become widely used applications of NGS (Goodwin et al., 2016), which has led to a spike in the number of Mendelian disease genes identified (Petersen et al., 2017). I have used data from both WGS (chapter 4) and WES (chapter 5) for studies in my thesis.

More recently developed third generation sequencing techniques allow the production of sequencing reads that are much longer than those from NGS, which can increase the quality of the assembly and deal more effectively with repetitive elements in the genome (Goodwin et al., 2016; van Dijk, Jaszczyszyn, Naquin, & Thermes, 2018). However these approaches are not yet widely used, and I have not made use of data from them in my thesis.

1.7.2.3. Genotyping with SNP chips

DNA microarrays are small chips with many spots, each containing DNA fragments referred to as probes, typically ranging from 25 to 80 bp in length (Levy & Burnside, 2019). The probes are designed to distinguish between the alternative alleles at single locations in the genome that are already known to be commonly polymorphic in the population – most often SNPs – by binding to a target DNA sequence (LaFramboise, 2009). While DNA microarray genotyping can be done for more than 1 million SNPs simultaneously, only SNPs that are known to exist *a priori* can be genotyped via this approach (LaFramboise, 2009).

Genotyping chips are designed in such a way that they can effectively cover the entire genome with only a limited number of SNPs, by making use of linkage disequilibrium (LD) patterns, i.e., the phenomenon that genetic variants in close proximity in the genome tend to be inherited together (Levy & Burnside, 2019). Therefore many SNPs that have not been directly genotyped can afterwards be imputed when they are known to be in high LD with the genotyped SNPs (Marchini & Howie, 2010). Genotype imputation

involves the prediction of the non-assayed genotypes, according to a reference set of genotype data from multiple individuals (e.g. HapMap (International HapMap, 2003)) (Marchini & Howie, 2010). The accuracy of imputation is dependent on how closely the reference population is matched to the study population, as well as the number of recombination events in a region, and the level of LD, where imputation within high LD regions can be more accurate than low LD regions (Marchini & Howie, 2010). In this way, most of the genome can be covered by directly genotyping only a selected number of known variants (500,000-1,000,000 SNPs) (Dehghan, 2018). Chapter 5 in this thesis makes use of imputed SNP data to examine common genetic variants.

Of note, the application of DNA microarrays has expanded, and now also includes e.g. the detection of copy number variants (CNVs), which are deletions, duplications, inversions or more complex rearrangements of larger genomic regions (LaFramboise, 2009). Furthermore, while most genotyping chips are designed to cover relatively common variants in the genome (e.g. above 0.1%-1% in population frequency), they can also be customized in order to improve the assessment of rarer variants (Bomba, Walter, & Soranzo, 2017). Reference panels with greater depth of rare and low-frequency haplotypes, such as those based on WGS data, can help improve imputation accuracy for rare variants (Bomba et al., 2017). Population-specific reference datasets can increase coverage and accuracy of imputation, particularly for low-frequency and rare variants (Huang et al., 2015), as those variants tend to be relatively younger than common variants and showing more ancestry-specific clustering (Bomba et al., 2017). Imputation of rare variants can result in genotypes from tens of millions of SNPs per individual in a study dataset. This improvement in coverage can enhance the statistical power to detect trait associations for rare variants that are in low LD with nearby SNPs (de Bakker et al., 2005). However, it also comes at the cost of increased multiple testing, and statistical correction for the increased testing can result in reduced power to detect trait associations with common variants, that are anyway well captured without rare variant imputation (de Bakker et al., 2005).

1.7.3 Neuro-imaging genetics

Neuroimaging genetic studies involve finding statistical relations between specific genomic variants and variations in brain structure or function (de Zubicaray et al., 2008). In the early 2000s, most neuroimaging genetic studies were based on candidate genes and small study samples, e.g. sometimes only 10-20 participants for functional neuroimaging studies (Glahn, Thompson, & Blangero, 2007). Candidate genes were chosen for study when they were suspected *a priori* to be involved in the phenotype of interest. For example, in one study, candidate genes for handedness were defined based on their known roles in visceral left-right asymmetry development (Van Agtmael, Forrest, & Williamson, 2002), and then linkage analysis was performed to test whether genomic regions containing those genes showed co-segregation with left-handedness in families (Van Agtmael et al., 2002). Note that linkage studies, which are based on inferred sharing of segments of the genome identical-by-descent among people who share a trait, should not be confused with genetic association studies, where the relationships of specific alleles to traits are tested (Pulst, 1999).

In the field of neuroimaging genetics, and more broadly across human genetics research, early findings from candidate gene studies or linkage studies could not be replicated using independent datasets (Thompson et al., 2020). In fact, these approaches have been widely replaced by genome-wide association scanning (GWAS), which offers a relatively hypothesis-free approach for examining associations between genetic variants and a phenotype (Duncan, Ostacher, & Ballon, 2019). The realization of GWAS, owing to advances in DNA microarray techniques for high-throughput genotyping of known SNPs (see above)(Dehghan, 2018), has led to an increase in novel, replicable findings, particularly with respect to common, non-Mendelian phenotypes linked to common genetic variation (i.e., phenotypes where many genetic variants, each with small effects, are involved)(Petersen et al., 2017). In GWAS, each genetic variant is tested separately for an association with the phenotype, so a stringent multiple-testing correction threshold needs to be applied. A genome-wide significance threshold $P < 5 \times 10^{-8}$ has been found to be suitable for modern GWAS studies in European-descent populations that test several million genotyped or

imputed SNPs with minor allele frequencies above 1% (Dehghan, 2018). This genome-wide significance threshold takes into account the patterns of LD between SNPs (Dehghan, 2018). Since variants with lower frequencies generally have lower LD, a more stringent threshold is recommended for GWAS that include rarer variants (Fadista, Manning, Florez, & Groop, 2016).

Adequate statistical power is crucial to reliably test millions of SNPs covering the genome, as for most complex traits, each SNP has at most only a small individual effect (Hong & Park, 2012). Statistical power in GWAS is dependent on many factors, including the effect sizes and MAFs of SNPs, disease prevalence, case-control ratio for binary traits, LD, and, last but not least, sample size (Hong & Park, 2012). The effective sample size, that is to say, the size that is minimally needed to obtain 80% statistical power, can be calculated under various assumptions (e.g. the genetic model, MAF, LD etc.) (Hong & Park, 2012). In recent GWAS studies of brain anatomy in tens of thousands of participants (Elliott et al., 2018; Grasby et al., 2020; Hibar et al., 2015; Vojinovic et al., 2018; Zhao et al., 2019), it has become clear that individual SNPs rarely have effect sizes that explain more than 1% of overall trait variance, and usually the effects of individual SNPs are lower than this. Nonetheless, the large sample size has permitted the identification of hundreds of loci at a genome-wide significant level, which affect diverse aspects of brain anatomy, such as specific subcortical volumes, cortical regional surface areas or thicknesses, total brain volume, or white matter metrics for specific tracts.

With the rise of NGS technology, methods complementary to GWAS have also been developed to increase power to assess rare variant-phenotype associations, which aggregate the combined putative effects of multiple rare variants into one score per genomic region (where regions are typically defined as genomic loci for individual genes)(Bomba et al., 2017). The directions of the allelic effects with respect to the phenotype can either be assumed to be the same (as in burden tests, where an increase in mutational load in a gene is assumed to increase disorder risk) or different (as in variance-component tests, which allow some variants to increase risk while others in the same gene might be protective). These types of tests have been developed

for both continuous and categorical phenotypes, and are used in chapter 5 of this thesis.

1.8 Large-scale imaging-genetics databases used in this thesis

The field of neuroimaging genetics has not only benefited from advancements in DNA sequencing and genotyping technologies, but also from the realization of large-scale consortia that combine data via meta-analyses (Thompson et al., 2020), or large-scale population Biobank projects (Bycroft et al., 2018). These efforts provide large enough sample sizes to achieve adequate statistical power to detect realistic effects (Munafo & Flint, 2010), for example in imaging case-control or GWAS studies. The move towards consortia efforts and large-scale resources has revolutionized the field (de Zubicaray et al., 2008).

In this thesis I make use of data from the ENIGMA (Enhancing NeuroImaging Genetics Through Meta-Analysis) consortium (chapters 2 and 3), and the UK Biobank (chapter 5), both of which have collected neuroimaging and genetic data from many thousands of individuals.

1.8.1 The ENIGMA consortium

The ENIGMA consortium, initiated in 2009, is a global partnership that currently comprises more than 1400 researchers from 43 different countries around the world (Thompson et al., 2020). More than 50 Working Groups are currently active, each having their own leadership, and they can be either clinical groups (e.g., studying a specific brain disorder from psychiatry or neurology) or non-clinical groups (e.g., testing harmonized methods, or examining general neurobiological features, like brain laterality) (Thompson et al., 2020). Combining data from multiple datasets worldwide into one mega- or meta-analysis has immensely increased sample sizes as compared to earlier studies based on single datasets (Glahn et al., 2007). However, there is inevitably substantial heterogeneity between datasets. That is to say, brain MRI measures, for example, are likely to vary between different datasets, e.g. due to differences in methodology across participating sites, or due to heterogeneity in recruitment and collection strategies (Feaster, Mikulich-

Gilbertson, & Brincks, 2011). Harmonized ENIGMA protocols for image processing and genetic analyses are used by each site, to reduce site-effects, but the actual MRI scanning and genotyping was typically done before membership of the consortium, and therefore according to varying procedures. Data are sometimes analyzed using meta-analytic techniques, when primary data are not available for sharing due to ethical or legal restrictions for legacy datasets. In meta-analytic approaches, each site runs its own analysis with its individual level data and then shares summary statistics with a central analysis team, which performs meta-analysis over the summary statistics from multiple sites. In this way, dataset heterogeneity is implicitly accounted for. However, sometimes sites are able to share their individual level data for joint analysis (which was the case for the present thesis – see Chapters 2 and 3), in which case it is important to correct for heterogeneity across datasets e.g. by modelling through random effects (Feaster et al., 2011).

1.8.2 The UK Biobank

The UK Biobank is another initiative to acquire a large sample size, this time from a general population cohort of middle- to older-age adults from the UK (aged between 40 and 69 at first recruitment). The UK Biobank has acquired data on a variety of phenotypes, ranging from genetic to imaging traits, as well as data related to health, including medical history, cognitive and sociodemographic measures (Bycroft et al., 2018). There are up to roughly 500,000 participants depending on the phenotype, of which currently around 40,000 have undergone multimodal MRI scanning, most of which are also genotyped with common SNPs genome-wide, and roughly 10,000 of which also have WES data available. In contrast to the ENIGMA consortium, which mostly involves post-hoc alignment of data from existing cohorts from many different research studies, data from the UK Biobank are acquired at a limited number of assessment centers with a uniform protocol (Bycroft et al., 2018).

1.9 Overview of chapters

1.9.1 General goals

In this thesis I aimed to increase knowledge of the extent to which altered human brain laterality is associated with neurodevelopmental disorders, and to better understand the genetic contributions to variation in human brain laterality. More specifically, chapters 2 and 3 aimed at quantifying the relations of ASD and ADHD to variation in structural brain asymmetries, while chapters 4 and 5 aimed at exploring the genetic contributions to left-handedness specifically in *situs inversus*, and atypical hemispheric lateralization of functional connectivity during resting-state in healthy individuals. These studies have increased our understanding of the genetic basis of brain asymmetries and their associations with disorders. Each of the studies is summarized below.

1.9.2 Chapter 2

In chapter 2, neuroimaging data from 54 different datasets were used to examine differences in brain anatomical asymmetries between people with ASD and healthy controls. Data were obtained via the ENIGMA ASD Working Group and were available from 1,774 individuals with ASD and 1,809 controls. Data included regional and total cortical thickness and surface area measures, as well as subcortical volume measures. For each of these measures, the AI was calculated, and a linear mixed regression model was fitted to test for associations with diagnosis. Effects of age, sex, IQ, severity and medication use on brain asymmetry were also examined, and various sensitivity analyses were also carried out. This was by far the largest ever study of brain asymmetry in ASD, and involved coordination with hundreds of scientific collaborators.

1.9.3 Chapter 3

In chapter 3, structural brain asymmetries in ADHD were examined using data from 39 datasets. These data were obtained via the ENIGMA ADHD Working group and comprised 1,736 individuals with ADHD and 1,654 controls. The data were divided into three age-groups (i.e., children, adolescents, adults) and analyses were conducted separately for each of these

groups. Again, the AI was calculated for each of the brain regional and total measures, and a linear mixed regression mega-analysis was performed on each of these measures separately. Additionally, effects of age, sex, IQ, comorbidity, severity and medication use were studied, and sensitivity analyses were performed. This was by far the largest ever study of brain asymmetry in ADHD, and again involved coordination with hundreds of collaborators.

Although we were interested in any commonalities in results from the studies of ASD and ADHD (see Chapter 6 for discussion), they were necessarily performed as separate studies, as the data were available from different teams of collaborators through distinct agreements.

1.9.4 Chapter 4

Chapter 4 describes the analysis of whole genome sequencing data from fifteen people with *situs inversus* (SI), of which nine did not have primary ciliary dyskinesia. Among these nine, five were left-handed, which raised the possibility that they carried rare, highly penetrant mutations affecting both brain and body laterality. Analysis of their genomes might therefore have helped to identify a developmental mechanism for left-right axis formation distinct from the classical ciliary-nodal pathway. Single nucleotide variants as well as larger structural genomic variations were investigated, using recessive and dominant genetic models and gene set enrichment analysis, and technical control data from 15 additional participants with *situs solitus* were also used. SI is a rare condition, affecting roughly 1 in 6,000 to 1 in 10,000 people, and this was the largest genome-sequencing study of the condition to date.

1.9.5 Chapter 5

In chapter 5, the genetic basis of brain functional asymmetry during the resting-state was examined in healthy individuals from the UK Biobank. Common genetic variation was investigated in 30,660 individuals with imaging and whole-genome imputed data, and rare variant data were studied in 16,949 individuals with imaging and whole exome sequencing data. Two measures of resting-state functional connectivity asymmetry were calculated that had previously shown associations with language dominance as assessed

with task fMRI (Joliot, Tzourio-Mazoyer, & Mazoyer, 2016; Labache et al., 2020). The first, called intra-hemispheric intrinsic connectivity asymmetry (HICAs_pop), was a hemisphere-wide index, whereas the second, called resting-state degree centrality asymmetry (Rs_DC_asymm), was based specifically on core regions of the language network (Labache et al., 2019).

1.9.6 Chapter 6

Chapter 6 summarizes the main findings of the experimental chapters (i.e., chapters 2 to 5) and discusses them in the context of their overall biological meaning, and their utility as the basis for potential directions of future research.

1.10 References

Abu-Rustum, R. S., Ziade, M. F., & Abu-Rustum, S. E. (2013). Reference values for the right and left fetal choroid plexus at 11 to 13 weeks: an early sign of "developmental" laterality? *J Ultrasound Med, 32*(9), 1623-1629. doi:10.7863/ultra.32.9.1623

Altarelli, I., Leroy, F., Monzalvo, K., Fluss, J., Billard, C., Dehaene-Lambertz, G., . . . Ramus, F. (2014). Planum temporale asymmetry in developmental dyslexia: Revisiting an old question. *Hum Brain Mapp, 35*(12), 5717-5735. doi:10.1002/hbm.22579

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). Washington, DC.

Annett, M. (1998). Handedness and cerebral dominance: the right shift theory. *J Neuropsychiatry Clin Neurosci, 10*(4), 459-469. doi:10.1176/jnp.10.4.459

Armour, J. A., Davison, A., & McManus, I. C. (2014). Genome-wide association study of handedness excludes simple genetic models. *Heredity (Edinb), 112*(3), 221-225. doi:10.1038/hdy.2013.93

Banks, S. J., Zhuang, X., Bayram, E., Bird, C., Cordes, D., Caldwell, J. Z. K., & Cummings, J. L. (2018). Default Mode Network Lateralization and Memory in Healthy Aging and Alzheimer's Disease. *J Alzheimers Dis, 66*(3), 1223-1234. doi:10.3233/jad-180541

Barth, K. A., Miklosi, A., Watkins, J., Bianco, I. H., Wilson, S. W., & Andrew, R. J. (2005). fsi zebrafish show concordant reversal of laterality of viscera, neuroanatomy, and a subset of behavioral responses. *Curr Biol, 15*(9), 844-850. doi:10.1016/j.cub.2005.03.047

Beking, T., Geuze, R. H., van Faassen, M., Kema, I. P., Kreukels, B. P. C., & Groothuis, T. G. G. (2018). Prenatal and pubertal testosterone affect brain lateralization. *Psychoneuroendocrinology, 88*, 78-91. doi:10.1016/j.psyneuen.2017.10.027

Biberacher, V., Schmidt, P., Keshavan, A., Boucard, C. C., Righart, R., Samann, P., . . . Muhlau, M. (2016). Intra- and interscanner variability of magnetic resonance imaging based volumetry in multiple sclerosis. *Neuroimage, 142*, 188-197. doi:10.1016/j.neuroimage.2016.07.035

Bisazza, A., Cantalupo, C., Robins, A., Rogers, L. J., & Vallortigara, G. (1997). Pawedness and motor asymmetries in toads. *Laterality, 2*(1), 49-64. doi:10.1080/713754252

Bishop, D. V. (2013). Cerebral asymmetry and language development: cause, correlate, or consequence? *Science, 340*(6138), 1230531. doi:10.1126/science.1230531

Bitar, R., Leung, G., Perng, R., Tadros, S., Moody, A. R., Sarrazin, J., . . . Roberts, T. P. (2006). MR pulse sequences: what every radiologist wants to know but is afraid to ask. *Radiographics, 26*(2), 513-537. doi:10.1148/rg.262055063

Bogen, J. E. (1969). The other side of the brain. II. An appositional mind. *Bull Los Angeles Neurol Soc, 34*(3), 135-162.

Bomba, L., Walter, K., & Soranzo, N. (2017). The impact of rare and low-frequency genetic variants in common disease. *Genome Biol, 18*(1), 77. doi:10.1186/s13059-017-1212-4

Brandler, W. M., & Paracchini, S. (2014). The genetic relationship between handedness and neurodevelopmental disorders. *Trends Mol Med, 20*(2), 83-90. doi:10.1016/j.molmed.2013.10.008

Broadhouse, K. (2019). The Physics of MRI and How We Use It to Reveal the Mysteries of the Mind. *Front. Young Minds*(7:23). doi:10.3389/frym.2019.00023

Broca, P. (1863). Localisations des fonctions cérébrales. Siège de la faculté du language articulé. Bulletin de la Société d'Anthropologie, 4, 200–208.

Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L. T., Sharp, K., . . . Marchini, J. (2018). The UK Biobank resource with deep phenotyping and genomic data. *Nature, 562*(7726), 203-209. doi:10.1038/s41586-018-0579-z

Carrion-Castillo, A., Pepe, A., Kong, X. Z., Fisher, S. E., Mazoyer, B., Tzourio-Mazoyer, N., . . . Francks, C. (2020). Genetic effects on planum temporale asymmetry and their limited relevance to neurodevelopmental disorders, intelligence or educational attainment. *Cortex, 124*, 137-153. doi:10.1016/j.cortex.2019.11.006

Concha, M. L., Burdine, R. D., Russell, C., Schier, A. F., & Wilson, S. W. (2000). A nodal signaling pathway regulates the laterality of neuroanatomical asymmetries in the zebrafish forebrain. *Neuron, 28*(2), 399-409. doi:10.1016/s0896-6273(00)00120-3

Concha, M. L., Signore, I. A., & Colombo, A. (2009). Mechanisms of directional asymmetry in the zebrafish epithalamus. *Semin Cell Dev Biol, 20*(4), 498-509. doi:10.1016/j.semcdb.2008.11.007

Corballis, M. C., & Haberling, I. S. (2017). The Many Sides of Hemispheric Asymmetry: A Selective Review and Outlook. *J Int Neuropsychol Soc, 23*(9-10), 710-718. doi:10.1017/S1355617717000376

Corey, D. M., Hurley, M. M., & Foundas, A. L. (2001). Right and left handedness defined: a multivariate approach using hand preference and hand performance measures. *Neuropsychiatry Neuropsychol Behav Neurol, 14*(3), 144-152.

Crow, T. J. (1990). Temporal-Lobe Asymmetries as the Key to the Etiology of Schizophrenia. *Schizophrenia Bulletin, 16*(3), 433-&.

Cuellar-Partida, G., Tung, J. Y., Eriksson, N., Albrecht, E., Aliev, F., Andreassen, O. A., . . . Medland, S. E. (2020). Genome-wide association study identifies 48 common genetic variants associated with handedness. *Nat Hum Behav*. doi:10.1038/s41562-020-00956-y

Dadda, M., Zandona, E., Agrillo, C., & Bisazza, A. (2009). The costs of hemispheric specialization in a fish. *Proc Biol Sci, 276*(1677), 4399-4407. doi:10.1098/rspb.2009.1406

Dassonville, P., Zhu, X. H., Uurbil, K., Kim, S. G., & Ashe, J. (1997). Functional activation in motor cortex reflects the direction and the degree of handedness. *Proc Natl Acad Sci U S A, 94*(25), 14015-14018. doi:10.1073/pnas.94.25.14015

de Bakker, P. I., Yelensky, R., Pe'er, I., Gabriel, S. B., Daly, M. J., & Altshuler, D. (2005). Efficiency and power in genetic association studies. *Nat Genet, 37*(11), 1217-1223. doi:10.1038/ng1669

de Kovel, C. G. F., Carrion-Castillo, A., & Francks, C. (2019). A large-scale population study of early life factors influencing left-handedness. *Sci Rep, 9*(1), 584. doi:10.1038/s41598-018-37423-8

de Kovel, C. G. F., & Francks, C. (2019). The molecular genetics of hand preference revisited. *Sci Rep, 9*(1), 5986. doi:10.1038/s41598-019-42515-0

de Kovel, C. G. F., Lisgo, S., Karlebach, G., Ju, J., Cheng, G., Fisher, S. E., & Francks, C. (2017). Left-Right Asymmetry of Maturation Rates in Human Embryonic Neural Development. *Biol Psychiatry, 82*(3), 204-212. doi:10.1016/j.biopsych.2017.01.016

de Kovel, C. G. F., Lisgo, S. N., Fisher, S. E., & Francks, C. (2018). Subtle left-right asymmetry of gene expression profiles in embryonic and foetal human brains. *Sci Rep, 8*(1), 12606. doi:10.1038/s41598-018-29496-2

de Zubicaray, G. I., Chiang, M. C., McMahon, K. L., Shattuck, D. W., Toga, A. W., Martin, N. G., . . . Thompson, P. M. (2008). Meeting the Challenges of Neuroimaging Genetics. *Brain Imaging Behav, 2*(4), 258-263. doi:10.1007/s11682-008-9029-0

Dehghan, A. (2018). Genome-Wide Association Studies. *Methods Mol Biol, 1793*, 37-49. doi:10.1007/978-1-4939-7868-7_4

Deng, H., Xia, H., & Deng, S. (2015). Genetic basis of human left-right asymmetry disorders. *Expert Rev Mol Med, 16*, e19. doi:10.1017/erm.2014.22

Doucet, G., Naveau, M., Petit, L., Delcroix, N., Zago, L., Crivello, F., . . . Joliot, M. (2011). Brain activity at rest: a multiscale hierarchical functional organization. *J Neurophysiol, 105*(6), 2753-2763. doi:10.1152/jn.00895.2010

Dougherty, C. C., Evans, D. W., Katuwal, G. J., & Michael, A. M. (2016). Asymmetry of fusiform structure in autism spectrum disorder: trajectory and association with symptom severity. *Mol Autism, 7*, 28. doi:10.1186/s13229-016-0089-5

Douglas, P. K., Gutman, B., Anderson, A., Larios, C., Lawrence, K. E., Narr, K., . . . Bookheimer, S. Y. (2018). Hemispheric brain asymmetry differences in youths with attention-deficit/hyperactivity disorder. *Neuroimage Clin, 18*, 744-752. doi:10.1016/j.nicl.2018.02.020

Duboc, V., Dufourcq, P., Blader, P., & Roussigne, M. (2015). Asymmetry of the Brain: Development and Implications. *Annu Rev Genet, 49*, 647-672. doi:10.1146/annurev-genet-112414-055322

Duncan, L. E., Ostacher, M., & Ballon, J. (2019). How genome-wide association studies (GWAS) made traditional candidate gene studies obsolete. *Neuropsychopharmacology, 44*(9), 1518-1523. doi:10.1038/s41386-019-0389-5

Elliott, L. T., Sharp, K., Alfaro-Almagro, F., Shi, S., Miller, K. L., Douaud, G., . . . Smith, S. M. (2018). Genome-wide association studies of brain imaging phenotypes in UK Biobank. *Nature, 562*(7726), 210-216. doi:10.1038/s41586-018-0571-7

Emser, T. S., Johnston, B. A., Steele, J. D., Kooij, S., Thorell, L., & Christiansen, H. (2018). Assessing ADHD symptoms in children and adults: evaluating the role of objective measures. *Behav Brain Funct, 14*(1), 11. doi:10.1186/s12993-018-0143-x

Esteves, M., Ganz, E., Sousa, N., & Leite-Almeida, H. (2020). Asymmetrical Brain Plasticity: Physiology and Pathology. *Neuroscience*. doi:10.1016/j.neuroscience.2020.01.022

Eyler, L. T., Pierce, K., & Courchesne, E. (2012). A failure of left temporal cortex to specialize for language is an early emerging and fundamental property of autism. *Brain, 135*, 949-960. doi:10.1093/brain/awr364

Fadista, J., Manning, A. K., Florez, J. C., & Groop, L. (2016). The (in)famous GWAS P-value threshold revisited and updated for low-frequency variants. *Eur J Hum Genet, 24*(8), 1202-1205. doi:10.1038/ejhg.2015.269

Feaster, D. J., Mikulich-Gilbertson, S., & Brincks, A. M. (2011). Modeling site effects in the design and analysis of multi-site trials. *Am J Drug Alcohol Abuse, 37*(5), 383-391. doi:10.3109/00952990.2011.600386

Fliegauf, M., Benzing, T., & Omran, H. (2007). When cilia go bad: cilia defects and ciliopathies. *Nat Rev Mol Cell Biol, 8*(11), 880-893. doi:10.1038/nrm2278

Francks, C. (2015). Exploring human brain lateralization with molecular genetics and genomics. *Ann N Y Acad Sci, 1359*, 1-13. doi:10.1111/nyas.12770

Frasnelli, E. (2013). Brain and behavioral lateralization in invertebrates. *Front Psychol, 4*, 939. doi:10.3389/fpsyg.2013.00939

Frasnelli, E., & Vallortigara, G. (2018). Individual-Level and Population-Level Lateralization: Two Sides of the Same Coin. *Symmetry-Basel, 10*(12). doi:ARTN 739

10.3390/sym10120739

Gazzaniga, M. S., Bogen, J. E., & Sperry, R. W. (1962). Some functional effects of sectioning the cerebral commissures in man. *Proc Natl Acad Sci U S A, 48*, 1765-1769. doi:10.1073/pnas.48.10.1765

Geschwind, N., & Levitsky, W. (1968). Human brain: left-right asymmetries in temporal speech region. *Science, 161*(3837), 186-187.

Ghirardi, L., Pettersson, E., Taylor, M. J., Freitag, C. M., Franke, B., Asherson, P., . . . Kuja-Halkola, R. (2019). Genetic and environmental contribution to the overlap between ADHD and ASD trait dimensions in young adults: a twin study. *Psychol Med, 49*(10), 1713-1721. doi:10.1017/s003329171800243x

Ghirlanda, S., Frasnelli, E., & Vallortigara, G. (2009). Intraspecific competition and coordination in the evolution of lateralization. *Philos Trans R Soc Lond B Biol Sci, 364*(1519), 861-866. doi:10.1098/rstb.2008.0227

Gingras, B., Honing, H., Peretz, I., Trainor, L. J., & Fisher, S. E. (2015). Defining the biological bases of individual differences in musicality. *Philos Trans R Soc Lond B Biol Sci, 370*(1664), 20140092. doi:10.1098/rstb.2014.0092

Glahn, D. C., Thompson, P. M., & Blangero, J. (2007). Neuroimaging endophenotypes: strategies for finding genes influencing brain structure and function. *Hum Brain Mapp, 28*(6), 488-501. doi:10.1002/hbm.20401

Glover, G. H. (2011). Overview of functional magnetic resonance imaging. *Neurosurg Clin N Am, 22*(2), 133-139, vii. doi:10.1016/j.nec.2010.11.001

Goodwin, S., McPherson, J. D., & McCombie, W. R. (2016). Coming of age: ten years of next-generation sequencing technologies. *Nat Rev Genet, 17*(6), 333-351. doi:10.1038/nrg.2016.49

Gotts, S. J., Jo, H. J., Wallace, G. L., Saad, Z. S., Cox, R. W., & Martin, A. (2013). Two distinct forms of functional lateralization in the human brain. *Proc Natl Acad Sci U S A, 110*(36), E3435-3444. doi:10.1073/pnas.1302581110

Grasby, K. L., Jahanshad, N., Painter, J. N., Colodro-Conde, L., Bralten, J., Hibar, D. P., . . . Medland, S. E. (2020). The genetic architecture of the human cerebral cortex. *Science, 367*(6484). doi:10.1126/science.aay6690

Gros, J., Feistel, K., Viebahn, C., Blum, M., & Tabin, C. J. (2009). Cell movements at Hensen's node establish left/right asymmetric gene expression in the chick. *Science, 324*(5929), 941-944. doi:10.1126/science.1172478

Grover, V. P., Tognarelli, J. M., Crossey, M. M., Cox, I. J., Taylor-Robinson, S. D., & McPhail, M. J. (2015). Magnetic Resonance Imaging: Principles and

Techniques: Lessons for Clinicians. *J Clin Exp Hepatol, 5*(3), 246-255. doi:10.1016/j.jceh.2015.08.001

Guadalupe, T., Kong, X.-Z., Akkermans, S. E. A., Fisher, S. E., & Francks, C. (2020). Relations between hemispheric asymmetries of grey matter and auditory processing of spoken syllables in 281 healthy adults. *bioRxiv*, 2020.2008.2012.247841. doi:10.1101/2020.08.12.247841

Guadalupe, T., Mathias, S. R., vanErp, T. G., Whelan, C. D., Zwiers, M. P., Abe, Y., . . . Francks, C. (2016). Human subcortical brain asymmetries in 15,847 people worldwide reveal effects of age and sex. *Brain Imaging Behav*. doi:10.1007/s11682-016-9629-z

Gunturkun, O., & Ocklenburg, S. (2017). Ontogenesis of Lateralization. *Neuron, 94*(2), 249-263. doi:10.1016/j.neuron.2017.02.045

Hamada, H., Meno, C., Watanabe, D., & Saijoh, Y. (2002). Establishment of vertebrate left-right asymmetry. *Nat Rev Genet, 3*(2), 103-113. doi:10.1038/nrg732

Hardie, S. M., Wright, L., & Clark, L. (2016). Handedness and social anxiety: Using Bryden's research as a catalyst to explore the influence of familial sinistrality and degree of handedness. *Laterality, 21*(4-6), 329-347. doi:10.1080/1357650x.2015.1131712

Hayden, E. C. (2014). Technology: The $1,000 genome. *Nature, 507*(7492), 294-295. doi:10.1038/507294a

Heilman, K. M., Bowers, D., Valenstein, E., & Watson, R. T. (1986). The right hemisphere: neuropsychological functions. *J Neurosurg, 64*(5), 693-704. doi:10.3171/jns.1986.64.5.0693

Hepper, P. G., McCartney, G. R., & Shannon, E. A. (1998). Lateralised behaviour in first trimester human foetuses. *Neuropsychologia, 36*(6), 531-534.

Hering-Hanit, R., Achiron, R., Lipitz, S., & Achiron, A. (2001). Asymmetry of fetal cerebral hemispheres: in utero ultrasound study. *Arch Dis Child Fetal Neonatal Ed, 85*(3), F194-196.

Hibar, D. P., Stein, J. L., Renteria, M. E., Arias-Vasquez, A., Desrivières, S., Jahanshad, N., . . . Medland, S. E. (2015). Common genetic variants influence human subcortical brain structures. *Nature, 520*(7546), 224-229. doi:10.1038/nature14101

Hirnstein, M., & Hugdahl, K. (2014). Excess of non-right-handedness in schizophrenia: meta-analysis of gender effects and potential biases in handedness assessment. *Br J Psychiatry, 205*(4), 260-267. doi:10.1192/bjp.bp.113.137349

Hirnstein, M., Hugdahl, K., & Hausmann, M. (2019). Cognitive sex differences and hemispheric asymmetry: A critical review of 40 years of research. *Laterality, 24*(2), 204-252. doi:10.1080/1357650x.2018.1497044

Holland, S. K., Vannest, J., Mecoli, M., Jacola, L. M., Tillema, J. M., Karunanayaka, P. R., . . . Byars, A. W. (2007). Functional MRI of language lateralization during development in children. *Int J Audiol, 46*(9), 533-551. doi:10.1080/14992020701448994

Holler-Wallscheid, M. S., Thier, P., Pomper, J. K., & Lindner, A. (2017). Bilateral recruitment of prefrontal cortex in working memory is associated with task demand but not with age. *Proc Natl Acad Sci U S A, 114*(5), E830-E839. doi:10.1073/pnas.1601983114

Hong, E. P., & Park, J. W. (2012). Sample size and statistical power calculation in genetic association studies. *Genomics Inform, 10*(2), 117-122. doi:10.5808/GI.2012.10.2.117

Hoogman, M., Bralten, J., Hibar, D. P., Mennes, M., Zwiers, M. P., Schweren, L. S., . . . Franke, B. (2017). Subcortical brain volume differences in participants with attention deficit hyperactivity disorder in children and adults: a cross-sectional mega-analysis. *Lancet Psychiatry, 4*(4), 310-319. doi:10.1016/s2215-0366(17)30049-4

Hoogman, M., Muetzel, R., Guimaraes, J. P., Shumskaya, E., Mennes, M., Zwiers, M. P., . . . Franke, B. (2019). Brain Imaging of the Cortex in ADHD: A Coordinated Analysis of Large-Scale Clinical and Population-Based Samples. *Am J Psychiatry*, appiajp201918091033. doi:10.1176/appi.ajp.2019.18091033

Hopkins, W. D. (2006). Comparative and familial analysis of handedness in great apes. *Psychol Bull, 132*(4), 538-559. doi:10.1037/0033-2909.132.4.538

Hopkins, W. D., Misiura, M., Pope, S. M., & Latash, E. M. (2015). Behavioral and brain asymmetries in primates: a preliminary evaluation of two evolutionary hypotheses. *Ann N Y Acad Sci, 1359*, 65-83. doi:10.1111/nyas.12936

Huang, J., Howie, B., McCarthy, S., Memari, Y., Walter, K., Min, J. L., . . . Soranzo, N. (2015). Improved imputation of low-frequency and rare variants using the UK10K haplotype reference panel. *Nat Commun, 6*, 8111. doi:10.1038/ncomms9111

Hugdahl, K., & Westerhausen, R. (2016). Speech processing asymmetry revealed by dichotic listening and functional brain imaging. *Neuropsychologia, 93*(Pt B), 466-481. doi:10.1016/j.neuropsychologia.2015.12.011

Hull, R., & Vaid, J. (2007). Bilingual language lateralization: a meta-analytic tale of two hemispheres. *Neuropsychologia, 45*(9), 1987-2008. doi:10.1016/j.neuropsychologia.2007.03.002

Husken, U., & Carl, M. (2013). The Wnt/beta-catenin signaling pathway establishes neuroanatomical asymmetries and their laterality. *Mech Dev, 130*(6-8), 330-335. doi:10.1016/j.mod.2012.09.002

Inaki, M., Liu, J., & Matsuno, K. (2016). Cell chirality: its origin and roles in left-right asymmetric development. *Philos Trans R Soc Lond B Biol Sci, 371*(1710). doi:10.1098/rstb.2015.0403

Inaki, M., Sasamura, T., & Matsuno, K. (2018). Cell Chirality Drives Left-Right Asymmetric Morphogenesis. *Front Cell Dev Biol, 6*, 34. doi:10.3389/fcell.2018.00034

International HapMap, C. (2003). The International HapMap Project. *Nature, 426*(6968), 789-796. doi:10.1038/nature02168

Jahanshad, N., Lee, A. D., Barysheva, M., McMahon, K. L., de Zubicaray, G. I., Martin, N. G., . . . Thompson, P. M. (2010). Genetic influences on brain asymmetry: a DTI study of 374 twins and siblings. *Neuroimage, 52*(2), 455-469. doi:10.1016/j.neuroimage.2010.04.236

Joliot, M., Tzourio-Mazoyer, N., & Mazoyer, B. (2016). Intra-hemispheric intrinsic connectivity asymmetry and its relationships with handedness and language Lateralization. *Neuropsychologia, 93*(Pt B), 437-447. doi:10.1016/j.neuropsychologia.2016.03.013

Josse, G., & Tzourio-Mazoyer, N. (2004). Hemispheric specialization for language. *Brain Res Brain Res Rev, 44*(1), 1-12. doi:10.1016/j.brainresrev.2003.10.001

Karlebach, G., & Francks, C. (2015). Lateralization of gene expression in human language cortex. *Cortex, 67*, 30-36. doi:10.1016/j.cortex.2015.03.003

Karolis, V. R., Corbetta, M., & Thiebaut de Schotten, M. (2019). The architecture of functional lateralisation and its relationship to callosal connectivity in the human brain. *Nat Commun, 10*(1), 1417. doi:10.1038/s41467-019-09344-1

Keenan, J. P., Thangaraj, V., Halpern, A. R., & Schlaug, G. (2001). Absolute pitch and planum temporale. *Neuroimage, 14*(6), 1402-1408. doi:10.1006/nimg.2001.0925

Kennedy, D. N., O'Craven, K. M., Ticho, B. S., Goldstein, A. M., Makris, N., & Henson, J. W. (1999). Structural and functional brain asymmetries in human situs inversus totalis. *Neurology, 53*(6), 1260-1265.

Kloppel, S., Vongerichten, A., van Eimeren, T., Frackowiak, R. S., & Siebner, H. R. (2007). Can left-handedness be switched? Insights from an early switch of handwriting. *J Neurosci, 27*(29), 7847-7853. doi:10.1523/JNEUROSCI.1299-07.2007

Kong, X.-Z., Postema, M., Carrión Castillo, A., Pepe, A., Crivello, F., Joliot, M., . . . Francks, C. (2020). Large-scale Phenomic and Genomic Analysis of Brain Asymmetrical Skew. *bioRxiv*, 756395. doi:10.1101/756395

Kong, X. Z., Mathias, S. R., Guadalupe, T., Glahn, D. C., Franke, B., Crivello, F., . . . Francks, C. (2018). Mapping cortical brain asymmetry in 17,141 healthy individuals worldwide via the ENIGMA Consortium. *Proc Natl Acad Sci U S A*. doi:10.1073/pnas.1718418115

Kong, X. Z., Postema, M. C., Guadalupe, T., de Kovel, C., Boedhoe, P. S. W., Hoogman, M., . . . Francks, C. (2020). Mapping brain asymmetry in health and disease through the ENIGMA consortium. *Hum Brain Mapp*. doi:10.1002/hbm.25033

Kundu, B., Rolston, J. D., & Grandhi, R. (2019). Mapping language dominance through the lens of the Wada test. *Neurosurg Focus, 47*(3), E5. doi:10.3171/2019.6.focus19346

Labache, L., Joliot, M., Saracco, J., Jobard, G., Hesling, I., Zago, L., . . . Tzourio-Mazoyer, N. (2019). A SENtence Supramodal Areas AtlaS (SENSAAS) based on multiple task-induced activation mapping and graph analysis of

intrinsic connectivity in 144 healthy right-handers. *Brain Struct Funct, 224*(2), 859-882. doi:10.1007/s00429-018-1810-2

Labache, L., Mazoyer, B., Joliot, M., Crivello, F., Hesling, I., & Tzourio-Mazoyer, N. (2020). Typical and atypical language brain organization based on intrinsic connectivity and multitask functional asymmetries. *eLife, 9*. doi:10.7554/eLife.58722

LaFramboise, T. (2009). Single nucleotide polymorphism arrays: a decade of biological, computational and technological advances. *Nucleic Acids Res, 37*(13), 4181-4193. doi:10.1093/nar/gkp552

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., . . . Szustakowki, J. (2001). Initial sequencing and analysis of the human genome. *Nature, 409*(6822), 860-921. doi:10.1038/35057062

Langel, J., Hakun, J., Zhu, D. C., & Ravizza, S. M. (2014). Functional specialization of the left ventral parietal cortex in working memory. *Front Hum Neurosci, 8*, 440. doi:10.3389/fnhum.2014.00440

Le Guen, Y., Auzias, G., Leroy, F., Noulhiane, M., Dehaene-Lambertz, G., Duchesnay, E., . . . Frouin, V. (2018). Genetic Influence on the Sulcal Pits: On the Origin of the First Cortical Folds. *Cereb Cortex, 28*(6), 1922-1933. doi:10.1093/cercor/bhx098

Le Guen, Y., Leroy, F., Philippe, C., Mangin, J. F., Dehaene-Lambertz, G., & Frouin, V. (2020). Enhancer Locus in ch14q23.1 Modulates Brain Asymmetric Temporal Regions Involved in Language Processing. *Cereb Cortex, 30*(10), 5322-5332. doi:10.1093/cercor/bhaa112

Lebreton, G., Geminard, C., Lapraz, F., Pyrpassopoulos, S., Cerezo, D., Speder, P., . . . Noselli, S. (2018). Molecular to organismal chirality is induced by the conserved myosin 1D. *Science, 362*(6417), 949-952. doi:10.1126/science.aat8642

Leitner, Y. (2014). The co-occurrence of autism and attention deficit hyperactivity disorder in children - what do we know? *Front Hum Neurosci, 8*, 268. doi:10.3389/fnhum.2014.00268

Leroy, F., Cai, Q., Bogart, S. L., Dubois, J., Coulon, O., Monzalvo, K., . . . Dehaene-Lambertz, G. (2015). New human-specific brain landmark: the depth asymmetry of superior temporal sulcus. *Proc Natl Acad Sci U S A, 112*(4), 1208-1213. doi:10.1073/pnas.1412389112

Letzkus, P., Ribi, W. A., Wood, J. T., Zhu, H., Zhang, S. W., & Srinivasan, M. V. (2006). Lateralization of olfaction in the honeybee Apis mellifera. *Curr Biol, 16*(14), 1471-1476. doi:10.1016/j.cub.2006.05.060

Levy, B., & Burnside, R. D. (2019). Are all chromosome microarrays the same? What clinicians need to know. *Prenat Diagn, 39*(3), 157-164. doi:10.1002/pd.5422

Li, X., Morgan, P. S., Ashburner, J., Smith, J., & Rorden, C. (2016). The first step for neuroimaging data analysis: DICOM to NIfTI conversion. *J Neurosci Methods, 264*, 47-56. doi:10.1016/j.jneumeth.2016.03.001

Liu, L., Li, Y., Li, S., Hu, N., He, Y., Pong, R., . . . Law, M. (2012). Comparison of next-generation sequencing systems. *J Biomed Biotechnol, 2012*, 251364. doi:10.1155/2012/251364

Logue, D. D., Logue, R. T., Kaufmann, W. E., & Belcher, H. M. (2015). Psychiatric disorders and left-handedness in children living in an urban environment. *Laterality, 20*(2), 249-256. doi:10.1080/1357650X.2014.961927

Lopes, S. S., Lourenco, R., Pacheco, L., Moreno, N., Kreiling, J., & Saude, L. (2010). Notch signalling regulates left-right asymmetry through ciliary length control. *Development, 137*(21), 3625-3632. doi:10.1242/dev.054452

Lust, J. M., Geuze, R. H., Van de Beek, C., Cohen-Kettenis, P. T., Bouma, A., & Groothuis, T. G. (2011). Differential effects of prenatal testosterone on lateralization of handedness and language. *Neuropsychology, 25*(5), 581-589. doi:10.1037/a0023293

Lust, J. M., Geuze, R. H., Van de Beek, C., Cohen-Kettenis, P. T., Groothuis, A. G., & Bouma, A. (2010). Sex specific effect of prenatal testosterone on language lateralization in children. *Neuropsychologia, 48*(2), 536-540. doi:10.1016/j.neuropsychologia.2009.10.014

Luys, J. B. (1881). Recherches nouvelles sur les hémiplégies émotives. . *Encéphale, 1*, 644-646.

Maguire, E. A., Woollett, K., & Spiers, H. J. (2006). London taxi drivers and bus drivers: a structural MRI and neuropsychological analysis. *Hippocampus, 16*(12), 1091-1101. doi:10.1002/hipo.20233

Marchini, J., & Howie, B. (2010). Genotype imputation for genome-wide association studies. *Nat Rev Genet, 11*(7), 499-511. doi:10.1038/nrg2796

Markou, P., Ahtam, B., & Papadatou-Pastou, M. (2017). Elevated Levels of Atypical Handedness in Autism: Meta-Analyses. *Neuropsychol Rev, 27*(3), 258-283. doi:10.1007/s11065-017-9354-4

Martinho, A., 3rd, Burns, Z. T., von Bayern, A. M., & Kacelnik, A. (2014). Monocular tool control, eye dominance, and laterality in New Caledonian crows. *Curr Biol, 24*(24), 2930-2934. doi:10.1016/j.cub.2014.10.035

Matsui, T., & Bessho, Y. (2012). Left-right asymmetry in zebrafish. *Cell Mol Life Sci, 69*(18), 3069-3077. doi:10.1007/s00018-012-0985-6

Mazoyer, B., Zago, L., Jobard, G., Crivello, F., Joliot, M., Perchey, G., . . . Tzourio-Mazoyer, N. (2014). Gaussian mixture modeling of hemispheric lateralization for language in a large sample of healthy individuals balanced for handedness. *PLoS One, 9*(6), e101165. doi:10.1371/journal.pone.0101165

Mazzio, E. A., & Soliman, K. F. (2012). Basic concepts of epigenetics: impact of environmental signals on gene expression. *Epigenetics, 7*(2), 119-130. doi:10.4161/epi.7.2.18764

McManus, I. C. (1985). Right- and left-hand skill: failure of the right shift model. *Br J Psychol, 76 (Pt 1)*, 1-34.

McManus, I. C., Martin, N., Stubbings, G. F., Chung, E. M., & Mitchison, H. M. (2004). Handedness and situs inversus in primary ciliary dyskinesia. *Proc Biol Sci, 271*(1557), 2579-2582. doi:10.1098/rspb.2004.2881

Medland, S. E., Duffy, D. L., Wright, M. J., Geffen, G. M., Hay, D. A., Levy, F., . . . Boomsma, D. I. (2009). Genetic influences on handedness: data from 25,732 Australian and Dutch twin families. *Neuropsychologia, 47*(2), 330-337. doi:10.1016/j.neuropsychologia.2008.09.005

Minkova, L., Habich, A., Peter, J., Kaller, C. P., Eickhoff, S. B., & Kloppel, S. (2017). Gray matter asymmetries in aging and neurodegeneration: A review and meta-analysis. *Hum Brain Mapp, 38*(12), 5890-5904. doi:10.1002/hbm.23772

Mitchell, K. (2018). Innate: How the Wiring of Our Brains Shapes Who We Are. *Princeton University Press*. doi:doi:10.2307/j.ctvc77m71

Munafo, M. R., & Flint, J. (2010). How reliable are scientific studies? *Br J Psychiatry, 197*(4), 257-258. doi:10.1192/bjp.bp.109.069849

Nakamura, T., & Hamada, H. (2012). Left-right patterning: conserved and divergent mechanisms. *Development, 139*(18), 3257-3262. doi:10.1242/dev.061606

Neale, B. (2017). Heritability of >2,000 traits and disorders in the UK Biobank.

Neugebauer, J. M., & Yost, H. J. (2014). FGF signaling is required for brain left-right asymmetry and brain midline formation. *Dev Biol, 386*(1), 123-134. doi:10.1016/j.ydbio.2013.11.020

Noli, S. A., Baini, I., Parazzini, F., Mauri, P. A., Vignali, M., Gerli, S., . . . Cipriani, S. (2019). Preterm Birth, Low Gestational Age, Low Birth Weight, Parity, and Other Determinants of Breech Presentation: Results from a Large Retrospective Population-Based Study. *Biomed Res Int, 2019*, 9581439. doi:10.1155/2019/9581439

Nonaka, S., Shiratori, H., Saijoh, Y., & Hamada, H. (2002). Determination of left-right patterning of the mouse embryo by artificial nodal flow. *Nature, 418*(6893), 96-99. doi:10.1038/nature00849

Nowakowska, C., Sachs, G. S., Zarate, C. A., Jr., Marangell, L. B., Calabrese, J. R., Goldberg, J. F., & Ketter, T. A. (2008). Increased rate of non-right-handedness in patients with bipolar disorder. *J Clin Psychiatry, 69*(5), 866-867.

Ocklenburg, S., Schmitz, J., Moinfar, Z., Moser, D., Klose, R., Lor, S., . . . Gunturkun, O. (2017). Epigenetic regulation of lateralized fetal spinal gene expression underlies hemispheric asymmetries. *eLife, 6*. doi:10.7554/eLife.22784

Oertel, V., Knochel, C., Rotarska-Jagiela, A., Schonmeyer, R., Lindner, M., van de Ven, V., . . . Linden, D. E. (2010). Reduced laterality as a trait marker of schizophrenia--evidence from structural and functional neuroimaging. *J Neurosci, 30*(6), 2289-2299. doi:10.1523/jneurosci.4575-09.2010

Ogawa, S., Lee, T. M., Kay, A. R., & Tank, D. W. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proc Natl Acad Sci U S A, 87*(24), 9868-9872. doi:10.1073/pnas.87.24.9868

Okada, Y., Nonaka, S., Tanaka, Y., Saijoh, Y., Hamada, H., & Hirokawa, N. (1999). Abnormal nodal flow precedes situs inversus in iv and inv mice. *Mol Cell, 4*(4), 459-468.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia, 9*(1), 97-113.

Ornstein, R. E. (1972). *The psychology of consciousness*. Oxford, England: Penguin.

Peters, M., Reimers, S., & Manning, J. T. (2006). Hand preference for writing and associations with selected demographic and behavioral variables in 255,100 subjects: the BBC internet study. *Brain Cogn, 62*(2), 177-189. doi:10.1016/j.bandc.2006.04.005

Petersen, B. S., Fredrich, B., Hoeppner, M. P., Ellinghaus, D., & Franke, A. (2017). Opportunities and challenges of whole-genome and -exome sequencing. *BMC Genet, 18*(1), 14. doi:10.1186/s12863-017-0479-5

Pulst, S. M. (1999). Genetic linkage analysis. *Arch Neurol, 56*(6), 667-672. doi:10.1001/archneur.56.6.667

Raemaekers, M., Schellekens, W., Petridou, N., & Ramsey, N. F. (2018). Knowing left from right: asymmetric functional connectivity during resting state. *Brain Struct Funct, 223*(4), 1909-1922. doi:10.1007/s00429-017-1604-y

Regan, J. C., Concha, M. L., Roussigne, M., Russell, C., & Wilson, S. W. (2009). An Fgf8-dependent bistable cell migratory event establishes CNS asymmetry. *Neuron, 61*(1), 27-34. doi:10.1016/j.neuron.2008.11.030

Renteria, M. E. (2012). Cerebral asymmetry: a quantitative, multifactorial, and plastic brain phenotype. *Twin Res Hum Genet, 15*(3), 401-413. doi:10.1017/thg.2012.13

Robins, A., & Rogers, L. J. (2006). Complementary and lateralized forms of processing in Bufo marinus for novel and familiar prey. *Neurobiol Learn Mem, 86*(2), 214-227. doi:10.1016/j.nlm.2006.03.002

Roe, J. M., Vidal-Piñeiro, D., Sørensen, Ø., Brandmaier, A. M., Düzel, S., Gonzalez, H. A., . . . Westerhausen, R. (2020). Asymmetric thinning of the cerebral cortex across the adult lifespan is accelerated in Alzheimer's Disease. *bioRxiv*, 2020.2006.2018.158980. doi:10.1101/2020.06.18.158980

Rogers, L., & Vallortigara, G. (2015). When and Why Did Brains Break Symmetry? *Symmetry, 7*(4), 2181-2194. doi:10.3390/sym7042181

Rogers, L. J., Zucca, P., & Vallortigara, G. (2004). Advantages of having a lateralized brain. *Proc Biol Sci, 271 Suppl 6*, S420-422. doi:10.1098/rsbl.2004.0200

Sa, M. J., Ruela, C., & Madeira, M. D. (2007). Dendritic right/left asymmetries in the neurons of the human hippocampal formation: a quantitative Golgi

study. *Arq Neuropsiquiatr, 65*(4b), 1105-1113. doi:10.1590/s0004-282x2007000700003

Sanger, F., & Coulson, A. R. (1975). A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J Mol Biol, 94*(3), 441-448. doi:10.1016/0022-2836(75)90213-2

Schmitz, J., Gunturkun, O., & Ocklenburg, S. (2019). Building an Asymmetrical Brain: The Molecular Perspective. *Front Psychol, 10*, 982. doi:10.3389/fpsyg.2019.00982

Sha, Z., Schijven, D., Carrion-Castillo, A., Joliot, M., Mazoyer, B., Fisher, S. E., . . . Francks, C. (2020). The genetic architecture of structural left-right asymmetry of the human brain. *bioRxiv*, 2020.2006.2030.179721. doi:10.1101/2020.06.30.179721

Shaw, P., Lalonde, F., Lepage, C., Rabin, C., Eckstrand, K., Sharp, W., . . . Rapoport, J. (2009). Development of cortical asymmetry in typically developing children and its disruption in attention-deficit/hyperactivity disorder. *Arch Gen Psychiatry, 66*(8), 888-896. doi:10.1001/archgenpsychiatry.2009.103

Staub, M. E. (2016). The other side of the brain: The politics of split-brain research in the 1970s-1980s. *Hist Psychol, 19*(4), 259-273. doi:10.1037/hop0000035

Tanaka, S., Kanzaki, R., Yoshibayashi, M., Kamiya, T., & Sugishita, M. (1999). Dichotic listening in patients with situs inversus: brain asymmetry and situs asymmetry. *Neuropsychologia, 37*(7), 869-874.

Thiebaut de Schotten, M., Dell'Acqua, F., Forkel, S. J., Simmons, A., Vergani, F., Murphy, D. G., & Catani, M. (2011). A lateralized brain network for visuospatial attention. *Nat Neurosci, 14*(10), 1245-1246. doi:10.1038/nn.2905

Thompson, P. M., Jahanshad, N., Ching, C. R. K., Salminen, L. E., Thomopoulos, S. I., Bright, J., . . . Zelman, V. (2020). ENIGMA and global neuroscience: A decade of large-scale studies of the brain in health and disease across more than 40 countries. *Transl Psychiatry, 10*(1), 100. doi:10.1038/s41398-020-0705-1

Thompson, P. M., Moussai, J., Zohoori, S., Goldkorn, A., Khan, A. A., Mega, M. S., . . . Toga, A. W. (1998). Cortical variability and asymmetry in normal aging and Alzheimer's disease. *Cereb Cortex, 8*(6), 492-509.

Toga, A. W., & Thompson, P. M. (2003). Mapping brain asymmetry. *Nat Rev Neurosci, 4*(1), 37-48. doi:10.1038/nrn1009

Vallortigara, G. (2006). The evolutionary psychology of left and right: costs and benefits of lateralization. *Dev Psychobiol, 48*(6), 418-427. doi:10.1002/dev.20166

Vallortigara, G., & Rogers, L. J. (2005). Survival with an asymmetrical brain: advantages and disadvantages of cerebral lateralization. *Behav Brain Sci, 28*(4), 575-589; discussion 589-633. doi:10.1017/S0140525X05000105

Van Agtmael, T., Forrest, S. M., & Williamson, R. (2002). Parametric and non-parametric linkage analysis of several candidate regions for genes for human handedness. *Eur J Hum Genet, 10*(10), 623-630. doi:10.1038/sj.ejhg.5200851

van Dijk, E. L., Jaszczyszyn, Y., Naquin, D., & Thermes, C. (2018). The Third Revolution in Sequencing Technology. *Trends Genet, 34*(9), 666-681. doi:10.1016/j.tig.2018.05.008

van Rooij, D., Anagnostou, E., Arango, C., Auzias, G., Behrmann, M., Busatto, G. F., . . . Buitelaar, J. K. (2018). Cortical and Subcortical Brain Morphometry Differences Between Patients With Autism Spectrum Disorder and Healthy Individuals Across the Lifespan: Results From the ENIGMA ASD Working Group. *Am J Psychiatry, 175*(4), 359-369. doi:10.1176/appi.ajp.2017.17010100

Vingerhoets, G. (2019). Phenotypes in hemispheric functional segregation? Perspectives and challenges. *Phys Life Rev, 30*, 1-18. doi:10.1016/j.plrev.2019.06.002

Vingerhoets, G., Li, X., Hou, L., Bogaert, S., Verhelst, H., Gerrits, R., . . . Roberts, N. (2018). Brain structural and functional asymmetry in human situs inversus totalis. *Brain Struct Funct*. doi:10.1007/s00429-017-1598-5

Vojinovic, D., Adams, H. H., Jian, X., Yang, Q., Smith, A. V., Bis, J. C., . . . Fornage, M. (2018). Genome-wide association study of 23,500 individuals identifies 7 loci associated with brain ventricular volume. *Nat Commun, 9*(1), 3945. doi:10.1038/s41467-018-06234-w

Wernicke, C. (1874). Der Aphasische Symptomencomplex. Eine psychologische Studie auf anatomische Basis. Breslau: Cohn and Wiegert.

Westerhausen, R., & Kompus, K. (2018). How to get a left-ear advantage: A technical review of assessing brain asymmetry with dichotic listening. *Scand J Psychol, 59*(1), 66-73. doi:10.1111/sjop.12408

Wiberg, A., Ng, M., Al Omran, Y., Alfaro-Almagro, F., McCarthy, P., Marchini, J., . . . Furniss, D. (2019). Handedness, language areas and neuropsychiatric diseases: insights from brain imaging and genetics. *Brain*. doi:10.1093/brain/awz257

Wiper, M. L. (2017). Evolutionary and mechanistic drivers of laterality: A review and new synthesis. *Laterality, 22*(6), 740-770. doi:10.1080/1357650X.2017.1291658

Wyczesany, M., Capotosto, P., Zappasodi, F., & Prete, G. (2018). Hemispheric asymmetries and emotions: Evidence from effective connectivity. *Neuropsychologia, 121*, 98-105. doi:10.1016/j.neuropsychologia.2018.10.007

Yaakub, S. N., Heckemann, R. A., Keller, S. S., McGinnity, C. J., Weber, B., & Hammers, A. (2020). On brain atlas choice and automatic segmentation methods: a comparison of MAPER & FreeSurfer using three atlas databases. *Sci Rep, 10*(1), 2837. doi:10.1038/s41598-020-57951-6

Yegnasubramanian, S. (2013). Explanatory chapter: next generation sequencing. *Methods Enzymol, 529*, 201-208. doi:10.1016/B978-0-12-418687-3.00016-1

Youssofzadeh, V., & Babajani-Feremi, A. (2019). Mapping critical hubs of receptive and expressive language using MEG: A comparison against fMRI. *Neuroimage, 201*, 116029. doi:10.1016/j.neuroimage.2019.116029

Yucel, K., McKinnon, M., Chahal, R., Taylor, V., Macdonald, K., Joffe, R., & MacQueen, G. (2009). Increased subgenual prefrontal cortex size in remitted patients with major depressive disorder. *Psychiatry Research-Neuroimaging, 173*(1), 71-76. doi:10.1016/j.pscychresns.2008.07.013

Yucel, M., Stuart, G. W., Maruff, P., Wood, S. J., Savage, G. R., Smith, D. J., . . . Pantelis, C. (2002). Paracingulate morphologic differences in males with established schizophrenia: a magnetic resonance imaging morphometric study. *Biol Psychiatry, 52*(1), 15-23.

Yucel, M., Wood, S. J., Phillips, L. J., Stuart, G. W., Smith, D. J., Yung, A., . . . Pantelis, C. (2003). Morphology of the anterior cingulate cortex in young men at ultra-high risk of developing a psychotic illness. *British Journal of Psychiatry, 182*, 518-524. doi:DOI 10.1192/bjp.182.6.518

Zhao, B., Zhang, J., Ibrahim, J. G., Luo, T., Santelli, R. C., Li, Y., . . . Zhu, H. (2019). Large-scale GWAS reveals genetic architecture of brain white matter microstructure and genetic overlap with cognitive and mental health traits (n = 17,706). *Mol Psychiatry*. doi:10.1038/s41380-019-0569-z

CHAPTER

Altered structural brain asymmetry in autism spectrum disorder in a study of 54 datasets

2

Abstract:

Altered structural brain asymmetry in autism spectrum disorder (ASD) has been reported. However, findings have been inconsistent, likely due to limited sample sizes. Here we investigated 1,774 individuals with ASD and 1,809 controls, from 54 independent datasets of the ENIGMA consortium. ASD was significantly associated with alterations of cortical thickness asymmetry in mostly medial frontal, orbitofrontal, cingulate and inferior temporal areas, and also with asymmetry of orbitofrontal surface area. These differences generally involved reduced asymmetry in individuals with ASD compared to controls. Furthermore, putamen volume asymmetry was significantly increased in ASD. The largest case-control effect size was Cohen's d=-0.13, for asymmetry of superior frontal cortical thickness. Most effects did not depend on age, sex, IQ, severity or medication use. Altered lateralized neurodevelopment may therefore be a feature of ASD, affecting widespread brain regions with diverse functions. Large-scale analysis was necessary to quantify subtle alterations of brain structural asymmetry in ASD.

Key words: autism, brain asymmetry, brain laterality, mega-analysis, structural imaging, cortical thickness

---

This chapter is published as: Postema, M.C., van Rooij, D., Anagnostou, E. *et al.* Altered structural brain asymmetry in autism spectrum disorder in a study of 54 datasets. *Nat Commun* **10,** 4958 (2019). https://doi.org/10.1038/s41467-019-13005-8

2.1 Introduction

Autism spectrum disorder (ASD) is an umbrella diagnosis, capturing several previously separate pervasive developmental disorders with various levels of symptom severity, including Autistic Disorder, Asperger's Syndrome, Childhood Disintegrative Disorder, and Pervasive Developmental Disorder – Not Otherwise Specified (PDD-NOS) (American Psychiatric Association, 2013). According to the Diagnostic and Statistical Manual of Mental Disorders (DSM) version 5, diagnosis of ASD requires the presence of at least three symptoms of impaired social communication and at least two symptoms of repetitive behaviours or restricted interests (American Psychiatric Association, 2013). ASD has a median prevalence of 1 out of 161 individuals in a study of worldwide data (Elsabbagh et al., 2012), with a higher diagnosis rate in some developed countries such as the United States(Xu et al., 2018).

Characterizing the neurobiology of ASD may eventually lead to improved diagnosis and clinical subgrouping, and the development of individually targeted treatment programs (Loth, Murphy, & Spooren, 2016). While much of the neurobiology of ASD remains unknown, subtle alterations of brain structure appear to be involved (reviewed in (Li, Karnath, & Xu, 2017; Rommelse, Buitelaar, & Hartman, 2017)). These include differences in total brain volume (children with ASD have shown a larger average volume (Courchesne et al., 2001; Courchesne et al., 2007; Lucibello et al., 2019; Retico et al., 2016)), as well as alterations of the medial and inferior frontal, anterior cingulate, superior temporal, and orbitofrontal cortices, and the caudate nucleus (Carlisi, Norman, Lukito, et al., 2017; Li et al., 2017; Rommelse et al., 2017). However, the results of structural magnetic resonance imaging (MRI) studies of ASD have often been inconsistent, potentially due to 1) small study sample sizes in relation to subtle effects, 2) differences across studies in terms of clinical characteristics, age, comorbidity and medication use, 3) methodological differences between studies, such as differences in hardware, software and distinct data processing pipelines (Biberacher et al., 2016), and 4) the etiological and neurobiological heterogeneity of ASD, which exists as a group of different syndromes rather than a single entity (Carlisi, Norman, Lukito, et al., 2017; Jeste & Geschwind, 2014).

In the ENIGMA (Enhancing Neuro-Imaging Genetics through Meta-Analysis) consortium (http://enigma.ini.usc.edu), researchers from around the world collaborate to analyse many separate datasets jointly, and to reduce some of the technical heterogeneity by using harmonized protocols for MRI data processing. A recent study by the ENIGMA consortium's ASD working group showed small average differences in bilateral cortical and subcortical brain measures between 1,571 cases and 1,650 healthy controls, in the largest study of brain structure in ASD yet performed (van Rooij et al., 2018). Relative to controls, ASD patients had significantly lower volumes of several subcortical structures, as well as greater thickness in various cortical regions - mostly in the frontal lobes - and lower thickness of temporal regions. No associations of diagnosis with regional cortical surface areas were found (van Rooij et al., 2018).

Left-right asymmetry is an important aspect of human brain organization, which may be altered in various psychiatric and neurocognitive conditions, including schizophrenia, dyslexia and ASD (Duboc, Dufourcq, Blader, & Roussigne, 2015; Renteria, 2012; Toga & Thompson, 2003). On a functional level, people with ASD demonstrate reduced leftward language lateralization more frequently than controls (Kleinhans, Muller, Cohen, & Courchesne, 2008; Knaus et al., 2010; Lindell & Hudry, 2013). Resting-state functional MRI data of people with ASD have also shown a generally rightward shift of asymmetry involving various functional networks of brain regions (Cardinale, Shih, Fishman, Ford, & Muller, 2013). In addition, people with ASD have a higher rate of left-handedness than the general population (Lindell & Hudry, 2013; Markou, Ahtam, & Papadatou-Pastou, 2017; Rysstad & Pedersen, 2018). Furthermore, an electroencephalography study reported that infants at high risk for ASD showed more rightward than leftward frontal alpha asymmetry at rest (Gabard-Durnam, Tierney, Vogel-Farley, Tager-Flusberg, & Nelson, 2015).

Brain structural imaging studies have also reported altered hemispheric asymmetry in ASD. Diffusion imaging studies indicated reduced asymmetry of a variety of different white matter tract metrics (Carper, Treiber, DeJesus, & Muller, 2016; Conti et al., 2016; Joseph et al., 2014), although in one study males with ASD lacked an age-dependent decrease in rightward asymmetry

of network global efficiency, compared to controls (Wei, Zhong, Nie, & Gong, 2018). A structural MRI study investigating grey matter reported lower leftward volume asymmetry of language-related cortical regions in ASD (i.e., *planum temporale*, Heschl's gyrus, posterior supramarginal gyrus and parietal operculum), as well as greater rightward asymmetry of the inferior parietal lobule (Floris et al., 2016). The volume and surface area of the fusiform gyrus also showed lower rightward asymmetry in ASD (Dougherty, Evans, Katuwal, & Michael, 2016). However, other studies did not find alterations of grey matter asymmetries in ASD (Joseph et al., 2014; Knaus, Tager-Flusberg, Mock, Dauterive, & Foundas, 2012).

Prior studies of structural brain asymmetry in ASD had sample sizes less than 128 cases and 127 controls. The previous ENIGMA consortium study of ASD (van Rooij et al., 2018) did not perform analyses of brain asymmetry, but reported bilateral effects only as strong as Cohen's d = -0.21 (for entorhinal thickness bilaterally) (van Rooij et al., 2018). Comparable bilateral effect sizes were also found in ENIGMA consortium studies of other disorders (Boedhoe et al., 2018; Boedhoe et al., 2017; Hoogman et al., 2017; Logue et al., 2018; Schmaal et al., 2017; Schmaal et al., 2016; van Erp et al., 2018; van Rooij et al., 2018). If effects on brain asymmetry are similarly subtle, then prior studies of this aspect of brain structure in ASD were likely underpowered. Low power not only reduces the chance of detecting true effects, but also the likelihood that a statistically significant result reflects a true effect (Button et al., 2013; Munafo & Flint, 2010). Therefore large scale analysis is needed to determine whether, and how, structural brain asymmetry might be altered in ASD, to better describe the neurobiology of the condition.

Here, we made use of MRI data from 54 datasets that were collected across the world by members of the ENIGMA consortium's ASD Working Group, to perform the first highly-powered study of structural brain asymmetry in ASD. Using a single, harmonized protocol for image analysis, we derived asymmetry indexes, AI= (Left-Right)/(Left+Right), for multiple brain regional and global hemispheric measures, in up to 1,778 individuals with ASD and 1,829 typically developing controls. The AI is a widely used index in brain asymmetry studies (Kurth, Gaser, & Luders, 2015; Leroy et al., 2015).

Age and sex are known to affect cortical-(Kong et al., 2018) as well as subcortical asymmetries(Guadalupe et al., 2016) in healthy individuals. In addition, a recent structural imaging study of roughly 500 individuals with ASD, and 800 controls, found that case-control differences of bilateral cortical thickness were greater in younger versus older individuals, while alsbeing related to ASD symptom severity, and with larger differences in individuals with lower versus higher full-scale IQ scores(Bedford et al., 2019). Other previous case-control MRI findings with respect to these indicators of clinical heterogeneity in ASD are also reviewed in that paper(Bedford et al., 2019). In the present study, we therefore carried out secondary analyses in which we tested brain asymmetries in relation to age- or sex-specific effects, IQ, and disorder severity. We also included an exploratory analysis of medication use.

2.2 Methods

2.2.1 Datasets

Structural MRI data were available for 57 different datasets (**Table 1**). Three datasets comprising either cases only, or controls only, were removed in this study (**Table 1**), as our analysis model included random intercepts for 'dataset' (below), and diagnosis was fully confounded with dataset for these three. The remaining 54 datasets comprised 1,778 people with ASD (N = 1,504 males; median age = 13 years; range = 2 to 64 years) and 1,829 typically developing controls (N = 1,400 males; median age = 13 years; range = 2 to 64 years).

All datasets were collected during the period when DSM-IV and DSM-IV-TR were the common classification systems, between 1994 and 2013, and the clinical diagnosis of ASD was made according to DSM-IV criteria (American Psychiatric Association, 2000). The datasets were collected in a variety of different countries, and intended originally as separate studies. Nonetheless, all subjects were diagnosed based on clinical diagnosis by a clinically experienced and board certified physician/psychiatrist/psychologist. This was a criterion for admission of a dataset into the ENIGMA-ASD database. For

each of the 54 datasets, all relevant ethical regulations were complied with, and appropriate informed consent was obtained for all individuals.

Total scores from the Autism Diagnostic Observation Schedule-Generic (ADOS), a standardized instrument commonly used in autism diagnosis (Lord et al., 2000), were available for a majority of cases (N=878). Cases from the entire ASD spectrum were included, but only 66 cases had IQ below 70 (cases: mean IQ = 104, SD = 19, min = 34, max = 149; see **Figure S1D**). The presence/absence of comorbid conditions had been recorded for 519 of the cases, but only 54 cases showed at least one comorbid condition (which could be ADHD, OCD, depression, anxiety, and/or Tourette's syndrome (van Rooij et al., 2018)). Numbers related to DSM-IV subtypes of ASD were not collated by the ENIGMA ASD working group, as this subtyping scheme has been dropped from DSM-V due to low reliability (Lord & Bishop, 2015).

There was not a homogeneous assessment/recruitment process for controls across the 54 datasets, but the overwhelming majority were typically developing/healthy at the time of MRI, and no controls showed features that might have met criteria for a diagnosis of ASD. Only 19 controls had IQ less than 70. In these subjects the exclusion of ASD diagnosis was performed by a senior child psychiatrist/physician. Eighteen of these were from the FSM dataset and were clinically diagnosed with idiopathic intellectual disability. Amongst all controls the mean IQ was 112, SD = 15, min = 31, max = 149; see **Figure S1D**.

2.2.2 Structural magnetic resonance imaging

Structural T1-weighted brain MRI scans were acquired at each study site. As shown in **Table 1**, images were acquired using different field strengths (1.5 T or 3 T) and scanner types. Each site used harmonized protocols from the ENIGMA consortium (http://enigma.ini.usc.edu/protocols/imaging-protocols) for data processing and quality control. The data used in the current study were thickness and surface area measures for each of 34 bilaterally paired cortical regions, as defined with the Desikan-Killiany atlas (Desikan et al., 2006), as well as the average cortical thickness and total surface area per entire hemisphere. In addition, left and right volumes of seven bilaterally

paired subcortical structures, plus the lateral ventricles, were analyzed. Cortical parcellations and subcortical segmentations were performed with the freely available and validated software FreeSurfer (versions 5.1 or 5.3) (Fischl, 2012), using the default 'recon-all' pipeline, which also incorporates renormalization. Parcellations of cortical grey matter regions and segmentations of subcortical structures were visually inspected following the standardized ENIGMA quality control protocol (http://enigma.ini.usc.edu/protocols/imaging-protocols). Exclusions on the basis of this quality control resulted in the sample sizes mentioned above (see Datasets). Briefly, cortical segmentations were overlayed on the T1 image of each subject. Web pages were generated with snapshots from internal slices, as well as external views of the segmentation from different angles. All sites were provided with the manual on how to judge these images, including the most common segmentation errors. For subcortical structures, the protocol again consisted of visually checking the individual images, plotted from a set of internal slices. Volume estimates derived from poorly segmented structures (i.e. where tissue labels were assigned incorrectly) were excluded from each site's datasets and subsequent analyses. In addition, any data points exceeding 1.5 times the interquartile range, as defined per site and diagnostic group, were visually inspected (3D). When identified as error, all values from the affected regions were excluded from further analysis.

2.2.3 Asymmetry measures

Separately for each structural measure and individual subject, left (L) and right (R) data were used in R (version 3.5.3) to calculate an asymmetry index (AI) with the following formula: $\mathrm{AI} = (L - R)/(L + R)$. Distributions of each of the AIs are plotted in **Figure S2**. Note that AIs do not necessarily scale with L, R, or brain size, due to their denominators.

2.2.4 Linear mixed effects random-intercept model mega-analysis

2.2.4.1 Model

Linear mixed effects models were fitted separately for each cortical regional surface and thickness AI, as well as the total hemispheric surface area and mean thickness AI, and the subcortical volume AIs. This was accomplished by means of mega-analysis incorporating data from all 54 datasets, using the 'nlme' package in R (Pinheiro J). All models included the same fixed- and random effects, and had the following formulation:

$$\mathrm{AI} = \mathrm{diagnosis} + \mathrm{age} + \mathrm{sex} + \mathrm{random}\ (= \mathrm{dataset})$$

where AI reflects the AI of a given brain structure, and diagnosis ('controls' (= reference), 'ASD'), sex ('males' (= reference), 'females') and dataset were coded as factor variables, with dataset having 54 different categories. Age was coded as a numeric variable.

The Maximum Likelihood (ML) method was used to fit the models. Subjects were omitted if data were missing for any of the predictor variables (method = na.omit). The *ggplot2* package in R was used to visualize residuals (**Figures S3-S5**). Collinearity of predictor variables was assessed using the 'usdm' package in R (version 3.5.3.).

2.2.4.2 Significance

Significance was assessed based on the P-values for the effects of diagnosis on AIs. The False Discovery Rate (FDR)(Benjamini & Hochberg, 1995) was estimated separately for the 35 cortical surface area AIs (i.e., 34 regional AIs and one hemispheric total AI) and the 35 cortical thickness AIs, and again for the seven subcortical structures plus lateral ventricles, each time with a FDR threshold of 0.05. Correlations between AI measures were calculated using Pearson's R and visualized using the 'corrplot' package in R (**Figures S6-S8**). Most pairwise correlations between AIs were low, with only 33/78 pairwise correlations either lower than -0.3 or greater than 0.3, with a minimum R = -0.351 between the inferior parietal surface area AI and supramarginal surface area AI, and maximum R = 0.487 between the cuneus surface area AI and pericalcarine surface area AI.

2.2.4.3 Cohen's d effect sizes

The *t*-statistic for the factor 'diagnosis' in each linear mixed effects model was used to calculate Cohen's *d* (Cohen, 1988), with

$$d = \frac{t * (\mathrm{n1}+\mathrm{n2})}{\sqrt{(\mathrm{n1}*\mathrm{n2})} * \sqrt{df}} \qquad (1)$$

where n1 and n2 are the number of cases and controls, and df the degrees of freedom.

The latter was derived from the lme summary table in R, but can also be calculated using $df = \mathrm{obs} - (x1 + x2)$, where obs equals the number of observations, x1 the number of groups and x2 the number of factors in the model.

The 95% confidence intervals for Cohen's *d* were calculated using $95\%\ \mathrm{CI} = d \pm 1.96\mathrm{SE}$, with the standard error (SE) around Cohen's *d* calculated according to:

$$\mathrm{SE} = \sqrt{\frac{\mathrm{n1}+\mathrm{n2}}{\mathrm{n1}*\mathrm{n2}} + \frac{d^2}{2*(\mathrm{n1}+\mathrm{n2}-2)}} \qquad (2)$$

For visualization of cerebral cortical results, Cohen's *d* values were loaded into Matlab (version R2016a), and 3D images of left hemisphere inflated cortical and subcortical structures were obtained using FreeSurfer-derived ply files.

2.2.4.4 Power analyses

As each linear model included multiple predictor variables, the power to detect an effect of diagnosis on AI could not be computed exactly, but we obtained an indication of the effect size that would be needed to provide 80% power, had we been using simple t-tests and Bonferroni correction for multiple testing, using the 'pwr' command in R. For this purpose, a significance level of 0.0014 (i.e. 0.05/35) was set in the context of multiple testing over the regional and total cortical surface area AIs (N = 35) or thickness AIs (N = 35), and 0.00625 (i.e., 0.05/8) for seven subcortical

volume plus lateral ventricle AIs (N = 8). This showed that a difference of roughly Cohen's $d = 0.13$ would be detectable with 80% power in the cortical analyses, and Cohen's $d = 0.12$ in the subcortical analyses.

2.2.4.5 Sensitivity analyses

No outliers were removed for the primary analysis, but to confirm that results were not dependent on outliers, all analyses were repeated after having winsorized using a threshold of k = 3, for each AI measure separately. That is to say, the two highest and two lowest values were assigned the value of the third highest or lowest value respectively, separately per AI. This threshold was chosen after visual inspection of frequency histograms.

The relationships between AIs and age showed no overt non-linearity (**Figures S9-S11**), so no polynomials for age were incorporated in the models for primary analysis. However, analyses were repeated using an additional non-linear term for age, to check whether this choice had affected the results. As Age and Age^2 are highly correlated, we made use of the poly()-function in R for these two predictors, which created a pair of uncorrelated variables to model age effects (so-called orthogonal polynomials) (Chambers & Hastie, 1992), where one variable was linear and one non-linear.

As our data included participants as young as 1.5 years of age, and segmentation of very young brains might be especially challenging for the FreeSurfer algorithms, we also repeated our primary analysis excluding all individuals aged below 6 years (N=64 controls, N=113 cases), to assess whether they might have impacted the findings substantially (although these had passed the same quality control procedures as all other datasets, and FreeSurfer segmentation in preschoolers is generally of good quality, even before visual QC (Gori et al., 2015).).

As adults aged over 40 years were relatively sparsely represented, we also repeated the primary analysis after removing any individuals aged ≥ 40, in case modelling of age as a continuous predictor might have been unduly affected by these individuals. (In addition, see below for further analysis of age, for the purposes of subset and interaction analyses).

Finally, we repeated the primary analysis using only the subset of 3T acquired data (45 out of 54 datasets), to test for possible sensitivity to this technical variable. The sample was reduced from 1778 cases and 1829 controls in the primary analysis to 1467 cases and 1574 controls in the 3T-only analysis.

2.2.4.6 Directions of asymmetry changes

For any AIs showing significant effects of diagnosis in the primary analysis, linear mixed effects modelling was also performed on the corresponding L and R measures separately, to understand the unilateral changes involved. The models included the same terms as were used in the main analysis of AIs (i.e., diagnosis, age and sex as fixed factors, and dataset as random factor). Again, the Cohen's *d* effect sizes for diagnosis were calculated based on the *t*-statistics. The raw mean AI values were calculated separately in controls and cases, to describe the reference direction of healthy asymmetry in controls, and whether cases showed reduced, increased, or reversed asymmetry relative to controls.

2.2.4.7 Age- or sex-specific effects

For all AIs, we carried out secondary analyses including $\text{age} * \text{diagnosis}$ and $\text{sex} * \text{diagnosis}$ interaction terms, in separate models. The models were as follows: $\text{AI} = \text{diagnosis} + \text{age} + \text{sex} + \text{age} * \text{diag} + \text{random} (= \text{dataset})$, and $\text{AI} = \text{diagnosis} + \text{age} + \text{sex} + \text{sex} * \text{diag} + \text{random} (= \text{dataset})$.

In addition, we separated the data into two subsets by age, i.e. children < 18 years and adults ≥ 18 years (using the same criteria as van Rooij *et al.* 2018), or else by sex (males, females). Models were then fitted separately for each AI within each subset, i.e. within each age subset $\text{AI} = \text{diagnosis} + \text{sex} + \text{random} (= \text{dataset})$, and within each sex subset $\text{AI} = \text{diagnosis} + \text{age} + \text{random} (= \text{dataset})$.

2.2.4.8 Analysis of IQ

For each AI that showed a significant effect of diagnosis in the primary analysis, we carried out exploratory analyses of IQ in cases and controls

separately, whereby IQ (as a continuous variable) was considered as a predictor variable for the AI, so that AI = IQ + age + sex + random (= dataset). This was done to understand whether individual differences in asymmetry might relate to IQ, and whether such relations might be specific to ASD.

2.2.4.9 Analysis of Autism Diagnostic Observation Schedule (ADOS) severity score

For each AI that showed a significant effect of diagnosis in the primary analysis, a within-case-only analysis was performed incorporating symptom severity based on ADOS score as a predictor variable for AI: AI = ADOS + age + sex + random (= dataset). This was to understand whether the observed asymmetry changes in cases were dependent on ASD severity. ADOS scores were first adjusted using $\log_{10}$ transformation to reduce skewing.

2.2.4.10 Analysis of medication use

Data on medication use (i.e., current use of psychiatric treatment drugs prescribed for ASD or comorbid psychiatric conditions) was available for 832 individuals with ASD, of which 214 were categorized as medication users. For each AI that showed a significant effect of diagnosis in the primary analysis, a linear mixed model analysis was performed within-cases only, AI = medication + age + sex + random (= dataset). 'Medication' was coded as a binary variable (0=no medication, 2=medication).

2.3 Results

2.3.1 Significant associations of ASD with brain asymmetry

Summary information for the datasets is in **Table 1**. Out of a total of 78 structural AIs that were investigated (**Tables S1-S3**), 10 showed a significant effect of diagnosis which survived multiple testing correction (**Table 2**). Among these were seven regional cortical thickness AIs, including frontal regions (superior frontal, rostral middle frontal, medial orbitofrontal),

temporal regions (fusiform, inferior temporal), and cingulate regions (rostral anterior, isthmus cingulate). Two cortical regional surface area AIs, namely of the medial- and lateral orbitofrontal cortex, were significantly associated with diagnosis (medial: β = 0.006, *t* = 3.2, *P* = 0.0015; lateral: β = -0.005, *t* = -3.3, *P* = 0.0010) (**Table 2, Table S2**), as well as one subcortical volume AI, namely that of the putamen (β = 0.00395, t = 3.4, *P* = 0.00069) (**Table 2, Table S3**).

Nominally significant effects of diagnosis on AIs (i.e., with *P* < 0.05, but which did not survive multiple comparison correction), were observed for the fusiform surface area AI (β = -0.005, *t* = -2.56, *P* = 0.010) (**Table S2**), pars orbitalis thickness AI (β = -0.003, *t* = -2.26, *P* = 0.024), posterior cingulate thickness AI (β = -0.003, *t* = -2.1, *P* = 0.034), superior temporal thickness (β = -0.002, *t* = -1.97, *P* = 0.049), and caudate nucleus volume (β = 0.003, *t* = 2.24, *P* = 0.025).

Table 1. Characteristics of the different datasets of the ENIGMA ASD working group.

Sample name	N total	N cases (M/F)	N controls (M/F)	Median age in years (range)	Scanner type	FS[1]
ABIDE_CALTECH	31	13/1	13/4	23.4 (17.5, 56.2)	Siemens Trio	3T
ABIDE_KKI	21	7/0	11/3	10.6 (8.4, 12.8)	Philips Achieva	3T
ABIDE_LEUVEN_1	29	14/0	15/0	22 (18, 32)	Philips Interna	3T
ABIDE_LEUVEN_2	35	12/3	15/5	14.2 (12.1, 16.9)	Philips Interna	3T
ABIDE_MAX_MUN	57	21/3	29/4	26 (7, 58)	Siemens Verio	3T
ABIDE_NYU	186	68/11	81/26	13.6 (6.5, 39.1)	Siemens Allegra	3T
ABIDE_OHSU	18	7/0	11/0	10 (8.2, 12.7)	Siemens Trio	3T
ABIDE_OLIN	36	17/3	14/2	17 (10, 24)	Siemens Allegra	3T
ABIDE_PITT	58	26/5	23/4	17 (9.3, 35.2)	Siemens Allegra	3T
ABIDE_SBL	30	15/0	15/0	33.5 (20, 64)	Philips Interna	3T
ABIDE_SDSU	37	14/1	16/6	14.8 (8.7, 37.7)	GE MR750	3T
ABIDE_STANFORD	40	16/4	16/4	9.4 (7.5, 12.9)	GR Signa	3T
ABIDE_TCD	55	24/1	30/0	15.9 (9.3, 25.9)	Philips Achieva	3T
ABIDE_UM_1	130	50/14	43/23	12.3 (8.1, 20.9)	GE Signa	3T

Sample name	N total	N cases (M/F)	N controls (M/F)	Median age in years (range)	Scanner type	FS[1]
ABIDE_UM_2	31	14/1	15/1	14.8 (11.1, 26.8)	GE Signa Siemens	3T
ABIDE_USM	101	59/0	42/0	19.6 (8.2, 50.2)	Trio Siemens	3T
ABIDE_YALE	55	20/8	19/8	12.8 (7, 17.8)	Magnetom Philips	3T
ABIDEII-BNI	58	29/0	29/0	43 (18, 64)	Ingenia	3T
ABIDEII-EMC	54	22/5	22/5	8.2 (6.2, 10.7)	GE MR750 Philips	3T
ABIDEII-ETH	37	13/0	24/0	22.3 (13.8, 30.7)	Achieva Siemens	3T
ABIDEII-GU	106	43/8	28/27	10.6 (8.1, 13.9)	TriTim Siemens	3T
ABIDEII-IP	56	14/8	12/22	18.4 (6.1, 46.6)	TriTim Philips	1.5T
ABIDEII-IU	40	16/4	15/5	22 (17, 54)	Achieva Philips	3T
ABIDEII-KKI	211	41/15	99/56	10.3 (8, 13)	Achieva Philips	3T
ABIDEII-KUL*	28	28/0	-	24 (18, 35)	Achieva Siemens	3T
ABIDEII-NYU_1	78	43/5	28/2	8.4 (5.2, 34.8)	Allegra Siemens	3T
ABIDEII-NYU_2*	27	24/3	-	7 (5.1, 8.8)	Allegra Siemens	3T
ABIDEII-OHSU	93	30/7	27/29	11 (7, 15)	Skyra Siemens	3T
ABIDEII-OILH	59	20/4	20/15	23 (18, 31)	TriTim	3T
ABIDEII-SDSU	57	26/7	22/2	13 (7.4, 18)	GE MR750 Philips	3T
ABIDEII-TCD	43	21/0	22/0	14.8 (10, 20)	Achieva Siemens	3T
ABIDEII-UCD	32	14/4	10/4	14.6 (12, 17.8)	TriTim Siemens	3T
ABIDEII-UCLA	32	15/1	11/5	9.6 (7.8, 15)	TriTim Siemens	3T
ABIDEII-USM	33	15/2	13/3	19.7 (9.1, 38.9)	TriTim Siemens	3T
Barcelona	76	39/4	32/1	12.3 (7.2, 17.1)	Trio GE Signa	3T
BRC	52	19/0	33/0	15 (10, 18)	HDx Siemens	3T
CMU	27	11/3	10/3	27 (19, 40)	Magnetom Siemens	3T
Dresden	45	18/3	20/4	31.2 (21.1, 56.8)	Trio Siemens	3T
FAIR	85	36/7	27/15	11.6 (7.2, 15.9)	Magnetom Siemens	3T
FRANKFURT	27	10/2	13/2	18 (18, 18)	Sonata	1.5T
FSM	80	20/20	20/20	4.1(1.8, 6)	GE Signa	1.5T

Sample name	N total	N cases (M/F)	N controls (M/F)	Median age in years (range)	Scanner type	FS[1]
MRC	148	74/0	74/0	25 (18, 45)	GE Signa HDx	3T
MYAD	73	59/0	14/0	4.5 (1.5, 9)	Siemens symphony	1.5T
NIJMEGEN1_1	33	14/3	14/2	15.1 (12.3, 18)	Siemens Trio	3T
NIJMEGEN1_2*	9	-	8/1	17.5(13.4, 18.5)	Siemens Trio	3T
NIJMEGEN2	72	29/19	16/8	26 (18, 40)	Siemens Avanto	1.5T
NIJMEGEN3	95	37/4	45/9	9.7 (6.1, 12.3)	Siemens Avanto	1.5T
ParelladaHGGM	66	33/2	30/1	13 (7, 18)	Philips Intera	1.5T
PITT_1	56	11/3	34/8	15 (8, 36)	Siemens Allegra	3T
PITT_2	90	39/6	39/6	15 (8, 36)	Siemens Allegra	3T
SAOPAULO	35	15/0	20/0	11 (6, 19)	Philps	3T
TCD_1	50	28/0	22/0	14.9 (10, 21.8)	Philips Achieva	3T
TCD_2	39	17/0	22/0	15.9 (12, 25.9)	Philips Achieva	3T
TORONTO_1	219	82/23	57/57	11.9 (3.3, 20.8)	Siemens Trio	3T
TORONTO_2	203	107/43	29/24	11 (2.5, 21.7)	Siemens Trio	3T
UMCU_1	88	41/7	37/3	12.1 (7.1, 24.7)	Philips	1.5T
UMCU_2	9	6/0	2/1	11.8 (10, 12.6)	Philips	1.5T
Total	**3671**	**1833**	**1838**			

*Excluded, as diagnosis would be completely confounded with random variable 'data set' in the analysis model. [1]FS = Field Strength.

Table 2. Linear mixed model results for regional AIs that survived multiple comparisons correction in the primary analysis.

AI region	N cases/ controls	*t*-value diag	*t*-value age	*t*-value sex	*P*-value diag	*P*-value age	*P*-value sex	Cohen's d (95% CI)
Fusiform thickness	1704/1767	3.20	-0.17	2.06	0.001	0.862	0.040	0.109 (0.04,0.18)
Inferior temporal thickness	1703/1768	2.97	0.07	0.73	0.003	0.948	0.467	0.102 (0.03,0.17)
Isthmus cingulate thickness	1699/1769	-2.58	2.81	1.02	0.010	0.005	0.310	-0.088 (-0.16,-0.02)

AI region	N cases/ controls	*t*-value diag	*t*-value age	*t*-value sex	*P*-value diag	*P*-value age	*P*-value sex	Cohen's d (95% CI)
Medial orbitofrontal thickness	1705/1769	-3.47	4.23	-0.31	0.001	$2.36 \cdot 10^{-5}$	0.758	-0.119 (-0.19,-0.05)
Rostral anterior cingulate thickness	1700/1765	-3.37	-0.69	1.45	0.001	0.491	0.147	-0.116 (-0.18,-0.05)
Rostral middle frontal thickness	1707/1769	-2.94	-2.87	1.47	0.003	0.004	0.141	-0.101 (-0.17,-0.03)
Superior frontal thickness	1706/1771	-3.92	0.48	-1.38	$8.86 \cdot 10^{-5}$	0.630	0.169	-0.134 (-0.2,-0.07)
Lateral orbitofrontal surface area	1701/1762	-3.28	0.81	-0.07	0.001	0.421	0.946	-0.112 (-0.18,-0.05)
Medial orbitofrontal surface area	1704/1763	3.17	-0.27	-0.87	0.002	0.787	0.385	0.109 (0.04,0.18)
Putamen	1712/1763	3.40	5.03	1.77	0.001	$5.21 \cdot 10^{-7}$	0.076	0.116 (0.05,0.18)

[a] Unadjusted P values are shown. Note that this table only includes AIs for which the effect of diagnosis was significant after FDR correction. Results for all AIs are in supplementary information (Supplementary Tables 1-3).

2.3.2 Sensitivity analyses

When we repeated the analysis after winsorizing outliers, the pattern of results remained the same (**Tables S4-S6**), except that a small change in *P* value for the effect of diagnosis on medial orbitofrontal surface area AI meant that it no longer survived FDR correction (**Table S5**).

When we added a non-linear effect for age, all of the 10 AIs that had shown significant effects of diagnosis in the primary analysis remained significant (**Tables S4-S6**).

When we excluded all individuals below 6 years of age, that may have been more difficult for FreeSurfer to segment, all AIs that had shown significant effects of diagnosis in the primary analysis remained significant, except for the isthmus cingulate thickness AI (**Tables S4-S6**). In addition, one

new association with diagnosis, of the fusiform surface area AI, now surpassed the multiple testing correction threshold. These subtle changes of *P* values do not necessarily indicate that exclusion of younger ages improved signal to noise in the data.

When excluding all individuals aged 40 years or older, the pattern of significant results stayed the same (**Tables S4-S6**).

Finally, when analyzing only the subset of 3T-acquired data, two of the diagnosis effects from the primary analysis (i.e., inferior temporal- and isthmus cingulate thickness AI) were no longer significant after false discovery rate correction, but three other effects now became significant (i.e., superior temporal thickness AI, fusiform surface area AI, and caudate nucleus AI) (**Tables S4-S6**). Again, slight changes in significance levels are expected when changing the sample, and do not necessarily indicate systematic differences of 3T and 1.5T data with respect to case-control asymmetry differences.

2.3.3 Magnitudes and directions of asymmetry changes

Cohen's *d* effect sizes of the associations between AIs and diagnosis, as derived from the primary analysis, are visualized in **Figure 1**. Effect sizes were low, ranging from -0.13 (superior frontal thickness AI) to 0.12 (Putamen AI) (**Table 2**, **Tables S1-S3**). All of the cortical AIs with significant effects of diagnosis in the primary analysis showed decreased asymmetry in ASD compared to controls, i.e. the AIs were closer to zero in individuals with ASD than in controls, regardless of whether the region was on average leftward or rightward asymmetrical in controls (**Table 3**). However, the putamen showed increased asymmetry in ASD (mean AI controls = 0.011, mean AI cases = 0.012) (**Table 3**).

Five of the seven significant changes in regional cortical thickness asymmetry involved left sided decreases accompanied by right sided increases of thickness (**Table 3**). For the other two significant effects on regional thickness asymmetry (the fusiform and inferior temporal cortex), thickness was decreased bilaterally in ASD, but more so in the right than the

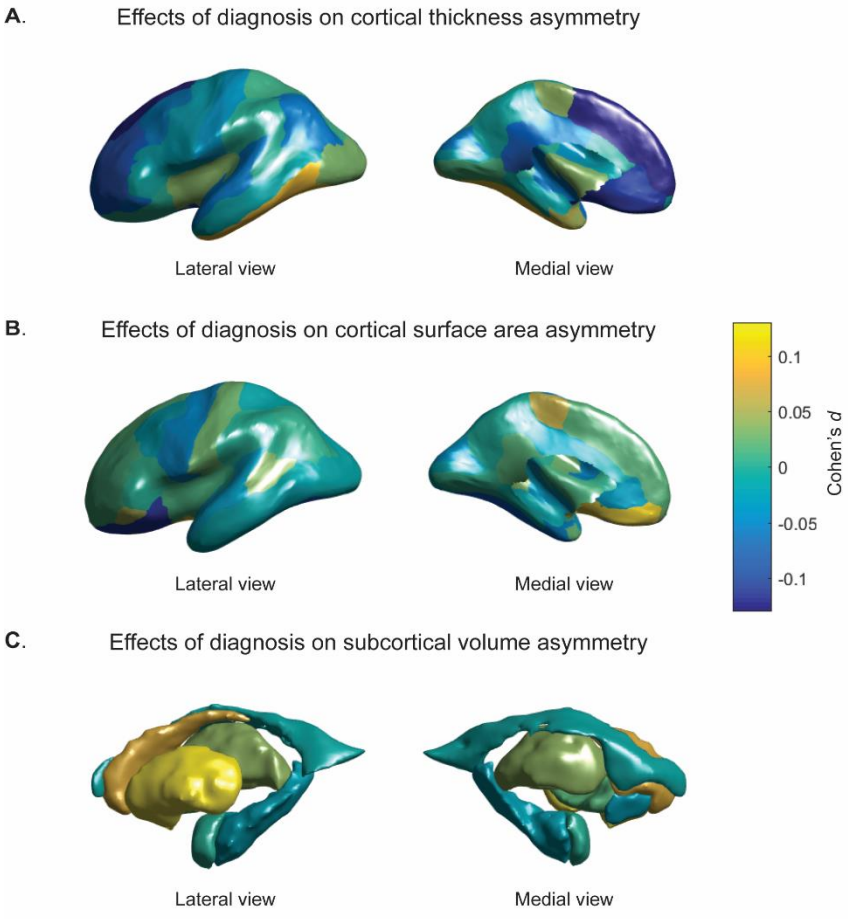


Figure 1. Cohen's *d* effect sizes of the associations between diagnosis and AIs. (**A**) regional cortical thickness measures, (**B**) cortical surface areas, (**C**) subcortical volumes. Values are overlaid on left hemisphere inflated brains. Positive Cohen's *d* values (*yellow*) indicate mean shifts towards greater leftward or reduced rightward asymmetry in cases, and negative Cohen's *d* values (*blue*) indicate mean shifts towards greater rightward asymmetry or reduced leftward asymmetry in individuals with ASD.

left hemisphere. For the significant changes in surface area asymmetry (lateral orbitofrontal and medial orbitofrontal cortex), surface area was altered in opposite directions in ASD in the two hemispheres, thus resulting in altered asymmetry (**Table 3**). Finally, the putamen showed a bilateral decrease in

volume in ASD that was more pronounced on the right, resulting in altered asymmetry (**Table 3**).

Table 3. Directions of asymmetry changes in cases versus controls

AI region	mean AI ± SD in controls	mean AI ± SD in ASD	Cohen's *d* (95% CI) Left hemisphere	Cohen's *d* (95% CI) Right hemisphere	Controls	ASD
Fusiform thickness	-0.004 ± 0.02	-0.002 ± 0.03	-0.152 (-0.22, -0.09)	-0.22 (-0.29, -0.15)	Rightward	Decreased
Inferior temporal thickness	-0.006 ± 0.03	-0.003 ± 0.03	-0.13 (-0.2, -0.06)	-0.194 (-0.26, -0.13)	Rightward	Decreased
Isthmus cingulate thickness	0.012 ± 0.04	0.007 ± 0.04	-0.014 (-0.08, 0.05)	0.07 (0, 0.14)	Leftward	Decreased
Medial orbitofrontal thickness	0.003 ± 0.04	-0.001 ± 0.04	-0.009 (-0.08, 0.06)	0.087 (0.02, 0.15)	Leftward	Reversed
Rostral anterior cingulate thickness	0.016 ± 0.05	0.011 ± 0.05	-0.116 (-0.18, -0.05)	0.009 (-0.06, 0.08)	Leftward	Decreased
Rostral middle frontal thickness	0.01 ± 0.03	0.008 ± 0.03	-0.016 (-0.08, 0.05)	0.046 (-0.02, 0.11)	Leftward	Decreased
Superior frontal thickness	0.006 ± 0.02	0.005 ± 0.02	-0.058 (-0.12, 0.01)	0.018 (-0.05, 0.08)	Leftward	Decreased
Lateral orbitofrontal surface area	0.014 ± 0.05	0.009 ± 0.05	-0.027 (-0.09, 0.04)	0.032 (-0.04, 0.1)	Leftward	Decreased
Medial orbitofrontal surface area	-0.012 ± 0.06	-0.004 ± 0.06	0.076 (0.01, 0.14)	-0.001 (-0.07, 0.07)	Rightward	Decreased
Putamen	0.011 ± 0.03	0.012 ± 0.04	-0.029 (-0.1, 0.04)	-0.095 (-0.16, -0.03)	Leftward	Increased

The raw means and standard deviations are indicated for each AI that showed a significant effect of diagnosis in the primary analysis. Also the Cohen's *d* effect sizes for left and right hemispheric measures are indicated (i.e., when left or right hemispheric measures were analyzed as dependent variables). In addition, the average direction of asymmetry in controls (derived from the raw mean AI) and its change in ASD is shown. Positive AI values indicate leftward asymmetry, negative AI values indicate rightward asymmetry.

2.3.4 Age or sex interaction effects

The distributions of age and sex across all datasets are plotted in Figure S1. In secondary analysis of interaction effects, there was only one significant

sex:diagnosis interaction effect after FDR correction, for the rostral anterior cingulate thickness AI (**Tables S7-S9**). This AI had shown a significant effect of diagnosis in the primary analysis. In analysis within the sexes separately, this AI was associated with diagnosis in males ($P = 1.4 \times 10^{-5}$) but not females ($p = 0.165$) (**Table S7**). For all of the AIs which showed significant effects of diagnosis in the primary analysis, adding sex:diagnosis interaction terms did not change the pattern of significant main effects of diagnosis, after FDR correction (**Tables S7-S9**).

There were no significant age:diagnosis interaction effects after FDR correction (**Tables S10-S12**). In general, for AIs which showed significant effects of diagnosis in the primary analysis, adding age:diagnosis interaction terms largely reduced the significance of the main effects of diagnosis, even though the age:diagnosis interaction terms were not significant (all $P > 0.05$) (**Tables S10-S12**). However, adding these interaction terms also increased the AIC and BIC scores compared to the primary analysis models without these terms, indicating poorer model fit when including these non-significant interaction terms (**Tables S10-S12**).

2.3.5 Exploratory analysis of IQ

The distributions of IQ within individuals with ASD and controls are shown in Figure S2. Out of the 10 AIs which showed significant case-control differences in the primary analysis, only one showed an association with IQ within individuals with ASD (uncorrected $P < 0.05$; **Table S13**). This was the rostral anterior cingulate thickness AI ($\beta = 0.00019$, $t = 2.49$, $p = 0.013$). The positive direction of this effect indicates that primarily those ASD individuals with lower IQ show reduced leftward asymmetry of the rostral anterior cingulate thickness. This regional asymmetry had also shown a significant sex*diagnosis interaction (see above). For this specific regional AI, we therefore added age:IQ, sex:IQ and age:sex:IQ interactions to the model, but none of these terms were significant (all uncorrected $P > 0.05$).

Within controls, only the superior frontal thickness AI was associated with IQ at uncorrected $P<0.05$ (**Table S13**) ($\beta = -0.00012$, $t = -3.41$, $p = 0.001$).This effect suggests that controls with lower IQ show

relatively increased asymmetry of superior frontal thickness, although this was *post hoc*, exploratory analysis without multiple testing correction.

2.3.6 ADOS severity scores

The distributions of ADOS severity scores are plotted in Figure S2. Out of the AIs that showed significant case-control differences in the primary analysis, only the isthmus cingulate thickness AI showed an association (uncorrected $P < 0.05$) with the ADOS severity score ($\beta = 0.0041$, $t = 2.6$, $p = 0.011$) (**Table S14**). The positive direction of the effect suggests that primarily cases with low ASD severity have reduced leftward asymmetry of this regional thickness.

2.3.7 Medication use

We found no significant effects of medication use (all uncorrected $P > 0.05$) (**Table S15**).

2.4 Discussion

In this, the largest study to date of brain asymmetry in ASD, we mapped differences in brain asymmetry between participants with ASD and controls, in a collection of 54 international datasets via the ENIGMA Consortium. We had 80% statistical power to detect Cohen's *d* effect sizes in the range of 0.12 to 0.13. We found significantly altered asymmetries of seven regional cortical thickness asymmetries in ASD compared to controls, predominantly involving medial frontal, orbitofrontal, inferior temporal and cingulate regions. The magnitude of all regional thickness asymmetries was decreased in ASD compared to controls, whether it was reduced leftward, reduced rightward, or reversed average asymmetry. Rightward asymmetry of the medial orbitofrontal surface area was also decreased in individuals with ASD, as was leftward asymmetry of the lateral orbitofrontal surface area. In addition, individuals with ASD showed an increase in leftward asymmetry of putamen volume, compared to controls.

Previous MRI studies of cerebral cortical asymmetries in ASD, based on much smaller datasets, and using diverse methods for image analysis, suggested variable case-control differences (Dougherty et al., 2016; Floris et al., 2016), or no differences (Joseph et al., 2014; Knaus et al., 2012). Our findings partly support a previously reported, generalized reduction of leftward asymmetry (Floris et al., 2016), as six of the nine significantly altered cortical regional asymmetries (thickness or surface area) involved decreased leftward asymmetries. However, three of the nine significantly altered cortical regional asymmetries involved shifts leftwards in ASD, either driven by a more prominent increase on the left side in ASD (i.e., medial orbitofrontal surface area), or by more prominent right- than left-side decreases in ASD (i.e., fusiform- and inferior temporal thickness). Thus the directional change of asymmetry can depend on the specific region, albeit that the overall magnitude of asymmetry is most likely to be reduced in ASD.

The significant associations of diagnosis with asymmetry in the present study were all weak (Cohen's *d* -0.13 to 0.12), indicating that altered structural brain asymmetry is unlikely to be a useful predictor for ASD. Prior studies using smaller samples were underpowered in this context. However, the effect sizes were comparable to those reported by recent, large-scale studies of bilateral disorder-related changes in brain structure, in which asymmetry was not studied, including for ASD (van Rooij et al., 2018) as well as attention-deficit hyperactivity disorder (ADHD) (Hoogman et al., 2017; Hoogman et al., 2019), schizophrenia (van Erp et al., 2018), obsessive compulsive disorder (OCD) (Boedhoe et al., 2018; Boedhoe et al., 2017), posttraumatic stress disorder (Logue et al., 2018), and major depressive disorder (Schmaal et al., 2017; Schmaal et al., 2016). It has become increasingly clear that anatomical differences between ASD and control groups are very small relative to the large within-group variability that is observed (Haar, Berman, Behrmann, & Dinstein, 2016).

Our findings may inform understanding of the neurobiology of ASD. Multi-regional reduction of cortical thickness asymmetry in ASD fits with the concept that laterality is an important organizing feature of the healthy human brain for multiple aspects of complex cognition, and is susceptible to disruption in disorders (e.g. (Francks, 2015; Renteria, 2012)). Left-right

asymmetry facilitates the development of localized and specialized modules in the brain, which can then have dominant control of behaviour (Geschwind & Galaburda, 1985; Rogers, Zucca, & Vallortigara, 2004). Notably, some of the cortical regions highlighted here are involved in diverse social cognitive processes, including perceptual processing (fusiform gyri), cognitive and emotional control (anterior cingulate) and reward evaluation (orbitofrontal cortex, ventral striatum) (Adolphs, 2009). However, the roles of these brain structures are by no means restricted to social behaviour. As we found altered asymmetry of various additional regions, our findings suggest broader disruption of lateralized neurodevelopment as part of the ASD phenotype. We note that many of the regions that showed significant case-control differences in asymmetry, including medial frontal, anterior cingulate and inferior temporal regions, overlap with the default mode network (DMN). The DMN comprises various cortical regions located in temporal (medial and lateral), parietal (medial and lateral) and prefrontal (medial) cortices(Raichle, 2015). DMN network organization has shown evidence for differences in ASD (Carlisi, Norman, Lukito, et al., 2017; Christakou et al., 2013; Nunes, Peatfield, Vakorin, & Doesburg, 2019; Uddin, 2011), including alterations in functional laterality (Nielsen et al., 2014). Our findings may therefore further support a role of altered lateralization of the DMN in ASD, warranting further investigations in this direction.

The medial orbitofrontal cortex was the only region that showed significantly altered asymmetry of both thickness and surface area in ASD, suggesting that disrupted laterality of this region might be particularly important in ASD. The orbitofrontal cortex may be involved in repetitive and stereotyped behaviours in ASD, due to its roles in executive functions (Ecker, Bookheimer, & Murphy, 2015). Prior studies have reported lower cortical thickness in the left medial orbitofrontal gyrus in ASD (Jiao et al., 2010), altered patterning of gyri and sulci in the right orbitofrontal cortex (Watanabe et al., 2014), and altered asymmetry in frontal regions globally (Conti et al., 2016; Knaus et al., 2012). These studies were in much smaller sample sizes than used here.

As regards the fusiform cortex, a previous study by Dougherty *et al.* reported an association between higher ASD symptom severity and increased

rightward surface area asymmetry, but not thickness asymmetry (Dougherty et al., 2016). The fusiform gyrus is involved in facial perception and memory among other functions, which are important for social interactions (Trontel et al., 2013). Here we report an asymmetry change in fusiform thickness in ASD that was significant after multiple testing correction, but there was also a nominally significant rightward change of surface area asymmetry in ASD (i.e., that did not survive multiple testing correction). This underlines that separate analyses of regional cortical thickness and surface area are well motivated, as they can vary relatively independently (Panizzon et al., 2009).

The altered volume asymmetry of the putamen in ASD may be related to its role in repetitive and restricted behaviours in ASD. One study reported that differences in striatal growth trajectories were correlated with circumscribed interests and insistence on sameness (Langen et al., 2014). The striatum is connected with lateral and orbitofrontal regions of the cortex via lateral-frontal-striatal reward and top-down cognitive control circuitry that might be dysfunctional in ASD (Carlisi, Norman, Murphy, et al., 2017). For example, individuals with ASD have shown decreased activation of the ventral striatum and lateral inferior/orbitofrontal cortex during outcome anticipation, and of dorsal striatum and lateral frontal regions during sustained attention and inhibitory control, compared with typically developing controls (Carlisi, Norman, Murphy, et al., 2017; Carlisi, Norman, Lukito, et al., 2017; Christakou et al., 2013).

Although the reasons for asymmetrical alterations in many of the structures implicated here are unclear, our findings suggest altered neurodevelopment affecting these structures in ASD. Further research is necessary to clarify the functional relevance and relationships between altered asymmetry and ASD. The findings we report in this large-scale study sometimes did not concur with prior, smaller studies. This may be due to limited statistical power in the earlier studies, while low power reduces the likelihood that a statistically significant result reflects a true effect (Button et al., 2013). However, the cortical atlas that we used did not have perfect equivalents for regions defined in many of the earlier studies, and we did not consider gyral/sulcal patterns, or grey matter volumes as such. Furthermore, discrepancies with earlier studies may be related to age differences, and

differences in clinical features of the disorder arising from case recruitment and diagnosis.

We included subjects from the entire ASD severity spectrum, with a broad range of ages, IQs, and of both sexes. Only one effect of diagnosis on regional asymmetry was influenced by sex, i.e. the rostral anterior cingulate thickness asymmetry, which was altered in males but not females. This same regional asymmetry was primarily altered in lower versus higher IQ cases. This may therefore be an alteration of cortical asymmetry that is relatively specific to an ASD subgroup, i.e., lower-performing males. In controls a different asymmetry (i.e., superior frontal thickness AI) showed a nominally significant association with IQ, which may point to different brain-IQ associations in ASD and controls. However, we cannot make strong interpretations based on these exploratory, secondary analyses without multiple testing correction.

As regards symptom severity, thickness asymmetry of the isthmus of the cingulate was associated with the ADOS score, such that the lower severity cases tended to have the most altered asymmetry. Again, this post hoc findings remains tentative in the context of multiple testing, and is reported here for descriptive purposes only. It is clear that most of the AIs that showed significant changes in ASD were not correlated with ADOS scores.

We found no evidence that medication use affected any of the asymmetries altered in ASD, although our medication variable was rudimentary. The role of specific medication usage should be investigated in future studies. As mentioned above, data on comorbidities were only available for 54 of the ASD subjects, precluding a high-powered analysis of this issue. This is a limitation of the study. We did not analyze handedness in the present study, as this had no significant effect on the same brain asymmetry measures as analyzed here, in studies of healthy individuals comprising more than 15,000 participants (Guadalupe et al., 2016; Kong et al., 2018).

In contrast to some prior studies of ASD, we did not adjust for IQ as a covariate effect in our main, case-control analysis. Rather we carried out post hoc analysis of possible associations between IQ and brain asymmetries,

separately in cases and controls. This was because lower average IQ was clearly part of the ASD phenotype in our total combined dataset (**Figure S1D**), so that including IQ as a confounding factor in case-control analysis might have reduced the power to detect an association of diagnosis with asymmetry. This would occur if underlying susceptibility factors contribute both to altered asymmetry and reduced IQ, as part of the ASD phenotype.

The Desikan-Killiany atlas (Desikan et al., 2006) was derived from manual segmentations of sets of reference brain images. Accordingly, the mean regional asymmetries in our samples partly reflect left-right differences present in the reference dataset used to construct the atlas. For detecting cerebral asymmetries with automated methods, some groups have chosen to work from artificially created, left-right symmetrical atlases, e.g. (Kawasaki et al., 2008). However, our study was focused on comparing *relative* asymmetry between groups. The use of a 'real-world' asymmetrical atlas had the advantage that regional identification was likely to be more accurate for structures that are asymmetrical both in the atlas and, on average, in our datasets. By defining the regions of interest in each hemisphere based on each hemisphere's own particular features, such as its sulcal and gyral geometry, we could then obtain the corresponding relationships between hemispheres. To this end, we used data from the automated labeling program within FreeSurfer for subdividing the human cerebral cortex. The labeling system incorporates hemisphere-specific information on sulcal and gyral geometry with spatial information regarding the locations of brain structures, and shows a high accuracy when compared to manual labeling results (Desikan et al., 2006). Thus, reliable measures of each region can be extracted for each subject, and regional asymmetries then accurately assessed.

Although a single image analysis pipeline was applied to all datasets, heterogeneity of imaging protocols was a feature of this study. There were substantial differences between datasets in the average asymmetry measured for some regions, which may be due in part to different scanner characteristics, as well as differences in patient profiles. We corrected for 'dataset' as a random effect in the analysis, and sensitivity analysis based on the subset of 3T acquired data showed similar results to the primary analysis. However, it is possible that between-dataset variability resulted in reduced

statistical power, relative to a hypothetical, equally-sized, single-centre study. In reality, no single centre has been able to collect such large samples alone. As long as researchers publish many separate papers based on single datasets, collected in particular ways, the field overall has the same problem. In this case, multi-centre studies can better represent the real-world heterogeneity, with more generalizable findings than single-centre studies (Costafreda, 2009). The primary purpose of our study, based on 54 datasets that were originally collected as separate studies, was to assess the total combined evidence for effects over all of these datasets, while allowing for heterogeneity between datasets through the use of random intercepts, and finally adding sensitivity and secondary analyses with respect to relevant variables.

The cross-sectional design limits our capacity to make causal inferences between diagnosis and asymmetry. ASD is highly heritable, with meta-analytic heritability estimates ranging from 64% to 91% (Tick, Bolton, Happe, Rutter, & Rijsdijk, 2016). Likewise, some of the brain asymmetry measures examined here have heritabilities as high as roughly 25% (Guadalupe et al., 2016; Kong et al., 2018). Future studies are required to investigate shared genetic contributions to ASD and variation in brain structural asymmetry. These could help to disentangle cause-effect relations between ASD and brain structural asymmetry. Given the high comorbidity of ASD with other disorders, such as ADHD, OCD, and schizophrenia (Sharma, Gonda, & Tarazi, 2018), cross-disorder analyses incorporating between-disorder genetic correlations may be informative.

In conclusion, large-scale analysis of brain asymmetry in ASD revealed primarily cortical thickness effects, but also effects on orbitofrontal cortex asymmetry, and putamen asymmetry, which were significant but very small. Our study illustrates how high-powered and systematic studies can yield much needed clarity in human clinical neuroscience, where prior smaller and more methodologically diverse studies were inconclusive.

2.5 References

Adolphs, R. (2009). The social brain: neural basis of social knowledge. *Annu Rev Psychol, 60*, 693-716. doi:10.1146/annurev.psych.60.110707.163514

American Psychiatric Association. (2000). Diagnostic and Statistical Manual of Mental disorders, 4th edition (DSM-IV), Washington DC.

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). Washington, DC.

Bedford, S. A., Park, M. T. M., Devenyi, G. A., Tullo, S., Germann, J., Patel, R., . . . Chakravarty, M. M. (2019). Large-scale analyses of the relationship between sex, age and intelligence quotient heterogeneity and cortical morphometry in autism spectrum disorder. *Mol Psychiatry*. doi:10.1038/s41380-019-0420-6

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate - A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B-Methodological, 57*(1), 289-300.

Biberacher, V., Schmidt, P., Keshavan, A., Boucard, C. C., Righart, R., Samann, P., . . . Muhlau, M. (2016). Intra- and interscanner variability of magnetic resonance imaging based volumetry in multiple sclerosis. *Neuroimage, 142*, 188-197. doi:10.1016/j.neuroimage.2016.07.035

Boedhoe, Schmaal, L., Abe, Y., Alonso, P., Ameis, S. H., Anticevic, A., . . . van den Heuvel, O. A. (2018). Cortical Abnormalities Associated With Pediatric and Adult Obsessive-Compulsive Disorder: Findings From the ENIGMA Obsessive-Compulsive Disorder Working Group. *Am J Psychiatry, 175*(5), 453-462. doi:10.1176/appi.ajp.2017.17050485

Boedhoe, Schmaal, L., Abe, Y., Ameis, S. H., Arnold, P. D., Batistuzzo, M. C., . . . van den Heuvel, O. A. (2017). Distinct Subcortical Volume Alterations in Pediatric and Adult OCD: A Worldwide Meta- and Mega-Analysis. *Am J Psychiatry, 174*(1), 60-69. doi:10.1176/appi.ajp.2016.16020201

Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., & Munafo, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nat Rev Neurosci, 14*(5), 365-376. doi:10.1038/nrn3475

Cardinale, R. C., Shih, P., Fishman, I., Ford, L. M., & Muller, R. A. (2013). Pervasive rightward asymmetry shifts of functional networks in autism spectrum disorder. *JAMA Psychiatry, 70*(9), 975-982. doi:10.1001/jamapsychiatry.2013.382

Carlisi, C. O., Norman, L., Murphy, C. M., Christakou, A., Chantiluke, K., Giampietro, V., . . . Rubia, K. (2017). Shared and Disorder-Specific

Neurocomputational Mechanisms of Decision-Making in Autism Spectrum Disorder and Obsessive-Compulsive Disorder. *Cereb Cortex, 27*(12), 5804-5816. doi:10.1093/cercor/bhx265

Carlisi, C. O., Norman, L. J., Lukito, S. S., Radua, J., Mataix-Cols, D., & Rubia, K. (2017). Comparative Multimodal Meta-analysis of Structural and Functional Brain Abnormalities in Autism Spectrum Disorder and Obsessive-Compulsive Disorder. *Biol Psychiatry, 82*(2), 83-102. doi:10.1016/j.biopsych.2016.10.006

Carper, R. A., Treiber, J. M., DeJesus, S. Y., & Muller, R. A. (2016). Reduced Hemispheric Asymmetry of White Matter Microstructure in Autism Spectrum Disorder. *J Am Acad Child Adolesc Psychiatry, 55*(12), 1073-1080. doi:10.1016/j.jaac.2016.09.491

Chambers, J. M., & Hastie, T. J. (1992). Statistical models in S. Pacific Grove, California, USA, *Wadsworth & Brooks/Cole*.

Christakou, A., Murphy, C. M., Chantiluke, K., Cubillo, A. I., Smith, A. B., Giampietro, V., . . . Rubia, K. (2013). Disorder-specific functional abnormalities during sustained attention in youth with Attention Deficit Hyperactivity Disorder (ADHD) and with autism. *Mol Psychiatry, 18*(2), 236-244. doi:10.1038/mp.2011.185

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Conti, E., Calderoni, S., Gaglianese, A., Pannek, K., Mazzotti, S., Rose, S., . . . Guzzetta, A. (2016). Lateralization of Brain Networks and Clinical Severity in Toddlers with Autism Spectrum Disorder: A HARDI Diffusion MRI Study. *Autism Res, 9*(3), 382-392. doi:10.1002/aur.1533

Costafreda, S. G. (2009). Pooling FMRI data: meta-analysis, mega-analysis and multi-center studies. *Front Neuroinform, 3*, 33. doi:10.3389/neuro.11.033.2009

Courchesne, E., Karns, C. M., Davis, H. R., Ziccardi, R., Carper, R. A., Tigue, Z. D., . . . Courchesne, R. Y. (2001). Unusual brain growth patterns in early life in patients with autistic disorder: an MRI study. *Neurology, 57*(2), 245-254. doi:10.1212/wnl.57.2.245

Courchesne, E., Pierce, K., Schumann, C. M., Redcay, E., Buckwalter, J. A., Kennedy, D. P., & Morgan, J. (2007). Mapping early brain development in autism. *Neuron, 56*(2), 399-413. doi:10.1016/j.neuron.2007.10.016

Desikan, R. S., Segonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., . . . Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of

interest. *Neuroimage, 31*(3), 968-980. doi:10.1016/j.neuroimage.2006.01.021

Dougherty, C. C., Evans, D. W., Katuwal, G. J., & Michael, A. M. (2016). Asymmetry of fusiform structure in autism spectrum disorder: trajectory and association with symptom severity. *Mol Autism, 7*, 28. doi:10.1186/s13229-016-0089-5

Duboc, V., Dufourcq, P., Blader, P., & Roussigne, M. (2015). Asymmetry of the Brain: Development and Implications. *Annu Rev Genet, 49*, 647-672. doi:10.1146/annurev-genet-112414-055322

Ecker, C., Bookheimer, S. Y., & Murphy, D. G. (2015). Neuroimaging in autism spectrum disorder: brain structure and function across the lifespan. *Lancet Neurol, 14*(11), 1121-1134. doi:10.1016/s1474-4422(15)00050-2

Elsabbagh, M., Divan, G., Koh, Y. J., Kim, Y. S., Kauchali, S., Marcin, C., . . . Fombonne, E. (2012). Global prevalence of autism and other pervasive developmental disorders. *Autism Res, 5*(3), 160-179. doi:10.1002/aur.239

Fischl, B. (2012). FreeSurfer. *Neuroimage, 62*(2), 774-781. doi:10.1016/j.neuroimage.2012.01.021

Floris, D. L., Lai, M. C., Auer, T., Lombardo, M. V., Ecker, C., Chakrabarti, B., . . . Suckling, J. (2016). Atypically rightward cerebral asymmetry in male adults with autism stratifies individuals with and without language delay. *Hum Brain Mapp, 37*(1), 230-253. doi:10.1002/hbm.23023

Francks, C. (2015). Exploring human brain lateralization with molecular genetics and genomics. *Ann N Y Acad Sci, 1359*, 1-13. doi:10.1111/nyas.12770

Gabard-Durnam, L., Tierney, A. L., Vogel-Farley, V., Tager-Flusberg, H., & Nelson, C. A. (2015). Alpha asymmetry in infants at risk for autism spectrum disorders. *J Autism Dev Disord, 45*(2), 473-480. doi:10.1007/s10803-013-1926-4

Geschwind, N., & Galaburda, A. M. (1985). Cerebral lateralization: Biological mechanisms, associations, and pathology: i. a hypothesis and a program for research. *Archives of Neurology, 42*(5), 428-459. doi:10.1001/archneur.1985.04060050026008

Gori, I., Giuliano, A., Muratori, F., Saviozzi, I., Oliva, P., Tancredi, R., . . . Retico, A. (2015). Gray Matter Alterations in Young Children with Autism Spectrum Disorders: Comparing Morphometry at the Voxel and Regional Level. *J Neuroimaging, 25*(6), 866-874. doi:10.1111/jon.12280

Guadalupe, T., Mathias, S. R., vanErp, T. G., Whelan, C. D., Zwiers, M. P., Abe, Y., . . . Francks, C. (2016). Human subcortical brain asymmetries in 15,847 people worldwide reveal effects of age and sex. *Brain Imaging Behav*. doi:10.1007/s11682-016-9629-z

Haar, S., Berman, S., Behrmann, M., & Dinstein, I. (2016). Anatomical Abnormalities in Autism? *Cereb Cortex, 26*(4), 1440-1452. doi:10.1093/cercor/bhu242

Hoogman, M., Bralten, J., Hibar, D. P., Mennes, M., Zwiers, M. P., Schweren, L. S., . . . Franke, B. (2017). Subcortical brain volume differences in participants with attention deficit hyperactivity disorder in children and adults: a cross-sectional mega-analysis. *Lancet Psychiatry, 4*(4), 310-319. doi:10.1016/s2215-0366(17)30049-4

Hoogman, M., Muetzel, R., Guimaraes, J. P., Shumskaya, E., Mennes, M., Zwiers, M. P., . . . Franke, B. (2019). Brain Imaging of the Cortex in ADHD: A Coordinated Analysis of Large-Scale Clinical and Population-Based Samples. *Am J Psychiatry*, appiajp201918091033. doi:10.1176/appi.ajp.2019.18091033

Jeste, S. S., & Geschwind, D. H. (2014). Disentangling the heterogeneity of autism spectrum disorder through genetic findings. *Nat Rev Neurol, 10*(2), 74-81. doi:10.1038/nrneurol.2013.278

Jiao, Y., Chen, R., Ke, X., Chu, K., Lu, Z., & Herskovits, E. H. (2010). Predictive models of autism spectrum disorder based on brain regional cortical thickness. *Neuroimage, 50*(2), 589-599. doi:10.1016/j.neuroimage.2009.12.047

Joseph, R. M., Fricker, Z., Fenoglio, A., Lindgren, K. A., Knaus, T. A., & Tager-Flusberg, H. (2014). Structural asymmetries of language-related gray and white matter and their relationship to language function in young children with ASD. *Brain Imaging Behav, 8*(1), 60-72. doi:10.1007/s11682-013-9245-0

Kawasaki, Y., Suzuki, M., Takahashi, T., Nohara, S., McGuire, P. K., Seto, H., & Kurachi, M. (2008). Anomalous cerebral asymmetry in patients with schizophrenia demonstrated by voxel-based morphometry. *Biol Psychiatry, 63*(8), 793-800. doi:10.1016/j.biopsych.2007.08.008

Kleinhans, N. M., Muller, R. A., Cohen, D. N., & Courchesne, E. (2008). Atypical functional lateralization of language in autism spectrum disorders. *Brain Res, 1221*, 115-125. doi:10.1016/j.brainres.2008.04.080

Knaus, T. A., Silver, A. M., Kennedy, M., Lindgren, K. A., Dominick, K. C., Siegel, J., & Tager-Flusberg, H. (2010). Language laterality in autism spectrum disorder and typical controls: a functional, volumetric, and diffusion tensor MRI study. *Brain Lang, 112*(2), 113-120. doi:10.1016/j.bandl.2009.11.005

Knaus, T. A., Tager-Flusberg, H., Mock, J., Dauterive, R., & Foundas, A. L. (2012). Prefrontal and occipital asymmetry and volume in boys with

autism spectrum disorder. *Cogn Behav Neurol, 25*(4), 186-194. doi:10.1097/WNN.0b013e318280e154

Kong, X. Z., Mathias, S. R., Guadalupe, T., Glahn, D. C., Franke, B., Crivello, F., . . . Francks, C. (2018). Mapping cortical brain asymmetry in 17,141 healthy individuals worldwide via the ENIGMA Consortium. *Proc Natl Acad Sci U S A*. doi:10.1073/pnas.1718418115

Kurth, F., Gaser, C., & Luders, E. (2015). A 12-step user guide for analyzing voxel-wise gray matter asymmetries in statistical parametric mapping (SPM). *Nat Protoc, 10*(2), 293-304. doi:10.1038/nprot.2015.014

Langen, M., Bos, D., Noordermeer, S. D., Nederveen, H., van Engeland, H., & Durston, S. (2014). Changes in the development of striatum are involved in repetitive behavior in autism. *Biol Psychiatry, 76*(5), 405-411. doi:10.1016/j.biopsych.2013.08.013

Leroy, F., Cai, Q., Bogart, S. L., Dubois, J., Coulon, O., Monzalvo, K., . . . Dehaene-Lambertz, G. (2015). New human-specific brain landmark: the depth asymmetry of superior temporal sulcus. *Proc Natl Acad Sci U S A, 112*(4), 1208-1213. doi:10.1073/pnas.1412389112

Li, D., Karnath, H. O., & Xu, X. (2017). Candidate Biomarkers in Children with Autism Spectrum Disorder: A Review of MRI Studies. *Neurosci Bull, 33*(2), 219-237. doi:10.1007/s12264-017-0118-1

Lindell, A. K., & Hudry, K. (2013). Atypicalities in cortical structure, handedness, and functional lateralization for language in autism spectrum disorders. *Neuropsychol Rev, 23*(3), 257-270. doi:10.1007/s11065-013-9234-5

Logue, M. W., van Rooij, S. J. H., Dennis, E. L., Davis, S. L., Hayes, J. P., Stevens, J. S., . . . Morey, R. A. (2018). Smaller Hippocampal Volume in Posttraumatic Stress Disorder: A Multisite ENIGMA-PGC Study: Subcortical Volumetry Results From Posttraumatic Stress Disorder Consortia. *Biol Psychiatry, 83*(3), 244-253. doi:10.1016/j.biopsych.2017.09.006

Lord, C., & Bishop, S. L. (2015). Recent advances in autism research as reflected in DSM-5 criteria for autism spectrum disorder. *Annu Rev Clin Psychol, 11*, 53-70. doi:10.1146/annurev-clinpsy-032814-112745

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Jr., Leventhal, B. L., DiLavore, P. C., . . . Rutter, M. (2000). The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism. *J Autism Dev Disord, 30*(3), 205-223.

Loth, E., Murphy, D. G., & Spooren, W. (2016). Defining Precision Medicine Approaches to Autism Spectrum Disorders: Concepts and Challenges. *Front Psychiatry, 7*, 188. doi:10.3389/fpsyt.2016.00188

Lucibello, S., Verdolotti, T., Giordano, F. M., Lapenta, L., Infante, A., Piludu, F., . . . Battini, R. (2019). Brain morphometry of preschool age children affected by autism spectrum disorder: Correlation with clinical findings. *Clin Anat, 32*(1), 143-150. doi:10.1002/ca.23252

Markou, P., Ahtam, B., & Papadatou-Pastou, M. (2017). Elevated Levels of Atypical Handedness in Autism: Meta-Analyses. *Neuropsychol Rev, 27*(3), 258-283. doi:10.1007/s11065-017-9354-4

Munafo, M. R., & Flint, J. (2010). How reliable are scientific studies? *Br J Psychiatry, 197*(4), 257-258. doi:10.1192/bjp.bp.109.069849

Nielsen, J. A., Zielinski, B. A., Fletcher, P. T., Alexander, A. L., Lange, N., Bigler, E. D., . . . Anderson, J. S. (2014). Abnormal lateralization of functional connectivity between language and default mode regions in autism. *Mol Autism, 5*(1), 8. doi:10.1186/2040-2392-5-8

Nunes, A. S., Peatfield, N., Vakorin, V., & Doesburg, S. M. (2019). Idiosyncratic organization of cortical networks in autism spectrum disorder. *Neuroimage, 190*, 182-190. doi:10.1016/j.neuroimage.2018.01.022

Panizzon, M. S., Fennema-Notestine, C., Eyler, L. T., Jernigan, T. L., Prom-Wormley, E., Neale, M., . . . Kremen, W. S. (2009). Distinct genetic influences on cortical surface area and cortical thickness. *Cereb Cortex, 19*(11), 2728-2735. doi:10.1093/cercor/bhp026

Pinheiro J, B. D., DebRoy S, Sarkar D and R Core Team (2018). *nlme: Linear and Nonlinear Mixed Effects Models.* R package version 3.1-137, https://CRAN.R-project.org/package=nlme.

Raichle, M. E. (2015). The brain's default mode network. *Annu Rev Neurosci, 38*, 433-447. doi:10.1146/annurev-neuro-071013-014030

Renteria, M. E. (2012). Cerebral asymmetry: a quantitative, multifactorial, and plastic brain phenotype. *Twin Res Hum Genet, 15*(3), 401-413. doi:10.1017/thg.2012.13

Retico, A., Giuliano, A., Tancredi, R., Cosenza, A., Apicella, F., Narzisi, A., . . . Calderoni, S. (2016). The effect of gender on the neuroanatomy of children with autism spectrum disorders: a support vector machine case-control study. *Mol Autism, 7*, 5. doi:10.1186/s13229-015-0067-3

Rogers, L. J., Zucca, P., & Vallortigara, G. (2004). Advantages of having a lateralized brain. *Proc Biol Sci, 271 Suppl 6*, S420-422. doi:10.1098/rsbl.2004.0200

Rommelse, N., Buitelaar, J. K., & Hartman, C. A. (2017). Structural brain imaging correlates of ASD and ADHD across the lifespan: a hypothesis-generating review on developmental ASD-ADHD subtypes. *J Neural Transm (Vienna), 124*(2), 259-271. doi:10.1007/s00702-016-1651-1

Rysstad, A. L., & Pedersen, A. V. (2018). There Are Indeed More Left-Handers Within the Autism Spectrum Disorder Compared with in the General Population, but the Many Mixed-Handers Is the More Interesting Finding. *J Autism Dev Disord*. doi:10.1007/s10803-018-3553-6

Schmaal, L., Hibar, D. P., Samann, P. G., Hall, G. B., Baune, B. T., Jahanshad, N., . . . Veltman, D. J. (2017). Cortical abnormalities in adults and adolescents with major depression based on brain scans from 20 cohorts worldwide in the ENIGMA Major Depressive Disorder Working Group. *Mol Psychiatry, 22*(6), 900-909. doi:10.1038/mp.2016.60

Schmaal, L., Veltman, D. J., van Erp, T. G., Samann, P. G., Frodl, T., Jahanshad, N., . . . Hibar, D. P. (2016). Subcortical brain alterations in major depressive disorder: findings from the ENIGMA Major Depressive Disorder working group. *Mol Psychiatry, 21*(6), 806-812. doi:10.1038/mp.2015.69

Sharma, S. R., Gonda, X., & Tarazi, F. I. (2018). Autism Spectrum Disorder: Classification, diagnosis and therapy. *Pharmacol Ther*. doi:10.1016/j.pharmthera.2018.05.007

Tick, B., Bolton, P., Happe, F., Rutter, M., & Rijsdijk, F. (2016). Heritability of autism spectrum disorders: a meta-analysis of twin studies. *J Child Psychol Psychiatry, 57*(5), 585-595. doi:10.1111/jcpp.12499

Toga, A. W., & Thompson, P. M. (2003). Mapping brain asymmetry. *Nat Rev Neurosci, 4*(1), 37-48. doi:10.1038/nrn1009

Trontel, H. G., Duffield, T. C., Bigler, E. D., Froehlich, A., Prigge, M. B., Nielsen, J. A., . . . Lainhart, J. E. (2013). Fusiform correlates of facial memory in autism. *Behav Sci (Basel), 3*(3), 348-371. doi:10.3390/bs3030348

Uddin, L. Q. (2011). The self in autism: an emerging view from neuroimaging. *Neurocase, 17*(3), 201-208. doi:10.1080/13554794.2010.509320

van Erp, T. G. M., Walton, E., Hibar, D. P., Schmaal, L., Jiang, W., Glahn, D. C., . . . Turner, J. A. (2018). Cortical Brain Abnormalities in 4474 Individuals With Schizophrenia and 5098 Control Subjects via the Enhancing Neuro Imaging Genetics Through Meta Analysis (ENIGMA) Consortium. *Biol Psychiatry, 84*(9), 644-654. doi:10.1016/j.biopsych.2018.04.023

van Rooij, D., Anagnostou, E., Arango, C., Auzias, G., Behrmann, M., Busatto, G. F., . . . Buitelaar, J. K. (2018). Cortical and Subcortical Brain Morphometry Differences Between Patients With Autism Spectrum

Disorder and Healthy Individuals Across the Lifespan: Results From the ENIGMA ASD Working Group. *Am J Psychiatry, 175*(4), 359-369. doi:10.1176/appi.ajp.2017.17010100

Watanabe, H., Nakamura, M., Ohno, T., Itahashi, T., Tanaka, E., Ohta, H., . . . Hashimoto, R. (2014). Altered orbitofrontal sulcogyral patterns in adult males with high-functioning autism spectrum disorders. *Soc Cogn Affect Neurosci, 9*(4), 520-528. doi:10.1093/scan/nst016

Wei, L., Zhong, S., Nie, S., & Gong, G. (2018). Aberrant development of the asymmetry between hemispheric brain white matter networks in autism spectrum disorder. *Eur Neuropsychopharmacol, 28*(1), 48-62. doi:10.1016/j.euroneuro.2017.11.018

Xu, G., Strathearn, L., Liu, B., O'Brien, M., Kopelman, T. G., Zhu, J., . . . Bao, W. (2018). Prevalence and Treatment Patterns of Autism Spectrum Disorder in the United States, 2016. *JAMA Pediatr*. doi:10.1001/jamapediatrics.2018.4208

2.6 Supplementary Information

List of tables

List of figures

Table S1. Full linear mixed model results for the cortical thickness AIs.

AI thickness region	N cases/ controls	β-value			Standard Error			*t*-value			*P*-value[1]			
		diag	age	sex	diag	age	sex	diag	age	sex	diag	age	sex	Cohen's *d* (95% CI)
bankssts	1691/1750	0.00009	0.00002	0.00298	0.0014	0.0001	0.0019	0.06	0.20	1.58	0.950	0.845	0.113	0.002(-0.06,0.07)
caudalanteriorcingulate	1699/1763	-0.00004	-0.00006	0.00234	0.0018	0.0001	0.0024	-0.02	-0.44	0.98	0.982	0.657	0.328	-0.001(-0.07,0.07)
caudalmiddlefrontal	1705/1770	-0.00166	0.00003	-0.00046	0.0010	0.0001	0.0013	-1.72	0.32	-0.36	0.086	0.749	0.717	-0.059(-0.13,0.01)
cuneus	1700/1767	-0.00215	0.00006	-0.00180	0.0012	0.0001	0.0015	-1.84	0.66	-1.17	0.066	0.508	0.240	-0.063(-0.13,0)
entorhinal	1688/1752	0.00360	-0.00032	-0.00190	0.0022	0.0001	0.0028	1.67	-2.17	-0.68	0.095	**0.030**	0.495	0.057(-0.01,0.12)
frontalpole	1704/1765	-0.00069	-0.00022	-0.00085	0.0023	0.0002	0.0030	-0.30	-1.37	-0.28	0.765	0.171	0.777	-0.01(-0.08,0.06)
fusiform	1704/1767	0.00277	-0.00001	0.00234	0.0009	0.0001	0.0011	3.20	-0.17	2.06	***0.001***	0.862	**0.040**	0.109(0.04,0.18)
inferiorparietal	1705/1769	-0.00146	-0.00002	-0.00089	0.0008	0.0001	0.0010	-1.85	-0.28	-0.86	0.065	0.782	0.392	-0.063(-0.13,0)
inferiortemporal	1703/1768	0.00305	0.00001	0.00098	0.0010	0.0001	0.0014	2.97	0.07	0.73	***0.003***	0.948	0.467	0.102(0.03,0.17)
insula	1700/1764	0.00107	0.00011	0.00256	0.0009	0.0001	0.0012	1.15	1.41	2.10	0.252	0.158	**0.036**	0.039(-0.03,0.11)
isthmuscingulate	1699/1769	-0.00365	0.00027	0.00186	0.0014	0.0001	0.0018	-2.58	2.81	1.02	***0.010***	**0.005**	0.310	-0.088(-0.16,-0.02)
lateraloccipital	1701/1766	0.00075	0.00009	0.00119	0.0008	0.0001	0.0011	0.92	1.40	1.12	0.359	0.160	0.264	0.031(-0.04,0.1)
lateralorbitofrontal	1704/1771	0.00017	-0.00003	-0.00066	0.0010	0.0001	0.0014	0.16	-0.33	-0.48	0.872	0.739	0.631	0.006(-0.06,0.07)
lingual	1704/1769	0.00051	0.00000	-0.00379	0.0010	0.0001	0.0013	0.53	0.01	-3.03	0.594	0.988	**0.002**	0.018(-0.05,0.08)
medialorbitofrontal	1705/1769	-0.00442	0.00043	-0.00052	0.0013	0.0001	0.0017	-3.47	4.23	-0.31	***0.001***	**0.00002**	0.758	-0.119(-0.19,-0.05)
middletemporal	1702/1764	-0.00034	0.00004	0.00106	0.0010	0.0001	0.0013	-0.35	0.47	0.84	0.726	0.641	0.403	-0.012(-0.08,0.05)
paracentral	1702/1771	0.00114	0.00002	-0.00037	0.0010	0.0001	0.0012	1.20	0.29	-0.29	0.231	0.772	0.769	0.041(-0.03,0.11)
parahippocampal	1702/1764	-0.00028	0.00004	-0.00264	0.0016	0.0001	0.0020	-0.18	0.36	-1.31	0.856	0.717	0.190	-0.006(-0.07,0.06)
parsopercularis	1703/1769	-0.00059	0.00001	0.00083	0.0010	0.0001	0.0014	-0.57	0.13	0.61	0.568	0.898	0.540	-0.02(-0.09,0.05)
parsorbitalis	1705/1770	-0.00348	0.00005	0.00139	0.0015	0.0001	0.0020	-2.26	0.43	0.69	**0.024**	0.670	0.491	-0.077(-0.14,-0.01)
parstriangularis	1703/1769	-0.00035	-0.00009	0.00144	0.0011	0.0001	0.0015	-0.31	-0.98	0.96	0.759	0.330	0.336	-0.011(-0.08,0.06)
pericalcarine	1699/1768	-0.00086	-0.00013	0.00063	0.0014	0.0001	0.0018	-0.62	-1.12	0.34	0.536	0.262	0.731	-0.021(-0.09,0.05)
postcentral	1704/1766	-0.00107	0.00007	-0.00040	0.0009	0.0001	0.0011	-1.25	1.16	-0.36	0.211	0.245	0.720	-0.043(-0.11,0.02)
posteriorcingulate	1706/1765	-0.00251	0.00007	-0.00079	0.0012	0.0001	0.0015	-2.13	0.80	-0.51	**0.034**	0.423	0.611	-0.073(-0.14,-0.01)
precentral	1704/1766	-0.00045	-0.00009	-0.00038	0.0007	0.0001	0.0010	-0.60	-1.52	-0.39	0.547	0.129	0.700	-0.021(-0.09,0.05)
precuneus	1705/1769	-0.00027	0.00005	-0.00188	0.0007	0.0001	0.0010	-0.37	0.81	-1.96	0.710	0.418	0.050	-0.013(-0.08,0.05)
rostralanteriorcingulate	1700/1765	-0.00568	-0.00010	0.00321	0.0017	0.0001	0.0022	-3.37	-0.69	1.45	***0.001***	0.491	0.147	-0.116(-0.18,-0.05)
rostralmiddlefrontal	1707/1769	-0.00255	-0.00022	0.00169	0.0009	0.0001	0.0011	-2.94	-2.87	1.47	***0.003***	**0.004**	0.141	-0.101(-0.17,-0.03)

Table S1. Full linear mixed model results for the cortical thickness AIs [continued]

AI thickness region	N cases/ controls	β-value			Standard Error			*t*-value			*P*-value[1]			Cohen's *d* (95% CI)
		diag	age	sex	diag	age	sex	diag	age	sex	diag	age	sex	
superiorfrontal	1706/1771	-0.00251	0.00003	-0.00116	0.0006	0.0001	0.0008	-3.92	0.48	-1.38	***0.0001***	0.630	0.169	-0.134(-0.2,-0.07)
superiorparietal	1702/1769	0.00023	0.00005	-0.00001	0.0007	0.0001	0.0009	0.35	0.89	-0.01	0.725	0.372	0.995	0.012(-0.05,0.08)
superiortemporal	1693/1760	-0.00168	0.00003	0.00177	0.0009	0.0001	0.0011	-1.97	0.47	1.59	**0.049**	0.635	0.112	-0.068(-0.13,0)
supramarginal	1698/1769	0.00002	-0.00009	0.00187	0.0009	0.0001	0.0011	0.03	-1.35	1.63	0.980	0.177	0.103	0.001(-0.07,0.07)
temporalpole	1697/1761	0.00300	-0.00002	0.00052	0.0020	0.0002	0.0026	1.49	-0.12	0.20	0.136	0.906	0.844	0.051(-0.02,0.12)
transversetemporal	1704/1771	0.00152	0.00025	0.00107	0.0017	0.0001	0.0022	0.91	1.95	0.49	0.364	0.051	0.626	0.031(-0.04,0.1)
total thickness	1706/1771	-0.00061	0.00001	0.00021	0.0003	0.0000	0.0004	-1.96	0.33	0.52	0.050	0.742	0.602	-0.067(-0.13,0)

[1]Unadjusted *P* values are shown, with in **bold** those that are significant ($P < 0.05$) at the uncorrected level. Those p-values for diagnosis that survived multiple testing

Table S2. Full linear mixed model results for the cortical surface area AIs.

AI surface arearegion	N cases/ controls	β-value			Standard Error			*t*-value			*P*-value[1]			Cohen's *d* (95% CI)
		diag	age	sex	diag	age	sex	diag	age	sex	diag	age	sex	
bankssts	1685/1745	0.00292	-0.00081	-0.00133	0.0030	0.0002	0.0038	0.97	-4.69	-0.35	0.332	**0.000003**	0.729	0.033(-0.03,0.1)
caudalanteriorcingulate	1701/1752	-0.00176	-0.00043	0.00742	0.0040	0.0002	0.0052	-0.44	-1.73	1.43	0.662	0.084	0.151	-0.015(-0.08,0.05)
caudalmiddlefrontal	1704/1763	-0.00186	0.00010	0.00642	0.0027	0.0002	0.0034	-0.70	0.65	1.89	0.483	0.516	0.059	-0.024(-0.09,0.04)
cuneus	1700/1754	-0.00303	0.00000	-0.00212	0.0024	0.0002	0.0032	-1.24	0.00	-0.66	0.215	0.998	0.506	-0.043(-0.11,0.02)
entorhinal	1687/1733	-0.00687	0.00054	-0.00009	0.0043	0.0003	0.0055	-1.61	1.96	-0.02	0.108	0.050	0.986	-0.055(-0.12,0.01)
frontalpole	1707/1757	0.00276	0.00014	0.00351	0.0032	0.0002	0.0042	0.86	0.58	0.84	0.388	0.560	0.400	0.03(-0.04,0.1)
fusiform	1702/1760	-0.00477	0.00007	0.00416	0.0018	0.0001	0.0024	-2.58	0.56	1.74	**0.010**	0.574	0.082	-0.089(-0.16,-0.02)
inferiorparietal	1705/1758	0.00034	-0.00009	0.00072	0.0019	0.0001	0.0024	0.18	-0.74	0.30	0.857	0.461	0.767	0.006(-0.06,0.07)
inferiortemporal	1698/1757	-0.00114	0.00006	-0.00066	0.0021	0.0001	0.0027	-0.55	0.45	-0.24	0.585	0.654	0.807	-0.019(-0.09,0.05)
insula	1699/1751	0.00063	-0.00004	-0.00279	0.0016	0.0001	0.0022	0.38	-0.34	-1.29	0.701	0.733	0.197	0.013(-0.05,0.08)
isthmuscingulate	1699/1753	0.00134	0.00006	0.00048	0.0026	0.0002	0.0034	0.51	0.34	0.14	0.611	0.734	0.888	0.017(-0.05,0.08)
lateraloccipital	1705/1754	-0.00080	-0.00009	0.00298	0.0018	0.0001	0.0023	-0.45	-0.80	1.30	0.654	0.424	0.193	-0.015(-0.08,0.05)
lateralorbitofrontal	1701/1762	-0.00500	0.00010	-0.00013	0.0015	0.0001	0.0020	-3.28	0.81	-0.07	***0.001***	0.421	0.946	-0.112(-0.18,-0.05)
lingual	1696/1758	0.00046	0.00017	0.00286	0.0020	0.0001	0.0026	0.23	1.14	1.08	0.820	0.252	0.279	0.008(-0.06,0.07)
medialorbitofrontal	1704/1763	0.00675	-0.00005	-0.00243	0.0021	0.0002	0.0028	3.17	-0.27	-0.87	***0.002***	0.787	0.385	0.109(0.04,0.18)
middletemporal	1698/1757	-0.00073	-0.00007	0.00166	0.0017	0.0001	0.0021	-0.44	-0.64	0.77	0.661	0.522	0.441	-0.015(-0.08,0.05)
paracentral	1703/1762	0.00408	-0.00018	0.00104	0.0023	0.0001	0.0029	1.81	-1.29	0.36	0.071	0.195	0.720	0.062(0,0.13)
parahippocampal	1692/1752	-0.00051	-0.00008	0.00168	0.0028	0.0002	0.0036	-0.19	-0.47	0.47	0.853	0.639	0.637	-0.006(-0.07,0.06)
parsopercularis	1704/1752	0.00042	0.00016	0.00050	0.0028	0.0002	0.0036	0.15	1.00	0.14	0.880	0.317	0.890	0.005(-0.06,0.07)
parsorbitalis	1699/1760	0.00321	0.00008	0.00054	0.0021	0.0001	0.0028	1.50	0.60	0.20	0.133	0.551	0.844	0.052(-0.02,0.12)
parstriangularis	1702/1758	0.00024	-0.00018	0.00459	0.0027	0.0002	0.0035	0.09	-1.02	1.32	0.929	0.309	0.188	0.003(-0.06,0.07)
pericalcarine	1699/1762	-0.00134	-0.00004	0.00129	0.0023	0.0002	0.0030	-0.58	-0.22	0.43	0.561	0.825	0.670	-0.02(-0.09,0.05)
postcentral	1706/1759	0.00135	0.00002	0.00082	0.0017	0.0001	0.0022	0.79	0.18	0.37	0.430	0.855	0.709	0.027(-0.04,0.09)
posteriorcingulate	1703/1758	-0.00189	-0.00023	0.00406	0.0026	0.0002	0.0034	-0.72	-1.41	1.20	0.475	0.160	0.232	-0.025(-0.09,0.04)
precentral	1703/1757	-0.00203	-0.00012	-0.00102	0.0014	0.0001	0.0018	-1.43	-1.38	-0.56	0.154	0.168	0.578	-0.049(-0.12,0.02)
precuneus	1702/1762	-0.00017	0.00009	0.00038	0.0016	0.0001	0.0020	-0.11	1.02	0.19	0.915	0.310	0.851	-0.004(-0.07,0.06)
rostralanteriorcingulate	1701/1760	-0.00344	0.00009	0.00583	0.0034	0.0003	0.0045	-1.00	0.36	1.30	0.317	0.719	0.193	-0.034(-0.1,0.03)
rostralmiddlefrontal	1705/1761	0.00061	0.00000	-0.00140	0.0016	0.0001	0.0020	0.39	0.01	-0.69	0.694	0.990	0.489	0.013(-0.05,0.08)
superiorfrontal	1710/1763	0.00071	-0.00009	-0.00168	0.0012	0.0001	0.0016	0.57	-1.18	-1.05	0.569	0.240	0.294	0.019(-0.05,0.09)

Table S2. Full linear mixed model results for the cortical surface area AIs [continued]

	N cases/ controls	β-value			Standard Error			*t*-value			*P*-value[1]			
AI surface area region		**diag**	**age**	**sex**	**diag**	**age**	**sex**	**diag**	**age**	**sex**	**diag**	**age**	**sex**	**Cohen's *d* (95% CI)**
superiorparietal	1702/1763	-0.00065	0.00002	-0.00233	0.0016	0.0001	0.0021	-0.40	0.16	-1.12	0.688	0.873	0.263	-0.014(-0.08,0.05)
superiortemporal	1691/1750	-0.00145	-0.00004	-0.00152	0.0015	0.0001	0.0020	-0.96	-0.34	-0.77	0.339	0.733	0.440	-0.033(-0.1,0.03)
supramarginal	1698/1759	0.00013	-0.00003	-0.00519	0.0022	0.0002	0.0029	0.06	-0.21	-1.80	0.954	0.831	0.072	0.002(-0.06,0.07)
temporalpole	1705/1751	0.00191	0.00050	0.00262	0.0030	0.0002	0.0039	0.65	2.29	0.68	0.517	**0.022**	0.497	0.022(-0.04,0.09)
transversetemporal	1704/1757	0.00046	-0.00023	-0.00068	0.0028	0.0002	0.0036	0.16	-1.23	-0.19	0.870	0.219	0.850	0.006(-0.06,0.07)
Total surface area	1706/1763	-0.00037	-0.00002	0.00008	0.0004	0.0000	0.0005	-0.95	-0.71	0.16	0.344	0.480	0.876	-0.032(-0.1,0.03)

with in **bold** those that are significant ($P < 0.05$) at the uncorrected level. Those p-values for diagnosis that survived multiple testing correction (FDR < 0.05) are indicated in *italic.*

Table S3. Full linear mixed model results for the subcortical volume AIs.

AI region	N cases/ controls	β-value			Standard Error			*t*-value			*P*-value[1]			Cohen's *d* (95% CI)
		diag	age	sex	diag	age	sex	diag	age	sex	diag	age	sex	
Accumbens	1716/1762	-0.00220	-0.00026	-0.00022	0.0028	0.0002	0.0036	-0.80	-1.09	-0.06	0.425	0.276	0.952	-0.027(-0.09,0.04)
Amygdala	1709/1768	0.00018	0.00002	0.00513	0.0022	0.0002	0.0029	0.08	0.13	1.77	0.934	0.893	0.077	0.003(-0.06,0.07)
CaudateNucleus	1711/1763	0.00293	0.00031	0.00223	0.0013	0.0001	0.0017	2.24	2.70	1.29	**0.025**	**0.007**	0.196	0.077(0.01,0.14)
GlobusPallidus	1707/1761	0.00113	-0.00053	0.00107	0.0021	0.0002	0.0028	0.53	-2.88	0.38	0.593	**0.004**	0.702	0.018(-0.05,0.08)
Hippocampus	1708/1762	-0.00106	0.00002	-0.00017	0.0015	0.0001	0.0020	-0.70	0.14	-0.08	0.485	0.885	0.934	-0.024(-0.09,0.04)
LateralVentricle	1660/1727	-0.00169	-0.00016	0.00306	0.0051	0.0003	0.0065	-0.33	-0.52	0.47	0.741	0.601	0.637	-0.011(-0.08,0.06)
Putamen	1712/1763	0.00395	0.00049	0.00273	0.0012	0.0001	0.0015	3.40	5.03	1.77	***0.001***	**$5.2 \cdot 10^{-7}$**	0.076	0.116(0.05,0.18)
Thalamus	1690/1763	0.00126	0.00063	0.00255	0.0011	0.0001	0.0015	1.15	6.36	1.75	0.252	**$2.3 \cdot 10^{-10}$**	0.081	0.039(-0.03,0.11)

[1]Unadjusted *P* values are shown, with in **bold** those that are significant ($P < 0.05$) at the uncorrected level. Those p-values for diagnosis that survived multiple testing correction (FDR < 0.05) are indicated in *italic*.

Table S4. Results of sensitivity analyses for cortical thickness AIs.

	winsorised		**non-linear age**		**3T**		**age ≥ 6y**		**age < 40y**	
AI thickness region	**P[1]**	***d***	**P[1]**	***d***	**P[1]**	***d***	**P[1]**	***d***	**P[1]**	***d***
bankssts	0.956	0.002	0.978	0.00	0.762	0.01	0.849	0.01	0.952	0.00
caudalanteriorcingulate	0.986	-0.001	0.965	0.00	0.578	0.02	0.769	0.01	0.867	-0.01
caudalmiddlefrontal	0.077	-0.06	0.090	-0.06	0.210	-0.05	0.149	-0.05	0.149	-0.05
cuneus	0.067	-0.06	0.065	-0.06	0.068	-0.07	0.093	-0.06		-0.06
entorhinal	0.091	0.06	0.103	0.06	0.287	0.04	0.184	0.05	0.096	0.06
frontalpole	0.761	-0.01	0.747	-0.01	0.979	0.00	0.914	0.00	0.786	-0.01
fusiform	***0.002***	0.11	***0.002***	0.11	***0.006***	0.10	***0.004***	0.10	***0.002***	0.11
inferiorparietal	0.052	-0.07	0.059	-0.06	0.131	-0.06	0.091	-0.06	**0.039**	-0.07
inferiortemporal	***0.003***	0.10	***0.003***	0.10	**0.016**	0.09	***0.004***	0.10	***0.002***	0.11
insula	0.259	0.04	0.262	0.04	0.303	0.04	0.231	0.04	0.224	0.04
isthmuscingulate	***0.010***	-0.09	***0.010***	-0.09	0.096	-0.06	**0.015**	-0.09	***0.006***	-0.10
lateraloccipital	0.355	0.03	0.368	0.03	0.459	0.03	0.308	0.04	0.462	0.03
lateralorbitofrontal	0.980	0.00	0.855	0.01	0.874	-0.01	0.944	0.00	0.996	0.00
lingual	0.625	0.02	0.607	0.02	0.316	0.04	0.759	0.01	0.633	0.02
medialorbitofrontal	***0.001***	-0.12	***0.001***	-0.12	***4.5·10^{-4}***	-0.13	***3.1·10^{-4}***	-0.13	***3.2·10^{-4}***	-0.13
middletemporal	0.684	-0.01	0.721	-0.01	0.530	-0.02	0.410	-0.03	0.829	-0.01
paracentral	0.204	0.04	0.231	0.04	0.116	0.06	0.247	0.04	0.325	0.03
parahippocampal	0.866	-0.01	0.835	-0.01	0.740	-0.01	0.940	0.00	0.930	0.00
parsopercularis	0.582	-0.02	0.584	-0.02	0.583	-0.02	0.601	-0.02	0.566	-0.02
parsorbitalis	**0.024**	-0.08	**0.024**	-0.08	**0.015**	-0.09	**0.021**	-0.08	**0.021**	-0.08
parstriangularis	0.746	-0.01	0.777	-0.01	0.746	-0.01	0.705	-0.01	0.682	-0.01
pericalcarine	0.544	-0.02	0.540	-0.02	0.854	0.01	0.829	-0.01	0.508	-0.02
postcentral	0.205	-0.04	0.197	-0.04	0.384	-0.03	0.309	-0.04	0.193	-0.05
posteriorcingulate	0.043	-0.07	**0.036**	-0.07	0.123	-0.06	**0.016**	-0.09	**0.023**	-0.08
precentral	0.482	-0.02	0.535	-0.02	0.763	-0.01	0.701	-0.01	0.774	-0.01
precuneus	0.744	-0.01	0.711	-0.01	0.931	0.00	0.858	-0.01	0.638	-0.02
rostralanteriorcingulate	***0.001***	-0.12	***0.001***	-0.12	***8.7·10^{-5}***	-0.15	***4.3·10^{-4}***	-0.12	***3.3·10^{-4}***	-0.13
rostralmiddlefrontal	***0.003***	-0.10	***0.004***	-0.10	***0.001***	-0.12	***0.001***	-0.12	***0.005***	-0.10
superiorfrontal	***8.7·10^{-5}***	-0.13	***1·10^{-4}***	-0.13	***7.4·10^{-6}***	-0.17	***5.9·10^{-5}***	-0.14	***7.2·10^{-5}***	-0.14
superiorparietal	0.736	0.01	0.740	0.01	0.498	0.03	0.663	0.02	0.586	0.02
superiortemporal	**0.049**	-0.07	**0.049**	-0.07	***0.004***	-0.11	0.051	-0.07	0.058	-0.07

	winsorised		**non-linear age**		**3T**		**age ≥ 6y**		**age < 40y**	
AI thickness region	**P[1]**	***d***	**P[1]**	***d***	**P[1]**	***d***	**P[1]**	***d***	**P[1]**	***d***
supramarginal	0.971	0.001	0.972	0.00	0.912	0.00	0.816	0.01	0.976	0.00
temporalpole	0.137	0.05	0.137	0.05	0.226	0.05	0.117	0.06	0.082	0.06
transversetemporal	0.374	0.03	0.375	0.03	0.504	0.03	0.701	0.01	0.303	0.04
Total average thickness	0.051	-0.07	**0.050**	-0.07	**0.039**	-0.08	**0.045**	-0.07	0.052	-0.07

Results are shown (1) after outliers were winsorized, (2) after a non-linear effect of age was added to the model, (3) after examining the subset of 3T acquired data, (4) after removing subjects below 6 year of age, and (5) after removing subjects aged 40 years or older. [1] Unadjusted p-values are shown, with in **bold** those that are significant ($P < 0.05$) at the uncorrected level. Those p-values for diagnosis that survived multiple testing correction (FDR < 0.05) are indicated in *italic.*

Table S5. Results of sensitivity analyses for cortical surface area AIs.

	winsorised		**non-linear age**		**3T**		**age ≥ 6y**		**age < 40y**	
AI surface area region	**P[1]**	***d***	**P[1]**	***d***	**P[1]**	***d***	**P[1]**	***d***	**P[1]**	***d***
bankssts	0.311	0.04	0.356	0.03	0.590	0.02	0.454	0.03	0.357	0.03
caudalanteriorcingulate	0.543	-0.02	0.643	-0.02	0.818	0.01	0.886	-0.01	0.636	-0.02
caudalmiddlefrontal	0.418	-0.03	0.500	-0.02	0.328	-0.04	0.232	-0.04	0.471	-0.03
cuneus	0.222	-0.04	0.232	-0.04	0.130	-0.06	0.309	-0.04	0.128	-0.05
entorhinal	0.117	-0.05	0.108	-0.06	**0.029**	-0.08	0.118	-0.06	0.153	-0.05
frontalpole	0.410	0.03	0.408	0.03	0.107	0.06	0.692	0.01	0.366	0.03
fusiform	**0.011**	-0.09	**0.010**	-0.09	***0.001***	-0.13	***0.002***	-0.11	**0.007**	-0.09
inferiorparietal	0.857	0.01	0.874	0.01	0.905	0.00	0.811	-0.01	0.961	0.00
inferiortemporal	0.594	-0.02	0.591	-0.02	0.519	-0.02	0.773	-0.01	0.420	-0.03
insula	0.675	0.01	0.703	0.01	0.508	0.03	0.592	0.02	0.741	0.01
isthmuscingulate	0.514	0.02	0.661	0.02	0.549	0.02	0.752	0.01	0.534	0.02
lateraloccipital	0.667	-0.02	0.620	-0.02	0.505	0.03	0.841	-0.01	0.776	-0.01
lateralorbitofrontal	***0.001***	-0.11	***0.001***	-0.11	***0.001***	-0.12	***0.001***	-0.12	***0.001***	-0.12
lingual	0.762	0.01	0.822	0.01	0.744	0.01	0.934	0.00	0.721	0.01
medialorbitofrontal	**0.003**	0.10	***0.002***	0.11	***0.001***	0.12	***0.002***	0.11	***0.001***	0.12
middletemporal	0.682	-0.01	0.672	-0.01	0.647	-0.02	0.850	-0.01	0.628	-0.02
paracentral	0.070	0.06	0.065	0.06	**0.047**	0.07	**0.048**	0.07	0.102	0.06
parahippocampal	0.808	-0.01	0.839	-0.01	0.960	0.00	0.947	0.00	0.620	-0.02
parsopercularis	0.895	0.01	0.932	0.00	0.945	0.00	0.740	0.01	0.680	0.01
parsorbitalis	0.143	0.05	0.144	0.05	0.207	0.05	0.099	0.06	0.196	0.05
parstriangularis	0.932	0.00	0.963	0.00	0.549	-0.02	0.992	0.00	0.668	0.02
pericalcarine	0.584	-0.02	0.562	-0.02	0.220	-0.05	0.476	-0.03	0.557	-0.02
postcentral	0.437	0.03	0.439	0.03	0.763	0.01	0.415	0.03	0.486	0.02
posteriorcingulate	0.466	-0.03	0.448	-0.03	0.808	-0.01	0.683	-0.01	0.486	-0.02
precentral	0.150	-0.05	0.149	-0.05	**0.030**	-0.08	0.059	-0.07	0.145	-0.05
precuneus	0.950	0.00	0.935	0.00	0.674	-0.02	0.601	-0.02	0.978	0.00

	winsorised		**non-linear age**		**3T**		**age ≥ 6y**		**age < 40y**	
AI surface area region	**P[1]**	***d***	**P[1]**	***d***	**P[1]**	***d***	**P[1]**	***d***	**P[1]**	***d***
rostralanteriorcingulate	0.329	-0.03	0.314	-0.03	0.323	-0.04	0.242	-0.04	0.245	-0.04
rostralmiddlefrontal	0.641	0.02	0.697	0.01	0.550	0.02	0.805	0.01	0.644	0.02
superiorfrontal	0.670	0.02	0.511	0.02	0.986	0.00	0.681	0.01	0.411	0.03
superiorparietal	0.717	-0.01	0.635	-0.02	0.779	-0.01	0.726	-0.01	0.609	-0.02
superiortemporal	0.356	-0.03	0.314	-0.03	0.215	-0.05	0.320	-0.04	0.603	-0.02
supramarginal	0.963	0.00	0.992	0.00	0.475	0.03	0.849	0.01	0.803	0.01
temporalpole	0.555	0.02	0.510	0.02	0.840	0.01	0.639	0.02	0.497	0.02
transversetemporal	0.887	0.01	0.864	0.01	0.937	0.00	0.901	0.00	0.906	0.00
totalsurf	0.431	-0.03	0.321	-0.03	0.196	-0.05	0.244	-0.04	0.411	-0.03

Results are shown (1) after outliers were winsorized, (2) after a non-linear effect of age was added to the model, (3) after examining the subset of 3T acquired data, (4) after removing subjects below 6 year of age, and (5) after removing subjects aged 40 year or older. [1]Unadjusted p-values are shown, with in **bold** those that are significant ($P < 0.05$) at the uncorrected level. Those p-values for diagnosis that survived multiple testing correction (FDR < 0.05) are indicated in *italic.*

Table S6. Results of sensitivity analyses in subcortical volume AIs.

	winsorised		**non-linear age**		**3T**		**age ≥ 6y**		**age < 40y**	
AI region	**P[1]**	***d***	**P[1]**	***d***	**P[1]**	***d***	**P[1]**	***d***	**P[1]**	***d***
Accumbens	0.386	-0.030	0.445	-0.026	0.560	-0.022	0.271	-0.039	0.544	-0.021
Amygdala	0.896	0.004	0.921	0.003	0.753	0.012	0.935	-0.003	0.765	-0.010
Caudate	**0.035**	0.072	**0.027**	0.076	***0.008***	0.099	0.014	0.086	0.031	0.075
Nucleus										
Globus Pallidus	0.666	0.015	0.640	0.016	0.300	0.039	0.446	0.027	0.844	0.007
Hippocampus	0.527	-0.022	0.504	-0.023	0.233	-0.044	0.455	-0.026	0.535	-0.022
Lateral Ventricle	0.741	-0.011	0.711	-0.013	0.988	-0.001	0.672	-0.015	0.759	-0.011
Putamen	***0.001***	0.118	***0.001***	0.117	***0.001***	0.125	***0.001***	0.121	***0.002***	0.106
Thalamus	0.251	0.039	0.217	0.042	**0.037**	0.078	0.134	0.053	0.104	0.057

Results are shown (1) after outliers were winsorized, (2) after a non-linear effect of age was added to the model, (3) after examining the subset of 3T acquired data, (4) after removing subjects below 6 year of age, and (5) after removing subjects aged 40 year or older.[1]Unadjusted *P* values are shown, with in **bold** those that are significant ($P < 0.05$) at the uncorrected level. Those p-values for diagnosis that survived multiple testing correction (FDR < 0.05) are also indicated in *italic.*

Table S7. Sex:diagnosis interaction effects for cortical thickness AIs.

	Primary analysis model				Sex:diagnosis interaction model					Males only		Females only	
Thickness region	**P_{diag}[1]**	***d***	**AIC**	**BIC**	**P_{diag}[1]**	***d***	**$P_{diag*sex}$[1]**	**AIC**	**BIC**	**P_{diag}[1]**	***d***	**P_{diag}[1]**	***d***
bankssts	0.950	0.002	-12107	-12070	0.672	-0.03	0.385	-12106	-12063	0.650	0.02	0.512	-0.05
caudalanteriorcingulate	0.982	-0.001	-10511	-10474	0.859	-0.02	0.639	-10509	-10466	0.970	0.00	0.722	-0.03
caudalmiddlefrontal	0.086	-0.06	-14975	-14938	0.447	-0.06	0.058	-14977	-14934	0.457	-0.03	**0.014**	-0.20
cuneus	0.066	-0.06	-13627	-13590	0.339	-0.06	0.092	-13628	-13585	0.382	-0.03	**0.015**	-0.20
entorhinal	0.095	0.06	-9348	-9311	0.087	-0.02	0.636	-9347	-9304	0.086	0.07	0.823	0.02
frontalpole	0.765	-0.01	-8929	-8892	0.616	0.02	0.586	-8927	-8884	0.630	-0.02	0.499	0.06
fusiform	***0.001***	0.11	-15724	-15687	***0.002***	-0.02	0.536	-15722	-15679	0.003	0.11	0.252	0.09
inferiorparietal	0.065	-0.06	-16323	-16286	0.271	-0.05	0.172	-16322	-16279	0.289	-0.04	0.061	-0.15
inferiortemporal	***0.003***	0.10	-14532	-14495	***0.002***	-0.03	0.314	-14531	-14488	***0.002***	0.12	0.663	0.04
insula	0.252	0.04	-15213	-15176	**0.030**	-0.09	**0.007**	-15218	-15175	0.037	0.08	0.181	-0.11
isthmuscingulate	***0.010***	-0.09	-12320	-12283	***0.006***	0.03	0.330	-12319	-12276	***0.007***	-0.10	0.666	-0.04
lateraloccipital	0.359	0.03	-16137	-16100	0.240	-0.03	0.414	-16135	-16092	0.259	0.04	0.797	-0.02
lateralorbitofrontal	0.872	0.01	-14413	-14376	0.930	0.02	0.576	-14411	-14368	0.930	0.00	0.783	0.02
lingual	0.594	0.02	-15045	-15008	0.305	-0.04	0.198	-15044	-15001	0.353	0.04	0.342	-0.08
medialorbitofrontal	***0.001***	-0.12	-13057	-13020	***0.0005***	0.03	0.420	-13055	-13012	***0.001***	-0.13	0.534	-0.05
middletemporal	***0.726***	-0.01	-14979	-14942	***0.567***	0.02	0.546	-14978	-14934	***0.535***	-0.02	0.564	0.05
paracentral	***0.231***	0.04	-15060	-15024	***0.074***	-0.06	0.095	-15061	-15018	***0.080***	0.07	0.171	-0.11
parahippocampal	***0.856***	-0.01	-11609	-11572	***0.558***	0.03	0.322	-11608	-11565	***0.533***	-0.02	0.210	0.10
parsopercularis	***0.568***	-0.02	-14483	-14446	***0.888***	-0.03	0.371	-14482	-14439	***0.901***	0.00	0.219	-0.10
parsorbitalis	***0.024***	-0.08	-11735	-11698	***0.112***	-0.04	0.278	-11734	-11691	***0.099***	-0.06	0.100	-0.05
parstriangularis	***0.759***	-0.01	-13813	-13776	***0.468***	0.04	0.288	-13812	-13769	***0.426***	-0.03	0.317	-0.03

Table S7. Sex:diagnosis interaction effects for cortical thickness AIs [continued]

	Primary analysis model				Sex:diagnosis interaction model					Males only		Females only	
Thickness region	**P_{diag}[1]**	***d***	**AIC**	**BIC**	**P_{diag}[1]**	***d***	**$P_{diag*sex}$[1]**	**AIC**	**BIC**	**P_{diag}[1]**	***d***	**P_{diag}[1]**	***d***
pericalcarine	0.536	-0.02	-12416	-12379	0.807	-0.03	0.451	-12414	-12371	0.822	-0.01	0.287	-0.20
postcentral	0.211	-0.04	-15802	-15765	0.462	-0.03	0.347	-15801	-15758	0.516	-0.02	0.194	-0.20
posteriorcingulate	**0.034**	-0.07	-13588	-13551	**0.019**	0.03	0.326	-13587	-13543	0.023	-0.09	0.813	0.02
precentral	0.547	-0.02	-16743	-16706	0.779	-0.02	0.528	-16742	-16699	0.770	-0.01	0.655	0.06
precuneus	0.710	-0.01	-16940	-16903	0.599	-0.07	0.040	-16942	-16899	0.669	0.02	0.033	0.09
rostralanteriorcingulate	***0.001***	-0.12	-11088	-11051	***$9.8{\cdot}10^{-6}$***	0.11	***0.001***	-11097	-11054	***$1.4{\cdot}10^{-5}$***	-0.17	0.165	-0.15
rostralmiddlefrontal	***0.003***	-0.10	-15671	-15634	***0.003***	0.02	0.512	-15669	-15626	***0.003***	-0.11	0.382	0.04
superiorfrontal	***$8.9{\cdot}10^{-5}$***	-0.13	-17795	-17758	***$2.9{\cdot}10^{-5}$***	0.05	0.135	-17795	-17752	***$1.5{\cdot}10^{-5}$***	-0.17	0.726	-0.11
superiorparietal	0.725	0.01	-17583	-17546	0.197	-0.08	**0.021**	-17586	-17543	0.236	0.05	**0.045**	-0.04
superiortemporal	**0.049**	-0.07	-15790	-15753	0.295	-0.06	0.079	-15791	-15748	0.292	-0.04	**0.017**	-0.02
supramarginal	0.980	0.00	-15627	-15590	0.587	-0.04	0.217	-15627	-15583	0.516	0.02	0.245	0.02
temporalpole	0.136	0.05	-9867	-9830	0.116	-0.02	0.602	-9865	-9822	0.155	0.05	0.708	-0.08
transversetemporal	0.364	0.03	-11161	-11124	0.275	-0.02	0.526	-11159	-11116	0.336	0.04	0.973	-0.05
Total average thickness	0.050	-0.07	-22896	-22859	0.249	-0.05	0.138	-22896	-22853	0.235	-0.05	**0.034**	0.05

P-values and Cohen's *d* effect sizes for the main effects of diagnosis are shown for the primary and sex:diagnosis interaction models, as well as for the stratification analyses separately in males and females. P-values for the sex:diagnosis interaction effects are also shown. Additionally, for the primary and interaction models the AIC and BIC model fit measures are shown. [1]Unadjusted *P* values are shown, with in **bold** those that are significant ($P < 0.05$) at the uncorrected level. Those that survived multiple testing correction (FDR < 0.05) are also indicated in *italic.*

Table S8. Sex:diagnosis interaction effects for cortical surface area AIs.

	Primary analysis model				Sex:diagnosis interaction model					Males only		Females only	
AI surface area region	**P_{diag}[1]**	***d***	**AIC**	**BIC**	**P_{diag}[1]**	***d***	**$P_{diag*sex}$[1]**	**AIC**	**BIC**	**P_{diag}[1]**	***d***	**P_{diag}[1]**	***d***
bankssts	0.332	0.03	-6952	-6915	0.565	0.02	0.488	-6951	-6908	0.586	0.02	0.331	0.08
caudalanteriorcingulate	0.662	-0.02	-5027	-4990	0.858	-0.02	0.610	-5025	-4982	0.891	-0.01	0.480	-0.06
caudalmiddlefrontal	0.483	-0.02	-7903	-7866	0.365	0.02	0.524	-7901	-7858	0.383	-0.03	0.818	0.02
cuneus	0.215	-0.04	-8508	-8471	0.680	-0.06	0.087	-8509	-8466	0.721	-0.01	**0.019**	-0.20
entorhinal	0.108	-0.06	-4612	-4576	0.180	-0.01	0.786	-4610	-4567	0.177	-0.05	0.242	-0.10
frontalpole	0.388	0.03	-6670	-6633	0.412	-0.003	0.932	-6668	-6625	0.388	0.03	0.827	0.02
fusiform	**0.010**	-0.09	-10454	-10417	**0.009**	0.02	0.534	-10452	-10409	**0.008**	-0.10	0.611	-0.04
inferiorparietal	0.857	0.01	-10288	-10251	0.982	0.02	0.660	-10286	-10243	0.979	-0.001	0.527	0.05
inferiortemporal	0.585	-0.02	-9610	-9573	0.947	-0.04	0.307	-9609	-9566	0.913	-0.004	0.214	-0.10
insula	0.701	0.01	-11222	-11185	0.930	0.02	0.532	-11220	-11177	0.997	-0.0001	0.561	0.05
isthmuscingulate	0.611	0.02	-7974	-7937	0.839	0.02	0.540	-7973	-7930	0.796	0.01	0.405	0.07
lateraloccipital	0.654	-0.02	-10709	-10672	0.492	0.02	0.504	-10707	-10664	0.422	-0.03	0.729	0.03
lateralorbitofrontal	***0.001***	-0.11	-11771	-11734	***0.001***	0.02	0.567	-11769	-11726	**0.003**	-0.12	0.166	-0.11
lingual	0.820	0.01	-9826	-9789	0.949	0.02	0.517	-9825	-9782	0.904	-0.005	0.594	0.04
medialorbitofrontal	***0.002***	0.11	-9475	-9438	**0.009**	0.02	0.526	-9474	-9431	**0.011**	0.10	**0.020**	0.19
middletemporal	0.661	-0.02	-11137	-11100	0.891	-0.02	0.537	-11135	-11092	0.835	-0.01	0.559	-0.05
paracentral	0.071	0.06	-9051	-9014	0.232	0.04	0.298	-9050	-9007	0.187	0.05	0.075	0.15
parahippocampal	0.853	-0.01	-7631	-7594	0.990	-0.01	0.668	-7629	-7586	0.934	0.003	0.628	-0.04
parsopercularis	0.880	0.01	-7494	-7457	0.912	0.02	0.564	-7492	-7449	0.982	0.001	0.628	0.04
parsorbitalis	0.133	0.05	-9436	-9399	0.233	0.01	0.688	-9434	-9391	0.202	0.05	0.363	0.08
parstriangularis	0.929	0.00	-7851	-7814	0.608	0.05	0.159	-7851	-7808	0.581	-0.02	0.182	0.11

Table S8. Sex:diagnosis interaction effects for cortical surface area AIs [continued]

	Primary analysis model				Sex:diagnosis interaction model					Males only		Females only	
AI surface area region	**P_{diag}[1]**	***d***	**AIC**	**BIC**	**P_{diag}[1]**	***d***	**$P_{diag*sex}$[1]**	**AIC**	**BIC**	**P_{diag}[1]**	***d***	**P_{diag}[1]**	***d***
pericalcarine	0.561	-0.02	-8905	-8868	0.933	-0.05	0.143	-8905	-8862	0.940	0.00	0.070	-0.15
postcentral	0.430	0.03	-10895	-10858	0.706	0.03	0.436	-10893	-10850	0.681	0.02	0.305	0.08
posteriorcingulate	0.475	-0.02	-7924	-7887	0.252	0.04	0.241	-7923	-7880	0.265	-0.04	0.425	0.07
precentral	0.154	-0.05	-12203	-12166	0.188	0.00	0.952	-12201	-12158	0.279	-0.04	0.622	-0.04
precuneus	0.915	-0.004	-11488	-11451	0.760	0.02	0.626	-11486	-11443	0.708	-0.01	0.723	0.03
rostralanteriorcingulate	0.317	-0.03	-6163	-6126	0.191	0.03	0.342	-6162	-6119	0.220	-0.05	0.633	0.04
rostralmiddlefrontal	0.694	0.01	-11642	-11605	0.859	0.04	0.204	-11642	-11599	0.753	-0.01	0.187	0.11
superiorfrontal	0.569	0.02	-13194	-13157	0.521	-0.01	0.766	-13192	-13149	0.508	0.03	0.951	0.01
superiorparietal	0.688	-0.01	-11377	-11340	0.699	0.00	0.955	-11375	-11332	0.736	-0.01	0.901	-0.01
superiortemporal	0.339	-0.03	-11787	-11750	0.244	0.02	0.478	-11786	-11743	0.259	-0.04	0.847	0.02
supramarginal	0.954	0.002	-9191	-9154	0.806	0.02	0.478	-9190	-9147	0.838	-0.01	0.498	0.06
temporalpole	0.517	0.02	-7204	-7168	0.350	-0.03	0.409	-7203	-7160	0.400	0.03	0.495	-0.06
transversetemporal	0.870	0.01	-7600	-7563	0.908	0.003	0.939	-7598	-7555	0.958	0.00	0.937	0.01
totalsurf	0.344	-0.03	-21189	-21152	0.116	0.06	0.088	-21190	-21147	0.127	-0.06	0.269	0.09

P-values and Cohen's d effect sizes for the main effects of diagnosis are shown for the primary and sex:diagnosis interaction models, as well as for the stratification analyses separately in males and females. P-values for the sex:diagnosis interaction effects are also shown. Additionally, for the primary and interaction models the AIC and BIC model fit measures are shown.[1] Unadjusted *P* values are shown, with in **bold** those that are significant ($P < 0.05$) at the uncorrected level. Those that survived multiple testing correction (FDR < 0.05) are also indicated in *italic*.

Table S9. Sex:diagnosis interaction effects for subcortical volume AIs.

	Primary analysis model				Sex:diagnosis interaction model					Males only		Females only	
region	**P_{diag}[1]**	***d***	**AIC**	**BIC**	**P_{diag}[1]**	***d***	**$P_{diag*sex}$[1]**	**AIC**	**BIC**	**P_{diag}[1]**	***d***	**P_{diag}[1]**	***d***
Accumbens	0.425	-0.03	-7684	-7647	0.231	0.04	0.26	-7683	-7640	0.234	-0.05	0.469	0.06
Amygdala	0.934	0.003	-9229	-9192	0.547	-0.04	0.21	-9229	-9186	0.561	0.02	0.167	-0.11
CaudateNucleus	**0.025**	0.08	-12825	-12788	***0.010***	-0.04	0.19	-12825	-12782	***0.011***	0.10	0.959	-0.004
GlobusPallidus	0.593	0.02	-9466	-9429	0.201	-0.06	0.06	-9467	-9424	0.204	0.05	0.209	-0.10
Hippocampus	0.485	-0.02	-11814	-11777	0.600	-0.01	0.79	-11812	-11769	0.579	-0.02	0.412	-0.07
LateralVentricle	0.741	-0.01	-3330	-3293	0.882	-0.04	0.30	-3329	-3286	0.879	0.01	0.264	-0.09
Putamen	***0.001***	0.12	-13668	-13632	***0.001***	-0.01	0.70	-13667	-13624	***0.001***	0.13	0.312	0.08
Thalamus	0.252	0.04	-13885	-13848	0.486	0.03	0.41	-13884	-13841	0.638	0.02	0.060	0.15

P-values and Cohen's d effect sizes for the main effects of diagnosis are shown for the primary and sex:diagnosis interaction models, as well as for the stratification analyses separately in males and females. P-values for the sex:diagnosis interaction effects are also shown. Additionally, for the primary and interaction models the AIC and BIC model fit measures are shown.[1]Unadjusted *P* values are shown, with in **bold** those that are significant ($P < 0.05$) at the uncorrected level. Those that survived multiple testing correction (FDR < 0.05) are also indicated in *italic*

Table S10. Age:diagnosis interaction effects for cortical thickness AIs.

	Primary analysis model				Age:diagnosis interaction model					Children only		Adults only	
AI thickness region	**P_{diag}[1]**	***d***	**AIC**	**BIC**	**P_{diag}[1]**	***d***	**$P_{diag*age}$[1]**	**AIC**	**BIC**	**P_{diag}[1]**	***d***	**P_{diag}[1]**	***d***
bankssts	0.950	0.002	-12107	-12070	0.819	0.01	0.767	-12105	-12062	0.542	-0.025	0.306	0.06
caudalanteriorcingulate	0.982	-0.001	-10511	-10474	0.291	0.04	0.232	-10510	-10467	0.610	-0.021	0.402	0.05
caudalmiddlefrontal	0.086	-0.06	-14975	-14938	0.840	-0.02	0.468	-14974	-14931	0.368	-0.037	0.078	-0.11
cuneus	0.066	-0.06	-13627	-13590	0.179	0.02	0.607	-13625	-13582	0.087	-0.071	0.337	-0.06
entorhinal	0.095	0.06	-9348	-9311	0.102	-0.03	0.348	-9347	-9304	0.288	0.044	0.131	0.10
frontalpole	0.765	-0.01	-8929	-8892	0.864	-0.01	0.716	-8927	-8884	0.665	0.018	0.235	-0.07
fusiform	***0.001***	0.11	-15724	-15687	**0.006**	-0.05	0.180	-15724	-15681	***0.001***	0.135	0.279	0.07
inferiorparietal	0.065	-0.06	-16323	-16286	0.171	0.02	0.589	-16321	-16278	0.140	-0.061	0.313	-0.06
inferiortemporal	***0.003***	0.10	-14532	-14495	**0.031**	-0.03	0.412	-14530	-14487	0.017	0.098	0.067	0.12
insula	0.252	0.04	-15213	-15176	0.206	-0.03	0.418	-15211	-15168	0.212	0.051	0.833	0.01
isthmuscingulate	***0.010***	-0.09	-12320	-12283	0.072	0.02	0.535	-12318	-12275	**0.011**	-0.104	0.342	-0.06
lateraloccipital	0.359	0.03	-16137	-16100	0.650	-0.0003	0.992	-16135	-16092	0.319	0.041	0.832	0.01
lateralorbitofrontal	0.872	0.01	-14413	-14376	0.558	0.03	0.447	-14411	-14368	0.998	0.000	0.808	0.02
lingual	0.594	0.02	-15045	-15008	0.294	-0.03	0.366	-15044	-15000	0.351	0.038	0.804	-0.02
medialorbitofrontal	***0.001***	-0.12	-13057	-13020	**0.023**	0.02	0.499	-13055	-13012	***0.001***	-0.143	0.241	-0.07
middletemporal	0.726	-0.01	-14979	-14942	0.219	-0.05	0.109	-14980	-14937	0.576	0.023	0.112	-0.10
paracentral	0.231	0.04	-15060	-15024	0.629	0.004	0.909	-15058	-15015	0.124	0.063	0.935	0.01
parahippocampal	0.856	-0.01	-11609	-11572	0.258	-0.05	0.162	-11609	-11566	0.221	0.050	0.065	-0.12
parsopercularis	0.568	-0.02	-14483	-14446	0.989	-0.01	0.739	-14482	-14438	0.653	-0.019	0.748	-0.02
parsorbitalis	**0.024**	-0.08	-11735	-11698	0.372	-0.01	0.814	-11733	-11690	**0.035**	-0.087	0.449	-0.05
parstriangularis	0.759	-0.01	-13813	-13776	0.817	-0.01	0.663	-13811	-13768	0.989	0.001	0.673	-0.03

Table S10. Age:diagnosis interaction effects for cortical thickness AIs [continued]

	Primary analysis model				Age:diagnosis interaction model					Children only		Adults only	
AI thickness region	**P_{diag}[1]**	***d***	**AIC**	**BIC**	**P_{diag}[1]**	***d***	**$P_{diag*age}$[1]**	**AIC**	**BIC**	**P_{diag}[1]**	***d***	**P_{diag}[1]**	***d***
pericalcarine	0.536	-0.02	-12416	-12379	**0.008**	0.09	**0.008**	-12421	-12378	0.081	-0.072	0.116	0.10
postcentral	0.211	-0.04	-15802	-15765	0.306	-0.06	0.061	-15804	-15761	0.669	0.018	**0.003**	-0.19
posteriorcingulate	**0.034**	-0.07	-13588	-13551	0.733	-0.03	0.427	-13586	-13543	0.232	-0.049	0.104	-0.10
precentral	0.547	-0.02	-16743	-16706	0.179	-0.06	0.061	-16745	-16702	0.533	0.026	0.043	-0.13
precuneus	0.710	-0.01	-16940	-16903	0.223	0.04	0.235	-16939	-16896	0.242	-0.048	0.275	0.07
rostralanteriorcingulate	***0.001***	-0.12	-11088	-11051	**0.015**	0.03	0.370	-11087	-11044	**0.007**	-0.111	0.058	-0.12
rostralmiddlefrontal	***0.003***	-0.10	-15671	-15634	0.579	-0.03	0.315	-15670	-15627	**0.019**	-0.097	0.131	-0.09
superiorfrontal	***$8.9{\cdot}10^{-5}$***	-0.13	-17795	-17758	0.056	0.0002	0.995	-17793	-17750	**0.009**	-0.107	**0.004**	-0.18
superiorparietal	0.725	0.01	-17583	-17546	0.255	-0.04	0.268	-17582	-17539	0.473	0.030	0.519	-0.04
superiortemporal	**0.049**	-0.07	-15790	-15753	0.439	-0.01	0.832	-15788	-15745	**0.040**	-0.085	0.474	-0.05
supramarginal	0.980	0.00	-15627	-15590	0.578	-0.02	0.533	-15625	-15582	0.490	0.028	0.374	-0.06
temporalpole	0.136	0.05	-9867	-9830	0.238	-0.02	0.604	-9865	-9822	0.119	0.064	0.614	0.03
transversetemporal	0.364	0.03	-11161	-11124	0.812	0.01	0.815	-11159	-11116	0.922	0.004	0.169	0.09
Total average thickness	0.050	-0.07	-22896	-22859	0.835	-0.05	0.183	-22895	-22852	0.358	-0.038	**0.043**	-0.13

P-values and Cohen's *d* effect sizes for the main effects of diagnosis are shown for the primary and age:diagnosis interaction models, as well as for the stratification analyses separately in children and adults. P-values for the age:diagnosis interaction effects are also shown. Additionally, for both primary and interaction models the AIC and BIC model fit measures are shown.[1]Unadjusted *P* values are shown, with in **bold** those that are significant ($P < 0.05$) at the uncorrected level. Those that survived multiple testing correction (FDR < 0.05) are also indicated in *italic*

Table S11. Age:diagnosis interaction effects for cortical surface area AIs.

	Primary analysis model				Age:diagnosis interaction model					Children only		Adults only	
AI surface area region	**P_{diag}[1]**	***d***	**AIC**	**BIC**	**P_{diag}[1]**	***d***	**$P_{diag*age}$[1]**	**AIC**	**BIC**	**P_{diag}[1]**	***d***	**P_{diag}[1]**	***d***
bankssts	0.332	0.03	-6952	-6915	0.365	-0.02	0.621	-6950	-6907	0.277	0.05	0.779	0.018
caudalanteriorcingulate	0.662	-0.02	-5027	-4990	0.076	0.06	0.074	-5028	-4985	0.355	-0.04	0.542	0.038
caudalmiddlefrontal	0.483	-0.02	-7903	-7866	0.571	-0.04	0.298	-7902	-7859	0.848	0.01	0.105	-0.102
cuneus	0.215	-0.04	-8508	-8471	0.032	0.06	0.079	-8509	-8466	0.190	-0.05	0.609	-0.032
entorhinal	0.108	-0.06	-4612	-4576	0.954	-0.03	0.406	-4611	-4568	0.230	-0.05	0.298	-0.066
frontalpole	0.388	0.03	-6670	-6633	0.151	-0.04	0.245	-6669	-6626	0.230	0.05	0.917	-0.007
fusiform	**0.010**	-0.09	-10454	-10417	0.157	0.01	0.860	-10452	-10409	**0.017**	-0.10	0.217	-0.077
inferiorparietal	0.857	0.01	-10288	-10251	0.361	-0.03	0.345	-10286	-10243	0.687	0.02	0.836	-0.013
inferiortemporal	0.585	-0.02	-9610	-9573	0.190	0.04	0.232	-9609	-9566	0.625	-0.02	0.739	-0.021
insula	0.701	0.01	-11222	-11185	0.229	0.05	0.111	-11222	-11179	0.652	-0.02	0.185	0.083
isthmuscingulate	0.611	0.02	-7974	-7937	0.089	-0.06	0.096	-7975	-7932	0.116	0.06	0.285	-0.067
lateraloccipital	0.654	-0.02	-10709	-10672	0.247	0.04	0.282	-10708	-10665	0.197	-0.05	0.272	0.069
lateralorbitofrontal	**0.001**	-0.11	-11771	-11734	0.042	0.02	0.622	-11769	-11726	***0.001***	-0.13	0.247	-0.073
lingual	0.820	0.01	-9826	-9789	0.609	-0.02	0.646	-9824	-9781	0.597	0.02	0.734	-0.021
medialorbitofrontal	0.002	0.11	-9475	-9438	**0.041**	-0.02	0.568	-9474	-9431	0.009	0.11	0.068	0.114
middletemporal	0.661	-0.02	-11137	-11100	0.478	-0.04	0.290	-11136	-11093	0.919	0.00	0.385	-0.054
paracentral	0.071	0.06	-9051	-9014	0.723	0.05	0.156	-9051	-9008	0.381	0.04	0.074	0.112
parahippocampal	0.853	-0.01	-7631	-7594	0.181	0.05	0.154	-7631	-7588	0.457	-0.03	0.405	0.052
parsopercularis	0.880	0.01	-7494	-7457	0.591	0.02	0.487	-7493	-7450	0.915	-0.004	0.670	0.027
parsorbitalis	0.133	0.05	-9436	-9399	0.979	0.03	0.383	-9435	-9392	0.212	0.05	0.617	0.031
parstriangularis	0.929	0.00	-7851	-7814	0.065	-0.07	**0.039**	-7853	-7810	0.516	0.03	0.267	-0.070

Table S11. Age:diagnosis interaction effects for cortical surface area AIs [continued]

	Primary analysis model				Age:diagnosis interaction model					Children only		Adults only	
AI surface area region	**P_{diag}[1]**	***d***	**AIC**	**BIC**	**P_{diag}[1]**	***d***	**$P_{diag*age}$[1]**	**AIC**	**BIC**	**P_{diag}[1]**	***d***	**P_{diag}[1]**	***d***
pericalcarine	0.561	-0.02	-8905	-8868	0.790	-0.001	0.984	-8903	-8860	1.000	$1.1 \cdot 10^{-5}$	0.232	-0.075
postcentral	0.430	0.03	-10895	-10858	0.463	-0.01	0.689	-10893	-10850	0.212	0.05	0.639	-0.029
posteriorcingulate	0.475	-0.02	-7924	-7887	0.551	0.01	0.777	-7922	-7879	0.520	-0.03	0.698	-0.024
precentral	0.154	-0.05	-12203	-12166	0.894	-0.03	0.344	-12202	-12159	0.528	-0.03	0.110	-0.100
precuneus	0.915	-0.004	-11488	-11451	0.120	-0.06	0.066	-11489	-11446	0.474	0.03	0.215	-0.078
rostralanteriorcingulate	0.317	-0.03	-6163	-6126	0.709	-0.005	0.895	-6161	-6118	0.974	0.00	0.056	-0.120
rostralmiddlefrontal	0.694	0.01	-11642	-11605	0.628	-0.01	0.737	-11640	-11597	0.977	0.00	0.426	0.050
superiorfrontal	0.569	0.02	-13194	-13157	0.090	-0.06	0.106	-13194	-13151	0.104	0.07	0.215	-0.078
superiorparietal	0.688	-0.01	-11377	-11340	0.353	0.03	0.400	-11375	-11332	0.632	-0.02	0.984	0.001
superiortemporal	0.339	-0.03	-11787	-11750	0.028	-0.11	**0.002**	-11795	-11752	0.327	0.04	**0.002**	-0.194
supramarginal	0.954	0.002	-9191	-9154	0.159	-0.05	0.113	-9192	-9149	0.437	0.03	0.354	-0.058
temporalpole	0.517	0.02	-7204	-7168	0.430	0.04	0.205	-7204	-7161	0.945	-0.003	0.273	0.069
transversetemporal	0.870	0.01	-7600	-7563	0.407	-0.03	0.391	-7598	-7555	0.721	0.01	0.807	-0.015
totalsurf	0.344	-0.03	-21189	-21152	0.592	-0.04	0.253	-21188	-21145	0.966	0.002	0.149	-0.090

P-values and Cohen's *d* effect sizes for the main effects of diagnosis are shown for the primary and age:diagnosis interaction models, as well as for the stratification analyses separately in children and adults. P-values for the age:diagnosis interaction effects are also shown. Additionally, for both primary and interaction models the AIC and BIC model fit measures are shown.[1]Unadjusted *P* values are shown, with in **bold** those that are significant ($P < 0.05$) at the uncorrected level. Those that survived multiple testing correction (FDR < 0.05) are also indicated in *italic*

Table S12. Age:diagnosis interaction effects for subcortical volume AIs.

	Primary analysis model				Age:diagnosis interaction model					Children only		Adults only	
region	**P_{diag}[1]**	***d***	**AIC**	**BIC**	**P_{diag}[1]**	***d***	**$P_{diag*age}$[1]**	**AIC**	**BIC**	**P_{diag}[1]**	***d***	**P_{diag}[1]**	***d***
Accumbens	0.425	-0.03	-7684	-7647	0.417	-0.05	0.169	-7684	-7641	0.990	0.001	0.180	-0.084
Amygdala	0.934	0.003	-9229	-9192	0.149	0.06	0.089	-9230	-9187	0.300	-0.04	0.109	0.101
CaudateNucleus	0.025	0.08	-12825	-12788	0.986	0.04	0.204	-12824	-12781	0.152	0.06	0.138	0.093
GlobusPallidus	0.593	0.02	-9466	-9429	0.056	0.09	**0.013**	-9470	-9427	0.818	-0.01	0.238	0.074
Hippocampus	0.485	-0.02	-11814	-11777	0.187	0.04	0.262	-11814	-11770	0.189	-0.05	0.572	0.035
LateralVentricle	0.741	-0.01	-3330	-3293	0.696	-0.02	0.529	-3329	-3286	0.892	-0.01	0.713	-0.024
Putamen	***0.001***	0.12	-13668	-13632	0.331	0.03	0.434	-13667	-13624	***0.001***	0.13	0.255	0.071
Thalamus	0.252	0.04	-13885	-13848	0.240	-0.02	0.480	-13884	-13840	0.400	0.03	0.609	0.032

P-values and Cohen's *d* effect sizes for the main effects of diagnosis are shown for the primary and age:diagnosis interaction models, as well as for the stratification analyses separately in children and adults. P-values for the age:diagnosis interaction effects are also shown. Additionally, for both primary and interaction models the AIC and BIC model fit measures are shown.[1]Unadjusted *P* values are shown, with in **bold** those that are significant ($P < 0.05$) at the uncorrected level. Those that survived multiple testing correction (FDR < 0.05) are also indicated in *italic*

Table S13. Association of IQ with asymmetry within cases (top) and within controls (bottom), for the AIs that showed significant effects of diagnosis in the primary analysis.

		β-value			Standard Error			*t*-value			*P*-value[1]		
AI region	**N total**	**IQ**	**sex**	**age**	**IQ**	**sex**	**age**	**IQ**	**sex**	**age**	**IQ**	**sex**	**age**
within cases:													
fusiform thick	1360	-0.00007	0.00054	-0.00001	0.00004	0.00198	0.00011	-1.75	0.27	-0.13	0.081	0.784	0.897
inferiortemporal thick	1361	-0.00008	-0.00124	0.00004	0.00005	0.00230	0.00012	-1.72	-0.54	0.36	0.085	0.591	0.718
isthmuscingulate thick	1358	-0.00007	0.00557	0.00038	0.00006	0.00304	0.00013	-1.25	1.83	2.88	0.212	0.067	**0.004**
medialorbitofrontal thick	1360	-0.00004	0.00073	0.00046	0.00005	0.00272	0.00014	-0.68	0.27	3.21	0.498	0.789	**0.001**
rostralanteriorcingulate thick	1357	0.00019	0.00950	0.00004	0.00008	0.00376	0.00020	2.49	2.53	0.20	**0.013**	**0.012**	0.844
rostralmiddlefrontal thick	1362	-0.00003	0.00245	-0.00031	0.00004	0.00187	0.00011	-0.75	1.31	-2.89	0.456	0.190	**0.004**
superiorfrontal thick	1362	-0.00003	-0.00001	0.00005	0.00003	0.00144	0.00008	-0.91	-0.01	0.56	0.365	0.996	0.577
lateralorbitofrontal surf	1351	-0.00001	0.00286	0.00003	0.00007	0.00343	0.00018	-0.21	0.83	0.19	0.833	0.404	0.847
medialorbitofrontal surf	1352	-0.00006	0.00055	0.00009	0.00010	0.00478	0.00024	-0.66	0.11	0.38	0.509	0.909	0.702
Putamen	1372	0.00003	0.00424	0.00052	0.00006	0.00278	0.00014	0.58	1.52	3.68	0.563	0.128	**0.0002**
within controls:													
fusiform thick	1432	-0.00002	0.00313	0.00009	0.00005	0.00162	0.00011	-0.42	1.94	0.84	0.677	0.053	0.404
inferiortemporal thick	1431	-0.00010	0.00262	0.00008	0.00006	0.00187	0.00013	-1.82	1.40	0.66	0.069	0.161	0.511
isthmuscingulate thick	1431	0.00001	0.00241	0.00026	0.00008	0.00261	0.00015	0.13	0.92	1.74	0.896	0.358	0.082
medialorbitofrontal thick	1431	0.00003	-0.00183	0.00036	0.00007	0.00247	0.00017	0.39	-0.74	2.12	0.695	0.458	**0.034**
rostralanteriorcingulate thick	1428	0.00009	-0.00262	-0.00020	0.00010	0.00321	0.00022	0.91	-0.82	-0.93	0.361	0.415	0.353
rostralmiddlefrontal thick	1431	-0.00008	0.00133	-0.00013	0.00005	0.00170	0.00013	-1.60	0.78	-1.04	0.111	0.434	0.299
superiorfrontal thick	1432	-0.00012	-0.00140	-0.00002	0.00004	0.00118	0.00009	-3.41	-1.18	-0.20	**0.001**	0.236	0.839
lateralorbitofrontal surf	1426	-0.00008	-0.00191	-0.00026	0.00008	0.00282	0.00018	-0.96	-0.68	-1.45	0.335	0.498	0.147

Table S13. Association of IQ with asymmetry within cases (top) and within controls (bottom), for the AIs that showed significant effects of diagnosis in the primary analysis [continued]

		β-value			Standard Error			*t*-value			*P*-value[1]		
AI region	**N total**	**IQ**	**sex**	**age**	**IQ**	**sex**	**age**	**IQ**	**sex**	**age**	**IQ**	**sex**	**age**
within controls:													
medialorbitofrontal surf	1426	-0.00018	-0.00285	0.00029	0.00012	0.00406	0.00026	-1.51	-0.70	1.09	0.132	0.482	0.276
Putamen	1426	0.00005	0.00145	0.00024	0.00006	0.00206	0.00015	0.82	0.71	1.67	0.414	0.480	

[1]Unadjusted *P* values are shown, with in **bold** those that are significant ($P < 0.05$) at the uncorrected level. Thick=thickness, surf=surface area

Table S14. Association of log_{10}-normalized ADOS severity scores with brain asymmetry within cases**,** for the AIs that showed significant effects of diagnosis in the primary analysis.

		β-value			Standard Error			*t*-value			*P*-value[1]		
AI region	**N cases**	**ADOS**	**sex**	**age**	**ADOS**	**sex**	**age**	**ADOS**	**sex**	**age**	**ADOS**	**sex**	**age**
fusiform thickness	855	0.0059	0.0036	-0.0001	0.0060	0.0031	0.0001	0.98	1.17	-0.82	0.325	0.242	0.415
inferiortemporal thickness	857	-0.0016	0.0007	$3.37 \cdot 10^{-5}$	0.0067	0.0034	0.0001	-0.24	0.20	0.25	0.807	0.841	0.805
isthmuscingulate thickness	853	0.0192	0.0040	0.0006	0.0069	0.0046	0.0001	2.76	0.87	3.97	**0.006**	0.382	$\mathbf{7.8 \cdot 10^{-5}}$
medialorbitofrontal thickness	856	0.0047	0.0002	0.0003	0.0078	0.0042	0.0002	0.61	0.04	2.22	0.543	0.970	**0.027**
rostralanteriorcingulate thickness	854	-0.0072	0.0065	$1.93 \cdot 10^{-5}$	0.0117	0.0060	0.0002	-0.61	1.07	0.08	0.539	0.284	0.936
rostralmiddlefrontal thickness	858	0.0015	0.0024	-0.0004	0.0059	0.0029	0.0001	0.26	0.84	-2.86	0.796	0.401	**0.004**
superiorfrontal thickness	858	-0.0009	-0.0001	$-5.32 \cdot 10^{-6}$	0.0043	0.0021	0.0001	-0.20	-0.07	-0.06	0.842	0.948	0.953
lateralorbitofrontal surface area	852	0.0039	0.0061	0.0001	0.0099	0.0050	0.0002	0.39	1.22	0.67	0.696	0.223	0.503
medialorbitofrontal surface area	854	-0.0032	0.0009	0.0002	0.0124	0.0066	0.0002	-0.26	0.13	0.84	0.797	0.894	0.401
Putamen	866	0.0077	0.0036	0.0004	0.0067	0.0035	0.0001	1.15	1.03	3.21	0.252	0.303	**0.001**

[1] Unadjusted *P* values are shown, with in **bold** those that are significant ($P < 0.05$) at the uncorrected level.

Table S15. Association of medication use with brain asymmetry within cases, for the AIs that showed significant effects of diagnosis in the primary analysis.

		β-value			Standard Error			*t*-value			*P*-value[1]		
AI region	**N cases**	**med**	**sex**	**age**	**med**	**sex**	**age**	**med**	**sex**	**age**	**med**	**sex**	**age**
fusiform thickness	804	-0.0023	0.0046	0.0001	0.0023	0.0029	0.0002	-1.00	1.60	0.77	0.316	0.110	0.444
inferiortemporal thickness	804	-0.0028	0.0022	0.0002	0.0026	0.0032	0.0002	-1.08	0.69	0.94	0.280	0.493	0.349
isthmuscingulate thickness	799	0.0031	-0.0013	0.0007	0.0035	0.0045	0.0002	0.88	-0.29	4.10	0.377	0.775	**$4.6 \cdot 10^{-5}$**
medialorbitofrontal thickness	805	0.0000	0.0008	0.0004	0.0031	0.0039	0.0002	0.00	0.21	2.05	0.997	0.837	**0.041**
rostral anterior cingulate thickness	803	0.0045	0.0077	-0.0001	0.0044	0.0054	0.0003	1.03	1.41	-0.36	0.302	0.159	0.718
rostralmiddlefrontal thickness	807	0.0018	0.0028	-0.0003	0.0021	0.0027	0.0002	0.84	1.05	-1.96	0.400	0.295	0.051
superiorfrontal thickness	807	0.0006	-0.0005	0.0000	0.0017	0.0021	0.0001	0.35	-0.24	-0.36	0.730	0.814	0.717
lateralorbitofrontal surface area	801	0.0030	0.0032	0.0001	0.0036	0.0045	0.0003	0.84	0.71	0.54	0.401	0.477	0.592
medialorbitofrontal surface area	803	0.0021	-0.0055	$4.4 \cdot 10^{-5}$	0.0048	0.0060	0.0003	0.43	-0.91	0.14	0.668	0.361	0.889
Putamen	813	0.0018	-0.0015	0.0005	0.0027	0.0034	0.0002	0.67	-0.43	2.71	0.502	0.666	**0.007**

[1] Unadjusted *P* values are shown, with in **bold** those that are significant ($P < 0.05$) at the uncorrected level.

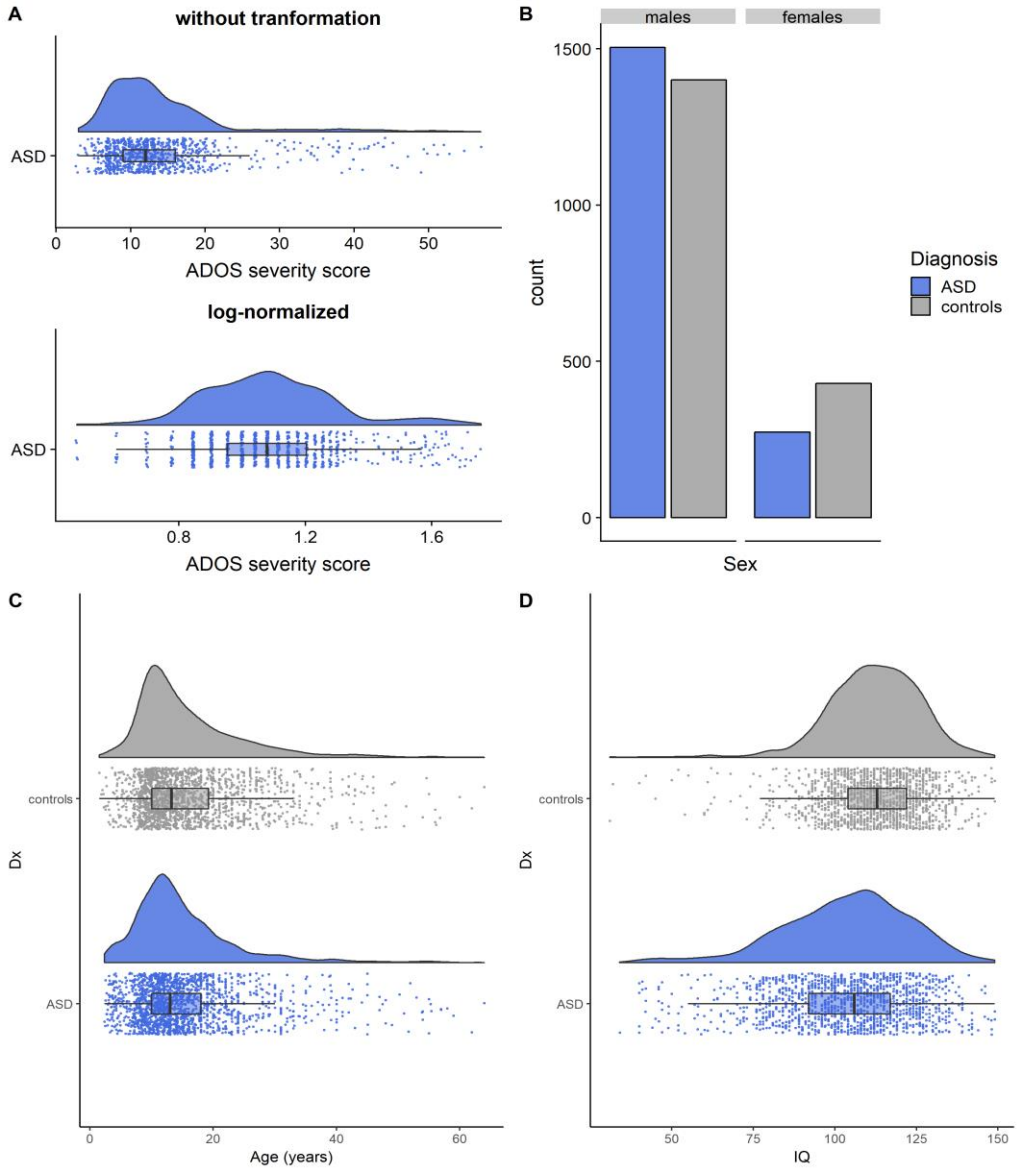


Figure S1. Distributions of variables of interest. (A) ADOS severity scores within cases, (B) sex, (C) age, (D) IQ. ADOS severity scores did not follow a normal distribution (top), so data were transformed using $\log_{10}$ (bottom). *Blue* = ASD group; *grey* = control group.

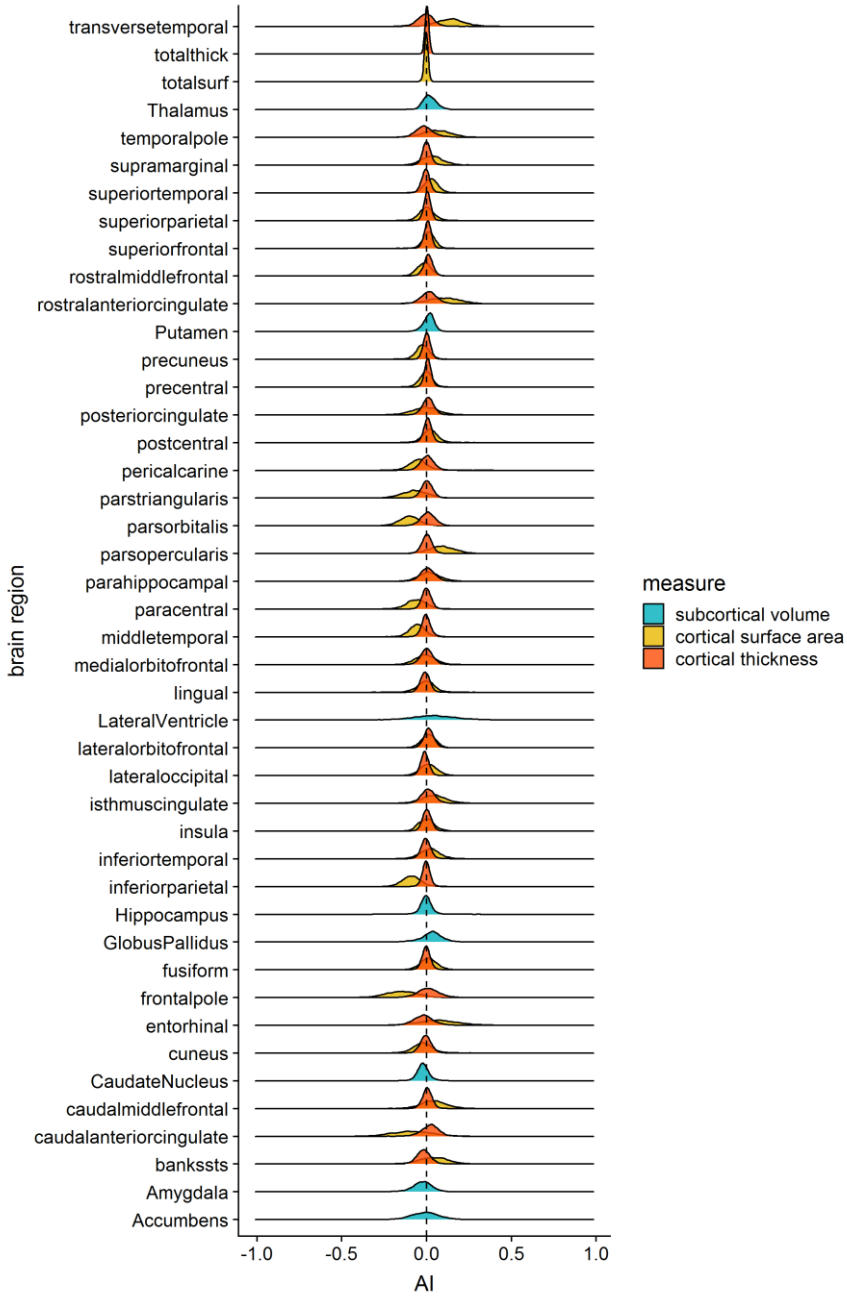


Figure S2. Joyplot of the distributions of AIs in the total sample (without winsorization). Shown for subcortical volumes (*cyan*), cortical surface areas

(*orange*), and cortical thicknesses (*red*). Lateral ventricles were categorized among subcortical measures

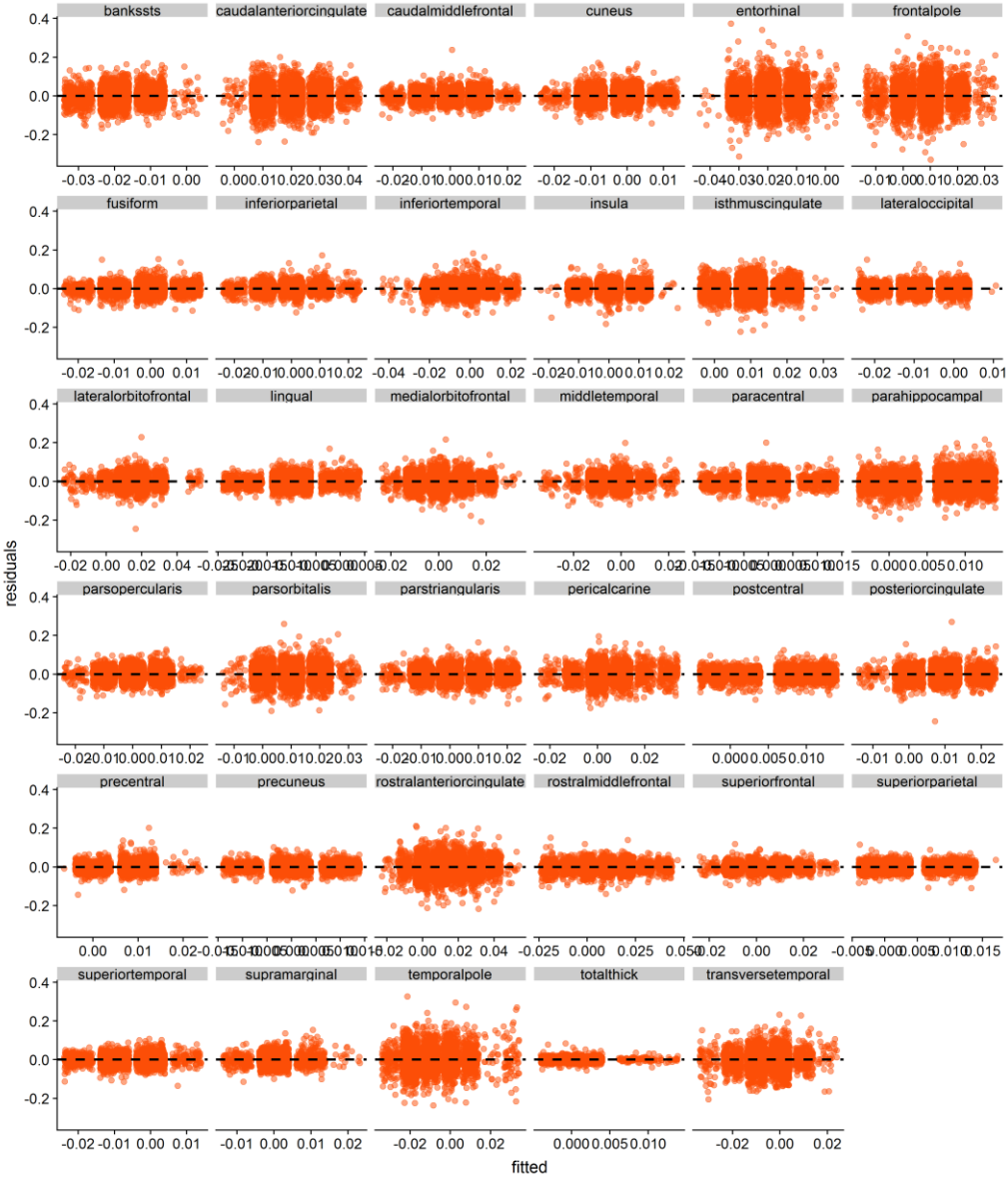


Figure S3. Residual plots of the linear mixed effects model analysis of cortical thickness AIs and the AI of the total average cortical thickness (totalthick).

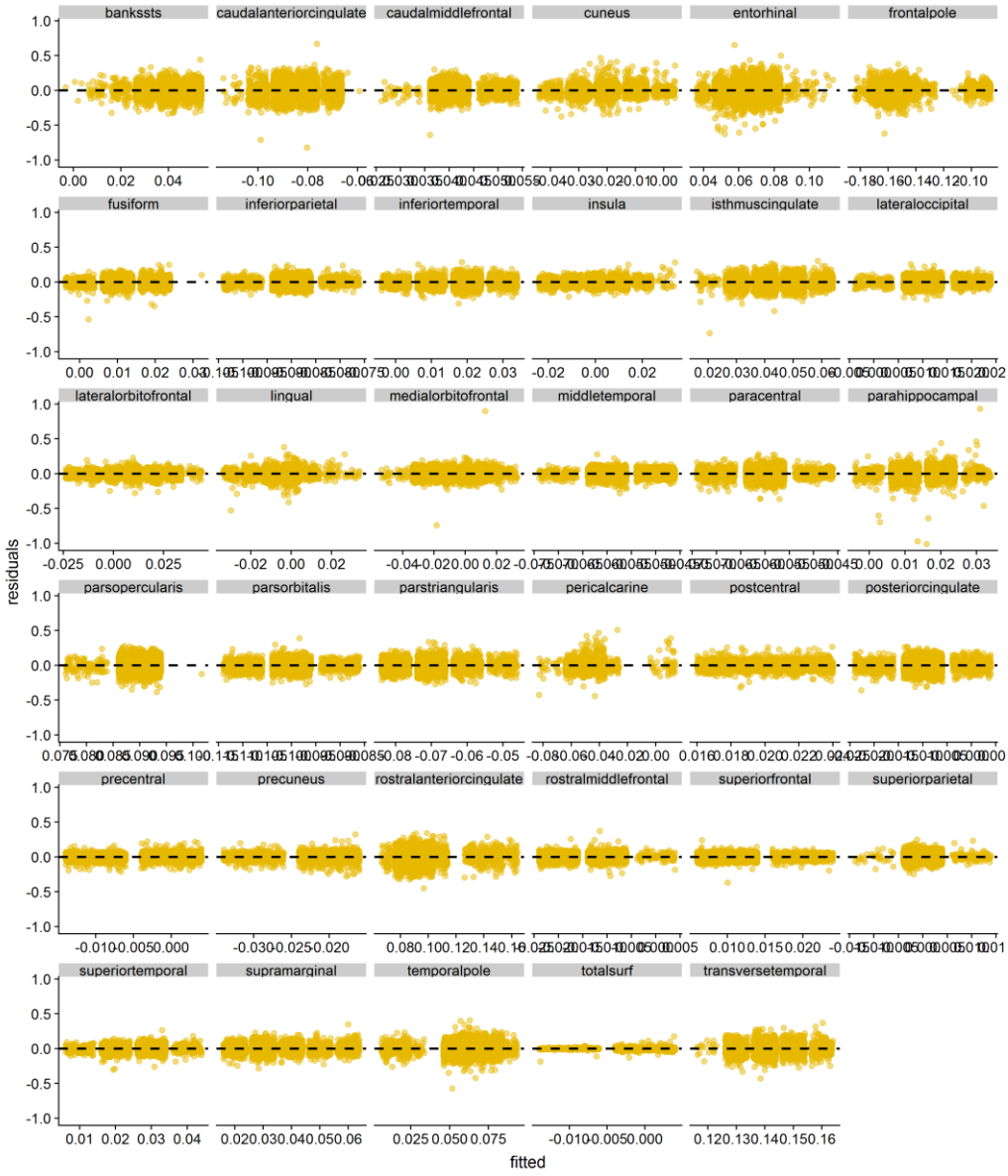


Figure S4. Residual plots of the linear mixed effects model analysis of cortical surface area AIs and the AI of the total cortical surface area (totalsurf).

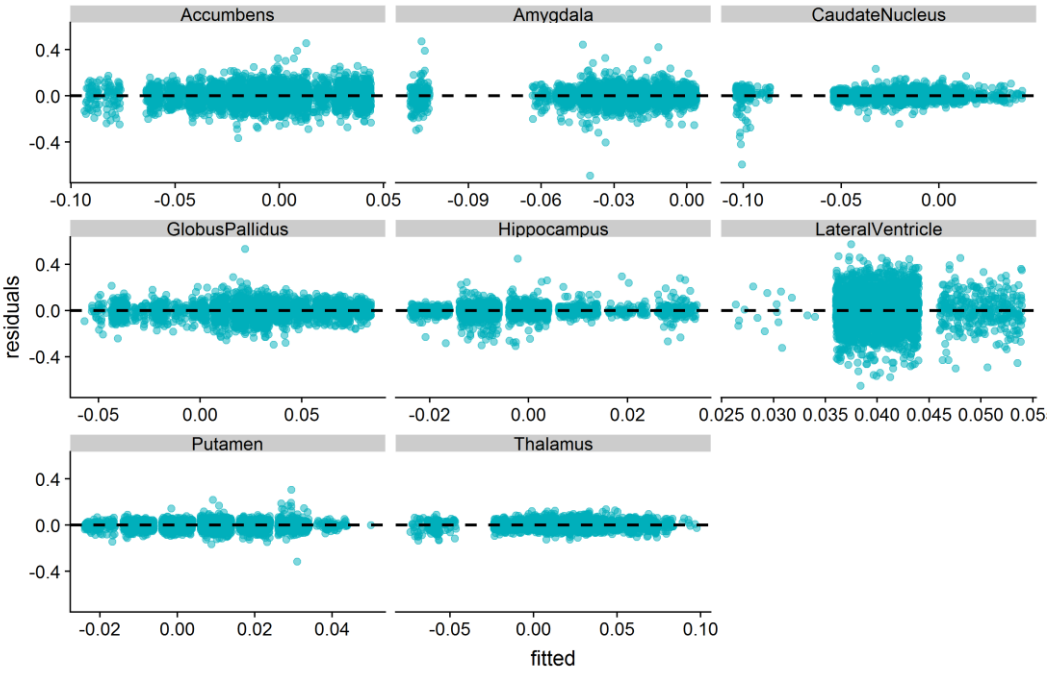


Figure S5. Residual plots of the linear mixed effects model analysis of subcortical volume AIs and the AI of the lateral ventricles.

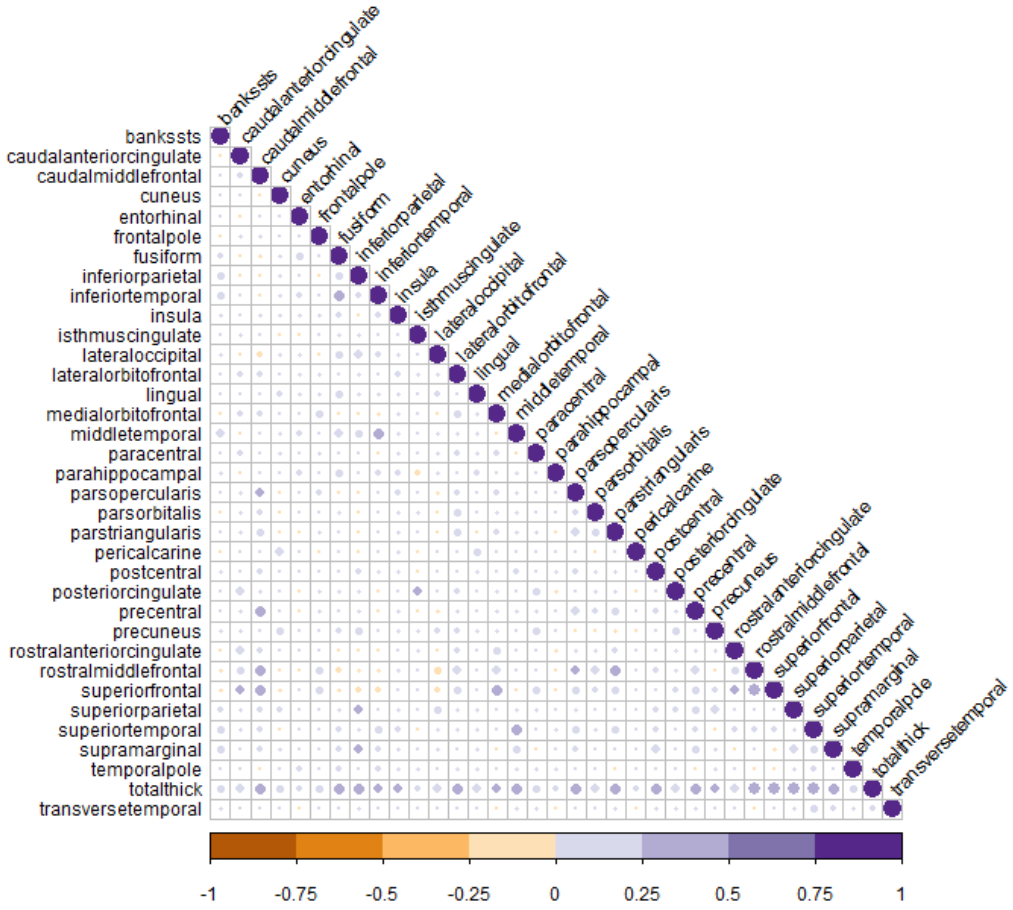


Figure S6. Correlations between AIs of cortical thickness. Correlations ranged from -015 to 0.47. Negative correlations are shown in *orange*, and positive correlations are shown in *purple*. Color intensities and circle sizes are proportional to the magnitudes of the correlation coefficients.

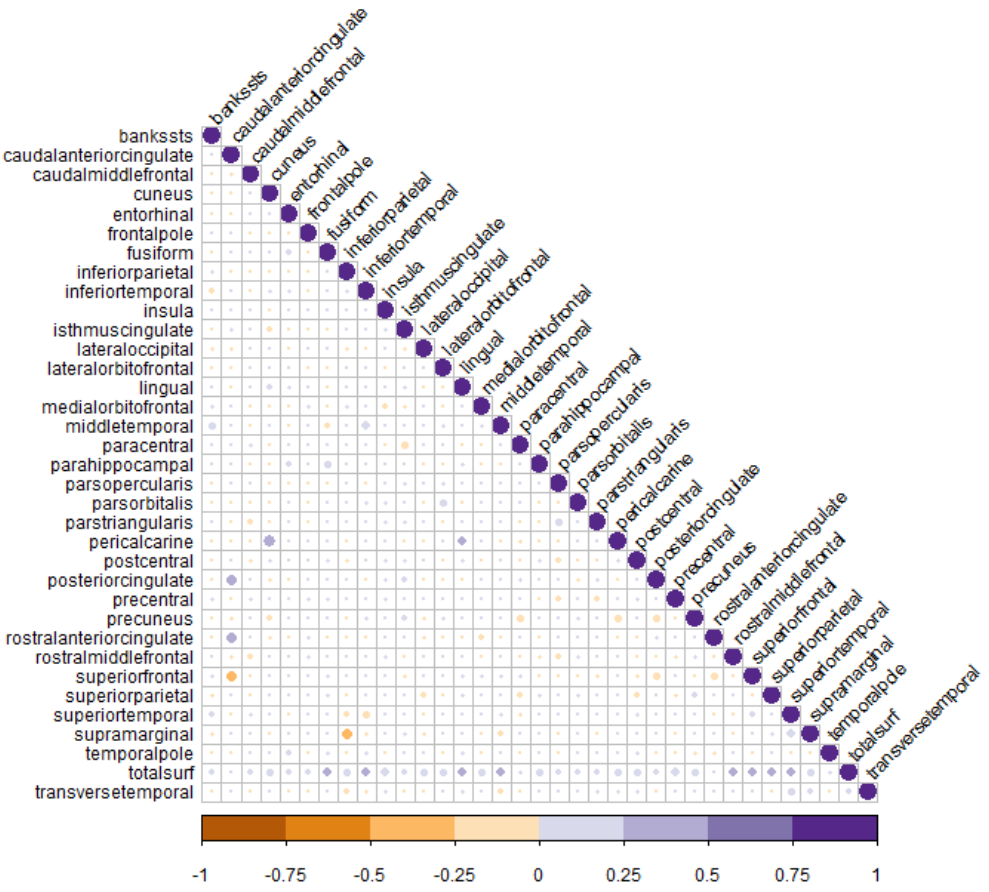


Figure S7. Correlations between AIs of cortical surface areas. Correlations ranged from -035 to 0.49. Negative correlations are shown in *orange*, and positive correlations are shown in *purple*. Color intensities and circle sizes are proportional to the magnitudes of the correlation coefficients.

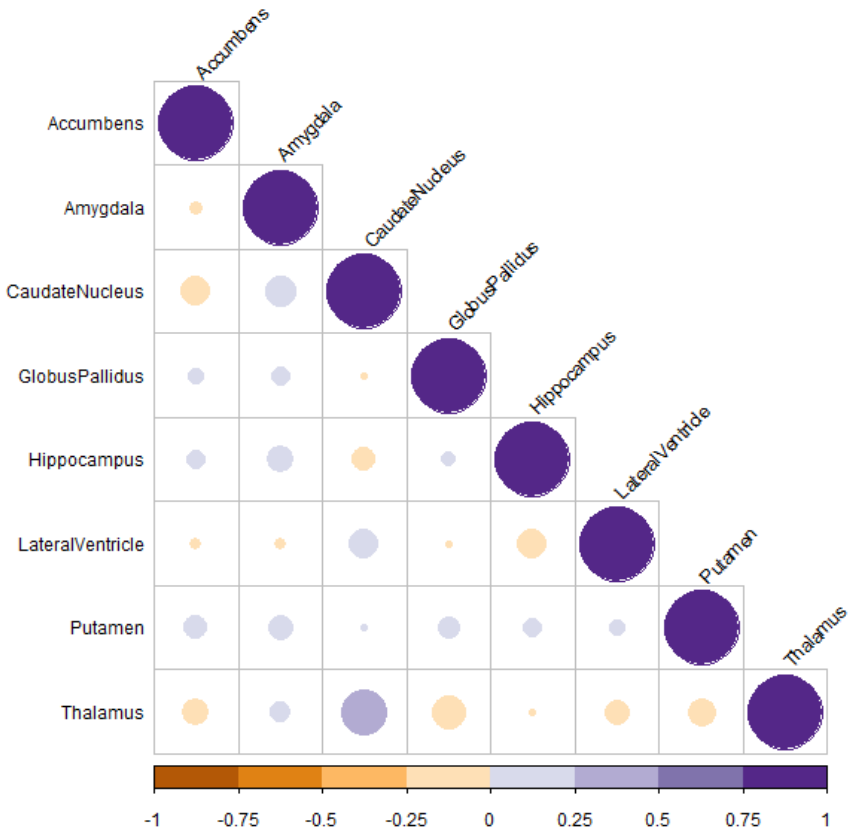


Figure S8. Correlations between AIs of subcortical volumes. Correlations ranged from -020 to 0.38. Negative correlations are shown in *orange*, and positive correlations are shown in *purple*. Color intensities and circle sizes are proportional to the magnitudes of the correlation coefficients.

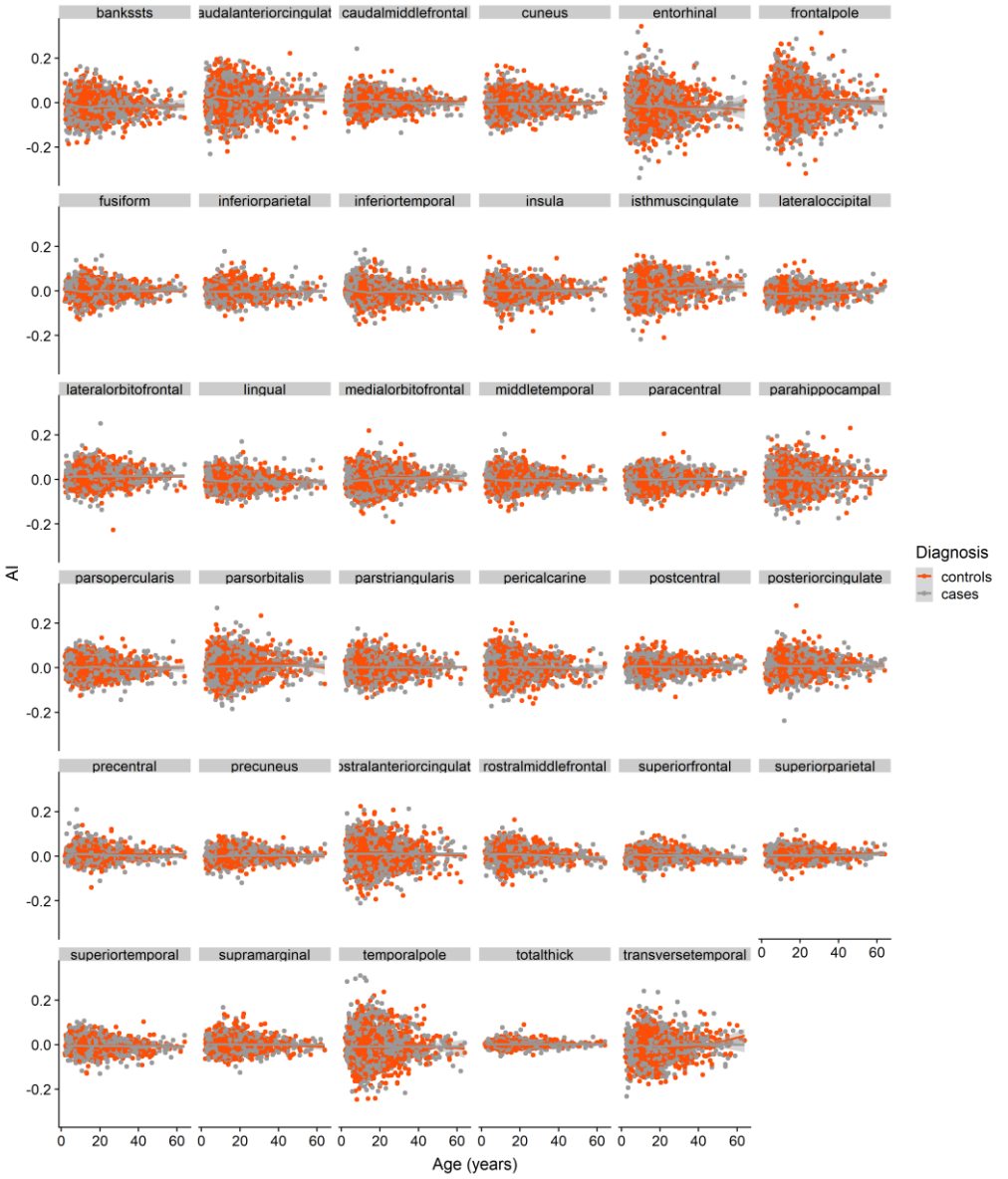


Figure S9. Scatter plots of the relationships between age and AIs of the total and regional cortical thicknesses.

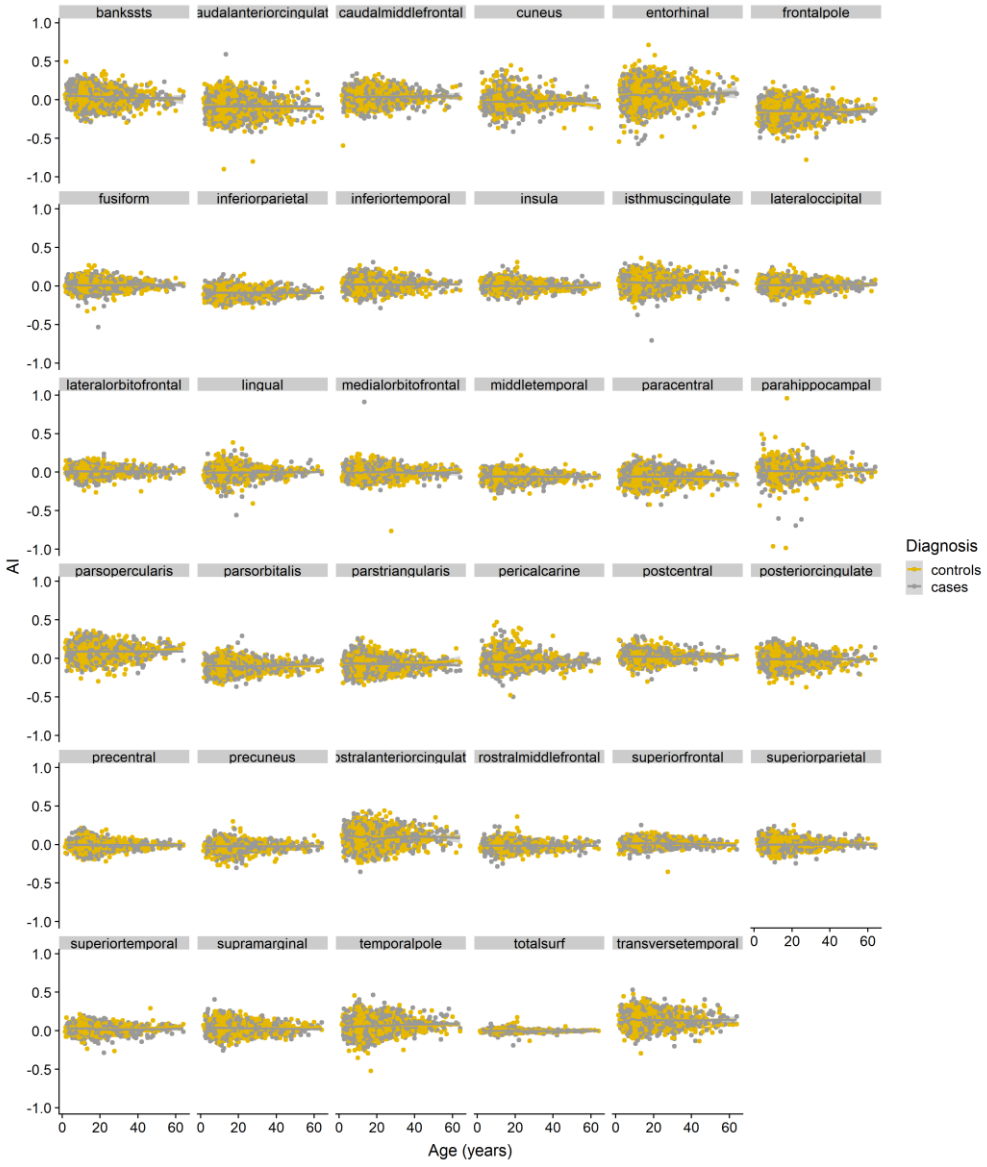


Figure S10. Scatter plots of the relationships between age and AIs of the total and regional cortical surface areas.

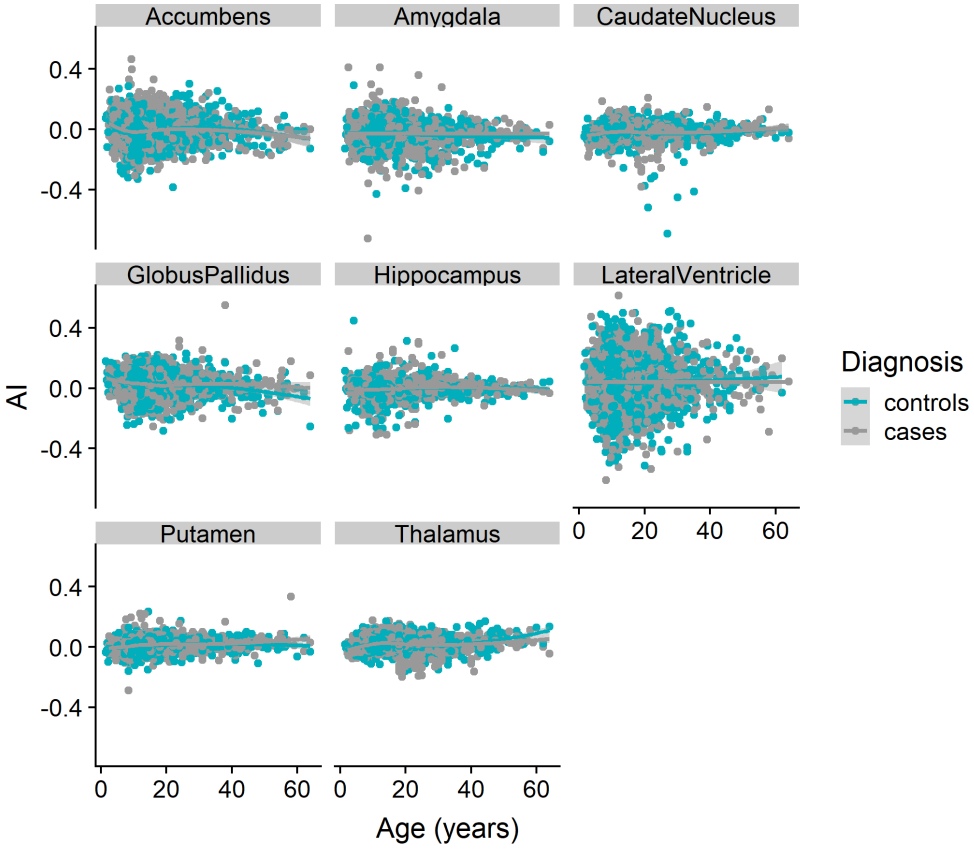


Figure S11. Scatter plots of the relationships between age and AIs of subcortical volumes and lateral ventricles.

CHAPTER

Analysis of structural brain asymmetries in Attention-Deficit/Hyperactivity Disorder in 39 datasets

3

Abstract:

Some studies have suggested alterations of structural brain asymmetry in attention-deficit/hyperactivity disorder (ADHD), but findings have been contradictory and based on small samples. Here we performed the largest-ever analysis of brain left-right asymmetry in ADHD, using 39 datasets of the ENIGMA consortium. We analyzed asymmetry of subcortical and cerebral cortical structures in up to 1,978 people with ADHD and unaffected 1,917 controls. Asymmetry Indexes (AIs) were calculated per participant for each bilaterally paired measure, and linear mixed effects modelling was applied separately in children, adolescents, adults, and the total sample, to test exhaustively for potential associations of ADHD with structural brain asymmetries. There was no evidence for altered caudate nucleus asymmetry in ADHD, in contrast to prior literature. In children, there was less rightward asymmetry of the total hemispheric surface area compared to controls (t=2.4, P=0.019). Lower rightward asymmetry of medial orbitofrontal cortex surface area in ADHD (t=2.4, P=0.007) was similar to a recent finding for autism spectrum disorder. There were also some differences in cortical thickness asymmetry across age groups. In adults with ADHD, globus pallidus asymmetry was altered compared to those without ADHD. However, all effects were small (Cohen's d from -0.18 to 0.18) and would not survive study-wide correction for multiple testing. Prior studies of altered structural brain asymmetry in ADHD were likely under-powered to detect the small effects reported here. Altered structural asymmetry is unlikely to provide a useful biomarker for ADHD, but may provide neurobiological insights into the trait.

Key words: ADHD, brain asymmetry, brain laterality, structural MRI, large-scale data

---

This chapter is published as: Postema, M. C., Hoogman, M., Ambrosino, S., Asherson, P., Banaschewski, T., Bandeira, C. E., . . . Francks, C. (2021). Analysis of structural brain asymmetries in attention-deficit/hyperactivity disorder in 39 datasets. *J Child Psychol Psychiatry*. doi:10.1111/jcpp.13396

3.1 Introduction

Attention-deficit/hyperactivity disorder (ADHD) is among the most frequently diagnosed childhood-onset mental disorders, affecting 5% of individuals worldwide (Polanczyk, de Lima, Horta, Biederman, & Rohde, 2007). ADHD is characterized by developmentally inappropriate and impairing levels of inattention and/or hyperactivity, impulsivity, and emotional dysregulation (American Psychiatric Association, 2013). At least 15% of children diagnosed with ADHD retain the diagnosis into adulthood (Faraone et al., 2015; Fayyad et al., 2017).

Left-right asymmetry (laterality) is an important feature of human brain organization (Duboc, Dufourcq, Blader, & Roussigne, 2015; Renteria, 2012; Toga & Thompson, 2003), and altered structural or functional asymmetry has been reported for a range of psychiatric conditions (Toga & Thompson, 2003). The right hemisphere is typically dominant for some aspects of attention and arousal (Heilman, Bowers, Valenstein, & Watson, 1986), and it was observed in the 1980s that people with unilateral lesions in the right hemisphere can show ADHD-like symptoms (Heilman et al., 1986). Since then, various neuropsychological and functional imaging studies have found differences between people with ADHD compared to controls (e.g. (Cortese et al., 2012)), with some pointing to a particular involvement of right hemisphere alterations (Geeraerts, Lafosse, Vaes, Vandenbussche, & Verfaillie, 2008; Hale et al., 2014; Hale et al., 2010; Langleben et al., 2001; Stefanatos & Wasserstein, 2001; Vance et al., 2007). However, not all functional data fit a primarily right-hemisphere model (Hale et al., 2009; Mohamed, Börger, Geuze, & van der Meere, 2016; Zou & Yang, 2019).

In terms of brain anatomy, several studies have reported altered asymmetry of the caudate nucleus in ADHD, although not consistently in the direction of effect (Castellanos et al., 1996; Dang et al., 2016; Filipek et al., 1997; Hynd et al., 1993; Schrimsher, Billingsley, Jackson, & Moore, 2002; Uhlikova et al., 2007). Altered asymmetry of grey matter volumes in the superior frontal and middle frontal gyri has been reported in ADHD (Cao et al., 2014), as well as decreased asymmetry of cortical convolution complexity in the prefrontal

cortex (X. Li et al., 2007). Reduced hemispheric asymmetry of white matter networks has also been reported in ADHD compared to controls (D. Li et al., 2018). Douglas *et al.* (Douglas et al., 2018) performed the largest study of brain anatomical asymmetry in ADHD to date, including 192 cases with ADHD with a history of pharmacotherapy, 149 medication-naïve cases with ADHD, and 508 typically developing controls (ages 6-21 years), from eight separate datasets. They calculated per-subject Asymmetry Indexes (AI) for various regional grey matter volumes, AI=(Left-Right)/((Left+Right)/2) (a widely used approach in studies of brain asymmetry (Kong et al., 2018; Kurth, Gaser, & Luders, 2015; Leroy et al., 2015; Postema et al., 2019)), but did not find any significant alterations of AIs in ADHD (Douglas et al., 2018). However, in a subset of their dataset (56 cases and 48 controls), Douglas *et al.* (Douglas et al., 2018) analyzed diffusion tensor imaging (DTI) data, including fractional anisotropy and mean diffusivity measures, and reported alterations in the asymmetry of six white matter tracts, again not specifically driven by alterations in the right hemisphere.

In the current study, we measured cortical regional AIs in 1,978 cases and 1,917 controls from 39 datasets, and subcortical AIs in 1,736 cases and 1,654 controls from 35 datasets, made available via the ADHD working group of the ENIGMA (Enhancing NeuroImaging Genetics through MetaAnalysis) consortium. The same datasets were recently analyzed in two other studies, by Hoogman *et al.*(Hoogman et al., 2017; Hoogman et al., 2019), that investigated bilateral changes in subcortical volumes and cortical measures, but not alterations of asymmetry. They found that ADHD was associated with lower average volumes of various subcortical structures (Hoogman et al., 2017), as well as lower total and regional cortical surface areas (including frontal, cingulate and temporal regions), and decreased cortical thickness in fusiform gyrus and temporal pole (Hoogman et al., 2019). These effects were largest in children, and even child-specific for the cortical findings, so that for the present study of asymmetries, we followed the age-group division of Hoogman *et al.* (Hoogman et al., 2019) into children (<15 years), adolescents (15-21 years) and adults (>21 years), as well as performing analysis in the total combined sample to explore age-general effects. Bilateral effect sizes reported by Hoogman *et al.* (Hoogman et al., 2017; Hoogman et al., 2019) were small, i.e. case-control Cohen's *d* values between -0.21 and 0.06. This

suggests that, if associations between ADHD diagnosis and regional brain asymmetries are similarly subtle, many previous studies of anatomical asymmetries in this disorder were underpowered, and the described effects may have been unreliable. Low statistical power in a study not only reduces the chance of detecting true effects, but also the likelihood that significant results reflect true effects (Munafo & Flint, 2010). It is important for the field of neuroimaging to mature around more highly powered analyses in relation to subtle effects. The current study aimed to provide detailed information on the extent to which laterality is affected in ADHD, based on the largest ever sample size for this question, comprised of multiple independent cohorts from around the world.

3.2 Methods

3.2.1 Ethical Considerations

This study made use of 39 pre-existing datasets from around the world. For all datasets, the participating sites had obtained ethical approval from local institutional review boards, as well as informed consent to participate.

3.2.2 Datasets

Bilateral brain measures derived from structural MRI were available from 39 different datasets via the ENIGMA-ADHD Working Group (**Table S1**). The 39 datasets comprised cortical data from a total of 1,933 participants with ADHD (1,392 males; median age = 15 y; range = 4 y to 62 y) and 1,829 healthy individuals (1,116 males; median age = 14 y; range = 4 y to 63 y). Subcortical data were available from 35 of the 39 datasets and comprised 1,691 cases (1,212 males, median age = 15 y; range = 5 y to 62 y) and 1,566 controls (953 males, median age = 14 y; range= 4 y to 63 y). A previous study by Douglas *et al.* (Douglas et al., 2018) (see introduction) included five datasets that were also analyzed in the present study (**Table S1**).

For all but 4 of the 39 datasets, ADHD diagnosis was based on the Diagnostic and Statistical Manual of Mental Disorders 4th Edition (DSM-IV) (American Psychiatric Association, 2000). Other instruments used were DSM5th Edition (DSM-5), or the International Classification of Diseases (ICD)10th Edition) (World Health Organization, 1992). For information per dataset see **Table S1**.

In terms of age groups, for children (<15 y) there were subcortical data from 802 cases and 842 controls, and cortical data from 912 cases and 950 controls; for adolescents (15 y – 21 y) there were subcortical data from 326 cases and 232 controls, and cortical data from 408 cases and 340 controls; for adults (> 21 y) there were subcortical data from 563 cases and 492 controls, and cortical data from 613 cases and 539 controls.

Eleven additional datasets, comprising cases-only or controls-only, were excluded for the purpose of the present study (these are not listed in **Table S1**). This was because our analysis models included random intercepts for 'dataset' (below), and diagnosis would be fully confounded with 'dataset' for case-only or control-only datasets.

3.2.3 MRI-based measures

Structural T1-weighted brain MRI scans had been acquired at each study site for each of the 39 pre-existing MRI datasets. MRI data within the ENIGMA consortium are typically processed separately at each participating site, due to varying restrictions on data sharing that apply to the many legacy datasets from different countries around the world. Images were obtained at different field strengths (1.5 T or 3 T: see **Table S1)**. Scanners and scanning sequences, recruitment criteria, and demographics differed between datasets, but all sites separately applied a single image processing and quality-control protocol from the ENIGMA consortium (http://enigma.ini.usc.edu/protocols/imaging-protocols), starting from their T1 image data. The harmonized processing was based on the freely available and validated software FreeSurfer (versions 5.1 or 5.3) (Fischl, 2012), with the default 'recon-all' pipeline (https://surfer.nmr.mgh.harvard.edu/fswiki/recon-all), which is a 29-step procedure that includes skull stripping, registration, subcortical segmentation,

normalization, white matter and pial surface creation, cortical parcellation according to the Desikan-Killiany atlas, and the output of region-specific measures of volume, average thickness and surface area. This was followed by visual inspection of both internal and external segmentations (**Supplementary Methods**). Exclusions on the basis of these quality control steps resulted in the sample sizes given above. The present study took as its starting point the FreeSurfer-derived measures of left and right volumes of seven bilaterally paired subcortical structures, and thickness and surface area measures for each of 34 bilaterally paired cortical regions, that were generated previously by each site. The cortical regions were defined by the Desikan-Killiany atlas (Desikan et al., 2006). In addition, the average cortical thickness and total surface area per hemisphere were analyzed. Freesurfer's measure of intracranial volume (ICV) was also considered as a covariate in sensitivity analyses (below).

The Desikan-Killiany atlas (Desikan et al., 2006) was derived from manual segmentations of reference brain images. The labeling system incorporates hemisphere-specific information on sulcal and gyral geometry with spatial information regarding the locations of brain structures (Desikan et al., 2006). Accordingly, the mean regional asymmetries in our data might be influenced by left-right differences present in the reference dataset used for constructing the atlas. Nonetheless, this approach was appropriate for our study focused on comparing relative asymmetry between groups. The use of an asymmetrical atlas has the advantage that regional identification is likely to be accurate for structures that are asymmetrical both in the atlas and, on average, in the study population.

3.2.4 Asymmetry indexes

Left and right data per brain region and individual participant were loaded into R (version 3.5.3), and null values were removed. An asymmetry index (AI) was calculated for each subject and each paired left-right measure using the following formula: (Left-Right)/(Left+Right). Negative AIs therefore indicate a right>left asymmetry, while positive AIs indicate a left>right asymmetry. In the AI formula, the L-R difference (numerator) is adjusted by the bilateral measure L+R (denominator), such that the AI does not scale with

the bilateral measure. We did not divide the denominator by 2, in contrast to some previous formulations of AIs (see Introduction), but this makes no difference in terms of deriving Cohen's *d* effect sizes and *P*-values for group comparisons. Distributions of each of the AIs in the total study sample are plotted in **Figure S1**.

Correlations between AI measures in the total study sample were calculated using Pearson's R and visualized using the *corrplot* package in R (**Figures S2-S4**). Most pairwise correlations between AIs were of low magnitude (median magnitude r = 0.024 for surface area AIs, 0.040 for thickness AIs, 0.091 for subcortical volume AIs), with a minimum r = -0.42 between caudal anterior cingulate surface area and superior frontal surface area, and maximum r = 0.49 between rostral middle frontal thickness and total average thickness.

3.2.5 Linear mixed effects random-intercept models

3.2.5.1 Main analysis

Linear mixed effects analyses were performed separately for each subcortical volume AI, cortical regional surface and thickness AI, and the total hemispheric surface area and average thickness AI, using the *nlme* package in R (version 3.5.3). Analyses were conducted separately within children, adolescents, and adults, as well as on the total study sample. All models included *diagnosis* (a binary variable; 0=control, 1=case), *sex* (binary; 0=female, 1=male) and *age* (numeric) as fixed factors, and *dataset* as a random factor (39 categories for cortical data, 35 categories for subcortical data):

$$AI \sim diagnosis + sex + age + \text{random } (\sim 1 \mid dataset) \quad (1)$$

The Maximum Likelihood (ML) method was used to fit the models. Whenever any of the predictor variables was missing in a given subject, the subject was omitted from the analysis (method = 'na.omit'). The 'optim' optimizer (lmeControl(opt='optim') was used for all models. Residual plots are in **Figures S5-S7**.

The t-statistic for the factor 'diagnosis' in each linear mixed effects model was derived and used to calculate Cohen's *d* (**Supplementary Methods**). For visualization of cerebral cortical results, Cohen's *d* values were loaded into Matlab (v. R2020a) and 3D images of left hemisphere inflated cortical and subcortical structures were obtained using Freesurfer-derived ply files.

Field strength was not included as a covariate because each dataset was scanned entirely at either 1.5 T or 3T (**Table S1**), and the models included 'dataset' as a random intercept effect, which adjusted for differences that applied to entire datasets.

3.2.5.2 Significance and detectable effect sizes

Significance was assessed based on the *P*-value for the diagnosis term within each model. Separately within each age group, and again within all age groups combined, we applied False Discovery Rate (FDR) correction (Benjamini & Hochberg, 1995) for multiple testing, separately across the seven subcortical structures, the 35 cortical surface area AIs (i.e. 34 regional AIs and one hemispheric total AI), and again for the 35 cortical thickness AIs, each time with an FDR threshold of 0.05. Therefore twelve separate FDR corrections were done. We also applied an additional FDR correction for the total combined analysis across all age groups and AIs of different types.

As each linear model included multiple predictor variables, the power to detect an effect of diagnosis on AI could not be computed exactly, but we obtained an indication of the effect size that would be needed to provide 80% power had we been using simple t-tests and Bonferroni correction for multiple testing, using the *pwr* command in R (**Supplementary Methods**). For this purpose, a significance level of 0.0071 (i.e., 0.05/7) or 0.0014 (i.e. 0.05/35) was set in the context of multiple testing over the seven subcortical volumes, or the regional and total cortical surface areas (N = 35) or thicknesses (N = 35). This showed that, in the total study sample, a case-control effect size of roughly Cohen's d=0.12 (subcortical), or d=0.13 (cortical), would be detectable with 80% power. For the analyses in the different age groups, this was, respectively, d=0.16 and d=0.19 in children, d=0.26 and d=0.30 in adolescents, and d=0.21 and d=0.24 in adults.

3.2.5.3 Directions of asymmetry changes

For any AIs showing nominally significant effects (i.e., unadjusted P<0.05) of diagnosis in any of the primary analyses, *post hoc* linear mixed effects modelling was also performed on the corresponding L and R measures separately, to understand the unilateral changes involved. The models included the same terms as were used in the main analysis of AIs (i.e., diagnosis, age and sex as fixed factors, and dataset as random factor). Again, the Cohen's *d* effect sizes for diagnosis were calculated based on the *t*-statistics. The raw mean AI values were calculated separately in controls and cases, to describe the reference direction of healthy asymmetry in controls, and whether cases showed lower, higher, or reversed asymmetry relative to controls.

3.2.5.4 Sensitivity analyses

The relationships between AIs and age appeared roughly linear across all age groups combined (**Figures S8-S10**). Therefore, no polynomials for age were incorporated in the main model (**Supplementary Methods**). However, analyses were repeated (only for all age groups combined) using an additional non-linear term for age, to check whether this choice had affected the results. The variables age and age^2 are inevitably highly correlated. To include linear and non-linear effects of age in the same model, we made use of the poly()-function in R for these two predictors, which created a pair of uncorrelated variables related to age (so-called orthogonal polynomials) (Chambers & Hastie, 1992), where one variable was linear and one non-linear:

$$AI \sim diagnosis + \text{poly}(age, 2) + sex + \text{random}\ (\sim 1 \mid dataset) \qquad (2)$$

Note that we were not interested to measure the effects of age or age-squared, but simply to correct for linear and non-linear effects related to age, as we measured the effects of diagnosis on brain asymmetry.

No AI outliers were removed for the main analysis, but to confirm that results were not dependent on outliers, the main analysis was also repeated (for all age groups combined) after having winsorized using a threshold of k= 3, for each AI measure separately in the total combined dataset.

3.2.5.5 Associations between brain asymmetries and IQ, comorbidity, ADHD severity and psychostimulant medication

Within the ADHD participants only (all age groups combined), brain asymmetries were tested in relation to several potentially associated variables (IQ, comorbidity, severity, medication use; see **Figures S11, S12**), using separate models in which each variable was considered as a fixed effect:

AI ~ variable + age + sex + random (~1 | *dataset*) (3)

See **Supplementary Methods** for the derivation of these variables. For binary variables, datasets were removed if they had < 1 subject per category, to avoid the random variable 'dataset' being fully confounded with the binary variable for any datasets. Depending on the availability of each specific AI, data for testing association with IQ were available for up to 1,719 ADHD individuals (exact numbers per AI depended on image quality control for that region and can be found in the relevant results tables, see below). For the presence/absence of comorbidities, four different binary variables were constructed: mood disorder (up to 179 yes, 384 no), anxiety disorder (up to 82 yes, 503 no), oppositional defiant disorder (ODD; up to 80 yes, 151 no), and substance use disorder (SUD; up to 77 yes, 335 no). For ADHD symptom severity, two continuous variables were used: hyperactivity/impulsivity (up to 1,009 ADHD participants) and inattention (1,006 ADHD participants). For psychostimulant medication use, two binary variables were constructed: lifetime use (up to 337 yes, 188 no), and current use (i.e., at the time of scanning, up to 361 yes, 377 no) (see **Figure S12** for the distributions, and **Supplementary Methods** for more explanation).

IQ was also examined in controls only (all age groups combined) to explore the relationships between IQ and brain asymmetries in typically developing individuals. IQ and AI data were available for up to 1,663 controls. The model for each AI was:

AI ~ IQ + age + sex + random (~1 | dataset)

3.2.5.6 IQ, handedness and intracranial volume as covariates in disorder case-control analysis

See the supplementary information for the derivation of IQ and handedness measures, and above for ICV. Distributions are in **Figure S11**. We did not adjust for IQ, handedness or ICV as covariate effects in our main, case-control analysis (above). This was because, *a priori*, there are various possible causal relations linking these traits to ADHD and brain asymmetry and other, possibly underlying factors shared between some or all of them. In this context, it is important not to bias associations between ADHD and brain asymmetry through correcting for these factors as covariates in primary analysis, as they may be colliders (Cole et al., 2010) (see the Discussion for more on this issue). However, we included a set of additional, secondary models to test for case-control effects in the presence of these variables as covariates:

AI ~ diagnosis + age + sex + + handedness + random (~1 | dataset)

*AI ~ diagnosis + age + sex + handedness + handedness*diagnosis + random (~1 | dataset)*

AI ~ diagnosis + age + sex + IQ + random (~1 | dataset)

*AI~ diagnosis + age + sex + IQ + IQ*diagnosis + random (~1 | dataset)*

AI ~ diagnosis + age + sex + + ICV + random (~1 | dataset)

*AI ~ diagnosis + age + sex + ICV + ICV*diagnosis + random (~1 | dataset)*

The analyses were also repeated after winsorization of outliers, as above.

3.3 Results

3.3.1 Associations of brain asymmetry with ADHD

Results for all AIs across the different age groups, and for all age groups combined, are listed in the supplement (**Tables S2-S13**), and are also available as supplementary comma-delimited text files.

3.3.1.1 Children

There were no associations of diagnosis with AIs that had FDR <0.05 in children (**Tables 1-3, Tables S2-S4**). The children showed nominally significant associations (unadjusted P<0.05) of diagnosis with the AIs of total hemispheric surface area (*t*=2.10, *P*=0.036), medial orbitofrontal cortex surface area (*t*=2.7, *P*=0.006), and paracentral lobule surface area (*t*=-2.16, p=0.031) (**Table 2, Table S3**). The Cohen's *d* for these effects were 0.11, 0.13 and -0.10 respectively (**Figure 1, Figure S13, Table S3**). *Post hoc* analysis showed that the effects on total hemispheric and medial orbitofrontal surface area asymmetries both involved relatively greater reductions on the right-side than left-side in ADHD compared to controls (**Table S14**). The effect on paracentral lobule surface area asymmetry was driven by a larger decrease of left compared to right-hemispheric surface area in this region (**Table S14**).

The children also showed nominally significant associations of diagnosis with four regional cortical thickness AIs, which were the banks of the superior temporal sulcus (*t*=-2.0, *P*=0.047; increased rightward asymmetry in ADHD), caudal middle frontal cortex (*t*=2.1, *P*=0.037; increased leftward asymmetry), precentral gyrus (*t*=2.4, *P*=0.019; increased leftward asymmetry) and insula (*t*=-2.0, *P*=0.047, decreased leftward asymmetry) (**Table 2, Table S14**).

3.3.1.2 Adolescents

There were two nominally significant associations between diagnosis and AIs in adolescents, but none with FDR <0.05 (**Tables 1-3, Tables S5-S7**). These involved the *pars orbitalis* of inferior frontal gyrus surface area (*t*=2.4, *P*=0.017), which showed lower rightward asymmetry in ADHD compared to controls, due to a smaller left than right sided decrease (**Table S14**), and cuneus thickness (*t*=-2.0, *P*=0.043), which showed greater rightward

Table 1. Linear mixed model results for subcortical volume AIs.

Subcortical volume AI	Children only		Adolescents only		Adults only		Total study sample	
	P[1]	*d*[2]	P[1]	*d*[2]	P[1]	*d*[2]	P[1]	*d*[2]
Accumbens	0.22	-0.06	0.40	-0.07	0.94	0.00	0.29	-0.04
Amygdala	0.96	-0.002	0.64	0.04	0.74	-0.02	0.79	-0.01
Caudate Nucleus	0.66	0.02	0.93	0.01	0.40	0.05	0.44	0.03
Globus Pallidus	0.91	0.01	0.44	-0.07	***0.003***	-0.18	**0.03**	-0.07
Hippocampus	0.71	-0.02	0.23	0.11	0.46	0.05	0.71	0.01
Putamen	0.37	-0.04	0.75	-0.03	0.53	-0.04	0.16	-0.05
Thalamus*	0.47	0.04	0.28	0.10	0.49	0.04	0.20	0.05

[1] Uncorrected P-values for diagnosis are indicated, with in **bold** those that are significant at the uncorrected level ($P < 0.05$), and in ***bold-italic*** those that survive multiple testing correction. [2] Cohen's *d* value for the effect of diagnosis. *****Thalamus volume was not available from the NIH sample.

Table 2. Linear mixed model results for the cortical surface area AIs.

Cortical surface area AI	Children only		Adolescents only		Adults only		Total study sample	
	P[1]	*d*[2]	P[1]	*d*[2]	P[1]	*d*[2]	P[1]	*d*[2]
Banks of superior temporal sulcus	0.68	-0.02	0.26	-0.09	0.82	0.01	0.38	-0.03
Caudal anterior cingulate cortex	0.65	-0.02	0.27	-0.08	0.71	0.02	0.58	-0.02
Caudal middle frontal cortex	0.37	0.04	0.82	-0.02	0.22	0.07	0.17	0.04
Cuneus	0.10	0.07	0.89	0.01	0.07	-0.11	0.87	-0.01
Entorhinal cortex	0.84	-0.01	0.66	-0.03	0.10	-0.10	0.27	-0.04
Frontal pole	0.07	-0.08	0.42	0.06	0.25	-0.07	0.12	-0.05
Fusiform gyrus	0.14	-0.07	0.32	0.07	0.11	-0.10	0.12	-0.05
Inferior parietal cortex	0.16	0.06	0.94	-0.01	0.89	-0.01	0.31	0.03
Inferior temporal gyrus	0.40	0.04	0.95	0.00	0.26	0.07	0.19	0.04
Insula	0.07	0.08	0.73	0.03	0.62	-0.03	0.24	0.04
Isthmus cingulate cortex	1.00	0.00	0.08	-0.13	0.48	0.04	0.72	-0.01
Lateral occipital cortex	0.38	-0.04	0.82	0.02	0.05	0.12	0.66	0.01
Lateral orbitofrontal cortex	0.27	-0.05	0.46	-0.06	0.40	-0.05	0.09	-0.06
Lingual gyrus	0.85	0.01	0.17	-0.10	0.49	0.04	0.89	0.00
Medial orbitofrontal cortex	**0.01**	0.12	0.35	0.07	0.77	-0.02	**0.03**	0.07

Table 2. Linear mixed model results for the cortical surface area AIs []

Cortical surface area AI	Children only		Adolescents only		Adults only		Total study sample	
	P[1]	*d*[2]	P[1]	*d*[2]	P[1]	*d*[2]	P[1]	*d*[2]
Middle temporal gyrus	0.16	0.07	0.63	-0.04	0.93	-0.01	0.44	0.03
Paracentral lobule	0.07	-0.08	0.98	0.00	0.29	-0.06	0.06	-0.06
Parahippocampal gyrus	0.29	0.05	0.43	-0.06	0.13	-0.09	0.85	-0.01
Pars opercularis of inferior frontal gyrus	0.88	-0.01	0.16	0.11	0.57	0.03	0.46	0.02
Pars orbitalis of inferior frontal gyrus	0.44	0.04	**0.02**	0.18	0.54	0.04	0.05	0.06
Pars triangularis of inferior frontal gyrus	0.30	0.05	0.13	0.11	0.57	-0.03	0.22	0.04
Pericalcarine cortex	0.22	0.06	0.11	-0.12	1.00	0.00	0.84	0.01
Postcentral gyrus	0.45	0.03	0.42	0.06	0.98	0.00	0.37	0.03
Posterior cingulate cortex	0.98	0.00	0.48	-0.05	0.82	0.01	0.88	-0.01
Precentral gyrus	0.80	0.01	0.07	-0.14	0.06	-0.11	0.09	-0.05
Precuneus	0.37	0.04	0.29	-0.08	0.65	0.03	0.57	0.02
Rostral anterior cingulate cortex	0.89	-0.01	0.93	-0.01	0.36	-0.05	0.46	-0.02
Rostral middle frontal gyrus	0.20	-0.06	0.89	-0.01	0.60	-0.03	0.21	-0.04
Superior frontal gyrus	0.29	0.05	0.07	0.13	0.12	-0.09	0.54	0.02
Superior parietal cortex	0.15	0.07	0.47	0.05	0.68	-0.02	0.34	0.03

Superior temporal gyrus	0.36	0.04	0.91	0.01	0.20	-0.08	0.99	0.00
Supramarginal gyrus	0.79	0.01	0.18	-0.10	0.22	-0.07	0.28	-0.04
Temporal pole	0.40	0.04	0.46	0.06	0.36	-0.06	0.73	0.01
Transverse temporal gyrus	0.56	-0.03	0.39	0.07	0.94	0.00	0.98	-0.0008
Total average surface area	**0.02**	0.11	0.90	0.01	0.24	-0.07	0.43	0.03

[1] Uncorrected P-values for diagnosis are indicated, with in **bold** those that are significant at the uncorrected level ($P < 0.05$). None survived multiple testing correction. [2] Cohen's *d* value for the effect of diagnosis.

Table 3. Linear mixed model results for the cortical thickness AIs.

Cortical surface area AI	Children only		Adolescents only		Adults only		Total study sample	
	P[1]	*d*[2]	P[1]	*d*[2]	P[1]	*d*[2]	P[1]	*d*[2]
Banks of superior temporal sulcus	0.04	-0.10	0.48	-0.06	0.64	-0.03	**0.05**	-0.07
Caudal anterior cingulate cortex	0.32	0.05	0.65	-0.03	0.06	0.11	0.15	0.05
Caudal middle frontal cortex	0.04	0.09	0.06	0.14	0.77	0.02	**0.02**	0.07
Cuneus	0.63	0.02	0.05	-0.15	0.07	0.11	0.59	0.02
Entorhinal cortex	0.14	-0.07	0.86	0.01	0.63	-0.03	0.27	-0.04
Frontal pole	0.22	0.05	0.19	-0.10	0.19	0.08	0.31	0.03
Fusiform gyrus	0.44	-0.04	0.91	0.01	0.75	0.02	0.85	-0.01
Inferior parietal cortex	0.81	-0.01	0.52	-0.05	0.54	0.04	0.92	0.00
Inferior temporal gyrus	0.30	-0.05	0.58	0.04	0.87	-0.01	0.77	-0.01
Insula	**0.04**	-0.09	0.19	-0.10	0.93	-0.01	**0.04**	-0.07
Isthmus cingulate cortex	0.53	-0.03	0.18	0.10	0.34	-0.06	0.73	-0.01
Lateral occipital cortex	0.55	0.03	0.22	-0.09	**0.03**	0.13	0.33	0.03
Lateral orbitofrontal cortex	0.73	-0.02	0.50	0.05	0.14	0.09	0.42	0.03
Lingual gyrus	0.41	-0.04	0.97	0.003	0.60	-0.03	0.34	-0.03
Medial orbitofrontal cortex	0.08	-0.08	0.23	0.09	**0.04**	0.12	0.82	0.01

Middle temporal gyrus	0.83	-0.01	0.74	-0.03	0.01	-0.16	0.15	-0.05
Paracentral lobule	0.11	-0.07	0.18	0.10	0.77	-0.02	0.39	-0.03
Parahippocampal gyrus	0.14	-0.07	0.06	-0.14	0.40	0.05	0.18	-0.04
Pars opercularis of inferior frontal gyrus	0.86	0.01	0.26	0.08	0.86	-0.01	0.44	0.03
Pars orbitalis of inferior frontal gyrus	0.27	0.05	0.93	-0.01	0.39	0.05	0.25	0.04
Pars triangularis of inferior frontal gyrus	0.84	-0.01	0.40	0.06	0.90	-0.01	0.72	0.01
Pericalcarine cortex	0.89	0.01	0.95	0.00	**0.004**	0.17	0.11	0.05
Postcentral gyrus	0.79	-0.01	0.79	-0.02	0.01	-0.15	0.08	-0.06
Posterior cingulate cortex	0.64	-0.02	0.54	-0.05	0.84	-0.01	0.34	-0.03
Precentral gyrus	**0.01**	0.12	0.39	-0.06	0.16	0.08	**0.03**	0.07
Precuneus	0.80	0.01	0.17	0.10	0.73	0.02	0.45	0.02
Rostral anterior cingulate cortex	0.82	-0.01	0.09	0.13	0.35	0.06	0.25	0.04
Rostral middle frontal gyrus	0.77	0.01	0.78	-0.02	0.39	-0.05	0.82	-0.01
Superior frontal gyrus	0.64	0.02	0.08	0.13	0.56	0.04	0.20	0.04
Superior parietal cortex	0.85	0.01	0.48	-0.05	0.88	0.01	0.82	-0.01

Table 3. Linear mixed model results for the cortical thickness AIs [continued]

Cortical surface area AI	Children only		Adolescents only		Adults only		Total study sample	
	P[1]	*d*[2]	P[1]	*d*[2]	P[1]	*d*[2]	P[1]	*d*[2]
Superior temporal gyrus	0.08	0.08	0.40	0.07	0.37	-0.06	0.24	0.04
Supramarginal gyrus	0.16	-0.06	0.41	-0.06	0.93	-0.01	0.15	-0.05
Temporal pole	0.53	0.03	0.66	0.03	0.63	-0.03	0.71	0.01
Transverse temporal gyrus	0.55	0.03	0.72	0.03	0.34	-0.06	0.90	0.00
Total average thickness	0.92	-0.004	0.75	0.02	0.78	0.02	0.79	0.01

[1] Uncorrected P-values for diagnosis are indicated, with in **bold** those that are significant at the uncorrected level ($P < 0.05$). None of the associations with diagnosis survived multiple testing correction. [2] Cohen's *d* value for the effect of diagnosis.

asymmetry in ADHD compared to controls, due to an increase in right- and a decrease in left hemispheric thickness (**Table S14**).

3.3.1.3 Adults

In adults, the globus pallidus AI was significantly associated with ADHD diagnosis with FDR <0.05 (*t*=-2.9, *P*=0.004, uncorrected) (**Table 1, Table S8**). The Cohen's *d* effect size for this association was -0.18 (**Table 1, Figure 1, Figure S13**). This effect involved a decrease in leftward asymmetry in ADHD compared to controls, driven by a larger reduction of left-side volume than right-side volume in ADHD compared to controls **(Table S14)**. Note this association was only significant in the context of FDR correction for 7 subcortical AIs within adults specifically. (No effects were significant at FDR-corrected *P*<0.05 when the correction was done across all age groups and AIs of different types, see below).

There were other nominally significant associations of AIs with diagnosis in adults: lateral occipital cortex surface area (*t*=2.0, *P*=0.049; increased leftward) (**Table 2, Table S9, Table S14**) and thickness (*t*=2.2, *P*=0.026; decreased rightward) (**Table 3, Table S10, Table S14**), medial orbitofrontal cortex thickness (*t*=2.0, *P*=0.045; increased leftward), middle temporal gyrus thickness (*t*=-2.6, *P*=0.009; increased rightward), pericalcarine cortex thickness (*t*=2.9, *P*=0.004; decreased rightward), and postcentral gyrus thickness (*t*=-2.5, *P*=0.013; decreased leftward). The corresponding unilateral effects are shown in **Table S14**.

3.3.1.4 All age groups combined

When combining all age groups, there were nominally significant associations of AIs with diagnosis for the medial orbitofrontal cortex surface area (*t*=2.2, *P*=0.029; decreased rightward), paracentral lobule surface area (*t*=-2.2, P=0.029; increased rightward), pars orbitalis of inferior frontal gyrus surface area (*t*=2.3, P=0.021; decreased rightward), caudal middle frontal thickness (*t*=2.2, *P*=0.027; increased leftward), insula thickness (*t*=-2.1, *P*=0.040; decreased leftward), as well as the volume of the globus pallidus

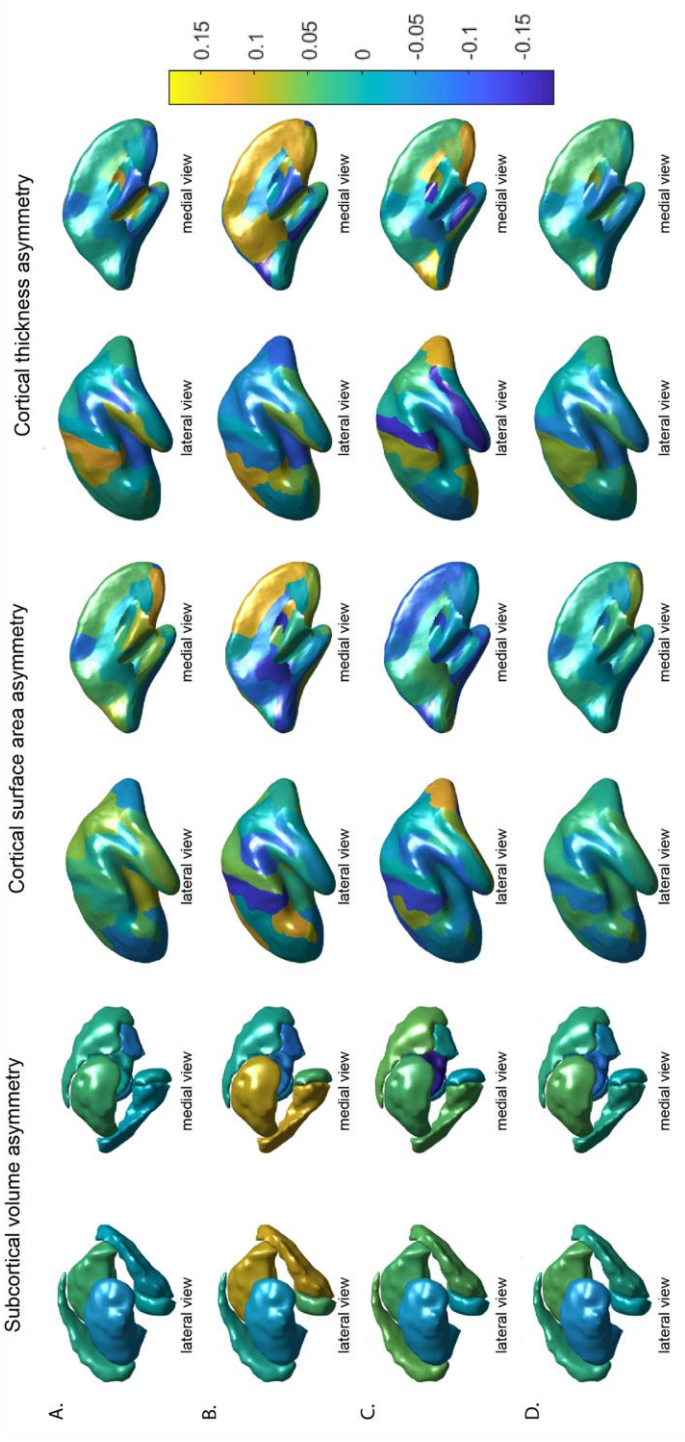


Figure 1. Cohen's d effect sizes of the associations between ADHD diagnosis and AIs of subcortical volumes (left), cortical surface areas (middle) and cortical thicknesses (right) for (A) children, (B) adolescents, (C) adults, and (D) all age groups combined. Cohen's *d* values are overlayed on left hemisphere inflated brains. Positive values (yellow) indicate mean shifts towards greater leftward or reduced rightward asymmetry in ADHD, and negative values (blue) indicate mean shifts towards greater rightward asymmetry or reduced leftward asymmetry in ADHD.

(*t*=-2.6, *P*=0.010; decreased leftward) (**Tables 1-3, Table S11-S13).** The corresponding unilateral effects are shown in **Table S14**.

No effects were significant at FDR-corrected *P*<0.05 when the correction was done across all age groups and AIs of different types.

The addition of non-linear effects of age to the model had negligible influences on the six nominally significant associations with diagnosis, all of which remained nominally significant except insula thickness (now *P*=0.050). Likewise, winsorizing outliers (using a threshold k=3, see **Methods**) also had little influence on the results (the effect on insula thickness asymmetry was no longer nominally significant, *P*=0.061) (**Tables S15-S17**).

3.3.2 Associations brain asymmetries with comorbidity, ADHD severity, psychostimulant medication, and IQ

Analyses in this section were carried out in all age groups combined.

When testing associations of comorbidity, ADHD severity, psychostimulant medication, or IQ with brain asymmetries within ADHD individuals (**Tables S18-S29**), only one significant association was found (FDR <0.05 within the particular type of AI and age-defined group), namely between comorbid mood disorder and the rostral middle frontal gyrus thickness AI (*P*=0.0002, *t*=3.70) (**Table S26**). Furthermore, various nominally significant (*P*<0.05) associations were observed: ADHD severity was associated with the AI of the entorhinal cortex surface area **(***t*=2.12, *P*=0.034; hyperactivity/impulsivity) (**Table S19**). ADHD severity was also associated with four regional cortical thickness asymmetries: the caudal anterior cingulate thickness AI (*t*=2.66, *P*=0.008; hyperactivity/impulsivity), *pars opercularis* of the inferior frontal gyrus thickness AI (*t*=2.12, *P*=0.034; hyperactivity/impulsivity, and *t*=2.04, *P*=0.04; inattention), and pericalcarine cortex thickness AI (*t*=2.04, *P*=0.04; hyperactivity/impulsivity) (**Table S20**).

Current psychostimulant medication use was associated with two cortical regional surface area asymmetries, i.e., precuneus (*t*=-2.25, *P*=0.025) and transverse temporal gyrus (*t*=-2.34, *P*=0.020) (**Table S22**), and with two

thickness asymmetries, i.e., inferior parietal cortex (*t*=-2.33, P=0.020) and precentral gyrus (*t*=-2.16, *P*=0.031) (**Table S23**). Lifetime psychostimulant medication use was associated with three cortical surface area asymmetries (insula (*t*=-2.03, *P*=0.043), supramarginal gyrus (*t*=-2.08, *P*=0.038), and rostral anterior cingulate cortex (*t*=1.97, *P*=0.049) (**Table S22**), and the thickness asymmetry of the paracentral lobule (*t*=2.15, *P*=0.032) (**Table S23**). Among the AIs which showed nominally significant associations with medication use, one had also shown a nominally significant association with diagnosis in all age groups combined, i.e., the AI of paracentral lobule surface area (see above). The direction of medication effect was positive, i.e. the opposite to the effect of diagnosis on this AI (see above).

For mood disorder, associations were observed with six thickness AIs (i.e., entorhinal cortex, pars triangularis of inferior frontal gyrus, pericalcarine cortex, precuneus, rostral middle frontal gyrus, and transverse temporal gyrus), and two surface area AIs (i.e., inferior temporal gyrus, and rostral anterior cingulate cortex), of which the association with rostral middle frontal thickness AI survived multiple testing correction (FDR < 0.05) (**Tables S25, S26**). Anxiety Disorder was associated with thickness AIs of the cuneus and lateral occipital cortex (**Table S26**). For ODD, associations were found with the AIs of medial orbitofrontal thickness (**Table S26**) and temporal pole surface area (**Table S25**). Additionally, SUD was associated with the thickness AIs of the cuneus and paracentral lobule (**Table S26**), and with surface area AIs of the postcentral gyrus and supramarginal gyrus (**Table S25**). None of these regions showed a nominally significant effect of diagnosis in the main analysis of all age groups combined.

Finally, within ADHD individuals, IQ was nominally associated with the accumbens volume AI (*t*=2.16, *P*=0.031), hippocampus volume AI (*t*=-2.06, *P*=0.039) (**Table S27**) and lateral occipital cortex surface area AI (*t*=-2.17, *P*=0.030) (**Table S28**). Within controls, IQ was associated with the middle temporal gyrus surface area AI (*t*=-2.52, *P*=0.012) (**Table S28**), rostral anterior cingulate thickness cortex AI (*t*=2.47, *P*=0.014), and supramarginal gyrus thickness AI (*t* = -2.55, *P*=0.011) (**Table S29**).

3.3.2.1 Including IQ, handedness or intracranial volume as covariates in case-control analysis

We carried out secondary analyses in which IQ, handedness or intracranial volume were included as covariates in case-control analysis, with or without interaction terms for these variables with diagnosis (i.e. case-control status) (see Methods for the models used). These extra models identified a small number of main effects of diagnosis, or interactions with diagnosis, that survived multiple testing correction at FDR<0.05 within the specific subset of AIs and ages being analyzed (but would not survive further correction for multiple testing). However, after winsorization of outliers (see Methods), only the diagnosis term for globus pallidus volume AI remained significant, in the model *AI ~ diagnosis + age + sex + ICV + random (~1|dataset)*, when analyzed in the total study sample (P=0.005, t=-2.75), or when analyzed in adults only (P=0.0035, t=-2.93). Complete model results from all of these secondary analyses can be found in supplementary comma-delimited text files.

3.4 Discussion

We conducted the largest study to date of associations between anatomical brain asymmetries and ADHD. Linear mixed effects model mega-analyses were carried out separately in children, adolescents, and adults, following previous ENIGMA ADHD working group studies of bilateral brain differences that showed contrasting effects in these age groups (Hoogman et al., 2017; Hoogman et al., 2019). We also analyzed the total study sample for age-general effects. All statistical effects of diagnosis on asymmetries were very small, with Cohen's *d* ranging from -0.18 to 0.18. Only one of these associations was significant with a false discovery rate <0.05 within the specific subset of AIs and age-defined subjects in which it was found (globus pallidus asymmetry in adults), and this effect was not significant in analysis of all age groups combined, with FDR correction across all AIs. Therefore, all effects remain tentative, even in this unprecedented sample size. The small effect sizes mean that altered brain asymmetry is unlikely, in itself, to be a

useful biomarker or clinical predictor of ADHD. In addition, our results suggest that significant effects reported in prior studies, based on much smaller samples, may have been unrealistically large. As noted in the Introduction, low power not only reduces the chance of detecting true effects, but also increases the likelihood that statistically significant results do not reflect true effects (Munafo & Flint, 2010). There were some notable associations of diagnosis with cortical asymmetry that reached nominal significance in our study. Among these, children with ADHD showed reduced rightward asymmetry of total hemispheric surface area, and medial orbitofrontal surface area. In a recent ENIGMA consortium study of autism spectrum disorder (ASD), medial orbitofrontal cortex surface area asymmetry was altered in the same direction, and to a similar extent, as in the present study (Postema et al., 2019). ADHD and ASD often co-occur (Leitner, 2014) and are known to share genetic influences (Ghirardi et al., 2019; Stergiakouli et al., 2017), such that the two diagnostic labels are likely to capture a partly overlapping spectrum of related disorders (Demopoulos, Hopkins, & Davis, 2013; van der Meer et al., 2012). Studies that aimed to identify shared brain structural traits between ADHD and ASD have found mixed results (Nevena V. Radonjić MD, 2019; Premika S.W. Boedhoe et al., 2019), with perhaps the greatest overlap involving regions of the 'social brain', including orbitofrontal cortex (Baribeau et al., 2019). However, laterality has not been specifically studied in this regard, so that our finding of reduced rightward medial orbitofrontal cortex surface area in both disorders may be a new insight into shared neurobiology between ADHD and ASD. Altered lateralized neurodevelopment may play a causal role in disorder susceptibility, or else may arise as a correlated trait due to other underlying susceptibility factors, or even be a downstream consequence of having the disorder (Bishop, 2013). Some aspects of brain asymmetry are partly heritable (Guadalupe et al., 2016; Kong et al., 2018), so that future gene mapping studies for brain asymmetry and disorder susceptibility may help to resolve causal relations underlying their associations.

One functional imaging study (94 cases, 85 controls) reported lower rightward lateralization in medial orbitofrontal cortex in ADHD compared to controls, based on temporal variability during resting-state (Zou & Yang, 2019). Furthermore, a study of 218 participants with ADHD and 358 healthy

controls reported that orbitofrontal cortex thickness, but not surface area, showed a left>right asymmetry in childhood controls that switched to right>left asymmetry by late adolescence, while this change did not occur to the same extent in ADHD (Shaw et al., 2009). However, in the present study, there was no effect of diagnosis on thickness asymmetry of this region in children or adolescents, while in adults, ADHD was associated with a relatively rightward shift of asymmetry compared to controls, i.e. opposite to what might be expected according to Shaw *et al*. For other cortical asymmetries too, our findings in this large-scale study were discrepant with what might have been expected from previous reports in smaller samples (see references in the Introduction). For example, a prior study reported reversed grey matter volume asymmetry (i.e., leftward instead of rightward) of the superior frontal gyrus in ADHD (Cao et al., 2014), but we saw no clear evidence of this in the present study.

The most often reported alteration of brain asymmetry in ADHD has involved the caudate nucleus, although the direction of the effect has not been consistent (Castellanos et al., 1996; Dang et al., 2016; Filipek et al., 1997; Hynd et al., 1993; Schrimsher et al., 2002; Uhlikova et al., 2007). We did not find evidence for altered asymmetry of caudate nucleus volume in the present study, again suggesting that prior findings were false positives in smaller samples. As mentioned above, we found a tentative association with ADHD for another regional asymmetry of the basal ganglia, namely of the globus pallidus, in adults-only. The globus pallidus is involved in movement and reward processing (Munte et al., 2017), both of which are involved in the symptomatology of ADHD. A previous meta-analysis comprising data from a total of 114 participants with ADHD (or a related disorder) and 143 control participants, noted a significantly lower average right putamen and right globus pallidus volumes in ADHD (Ellison-Wright, Ellison-Wright, & Bullmore, 2008), although asymmetry was not quantified in that study. Regardless, our finding of lower leftward asymmetry seems discrepant with this earlier report.

We have already remarked on the limited statistical power of previous studies as a likely explanation for their findings being discrepant with the current study. Low sample sizes in relation to subtle effects can result in poor

reproducibility (Button et al., 2013; Munafo & Flint, 2010). Here, we had 80% power to detect case-control Cohen's *d* effect sizes as low as roughly 0.12, or as high as 0.3 in the smallest subset by age (see **Methods**). In addition to limited sample sizes, there are various other possible explanations for discrepancies with previous studies. Methodological differences in hardware, software, and data processing pipelines can influence results (Biberacher et al., 2016), although our focus on asymmetry through use of the AI is likely to have reduced the impact of heterogeneity factors that affect both hemispheres equally. In contrast to some previous studies mentioned above, we did not consider gyral/sulcal patterns or cortical grey matter volumes as such. Rather, we studied regional cortical thicknesses and surface areas as distinct measures, which together drive grey matter volumetric measures. Since area and thickness have been shown to vary relatively independently (Panizzon et al., 2009), separate analyses are advisable, although cortical thickness measures are particularly prone to effects related to site-, scanner- or protocol differences (Chung et al., 2017; Fortin et al., 2018). Likewise, the choice of brain atlas can influence results, as each atlas has its own properties that impact brain segmentation (Yaakub et al., 2020). In addition, the approach we used is based on hemisphere-specific definitions of regional anatomy, because each hemisphere has its own atlas, based on its own average distribution of features (Desikan et al., 2006). Correspondence between hemispheres is then achieved at the regional level, based on expert neuroanatomical regional segmentation that was adapted to each hemisphere's distinct features when constructing the atlas. However, future studies using higher-resolution atlases, or vertex-based analysis using hemispheric co-registration (Kang, Herron, Cate, Yund, & Woods, 2012; Maingault, Tzourio-Mazoyer, Mazoyer, & Crivello, 2016), may identify restricted regions showing stronger associations between ADHD and cortical asymmetry than we report here. Furthermore, for subcortical volume asymmetries, discrepancies between the findings of our study and previous studies could be due to differences in parcellation methods, which can perform with varying accuracy (Guadalupe et al., 2014; Pardoe, Pell, Abbott, & Jackson, 2009; Perlaki et al., 2017).

The conceptualization of laterality can also differ across studies. In terms of AIs, our cortical results are largely in line with a previous report based on measuring grey matter volume asymmetries in 192 participants with ADHD and 508 controls (Douglas et al., 2018), insofar as no FDR-significant results were found (five of those datasets were in common with the present study, see **Methods**). However, the authors of that study also calculated the unsigned magnitudes of the AIs (i.e., absolute degrees of asymmetry, regardless of directions). They reported significant differences in absolute asymmetry for various cortical and subcortical structures (Douglas et al., 2018). In the present study, we did not calculate absolute AIs, in order not to compound multiple testing, and because these measures are highly non-normal with a floor effect at value zero, which would violate the assumptions of the modelling that we applied. It is not clear whether this issue may have affected the results in the earlier study (Douglas et al., 2018). Future studies may consider the unsigned magnitude of brain asymmetry indexes further in ADHD, but it will be necessary to use statistical methods that can account for non-normal distributions.

Discrepancies with earlier studies may also be due to differences in clinical features of the disorder that arise from case recruitment and diagnosis, for example with respect to medication use (which has been suggested to partly normalize brain structural abnormalities, although the previous ENIGMA studies of bilateral changes in ADHD did not support this) (Nakao, Radua, Rubia, & Mataix-Cols, 2011; Pretus et al., 2017), comorbidities (Reale et al., 2017), symptom severity, and/or IQ. Some asymmetries showed tentative associations with some of these clinical variables in the present study, although none of these results survived correction for multiple testing, apart from mood disorder with the rostral middle frontal thickness AI. Also, some of the clinical variables (medication, comorbidity) were missing for many ADHD individuals in this study. Regardless, it remains possible that certain subsets of ADHD might be associated more strongly with altered brain asymmetry than was apparent in our large-scale analysis of average changes over many datasets, comprising many and varied collections of ADHD individuals and controls.

In general, between-centre heterogeneity (in terms of scanning setup, patient subgroups, demographics) may result in reduced statistical power to detect effects that are specific to certain subgroups of datasets, or to individual datasets, when tested in mega-analysis over all datasets. For example, harmonization of scanning protocols might lead to stronger effects being found, as heterogeneity of this aspect would be reduced. Here we used random-intercept models to adjust for heterogeneity between datasets. This was a strong correction for cross-dataset heterogeneity, as it removed mean differences between datasets, although between-dataset heterogeneity that affected model coefficients within datasets would not be fully accounted by this approach. While the random-intercept model cannot fully rescue power in the case that effects are truly specific to certain subsets, no single centre has been able to collect such a large ADHD-control sample alone. Our large sample size yields accurate estimates of effect sizes with respect to the overall case-control population, as represented across many research centres. In this way, the findings from multi-centre studies such as ours can be considered more generalizable than single-centre studies (Costafreda, 2009). In any case, as long as researchers publish separate papers based on many single, smaller datasets, collected in particular ways, the field overall has the same issue of heterogeneity. Of note, the ENIGMA consortium previously showed that using the random intercept approach to account for dataset heterogeneity is similar to random effect meta-analysis across datasets, but preferable because it produces lower standard errors and narrower confidence intervals than meta-analysis (Boedhoe et al., 2018).

Although not a longitudinal study, our data spanned a wide age range from childhood through to older adulthood, which allowed us to study different age groups separately, as the disorder may be neurobiologically distinct in different age groups (Alexander & Farrelly, 2018; Hoogman et al., 2019). The previous ENIGMA study of bilateral cortical differences in ADHD found children to be most affected, particularly in frontal, cingulate, and temporal regions, as well as the total hemispheric surface area, which was lower in ADHD (Hoogman et al., 2019). In the children-only analysis in our present study of asymmetries, we also found associations with diagnosis for some frontal and temporal regions (including the caudal middle frontal cortex thickness, precentral gyrus thickness, medial orbitofrontal cortex surface

area, banks of the superior temporal sulcus thickness), as well as a change in the asymmetry of total hemispheric surface area, driven by a greater decrease of area in ADHD on the right-side than the left-side. These findings offer a more nuanced description of brain changes in childhood ADHD, which may involve altered lateralized neurodevelopment.

However, considering all brain asymmetry measures, the effect sizes in the present study were not stronger in children as compared to adolescents or adults. Furthermore, bilateral case-control differences are not necessarily a good guide to case-control differences in asymmetry, since a difference in asymmetry can arise, for example, from a simultaneous left-sided increase and right-sided decrease in a brain measure, which can involve no change at all in the bilateral measure. Hence, we took a screening approach to the present study, rather than constraining our search on prior bilateral findings. It is also not entirely clear how/whether to statistically adjust the test for total hemispheric surface asymmetry, in the context of also testing multiple sub-regions, and also with respect to study-wide multiple testing. Therefore, we present all *P*-values unadjusted, while also being mindful of the tentative nature of these findings in the context of our survey across many brain asymmetry measures. Together with the corresponding effect size estimates, this mapping information should be useful for the field.

We did not include handedness, IQ or brain size as covariates in our primary analysis, in order to avoid possible collider bias (Cole et al., 2010), as there are various plausible causal relations linking these traits with ADHD, brain asymmetry and other, possibly underlying factors shared between some or all of them. For example having the disorder for other underlying reasons may lead to altered asymmetry and brain size, or altered asymmetry and brain size may contribute to having the disorder. A priori, altered asymmetry may not be associated with the disorder, but be associated with brain size, which can be associated with disorder. In this latter case, correcting for brain size can induce spurious associations between asymmetry and disorder. Collider bias is under-appreciated in the field, perhaps because it is not intuitive. Alternatively, including brain size as a covariate in case-control analysis might have reduced the power to detect an association of diagnosis with asymmetry. This would occur if underlying susceptibility factors contribute

both to altered asymmetry and reduced brain size, as part of the ADHD phenotype. Regardless, our primary interest was to detect associations of diagnosis with asymmetry regardless of other brain features such as overall size. We have made available, in supplementary csv files, the results from secondary analyses in which we included handedness, IQ or intracranial volume as covariates, with or without interaction terms with case-control status. As regards handedness specifically, previous studies of subcortical and cortical anatomical asymmetry in over 15,000 subjects from healthy control and population datasets, also performed by the ENIGMA consortium (Guadalupe et al., 2016; Kong et al., 2018), found no significant effects of handedness.

Our study was limited to macro-anatomical asymmetries of cortical grey matter and subcortical volumes. It is possible that altered brain asymmetry in ADHD will be more apparent in different structural or functional modalities, or at different scales. For example, cortical thickness measures can correlate with the degree of myelination (Natu et al., 2019), such that quantitative neuroimaging methods that are sensitive to microstructural tissue content may reveal further alterations in ADHD. At a larger scale, asymmetries of white matter tracts (Wu et al., 2020) may also benefit from the large-scale approach that we have used here. Asymmetries of functional asymmetry, particularly linked to attentional tasks, may also reveal stronger case-control differences than the structural effects we observed (see Introduction).

To conclude, we carried out the largest case-control study of structural brain asymmetry in ADHD. We described average changes of asymmetry that are small, but helpful towards a more complete description of brain anatomical changes in this disorder. Results were largely discrepant with earlier, inconsistent findings from smaller-scale studies, which illustrates the value of taking a large-scale approach to human clinical neuroscience. The small effects that we found remain statistically tentative in the context of multiple testing, even in this unprecedented sample size. Future longitudinal and genetic studies may probe causative relations between ADHD and brain asymmetry, focused on measures defined from this study, such as total hemispheric surface area asymmetry, medial orbitofrontal area asymmetry, or globus pallidus volume asymmetry.

3.5 References

Alexander, L., & Farrelly, N. (2018). Attending to adult ADHD: a review of the neurobiology behind adult ADHD. *Ir J Psychol Med, 35*(3), 237-244. doi:10.1017/ipm.2017.78

American Psychiatric Association. (2000). Diagnostic and Statistical Manual of Mental disorders, 4th edition (DSM-IV), Washington DC.

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). Washington, DC.

Baribeau, D. A., Dupuis, A., Paton, T. A., Hammill, C., Scherer, S. W., Schachar, R. J., . . . Anagnostou, E. (2019). Structural neuroimaging correlates of social deficits are similar in autism spectrum disorder and attention-deficit/hyperactivity disorder: analysis from the POND Network. *Transl Psychiatry, 9*(1), 72. doi:10.1038/s41398-019-0382-0

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate - A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B-Methodological, 57*(1), 289-300.

Biberacher, V., Schmidt, P., Keshavan, A., Boucard, C. C., Righart, R., Samann, P., . . . Muhlau, M. (2016). Intra- and interscanner variability of magnetic resonance imaging based volumetry in multiple sclerosis. *Neuroimage, 142*, 188-197. doi:10.1016/j.neuroimage.2016.07.035

Bishop, D. V. (2013). Cerebral asymmetry and language development: cause, correlate, or consequence? *Science, 340*(6138), 1230531. doi:10.1126/science.1230531

Boedhoe, P. S. W., Heymans, M. W., Schmaal, L., Abe, Y., Alonso, P., Ameis, S. H., . . . Twisk, J. W. R. (2018). An Empirical Comparison of Meta- and Mega-Analysis With Data From the ENIGMA Obsessive-Compulsive Disorder Working Group. *Front Neuroinform, 12*, 102. doi:10.3389/fninf.2018.00102

Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., & Munafo, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nat Rev Neurosci, 14*(5), 365-376. doi:10.1038/nrn3475

Cao, Q., Wang, J., Sun, L., Wang, P., Wu, Z., & Wang, Y. (2014). [Altered anatomical asymmetry in children with attention deficit/hyperactivity disorder: a pilot optimized voxel-based morphometric study]. *Zhonghua Yi Xue Za Zhi, 94*(43), 3387-3391.

Castellanos, F. X., Giedd, J. N., Marsh, W. L., Hamburger, S. D., Vaituzis, A. C., Dickstein, D. P., . . . Rapoport, J. L. (1996). Quantitative brain magnetic resonance imaging in attention-deficit hyperactivity disorder. *Arch Gen Psychiatry, 53*(7), 607-616.

Chambers, J. M., & Hastie, T. J. (1992). Statistical models in S. Pacific Grove, California, USA, *Wadsworth & Brooks/Cole*.

Chung, J., Yoo, K., Lee, P., Kim, C. M., Roh, J. H., Park, J. E., . . . Jeong, Y. (2017). Normalization of cortical thickness measurements across different T1 magnetic resonance imaging protocols by novel W-Score standardization. *Neuroimage, 159*, 224-235. doi:10.1016/j.neuroimage.2017.07.053

Cole, S. R., Platt, R. W., Schisterman, E. F., Chu, H., Westreich, D., Richardson, D., & Poole, C. (2010). Illustrating bias due to conditioning on a collider. *Int J Epidemiol, 39*(2), 417-420. doi:10.1093/ije/dyp334

Cortese, S., Kelly, C., Chabernaud, C., Proal, E., Di Martino, A., Milham, M. P., & Castellanos, F. X. (2012). Toward systems neuroscience of ADHD: a meta-analysis of 55 fMRI studies. *Am J Psychiatry, 169*(10), 1038-1055. doi:10.1176/appi.ajp.2012.11101521

Costafreda, S. (2009). Pooling fMRI data: meta-analysis, mega-analysis and multi-center studies. *Frontiers in Neuroinformatics, 3*(33). doi:10.3389/neuro.11.033.2009

Dang, L. C., Samanez-Larkin, G. R., Young, J. S., Cowan, R. L., Kessler, R. M., & Zald, D. H. (2016). Caudate asymmetry is related to attentional impulsivity and an objective measure of ADHD-like attentional problems in healthy adults. *Brain Struct Funct, 221*(1), 277-286. doi:10.1007/s00429-014-0906-6

Demopoulos, C., Hopkins, J., & Davis, A. (2013). A comparison of social cognitive profiles in children with autism spectrum disorders and attention-deficit/hyperactivity disorder: a matter of quantitative but not qualitative difference? *J Autism Dev Disord, 43*(5), 1157-1170. doi:10.1007/s10803-012-1657-y

Desikan, R. S., Segonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., . . . Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage, 31*(3), 968-980. doi:10.1016/j.neuroimage.2006.01.021

Douglas, P. K., Gutman, B., Anderson, A., Larios, C., Lawrence, K. E., Narr, K., . . . Bookheimer, S. Y. (2018). Hemispheric brain asymmetry differences in

youths with attention-deficit/hyperactivity disorder. *Neuroimage Clin, 18*, 744-752. doi:10.1016/j.nicl.2018.02.020

Duboc, V., Dufourcq, P., Blader, P., & Roussigne, M. (2015). Asymmetry of the Brain: Development and Implications. *Annu Rev Genet, 49*, 647-672. doi:10.1146/annurev-genet-112414-055322

Ellison-Wright, I., Ellison-Wright, Z., & Bullmore, E. (2008). Structural brain change in Attention Deficit Hyperactivity Disorder identified by meta-analysis. *BMC Psychiatry, 8*, 51. doi:10.1186/1471-244X-8-51

Faraone, S. V., Asherson, P., Banaschewski, T., Biederman, J., Buitelaar, J. K., Ramos-Quiroga, J. A., . . . Franke, B. (2015). Attention-deficit/hyperactivity disorder. *Nat Rev Dis Primers, 1*, 15020. doi:10.1038/nrdp.2015.20

Fayyad, J., Sampson, N. A., Hwang, I., Adamowski, T., Aguilar-Gaxiola, S., Al-Hamzawi, A., . . . Kessler, R. C. (2017). The descriptive epidemiology of DSM-IV Adult ADHD in the World Health Organization World Mental Health Surveys. *Atten Defic Hyperact Disord, 9*(1), 47-65. doi:10.1007/s12402-016-0208-3

Filipek, P. A., SemrudClikeman, M., Steingard, R. J., Renshaw, P. F., Kennedy, D. N., & Biederman, J. (1997). Volumetric MRI analysis comparing subjects having attention-deficit hyperactivity disorder with normal controls. *Neurology, 48*(3), 589-601.

Fischl, B. (2012). FreeSurfer. *Neuroimage, 62*(2), 774-781. doi:10.1016/j.neuroimage.2012.01.021

Fortin, J. P., Cullen, N., Sheline, Y. I., Taylor, W. D., Aselcioglu, I., Cook, P. A., . . . Shinohara, R. T. (2018). Harmonization of cortical thickness measurements across scanners and sites. *Neuroimage, 167*, 104-120. doi:10.1016/j.neuroimage.2017.11.024

Geeraerts, S., Lafosse, C., Vaes, N., Vandenbussche, E., & Verfaillie, K. (2008). Dysfunction of right-hemisphere attentional networks in attention deficit hyperactivity disorder. *J Clin Exp Neuropsychol, 30*(1), 42-52. doi:10.1080/13803390601186676

Ghirardi, L., Pettersson, E., Taylor, M. J., Freitag, C. M., Franke, B., Asherson, P., . . . Kuja-Halkola, R. (2019). Genetic and environmental contribution to the overlap between ADHD and ASD trait dimensions in young adults: a twin study. *Psychol Med, 49*(10), 1713-1721. doi:10.1017/s003329171800243x

Guadalupe, T., Mathias, S. R., vanErp, T. G., Whelan, C. D., Zwiers, M. P., Abe, Y., . . . Francks, C. (2016). Human subcortical brain asymmetries in 15,847

people worldwide reveal effects of age and sex. *Brain Imaging Behav*. doi:10.1007/s11682-016-9629-z

Guadalupe, T., Zwiers, M. P., Teumer, A., Wittfeld, K., Vasquez, A. A., Hoogman, M., . . . Francks, C. (2014). Measurement and genetics of human subcortical and hippocampal asymmetries in large datasets. *Hum Brain Mapp, 35*(7), 3277-3289. doi:10.1002/hbm.22401

Hale, T. S., Kane, A. M., Kaminsky, O., Tung, K. L., Wiley, J. F., McGough, J. J., . . . Kaplan, J. T. (2014). Visual Network Asymmetry and Default Mode Network Function in ADHD: An fMRI Study. *Front Psychiatry, 5*, 81. doi:10.3389/fpsyt.2014.00081

Hale, T. S., Loo, S. K., Zaidel, E., Hanada, G., Macion, J., & Smalley, S. L. (2009). Rethinking a right hemisphere deficit in ADHD. *J Atten Disord, 13*(1), 3-17. doi:10.1177/1087054708323005

Hale, T. S., Smalley, S. L., Walshaw, P. D., Hanada, G., Macion, J., McCracken, J. T., . . . Loo, S. K. (2010). Atypical EEG beta asymmetry in adults with ADHD. *Neuropsychologia, 48*(12), 3532-3539. doi:10.1016/j.neuropsychologia.2010.08.002

Heilman, K. M., Bowers, D., Valenstein, E., & Watson, R. T. (1986). The right hemisphere: neuropsychological functions. *J Neurosurg, 64*(5), 693-704. doi:10.3171/jns.1986.64.5.0693

Hoogman, M., Bralten, J., Hibar, D. P., Mennes, M., Zwiers, M. P., Schweren, L. S., . . . Franke, B. (2017). Subcortical brain volume differences in participants with attention deficit hyperactivity disorder in children and adults: a cross-sectional mega-analysis. *Lancet Psychiatry, 4*(4), 310-319. doi:10.1016/s2215-0366(17)30049-4

Hoogman, M., Muetzel, R., Guimaraes, J. P., Shumskaya, E., Mennes, M., Zwiers, M. P., . . . Franke, B. (2019). Brain Imaging of the Cortex in ADHD: A Coordinated Analysis of Large-Scale Clinical and Population-Based Samples. *Am J Psychiatry*, appiajp201918091033. doi:10.1176/appi.ajp.2019.18091033

Hynd, G. W., Hern, K. L., Novey, E. S., Eliopulos, D., Marshall, R., Gonzalez, J. J., & Voeller, K. K. (1993). Attention deficit-hyperactivity disorder and asymmetry of the caudate nucleus. *J Child Neurol, 8*(4), 339-347. doi:10.1177/088307389300800409

Kang, X., Herron, T. J., Cate, A. D., Yund, E. W., & Woods, D. L. (2012). Hemispherically-unified surface maps of human cerebral cortex: reliability and hemispheric asymmetries. *PLoS One, 7*(9), e45582. doi:10.1371/journal.pone.0045582

Kong, X. Z., Mathias, S. R., Guadalupe, T., Glahn, D. C., Franke, B., Crivello, F., . . . Francks, C. (2018). Mapping cortical brain asymmetry in 17,141 healthy individuals worldwide via the ENIGMA Consortium. *Proc Natl Acad Sci U S A*. doi:10.1073/pnas.1718418115

Kurth, F., Gaser, C., & Luders, E. (2015). A 12-step user guide for analyzing voxel-wise gray matter asymmetries in statistical parametric mapping (SPM). *Nat Protoc, 10*(2), 293-304. doi:10.1038/nprot.2015.014

Langleben, D. D., Austin, G., Krikorian, G., Ridlehuber, H. W., Goris, M. L., & Strauss, H. W. (2001). Interhemispheric asymmetry of regional cerebral blood flow in prepubescent boys with attention deficit hyperactivity disorder. *Nucl Med Commun, 22*(12), 1333-1340.

Leitner, Y. (2014). The co-occurrence of autism and attention deficit hyperactivity disorder in children - what do we know? *Front Hum Neurosci, 8*, 268. doi:10.3389/fnhum.2014.00268

Leroy, F., Cai, Q., Bogart, S. L., Dubois, J., Coulon, O., Monzalvo, K., . . . Dehaene-Lambertz, G. (2015). New human-specific brain landmark: the depth asymmetry of superior temporal sulcus. *Proc Natl Acad Sci U S A, 112*(4), 1208-1213. doi:10.1073/pnas.1412389112

Li, D., Li, T., Niu, Y., Xiang, J., Cao, R., Liu, B., . . . Wang, B. (2018). Reduced hemispheric asymmetry of brain anatomical networks in attention deficit hyperactivity disorder. *Brain Imaging Behav*. doi:10.1007/s11682-018-9881-5

Li, X., Jiang, J., Zhu, W., Yu, C., Sui, M., Wang, Y., & Jiang, T. (2007). Asymmetry of prefrontal cortical convolution complexity in males with attention-deficit/hyperactivity disorder using fractal information dimension. *Brain Dev, 29*(10), 649-655. doi:10.1016/j.braindev.2007.04.008

Maingault, S., Tzourio-Mazoyer, N., Mazoyer, B., & Crivello, F. (2016). Regional correlations between cortical thickness and surface area asymmetries: A surface-based morphometry study of 250 adults. *Neuropsychologia, 93*(Pt B), 350-364. doi:10.1016/j.neuropsychologia.2016.03.025

Mohamed, S. M., Börger, N. A., Geuze, R. H., & van der Meere, J. J. (2016). Linking state regulation, brain laterality, and self-reported attention-deficit/hyperactivity disorder (ADHD) symptoms in adults. *J Clin Exp Neuropsychol, 38*(8), 831-843. doi:10.1080/13803395.2016.1167174

Munafo, M. R., & Flint, J. (2010). How reliable are scientific studies? *Br J Psychiatry, 197*(4), 257-258. doi:10.1192/bjp.bp.109.069849

Munte, T. F., Marco-Pallares, J., Bolat, S., Heldmann, M., Lutjens, G., Nager, W., . . . Krauss, J. K. (2017). The human globus pallidus internus is sensitive to

rewards - Evidence from intracerebral recordings. *Brain Stimul, 10*(3), 657-663. doi:10.1016/j.brs.2017.01.004

Nakao, T., Radua, J., Rubia, K., & Mataix-Cols, D. (2011). Gray matter volume abnormalities in ADHD: voxel-based meta-analysis exploring the effects of age and stimulant medication. *Am J Psychiatry, 168*(11), 1154-1163. doi:10.1176/appi.ajp.2011.11020281

Natu, V. S., Gomez, J., Barnett, M., Jeska, B., Kirilina, E., Jaeger, C., . . . Grill-Spector, K. (2019). Apparent thinning of human visual cortex during childhood is associated with myelination. *Proceedings of the National Academy of Sciences, 116*(41), 20750-20759. doi:10.1073/pnas.1904931116

Nevena V. Radonjić MD, P., Jonathan L. Hess PhD, Paula Rovira, Ole Andreassen PhD, Jan K. Buitelaar MD, PhD, Christopher R. K. Ching PhD, Barbara Franke PhD, Martine Hoogman PhD10, Neda Jahanshad PhD, Carrie McDonald PhD, Lianne Schmaal PhD, Sanjay M. Sisodiya PhD, Dan J. Stein PhD, Odile A. van den Heuvel MD, PhD, Theo G.M. van Erp PhD, Daan van Rooij PhD, Dick J. Veltman MD, PhD, Paul Thompson PhD, Stephen V. Faraone PhD. (2019). Structural Brain Imaging Studies Offer Clues about the Effects of the Shared Genetic Etiology among Neuropsychiatric Disorders. *bioRxiv*.

Panizzon, M. S., Fennema-Notestine, C., Eyler, L. T., Jernigan, T. L., Prom-Wormley, E., Neale, M., . . . Kremen, W. S. (2009). Distinct genetic influences on cortical surface area and cortical thickness. *Cereb Cortex, 19*(11), 2728-2735. doi:10.1093/cercor/bhp026

Pardoe, H. R., Pell, G. S., Abbott, D. F., & Jackson, G. D. (2009). Hippocampal volume assessment in temporal lobe epilepsy: How good is automated segmentation? *Epilepsia, 50*(12), 2586-2592. doi:10.1111/j.1528-1167.2009.02243.x

Perlaki, G., Horvath, R., Nagy, S. A., Bogner, P., Doczi, T., Janszky, J., & Orsi, G. (2017). Comparison of accuracy between FSL's FIRST and Freesurfer for caudate nucleus and putamen segmentation. *Sci Rep, 7*(1), 2418. doi:10.1038/s41598-017-02584-5

Polanczyk, G., de Lima, M. S., Horta, B. L., Biederman, J., & Rohde, L. A. (2007). The worldwide prevalence of ADHD: a systematic review and metaregression analysis. *Am J Psychiatry, 164*(6), 942-948. doi:10.1176/ajp.2007.164.6.942

Postema, M. C., van Rooij, D., Anagnostou, E., Arango, C., Auzias, G., Behrmann, M., . . . Francks, C. (2019). Altered structural brain asymmetry in autism

spectrum disorder in a study of 54 datasets. *Nat Commun, 10*(1), 4958. doi:10.1038/s41467-019-13005-8

Premika S.W. Boedhoe, M. S., Daan van Rooij, Ph.D. *, Martine Hoogman, Ph.D.*, Jos W.R. Twisk Ph.D., Lianne, Schmaal, P. D., Yoshinari Abe, M.D., Ph.D., Pino Alonso, M.D., Ph.D., Stephanie H. Ameis, M.D., M.Sc., Anatoly, Anikin, P. D., Alan Anticevic, Ph.D., Philip Aherson, Ph.D., Celso Arango, M.D., Ph.D., Paul D. Arnold, M.D.,, Ph.D., F. A., Ph.D., Guillaume Auzias, Ph.D., Tobias Banaschewski, M.D., Ph.D., Alexander, Baranov, P. D., Marcelo C. Batistuzzo, Ph.D., Sarah Baumeister, Ph.D., Ramona Baur-Streubel, Ph.D., Marlene, Behrmann, P. D., Mark A. Bellgrove, Ph.D., Francesco Benedetti, M.D., Jan C. Beucke, Ph.D., Joseph Biederman,, . . . Brandeis, P. D., Silvia Brem, Ph.D., Brian P. Brennan, M.D., M.M.Sc., Geraldo F. Busatto, Ph.D., Sara Calderoni. (2019). Subcortical brain volume, regional cortical thickness and cortical surface area across attention-deficit/hyperactivity disorder (ADHD), autism spectrum disorder (ASD), and obsessive-compulsive disorder (OCD). *bioRxiv*. doi:http://dx.doi.org/10.1101/673012.

Pretus, C., Ramos-Quiroga, J. A., Richarte, V., Corrales, M., Picado, M., Carmona, S., & Vilarroya, O. (2017). Time and psychostimulants: Opposing long-term structural effects in the adult ADHD brain. A longitudinal MR study. *Eur Neuropsychopharmacol*. doi:10.1016/j.euroneuro.2017.10.035

Reale, L., Bartoli, B., Cartabia, M., Zanetti, M., Costantino, M. A., Canevini, M. P., . . . Bonati, M. (2017). Comorbidity prevalence and treatment outcome in children and adolescents with ADHD. *Eur Child Adolesc Psychiatry*. doi:10.1007/s00787-017-1005-z

Renteria, M. E. (2012). Cerebral asymmetry: a quantitative, multifactorial, and plastic brain phenotype. *Twin Res Hum Genet, 15*(3), 401-413. doi:10.1017/thg.2012.13

Schrimsher, G. W., Billingsley, R. L., Jackson, E. F., & Moore, B. D., 3rd. (2002). Caudate nucleus volume asymmetry predicts attention-deficit hyperactivity disorder (ADHD) symptomatology in children. *J Child Neurol, 17*(12), 877-884. doi:10.1177/08830738020170122001

Shaw, P., Lalonde, F., Lepage, C., Rabin, C., Eckstrand, K., Sharp, W., . . . Rapoport, J. (2009). Development of cortical asymmetry in typically developing children and its disruption in attention-deficit/hyperactivity disorder. *Arch Gen Psychiatry, 66*(8), 888-896. doi:10.1001/archgenpsychiatry.2009.103

Stefanatos, G. A., & Wasserstein, J. (2001). Attention deficit/hyperactivity disorder as a right hemisphere syndrome. Selective literature review and detailed neuropsychological case studies. *Ann N Y Acad Sci, 931*, 172-195.

Stergiakouli, E., Davey Smith, G., Martin, J., Skuse, D. H., Viechtbauer, W., Ring, S. M., . . . St Pourcain, B. (2017). Shared genetic influences between dimensional ASD and ADHD symptoms during child and adolescent development. *Mol Autism, 8*, 18. doi:10.1186/s13229-017-0131-2

Toga, A. W., & Thompson, P. M. (2003). Mapping brain asymmetry. *Nat Rev Neurosci, 4*(1), 37-48. doi:10.1038/nrn1009

Uhlikova, P., Paclt, I., Vaneckova, M., Morcinek, T., Seidel, Z., Krasensky, J., & Danes, J. (2007). Asymmetry of basal ganglia in children with attention deficit hyperactivity disorder. *Neuro Endocrinol Lett, 28*(5), 604-609.

van der Meer, J. M., Oerlemans, A. M., van Steijn, D. J., Lappenschaar, M. G., de Sonneville, L. M., Buitelaar, J. K., & Rommelse, N. N. (2012). Are autism spectrum disorder and attention-deficit/hyperactivity disorder different manifestations of one overarching disorder? Cognitive and symptom evidence from a clinical and population-based sample. *J Am Acad Child Adolesc Psychiatry, 51*(11), 1160-1172.e1163. doi:10.1016/j.jaac.2012.08.024

Vance, A., Silk, T. J., Casey, M., Rinehart, N. J., Bradshaw, J. L., Bellgrove, M. A., & Cunnington, R. (2007). Right parietal dysfunction in children with attention deficit hyperactivity disorder, combined type: a functional MRI study. *Mol Psychiatry, 12*(9), 826-832, 793. doi:10.1038/sj.mp.4001999

World Health Organization. (1992). International Classification of Diseases and Related Health Problems, 10th Revision. World Health Organization: Geneva.

Wu, Z. M., Wang, P., Yang, L., Liu, L., Sun, L., An, L., . . . Wang, Y. F. (2020). Altered brain white matter microstructural asymmetry in children with ADHD. *Psychiatry Res, 285*, 112817. doi:10.1016/j.psychres.2020.112817

Yaakub, S. N., Heckemann, R. A., Keller, S. S., McGinnity, C. J., Weber, B., & Hammers, A. (2020). On brain atlas choice and automatic segmentation methods: a comparison of MAPER & FreeSurfer using three atlas databases. *Sci Rep, 10*(1), 2837. doi:10.1038/s41598-020-57951-6

Zou, H., & Yang, J. (2019). Temporal Variability-Based Functional Brain Lateralization Study in ADHD. *J Atten Disord*, 1087054719859074. doi:10.1177/1087054719859074

3.6 Supplementary Information

Table of Contents

Supplementary Methods

ENIGMA MRI quality control

Visual inspection of both internal and external Freesurfer segmentations was done per site. All sites followed the standardized ENIGMA protocols that are publicly available on http://enigma.ini.usc.edu/protocols/imaging-protocols. In short, outliers were determined by calculating the interquartile range (IQR) for each of the values per cohort and per diagnostic group (ADHD and Controls). Values that were above or below 1.5 times the IQR were identified as an outlier, and were visually inspected (3D) by the researcher. When a segmentation failure was identified, all values from the affected regions were excluded from further analyses. Additionally, cortical segmentations were overlayed on T1 images of the subjects. Webpages were generated with snapshots from internal slices, and also with external views of the segmentation from different angles. All sites were provided with the manual on how to judge these images, including the most common segmentation errors.

Cohen's *d* calculation

The t-statistic for the factor 'diagnosis' in each linear mixed effects model was derived and used to calculate Cohen's *d*, with

$$d = \frac{\mathrm{t} * (\mathrm{n1}+\mathrm{n2})}{\sqrt{(n1 * n2)} * \sqrt{df}}$$

where $n1$ and $n2$ are the number of cases and controls, and df the degrees of freedom. The latter was derived from the lme summary table in R, but can also be calculated using $df = obs - (x1 + x2)$, wherein obs equals the number of observations, $x1$ the number of groups and $x2$ the number of factors in the model.

The 95% confidence intervals for Cohen's d were calculated using $95\%\ CI = d \pm 1.96 * SE$, with the standard error (SE) around Cohen's d calculated according to:

$$SE = \sqrt{\frac{n1+n2}{n1*n2} + \frac{d^2}{2*(n1+n2-2)}}$$

Additional traits

IQ was assessed differently per dataset, but most frequently using an age-appropriate version of the Wechsler intelligence scales (**Table S1**).

Comorbidity was most often assessed by means of the Structural Clinical Interview for DSM-IV Axis I Disorders (SCID) (1), or using the Schedule for Affective Disorders and Schizophrenia for School-Age Children Present and Lifetime Version (KSADS-PL) (2) (**Table S1**).

ADHD severity was assessed based on the Conners questionnaires (3), and included hyperactivity/impulsivity and inattention scores, which were tested for associations with brain asymmetries in separate models.

The use of psychostimulant medication was recorded in two different ways: current use at the time of scanning (‘currently using stimulants’ versus ‘not currently using stimulants’), or else in terms of lifetime use (‘ever used stimulants’ versus ‘never used stimulants’).

Handedness was assessed in different ways across datasets, most often using the Edinburgh handedness questionnaire (4) (**Table S1**). Most sites provided handedness data as a binary variable, which we coded 0=left, 1=right. For the NY200 dataset we binarized continuous data from the Edinburgh inventory: left-handers < -0.25 and right-handers >0.25, with those in between assigned as missi

Table S1. Characteristics of the different datasets.

Sample name	N total	N cases (M/F)	N controls (M/F)	median age (range)	F	FS (T)	Diagnostic instrument	IQ instrument	Handedness instrument	Comorbidity instrument	Instrument for symptom rating	Scanner	Voxel size (mm)	c/r/s
ACPU[2]	67	39/0	28/0	13 (9, 18)	5.3	3	DSM-IV	WISC subtests and full	Edinburgh handedness questionnaire (Oldfield, 1971)	DISC-IV	Conners parent long version	3T Siemens TIM Trio	0.9x0.9x0.9	208/230/192
Amsterdam Neuroimage	173	68/23	54/28	17 (11, 26)	5.3	1.5	DSM-IV	Vocabulary and block design subtest of WAIS/WIC	Self-report	K-SADS-PL	Algorithm Von Rhein (von Rhein et al, ECAP 2014)	1.5T Siemens Sonata	1x1x1	256/256/176
BergenADHD	81	21/17	16/27	29 (21, 48)	5.3	3	ICD10 or DSM-IV	WASI	Self-report, drawing and writing hand	MINI Plus	ASRS	3T GE Signa	1x1x1	256/256/180
CAPSUZH scan1	57	15/6	21/15	11 (8, 18)	5.3	3	ICD10 and DSM-IV	WISC subtests block design, similarities, digit span	Edinburgh handedness questionnaire (Oldfield, 1971)	K-SADS-PL	NA	3T Philips Achieva XT	1x1x1.1	240/240/160
DATlondon	56	27/0	29/0	16 (12, 21)	5.3	3	DSM-IV	Vocabulary and block design subtests of WAIS	adapted Hollingshead	NA	NA	3T GE Signa	1.1x1.1x1.1	256/256/196
Dublin1	80	30/9	32/9	20.5 (18, 49)	5.3	3	DSM-IV	Verbal comprehension, perceptual reasoning, working memory and processing speed subtests of WAIS-IV	Edinburgh handedness questionnaire (Oldfield, 1971)	SCID-I	Conners Adult ADHD rating scale observer	3T Philips Achieva	1x1x1	256/240/NA

Table S1. Characteristics of the different datasets [continued]

Sample name	N total	N cases (M/F)	N controls (M/F)	median age (range)	F	FS (T)	Diagnostic instrument	IQ instrument	Handedness instrument	Comorbidity instrument	Instrument for symptom rating	Scanner	Voxel size (mm)	c/r/s
NICAP	146	53/12	47/34	10 (9, 11)	5.3	3	DSM-IV	WASI: vocabulary, matrix reasoning	Self-report	DISC-IV	not applicable	3T Siemens TIM Trio	0.9x0.9x0.9	256/232/176
Niche scan1	108	49/6	44/9	11 (7, 16)	5.1	1.5	DSM-IV	Vocabulary and block design WISC-III	Self-report	DISC-IV	NA	1.5T Philips Achieva	1x1x1.2	246/160-180/NA
Niche scan2	47	17/4	22/4	9 (7, 16)	5.1	1.5	DSM-IV	Vocabulary and block design WISC-III	Self-report	DISC-IV	NA	1.5T Philips Achieva	1x1x1.5	256/130-150/NA
NIH	331	111/55	110/55	11 (4, 18)	5.3	1.5	DSM-IV	Subtests of WISC	NA	DICA	DSM-IV interview	1.5T GE Signa	0.9375x0.9375x1.5	256/192/124
Nijmegen Neuroimage	158	82/38	23/15	18 (9, 24)	5.3	1.5	DSM-IV	Vocabulary and block design substest of WAIS/WIC	Self-report	K-SADS-PL	Algorithm Von Rhein (von Rhein et al, ECAP 2014)	1.5T Siemens Avanto	1x1x1	256/256/176
NYU	80	22/18	22/18	30 (18, 53)	5.3	3	DSM-IV	WASI	Edinburgh handedness questionnaire (Oldfield, 1971)	SCID-I	CAARS	3T Siemens Allegra	1x1x1.3	256/256/128
NYUADHD200[1]	228	94/35	48/51	11 (7, 18)	5.3	3	DSM-IV	WASI	Edinburgh handedness questionnaire (Oldfield, 1971)	KSADS	KSADS-PL	3T Siemens Allegra	1.3x1.0x1.3	256/256/128

OHSUADHD200[1]	89	19/7	28/35	9 (7, 13)	5.3	3	DSM-IV	Block Design, Vocabulary and Information subtests of WISC-IV	NA	KSADS	Parent/Teacher Conners rating scale 3rd edition, Parent Teacher ADHD Rating Scale K-SADS	3T Siemens TIM Trio	1.0 x 1.0 x 1.1	256/240/160
OHSU	229	81/39	59/50	9 (7, 13)	5.3	3	DSM-IV and DSM-V	WISC subtests: block design, vocabulary, and information	Edinburgh handedness questionnaire (modified version)	KSADS-PL	KSADS-PL	3T Siemens TIM Trio	1x1x1.1	256/256/160
Olin Neuropsychiatry Research Centre[2]	181	59/14	58/50	15 (12, 19)	5.3	3 T	DSM-IV	WASI Full Scale	Edinburgh handedness questionnaire (Oldfield, 1971)	KSADS-PL	KSADS-PL	3T Siemens Allegra	1x1x1	208/256/176
PekingADHD200 scan1[1]	32	14/0	18/0	13.5 (12, 16)	5.3	3 T	DSM-IV	WISCC-R	Not available	NA	ADHD rating scale	3T Siemens TIM Trio	1.3×1.0×1.3	256/256/128
PekingADHD200 scan2[1]	64	32/0	31/1	12 (9, 16)	5.3	3 T	DSM-IV	WISCC-R	Not available	NA	ADHD rating scale	3T Siemens TIM Trio	1x1x1	256/256/176
Rubia ADHD	71	41/0	30/0	14 (10, 18)	5.3	3	DSM-IV	WASI	Edinburgh handedness questionnaire (Oldfield, 1971)	Co-morbid disorders were exclusion criteria	SDQ for Hyperactive Impulsive symptoms and Conners parent Rater scale revised for Inattentive symptoms	3T GE Signa	1.0156x1.0156x1.2	256/256/166
SãoPaulo1 – Estado	147	57/24	44/22	27 (17, 50)	5.3	3	DSM-IV	WASI	Edinburgh handedness questionnaire (Oldfield, 1971)	SCID	NA	1.5T Siemens Espree	1.36x1.36x1.2	192/192/160

Table S1. Characteristics of the different datasets [continued]

Sample name	N total	N cases (M/F)	N controls (M/F)	median age (range)	F	FS (T)	Diagnostic instrument	IQ instrument	Handedness instrument	Comorbidity instrument	Instrument for symptom rating	Scanner	Voxel size (mm)	c/r/s
Sussex[2S]	60	19/11	19/11	31 (19, 59)	5.3	1.5	DSM-IV	NART	NA	NA	NA	1.5T Siemens Avanto	1x1x1	256/240/192
SVG Bergen	51	19/4	20/8	10 (8, 12)	5.3	3	DSM-IV	WISC-IV	Self-report, writing hand	K-SADS-PL	K-SADS PL	3T GE Signa	1x1x1	256/256/180
UAB scan1	125	63/6	46/10	15 (6,51)	5.3	3	DSM-IV	WISC	NA	SCID-I	CAARS	3T Philips Achieva	0.94x0.94x1	256/256/180
UAB scan2	73	19/15	18/21	33 (2,52)		1.5	DSM-IV	WISC	NA	SCID-I	CAARS	1.5 T GE Signa	0.86x0.86x1.6	256/256/96
UCHZ	78	20/19	21/18	14.5 (9, 61)	5.3	3	DSM-IV	HAWIK	Edinburgh handedness questionnaire (Oldfield, 1971)	K-SADS-PL	Adults; adult conners; children conners -3d	3T GE Signa	1x1x1	192/256/172
UKA scan4	59	21/7	13/18	9 (4, 15)	5.3	3	ICD10 or DSM-IV	CPM (N=30) or WASI (N=49) or WISC-IV (N=14)	NA	K-SADS and German K-Dips	German Parental and Teacher Report on ADHD	3T Siemens Trio	0.98x0.98x1	256/256/176
Wurzburg ADHD	107	30/25	24/28	43 (18, 62)	5.3	1.5	DSM-IV	MWT-B	Self-report	SCID-I	DMS-IV interview	3T Siemens Avanto	1 × 1 × 1	256/256/160
ZiCAPS	34	17/4	7/6	13 (9, 15)	5.3	3	DSM-IV	Subscales of WICS-IV	Edinburgh handedness questionnaire (Oldfield, 1971)	ODD and CD with structured clinical interview	KSADS-PL; self-ratings: Conners-3	3T Siemens TIM Trio	1x1x1	256/256/192
Total	**3762**	**1933**	**1829**											

[1] Included in the mega-analysis analysis by Douglas *et al.* (2018), mentioned in the introduction. [2] Only cortical data available for these datasets. F=FreeSurfer version; FS= Field Strength. c/s/r=cols/rows/slices.

Table S2. Full linear model results for the subcortical volume AIs in children.

AI region		beta-coefficient			Standard Error			t-value			p-value			Cohen's d (95% CI)
	N cases/ controls	diag	sex	age	diag	sex	age	diag	sex	age	diag	sex	age	diag
Accumbens	801/842	-0.0041	-0.0003	-0.0010	0.0036	0.0041	0.0010	-1.13	-0.07	-1.01	0.260	0.942	0.312	-0.06 (-0.15,0.04)
Amygdala	800/842	-0.0007	0.0042	0.0014	0.0027	0.0030	0.0007	-0.28	1.39	1.85	0.782	0.164	0.064	-0.01 (-0.11,0.08)
Caudate Nucleus	802/842	0.0008	-0.0029	0.0003	0.0015	0.0016	0.0004	0.52	-1.79	0.82	0.600	0.074	0.411	0.03 (-0.07,0.12)
Globus Pallidus	802/840	-0.0012	-0.0002	-0.0017	0.0026	0.0029	0.0007	-0.45	-0.05	-2.31	0.655	0.959	0.021	-0.02 (-0.12,0.07)
Hippocampus	800/840	-0.0004	0.0002	0.0010	0.0021	0.0023	0.0006	-0.21	0.10	1.87	0.837	0.924	0.062	-0.01 (-0.11,0.09)
Putamen	802/839	-0.0008	0.0024	0.0010	0.0014	0.0015	0.0004	-0.61	1.57	2.65	0.540	0.117	**0.008**	-0.03 (-0.13,0.07)
Thalamus[1]	706/733	0.0012	0.0011	0.0037	0.0014	0.0016	0.0004	0.81	0.68	8.60	0.418	0.494	**2.02×10^{-17}**	0.04 (-0.06,0.15)

P-values in **bold** are significant at the uncorrected level ($P < 0.05$ and in ***bold-italic*** those that survive multiple testing correction within the particular analysis indicated (see text).[1] Thalamus volume was not available from the NIH dataset.

Table S3. Full linear model results for the cortical surface area AIs in children

Cortical surface area AI		beta-coefficient			Standard Error			t-value			p-value			Cohen's d (95% CI)
	N cases / controls	diag	sex	age	diag	sex	age	diag	sex	age	diag	sex	age	diag
bankssts	850/891	-0.0010	0.0017	0.0031	0.0041	0.0046	0.0011	-0.25	0.37	2.81	0.800	0.712	**0.005**	-0.01 (-0.11,0.08)
caudalanteriorcingulate	910/943	-0.0018	0.0065	-0.0006	0.0055	0.0061	0.0014	-0.32	1.06	-0.46	0.750	0.291	0.648	-0.01 (-0.11,0.08)
caudalmiddlefrontal	912/948	0.0030	-0.0036	-0.0009	0.0036	0.0040	0.0010	0.82	-0.90	-0.91	0.410	0.369	0.364	0.04 (-0.05,0.13)
cuneus	906/948	0.0045	0.0033	0.0018	0.0032	0.0036	0.0009	1.39	0.91	2.04	0.163	0.365	0.041	0.07 (-0.03,0.16)
entorhinal	869/915	0.0003	-0.0063	0.0036	0.0058	0.0065	0.0015	0.06	-0.96	2.40	0.953	0.336	**0.017**	0 (-0.09,0.1)
frontalpole	912/949	-0.0084	0.0053	-0.0012	0.0043	0.0048	0.0010	-1.94	1.10	-1.16	0.052	0.274	0.245	-0.09 (-0.18,0)
fusiform	906/945	-0.0034	0.0021	0.0003	0.0025	0.0027	0.0006	-1.36	0.75	0.50	0.173	0.455	0.616	-0.06 (-0.15,0.03)
inferiorparietal	909/943	0.0029	0.0133	-0.0003	0.0026	0.0029	0.0007	1.11	4.59	-0.39	0.265	$\mathbf{4.6\times10^{-6}}$	0.697	0.05 (-0.04,0.14)
inferiortemporal	896/943	0.0015	-0.0006	-0.0004	0.0027	0.0030	0.0007	0.57	-0.22	-0.61	0.568	0.828	0.543	0.03 (-0.06,0.12)
insula	909/944	0.0038	0.0028	-0.0015	0.0023	0.0025	0.0006	1.66	1.10	-2.39	0.096	0.272	**0.017**	0.08 (-0.01,0.17)
isthmuscingulate	910/942	-0.0002	-0.0088	-0.0005	0.0036	0.0040	0.0009	-0.06	-2.18	-0.51	0.949	0.029	0.612	0 (-0.09,0.09)
lateraloccipital	911/949	-0.0013	-0.0020	0.0001	0.0024	0.0027	0.0006	-0.53	-0.77	0.24	0.594	0.440	0.810	-0.02 (-0.12,0.07)
lateralorbitofrontal	912/949	-0.0027	0.0050	-0.0002	0.0020	0.0023	0.0006	-1.35	2.19	-0.33	0.176	0.029	0.745	-0.06 (-0.15,0.03)
lingual	912/949	-0.0004	0.0039	-0.0004	0.0024	0.0027	0.0006	-0.15	1.41	-0.68	0.879	0.159	0.499	-0.01 (-0.1,0.08)
medialorbitofrontal	905/941	0.0076	-0.0040	-0.0004	0.0028	0.0031	0.0008	2.74	-1.26	-0.52	**0.006**	0.208	0.605	0.13 (0.04,0.22)
middletemporal	865/913	0.0032	0.0020	0.0010	0.0022	0.0025	0.0006	1.43	0.80	1.62	0.153	0.425	0.106	0.07 (-0.02,0.16)

Table S3. Full linear model results for the cortical surface area AIs in children [continued]

Cortical surface area AI		beta-coefficient			Standard Error			t-value			p-value			Cohen's d (95% CI)
	N cases / controls	diag	sex	age	diag	sex	age	diag	sex	age	diag	sex	age	diag
paracentral	912/949	-0.0067	0.0028	0.0005	0.0031	0.0034	0.0007	-2.16	0.81	0.70	**0.031**	0.417	0.485	-0.1 (-0.19,-0.01)
parahippocampal	906/942	0.0032	0.0123	0.0008	0.0036	0.0040	0.0009	0.89	3.03	0.82	0.373	**0.002**	0.412	0.04 (-0.05,0.13)
parsopercularis	906/946	0.0006	0.0042	0.0003	0.0039	0.0043	0.0009	0.15	0.98	0.31	0.880	0.330	0.759	0.01 (-0.08,0.1)
parsorbitalis	912/948	0.0037	0.0032	-0.0012	0.0029	0.0032	0.0007	1.29	1.00	-1.66	0.198	0.316	0.097	0.06 (-0.03,0.15)
parstriangularis	905/950	0.0035	0.0006	-0.0009	0.0035	0.0039	0.0009	1.00	0.14	-1.04	0.319	0.889	0.300	0.05 (-0.04,0.14)
pericalcarine	910/948	0.0028	-0.0059	-0.0008	0.0027	0.0031	0.0007	1.03	-1.91	-1.15	0.301	0.056	0.252	0.05 (-0.04,0.14)
postcentral	898/934	0.0016	-0.0045	-0.0003	0.0021	0.0024	0.0006	0.77	-1.88	-0.45	0.440	0.060	0.655	0.04 (-0.06,0.13)
posteriorcingulate	909/944	-0.0018	0.0050	0.0006	0.0036	0.0040	0.0009	-0.50	1.26	0.64	0.618	0.209	0.521	-0.02 (-0.11,0.07)
precentral	901/943	0.0003	-0.0011	-0.0006	0.0019	0.0021	0.0005	0.18	-0.53	-1.13	0.855	0.594	0.257	0.01 (-0.08,0.1)
precuneus	912/946	0.0021	0.0052	0.0006	0.0019	0.0022	0.0005	1.06	2.38	1.23	0.289	**0.017**	0.220	0.05 (-0.04,0.14)
rostralanteriorcingulate	902/943	-0.0002	0.0052	-0.0015	0.0047	0.0053	0.0012	-0.04	0.98	-1.20	0.971	0.326	0.231	0 (-0.09,0.09)
rostralmiddlefrontal	911/946	-0.0032	-0.0004	0.0006	0.0019	0.0021	0.0005	-1.63	-0.17	1.33	0.103	0.866	0.184	-0.08 (-0.17,0.01)
superiorparietal	909/947	0.0034	0.0023	-0.0004	0.0020	0.0023	0.0005	1.67	1.02	-0.81	0.094	0.306	0.417	0.08 (-0.01,0.17)
superiortemporal	839/904	0.0033	-0.0103	0.0001	0.0019	0.0021	0.0005	1.70	-4.83	0.22	0.090	**1.5×10-6**	0.829	0.08 (-0.01,0.18)
supramarginal	897/938	0.0005	-0.0068	0.0007	0.0030	0.0034	0.0008	0.18	-2.00	0.91	0.855	**0.046**	0.361	0.01 (-0.08,0.1)
temporalpole	906/945	0.0018	-0.0099	0.0011	0.0040	0.0044	0.0010	0.45	-2.26	1.09	0.651	**0.024**	0.276	0.02 (-0.07,0.11)

temporalpole	906/945	0.0018	-0.0099	0.0011	0.0040	0.0044	0.0010	0.45	-2.26	1.09	0.651	**0.024**	0.276	0.02 (-0.07,0.11)
transversetemporal	909/948	-0.0017	0.0032	0.0001	0.0037	0.0042	0.0010	-0.45	0.77	0.12	0.656	0.441	0.903	-0.02 (-0.11,0.07)
Total surface area	912/950	0.0008	0.0004	0.0001	0.0004	0.0004	0.0001	2.10	0.85	0.90	0.036	0.395	0.366	0.1 (0.01,0.19)

P- values in **bold** are significant at the uncorrected level ($P < 0.05$), and in ***bold-italic*** those that survive multiple testing correction within the particular analysis indicated (see text).

Table S4. Full linear model results for the cortical thickness AIs in children

Cortical surface area AI		beta-coefficient			Standard Error			t-value			p-value			Cohen's d (95% CI)
	N cases/ controls	diag	sex	age	diag	sex	age	diag	sex	age	diag	sex	age	diag
bankssts	851/891	-0.0037	0.0021	-0.0003	0.0019	0.0021	0.0005	-1.99	1.03	-0.65	**0.047**	0.305	0.514	-0.1 (-0.19,0)
caudalanteriorcingulate	910/943	0.0029	-0.0008	-0.0003	0.0025	0.0028	0.0007	1.16	-0.29	-0.44	0.245	0.775	0.662	0.05 (-0.04,0.15)
caudalmiddlefrontal	911/949	0.0025	-0.0019	-0.0007	0.0012	0.0013	0.0003	2.09	-1.42	-2.09	**0.037**	0.155	0.037	0.1 (0.01,0.19)
cuneus	907/949	0.0007	0.0016	-0.0003	0.0016	0.0018	0.0004	0.40	0.88	-0.62	0.689	0.378	0.538	0.02 (-0.07,0.11)
entorhinal	869/916	-0.0046	-0.0086	-0.0008	0.0029	0.0032	0.0007	-1.58	-2.66	-1.16	0.115	**0.008**	0.246	-0.08 (-0.17,0.02)
frontalpole	911/949	0.0035	-0.0002	-0.0003	0.0032	0.0036	0.0008	1.11	-0.07	-0.39	0.267	0.947	0.694	0.05 (-0.04,0.14)
fusiform	908/946	-0.0006	0.0003	-0.0004	0.0010	0.0011	0.0003	-0.59	0.23	-1.46	0.555	0.817	0.143	-0.03 (-0.12,0.06)
inferiorparietal	910/945	0.0001	-0.0012	-0.0007	0.0009	0.0011	0.0003	0.06	-1.12	-2.72	0.955	0.264	**0.007**	0 (-0.09,0.09)
inferiortemporal	896/942	-0.0015	0.0015	0.0005	0.0013	0.0015	0.0004	-1.17	1.04	1.48	0.243	0.297	0.139	-0.05 (-0.15,0.04)
insula	910/945	-0.0025	0.0006	0.0000	0.0013	0.0014	0.0003	-1.98	0.45	0.11	**0.047**	0.652	0.913	-0.09 (-0.18,0)
isthmuscingulate	909/943	-0.0005	0.0017	0.0011	0.0019	0.0021	0.0005	-0.24	0.79	2.42	0.810	0.430	**0.016**	-0.01 (-0.1,0.08)
lateraloccipital	911/950	0.0003	-0.0003	0.0003	0.0011	0.0012	0.0003	0.31	-0.25	1.14	0.760	0.800	0.254	0.01 (-0.08,0.11)
lateralorbitofrontal	912/949	-0.0004	-0.0033	-0.0002	0.0014	0.0016	0.0004	-0.32	-2.09	-0.45	0.750	**0.037**	0.651	-0.01 (-0.11,0.08)
lingual	912/948	-0.0012	-0.0004	0.0009	0.0012	0.0014	0.0003	-0.96	-0.32	2.71	0.339	0.751	0.007	-0.04 (-0.14,0.05)
medialorbitofrontal	905/942	-0.0034	-0.0060	-0.0003	0.0018	0.0020	0.0005	-1.88	-2.98	-0.69	0.060	0.003	0.488	-0.09 (-0.18,0)
middletemporal	866/913	-0.0004	0.0009	0.0005	0.0012	0.0014	0.0003	-0.32	0.62	1.43	0.747	0.535	0.154	-0.02 (-0.11,0.08)
paracentral	912/949	-0.0018	-0.0005	-0.0008	0.0012	0.0014	0.0003	-1.45	-0.37	-2.56	0.148	0.711	**0.010**	-0.07 (-0.16,0.02)

parahippocampal	907/943	-0.0039	-0.0047	0.0009	0.0021	0.0024	0.0005	-1.84	-1.98	1.70	0.066	0.048	0.089	-0.09 (-0.18,0)
parsopercularis	906/946	0.0003	0.0018	-0.0008	0.0013	0.0015	0.0004	0.26	1.23	-2.12	0.798	0.221	0.034	0.01 (-0.08,0.1)
parsorbitalis	912/948	0.0020	0.0012	0.0003	0.0022	0.0025	0.0006	0.91	0.47	0.46	0.364	0.635	0.647	0.04 (-0.05,0.13)
parstriangularis	905/950	-0.0006	-0.0011	-0.0001	0.0015	0.0017	0.0004	-0.42	-0.65	-0.32	0.674	0.516	0.752	-0.02 (-0.11,0.07)
pericalcarine	908/947	-0.0002	0.0009	-0.0008	0.0019	0.0022	0.0005	-0.09	0.41	-1.60	0.925	0.681	0.111	-0.004 (-0.1,0.09)
postcentral	897/936	-0.0001	-0.0019	0.0000	0.0011	0.0012	0.0003	-0.08	-1.66	-0.13	0.937	0.097	0.896	-0.004 (-0.1,0.09)
posteriorcingulate	908/947	-0.0009	-0.0010	0.0003	0.0016	0.0017	0.0004	-0.57	-0.57	0.73	0.569	0.570	0.463	-0.03 (-0.12,0.06)
precentral	901/942	0.0021	-0.0007	-0.0001	0.0009	0.0010	0.0002	2.35	-0.66	-0.32	0.019	0.511	0.753	0.11 (0.02,0.2)
precuneus	912/946	0.0003	-0.0001	-0.0001	0.0009	0.0010	0.0003	0.35	-0.13	-0.45	0.728	0.893	0.652	0.02 (-0.07,0.11)
rostralanteriorcingulate	902/941	-0.0002	0.0029	-0.0003	0.0022	0.0025	0.0006	-0.09	1.16	-0.54	0.925	0.247	0.588	-0.004 (-0.1,0.09)
rostralmiddlefrontal	911/947	0.0005	-0.0003	-0.0006	0.0011	0.0013	0.0003	0.42	-0.26	-2.00	0.676	0.797	0.046	0.02 (-0.07,0.11)
superiorfrontal	909/946	0.0002	-0.0010	0.0002	0.0008	0.0009	0.0002	0.30	-1.16	0.76	0.767	0.246	0.450	0.01 (-0.08,0.1)
superiorparietal	909/947	0.0000	-0.0008	-0.0002	0.0009	0.0010	0.0002	-0.02	-0.81	-0.90	0.982	0.418	0.371	-0.001 (-0.09,0.09)
superiortemporal	840/907	0.0021	0.0026	0.0009	0.0012	0.0013	0.0003	1.85	2.02	2.87	0.065	0.044	0.004	0.09 (0,0.18)
supramarginal	897/941	-0.0014	0.0004	-0.0004	0.0011	0.0012	0.0003	-1.33	0.32	-1.50	0.183	0.748	0.134	-0.06 (-0.15,0.03)
temporalpole	906/944	0.0017	0.0001	0.0000	0.0029	0.0032	0.0008	0.59	0.03	0.04	0.555	0.973	0.968	0.03 (-0.06,0.12)
transversetemporal	909/948	0.0010	0.0039	0.0006	0.0023	0.0025	0.0006	0.44	1.56	1.01	0.664	0.120	0.313	0.02 (-0.07,0.11)
total thickness	912/950	0.0000	-0.0003	0.0000	0.0004	0.0004	0.0001	-0.10	-0.79	0.03	0.918	0.430	0.980	-0.005 (-0.1,0.09)

P-values in **bold** are significant at the uncorrected level (P < 0.05), and in ***bold-italic*** those that survive multiple testing correction within the particular analysis indicated (see text).

Table S5. Full linear model results for the subcortical volume AIs in adolescents.

AI region	N cases/ controls	beta-coefficient			Standard Error			t-value			p-value			Cohen's d (95% CI)
		diag	sex	age	diag	sex	age	diag	sex	age	diag	sex	age	diag
Accumbens	314/214	-0.0060	0.0080	0.0013	0.0065	0.0077	0.0020	-0.91	1.03	0.67	0.361	0.303	0.501	-0.08 (-0.26,0.09)
Amygdala	314/213	0.0016	0.0051	-0.0016	0.0045	0.0052	0.0013	0.35	0.98	-1.27	0.724	0.328	0.204	0.03 (-0.14,0.21)
Caudate Nucleus	313/214	0.0004	-0.0041	0.0011	0.0025	0.0030	0.0008	0.15	-1.35	1.36	0.877	0.178	0.173	0.01 (-0.16,0.19)
Globus Pallidus	314/214	-0.0039	-0.0019	-0.0022	0.0046	0.0054	0.0014	-0.86	-0.36	-1.54	0.391	0.723	0.125	-0.08 (-0.25,0.1)
Hippocampus	314/214	0.0058	0.0006	-0.0003	0.0034	0.0039	0.0009	1.68	0.15	-0.31	0.093	0.879	0.758	0.15 (-0.02,0.33)
Putamen	313/212	-0.0004	0.0012	0.0019	0.0027	0.0031	0.0008	-0.17	0.38	2.45	0.867	0.707	0.015	-0.02 (-0.19,0.16)
Thalamus[1]	293/207	0.0029	-0.0021	0.0005	0.0027	0.0032	0.0008	1.07	-0.64	0.65	0.285	0.521	0.515	0.1 (-0.08,0.28)

P-values in **bold** are significant at the uncorrected level ($P < 0.05$ and in ***bold-italic*** those that survive multiple testing correction within the particular analysis indicated (see text). [1]Thalamus volume was not available from the NIH dataset.

Table S6. Full linear model results for the cortical surface area AIs in adolescents.

Cortical surface area AI		beta-coefficient			Standard Error			t-value			p-value			Cohen's d (95% CI)
	N cases / controls	diag	sex	age	diag	sex	age	diag	sex	age	diag	sex	age	diag
bankssts	368/301	-0.0042	0.0002	0.0009	0.0067	0.0074	0.0018	-0.62	0.03	0.47	0.533	0.980	0.637	-0.05 (-0.2,0.1)
caudalanteriorcingulate	396/322	-0.0094	0.0041	0.0018	0.0088	0.0098	0.0023	-1.07	0.42	0.80	0.286	0.674	0.424	-0.08 (-0.23,0.07)
caudalmiddlefrontal	396/320	-0.0033	-0.0013	0.0015	0.0055	0.0061	0.0014	-0.60	-0.20	1.01	0.547	0.838	0.314	-0.05 (-0.19,0.1)
cuneus	396/322	-0.0005	0.0026	-0.0002	0.0047	0.0052	0.0012	-0.10	0.50	-0.17	0.921	0.620	0.866	-0.01 (-0.15,0.14)
entorhinal	386/318	-0.0068	-0.0161	0.0015	0.0084	0.0093	0.0022	-0.81	-1.72	0.70	0.417	0.085	0.486	-0.06 (-0.21,0.09)
frontalpole	396/322	0.0084	0.0041	0.0039	0.0069	0.0076	0.0019	1.23	0.54	2.04	0.220	0.590	**0.041**	0.09 (-0.05,0.24)
fusiform	392/320	0.0035	0.0063	-0.0010	0.0037	0.0041	0.0010	0.94	1.56	-1.04	0.346	0.120	0.299	0.07 (-0.08,0.22)
inferiorparietal	394/321	-0.0001	0.0006	0.0005	0.0038	0.0042	0.0010	-0.03	0.14	0.55	0.978	0.889	0.584	0 (-0.15,0.15)
inferiortemporal	372/311	0.0008	-0.0118	0.0011	0.0041	0.0046	0.0011	0.20	-2.54	0.94	0.842	**0.011**	0.349	0.02 (-0.13,0.17)
insula	393/319	0.0020	0.0050	-0.0012	0.0033	0.0038	0.0010	0.59	1.32	-1.16	0.558	0.187	0.247	0.04 (-0.1,0.19)
isthmuscingulate	396/322	-0.0072	-0.0126	0.0021	0.0055	0.0061	0.0015	-1.32	-2.06	1.42	0.188	**0.040**	0.156	-0.1 (-0.25,0.05)
lateraloccipital	396/322	-0.0002	0.0094	-0.0011	0.0035	0.0038	0.0009	-0.05	2.46	-1.28	0.961	**0.014**	0.200	0 (-0.15,0.14)
lateralorbitofrontal	395/322	-0.0019	0.0020	0.0004	0.0030	0.0034	0.0009	-0.61	0.58	0.43	0.540	0.561	0.668	-0.05 (-0.19,0.1)
lingual	393/320	-0.0057	-0.0031	0.0004	0.0039	0.0044	0.0012	-1.46	-0.70	0.38	0.144	0.482	0.706	-0.11 (-0.26,0.04)

Table S6. Full linear model results for the cortical surface area AIs in adolescents [continued]

Cortical surface area AI		beta-coefficient			Standard Error			t-value			p-value			Cohen's d (95% CI)
	N cases / controls	diag	sex	age	diag	sex	age	diag	sex	age	diag	sex	age	diag
medialorbitofrontal	394/321	0.0046	-0.0047	-0.0007	0.0042	0.0047	0.0012	1.10	-0.98	-0.60	0.272	0.326	0.546	0.08 (-0.06,0.23)
middletemporal	355/294	-0.0027	-0.0019	0.0007	0.0036	0.0040	0.0009	-0.76	-0.48	0.76	0.446	0.631	0.445	-0.06 (-0.22,0.09)
paracentral	395/322	-0.0002	0.0086	0.0003	0.0047	0.0052	0.0012	-0.05	1.66	0.21	0.960	0.098	0.833	0 (-0.15,0.14)
parahippocampal	393/320	-0.0060	0.0125	-0.0035	0.0052	0.0059	0.0015	-1.15	2.13	-2.30	0.252	**0.034**	**0.022**	-0.09 (-0.24,0.06)
parsopercularis	393/321	0.0082	0.0028	-0.0018	0.0062	0.0069	0.0017	1.33	0.41	-1.07	0.185	0.684	0.285	0.1 (-0.05,0.25)
parsorbitalis	396/321	0.0113	0.0031	-0.0014	0.0047	0.0052	0.0012	2.39	0.58	-1.14	0.017	0.560	0.256	0.18 (0.04,0.33)
parstriangularis	396/322	0.0082	-0.0058	-0.0006	0.0056	0.0062	0.0014	1.46	-0.94	-0.44	0.145	0.348	0.660	0.11 (-0.04,0.26)
pericalcarine	395/322	-0.0064	-0.0057	-0.0007	0.0043	0.0047	0.0012	-1.51	-1.20	-0.56	0.131	0.232	0.578	-0.12 (-0.26,0.03)
postcentral	392/320	0.0035	-0.0019	0.0001	0.0033	0.0037	0.0009	1.06	-0.52	0.10	0.289	0.606	0.918	0.08 (-0.07,0.23)
posteriorcingulate	396/322	-0.0044	0.0042	-0.0001	0.0059	0.0065	0.0015	-0.75	0.64	-0.09	0.456	0.525	0.929	-0.06 (-0.2,0.09)
precentral	391/320	-0.0051	0.0017	-0.0007	0.0030	0.0033	0.0009	-1.71	0.51	-0.76	0.088	0.612	0.445	-0.13 (-0.28,0.02)
precuneus	395/322	-0.0021	0.0090	0.0004	0.0029	0.0032	0.0008	-0.73	2.78	0.50	0.466	0.006	0.617	-0.06 (-0.2,0.09)
rostralanteriorcingulate	394/321	0.0002	-0.0010	0.0010	0.0073	0.0081	0.0019	0.03	-0.12	0.52	0.977	0.906	0.603	0 (-0.15,0.15)
rostralmiddlefrontal	396/322	-0.0009	-0.0018	-0.0008	0.0031	0.0034	0.0008	-0.29	-0.51	-0.89	0.769	0.608	0.376	-0.02 (-0.17,0.12)
superiorfrontal	395/322	0.0040	0.0043	-0.0008	0.0024	0.0026	0.0006	1.70	1.64	-1.35	0.089	0.100	0.179	0.13 (-0.02,0.28)

superiorparietal	394/322	0.0031	0.0057	-0.0008	0.0032	0.0036	0.0009	0.97	1.59	-0.89	0.331	0.113	0.373	0.07 (-0.07,0.22)
superiortemporal	346/297	0.0005	-0.0135	0.0005	0.0030	0.0033	0.0008	0.16	-4.05	0.63	0.872	0.000	0.530	0.01 (-0.14,0.17)
supramarginal	393/321	-0.0050	-0.0085	0.0009	0.0043	0.0047	0.0011	-1.16	-1.79	0.82	0.245	0.073	0.415	-0.09 (-0.24,0.06)
temporalpole	393/320	0.0025	-0.0027	-0.0005	0.0062	0.0068	0.0016	0.40	-0.39	-0.34	0.689	0.695	0.731	0.03 (-0.12,0.18)
transversetemporal	393/320	0.0046	-0.0010	0.0020	0.0060	0.0068	0.0018	0.77	-0.14	1.09	0.444	0.886	0.275	0.06 (-0.09,0.21)
Total surface area	396/322	0.0002	0.0007	-0.0001	0.0005	0.0006	0.0002	0.35	1.16	-0.83	0.730	0.247	0.405	0.03 (-0.12,0.17)

P-values in **bold** are significant at the uncorrected level (P < 0.05), and in ***bold-italic*** those that survive multiple testing correction within the particular analysis indicated (see text).

Table S7. Full linear model results for the cortical thickness AIs in adolescents.

Cortical thickness AI		beta-coefficient			Standard Error			t-value			p-value			Cohen's d (95% CI)
	N cases / controls	diag	sex	age	diag	sex	age	diag	sex	age	diag	sex	age	diag
bankssts	367/301	-0.0019	0.0042	-0.0012	0.0031	0.0035	0.0009	-0.62	1.22	-1.36	0.539	0.223	0.175	-0.05 (-0.2,0.1)
caudalanteriorcingulate	395/322	-0.0022	-0.0041	0.0003	0.0042	0.0047	0.0013	-0.53	-0.86	0.20	0.597	0.388	0.843	-0.04 (-0.19,0.11)
caudalmiddlefrontal	395/320	0.0031	-0.0014	-0.0009	0.0018	0.0020	0.0005	1.70	-0.70	-2.00	0.089	0.484	**0.046**	0.13 (-0.02,0.28)
cuneus	395/322	-0.0049	-0.0032	-0.0002	0.0024	0.0027	0.0007	-2.03	-1.19	-0.26	**0.043**	0.235	0.795	-0.15 (-0.3,-0.01)
entorhinal	386/318	0.0012	-0.0093	-0.0012	0.0044	0.0050	0.0013	0.27	-1.87	-0.93	0.788	0.062	0.352	0.02 (-0.13,0.17)
frontalpole	395/322	-0.0061	-0.0036	-0.0005	0.0048	0.0053	0.0012	-1.29	-0.68	-0.40	0.197	0.495	0.687	-0.1 (-0.25,0.05)
fusiform	391/320	0.0000	-0.0014	-0.0001	0.0017	0.0020	0.0005	0.02	-0.74	-0.24	0.984	0.460	0.809	0 (-0.15,0.15)
inferiorparietal	393/321	-0.0008	-0.0007	-0.0002	0.0015	0.0017	0.0005	-0.53	-0.41	-0.43	0.594	0.680	0.665	-0.04 (-0.19,0.11)
inferiortemporal	371/311	0.0006	-0.0045	0.0005	0.0022	0.0026	0.0007	0.27	-1.75	0.77	0.786	0.080	0.441	0.02 (-0.13,0.17)
insula	392/319	-0.0018	0.0003	-0.0006	0.0018	0.0020	0.0005	-1.00	0.17	-1.25	0.318	0.864	0.211	-0.08 (-0.22,0.07)
isthmuscingulate	394/321	0.0035	-0.0029	0.0009	0.0029	0.0032	0.0007	1.22	-0.92	1.20	0.222	0.355	0.231	0.09 (-0.05,0.24)
lateraloccipital	395/322	-0.0014	0.0005	0.0005	0.0017	0.0019	0.0005	-0.84	0.26	1.00	0.401	0.796	0.318	-0.06 (-0.21,0.08)
lateralorbitofrontal	394/322	0.0014	-0.0033	-0.0007	0.0021	0.0024	0.0007	0.67	-1.38	-1.08	0.506	0.167	0.280	0.05 (-0.1,0.2)
lingual	392/320	0.0004	-0.0031	0.0001	0.0020	0.0022	0.0005	0.19	-1.39	0.19	0.848	0.163	0.846	0.01 (-0.13,0.16)

medialorbitofrontal	394/321	0.0028	-0.0011	0.0005	0.0027	0.0031	0.0008	1.02	-0.37	0.60	0.306	0.710	0.550	0.08 (-0.07,0.23)
middletemporal	354/294	-0.0010	-0.0049	0.0001	0.0020	0.0023	0.0006	-0.50	-2.13	0.09	0.620	0.034	0.925	-0.04 (-0.19,0.11)
paracentral	394/322	0.0030	0.0036	0.0010	0.0020	0.0022	0.0006	1.55	1.63	1.69	0.121	0.103	0.091	0.12 (-0.03,0.27)
parahippocampal	392/320	-0.0062	-0.0060	-0.0015	0.0037	0.0041	0.0010	-1.68	-1.47	-1.57	0.094	0.143	0.117	-0.13 (-0.28,0.02)
parsopercularis	392/321	0.0018	-0.0025	-0.0008	0.0021	0.0023	0.0005	0.85	-1.06	-1.45	0.395	0.291	0.148	0.07 (-0.08,0.21)
parsorbitalis	395/321	-0.0002	0.0020	-0.0001	0.0040	0.0045	0.0013	-0.06	0.44	-0.08	0.954	0.662	0.935	0 (-0.15,0.14)
parstriangularis	394/322	0.0021	0.0011	-0.0004	0.0023	0.0026	0.0007	0.91	0.43	-0.62	0.364	0.669	0.538	0.07 (-0.08,0.22)
pericalcarine	394/322	-0.0001	-0.0028	-0.0003	0.0030	0.0033	0.0009	-0.02	-0.82	-0.39	0.983	0.410	0.695	0 (-0.15,0.15)
postcentral	390/320	-0.0002	-0.0020	-0.0003	0.0017	0.0019	0.0005	-0.10	-1.08	-0.69	0.923	0.280	0.492	-0.01 (-0.16,0.14)
posteriorcingulate	395/322	-0.0017	-0.0006	0.0004	0.0024	0.0027	0.0007	-0.72	-0.21	0.49	0.474	0.837	0.621	-0.05 (-0.2,0.09)
precentral	389/319	-0.0014	-0.0027	-0.0008	0.0014	0.0016	0.0004	-0.99	-1.71	-2.10	0.320	0.088	**0.036**	-0.08 (-0.22,0.07)
precuneus	394/322	0.0019	0.0003	0.0005	0.0015	0.0017	0.0005	1.22	0.20	1.03	0.223	0.841	0.303	0.09 (-0.05,0.24)
rostralanteriorcingulate	393/321	0.0071	-0.0026	0.0004	0.0037	0.0042	0.0011	1.90	-0.61	0.36	0.058	0.544	0.721	0.15 (0,0.29)
rostralmiddlefrontal	395/322	-0.0004	-0.0014	-0.0006	0.0015	0.0018	0.0005	-0.27	-0.79	-1.25	0.784	0.427	0.211	-0.02 (-0.17,0.13)
precuneus	394/322	0.0019	0.0003	0.0005	0.0015	0.0017	0.0005	1.22	0.20	1.03	0.223	0.841	0.303	0.09 (-0.05,0.24)
rostralanteriorcingulate	393/321	0.0071	-0.0026	0.0004	0.0037	0.0042	0.0011	1.90	-0.61	0.36	0.058	0.544	0.721	0.15 (0,0.29)
rostralmiddlefrontal	395/322	-0.0004	-0.0014	-0.0006	0.0015	0.0018	0.0005	-0.27	-0.79	-1.25	0.784	0.427	0.211	-0.02 (-0.17,0.13)

Table S7. Full linear model results for the cortical thickness AIs in adolescents [continued]

Cortical thickness AI		beta-coefficient			Standard Error			t-value			p-value			Cohen's d (95% CI)
	N cases / controls	diag	sex	age	diag	sex	age	diag	sex	age	diag	sex	age	diag
superiorfrontal	394/322	0.0019	0.0012	-0.0004	0.0012	0.0013	0.0004	1.66	0.90	-1.08	0.098	0.369	0.281	0.13 (-0.02,0.27)
superiorparietal	393/322	-0.0010	-0.0007	-0.0005	0.0014	0.0015	0.0004	-0.73	-0.45	-1.19	0.466	0.654	0.234	-0.06 (-0.2,0.09)
superiortemporal	345/297	0.0015	0.0006	-0.0001	0.0018	0.0020	0.0005	0.82	0.31	-0.25	0.415	0.757	0.801	0.07 (-0.09,0.22)
supramarginal	392/321	-0.0011	0.0002	-0.0005	0.0017	0.0019	0.0005	-0.66	0.08	-0.93	0.508	0.933	0.355	-0.05 (-0.2,0.1)
temporalpole	391/320	0.0012	0.0031	-0.0006	0.0041	0.0046	0.0011	0.30	0.68	-0.52	0.766	0.495	0.602	0.02 (-0.12,0.17)
transversetemporal	392/320	0.0018	0.0008	-0.0003	0.0040	0.0045	0.0011	0.46	0.19	-0.32	0.649	0.853	0.746	0.03 (-0.11,0.18)
total thickness	395/322	0.0002	-0.0011	-0.0001	0.0006	0.0006	0.0002	0.29	-1.72	-0.90	0.775	0.086	0.370	0.02 (-0.13,0.17)

P-values in **bold** are significant at the uncorrected level ($P < 0.05$), and in ***bold-italic*** those that survive multiple testing correction within the particular analysis indicated (see text).

Table S8. Full linear model results for the subcortical volume AIs in adults

AI region		beta-coefficient			Standard Error			t-value			p-value			Cohen's d (95% CI)
	N cases/ controls	diag	sex	age	diag	sex	age	diag	sex	age	diag	sex	age	diag
Accumbens	562/492	0.0006	0.0054	0.0000	0.0050	0.0052	0.0003	0.12	1.03	-0.14	0.901	0.304	0.888	0.01 (-0.11,0.13)
Amygdala	562/492	-0.0016	0.0066	0.0000	0.0039	0.0041	0.0002	-0.40	1.63	0.02	0.686	0.104	0.983	-0.03 (-0.15,0.1)
Caudate Nucleus	562/492	0.0013	-0.0008	0.0001	0.0016	0.0017	0.0001	0.76	-0.47	1.54	0.445	0.641	0.123	0.05 (-0.07,0.17)
Globus Pallidus	563/491	-0.0101	-0.0057	-0.0003	0.0035	0.0037	0.0002	-2.90	-1.55	-1.41	***0.004***	0.121	0.160	-0.18 (-0.3,-0.06)
Hippocampus	562/491	0.0019	0.0001	-0.0002	0.0026	0.0026	0.0001	0.73	0.04	-1.83	0.463	0.969	0.067	0.05 (-0.08,0.17)
Putamen	563/492	-0.0014	-0.0018	0.0003	0.0022	0.0023	0.0001	-0.65	-0.79	2.67	0.516	0.431	0.008	-0.04 (-0.16,0.08)
Thalamus[1]	563/490	0.0013	-0.0013	0.0002	0.0018	0.0019	0.0001	0.71	-0.68	1.92	0.480	0.497	0.055	0.04 (-0.08,0.17)

P-values in **bold** are significant at the uncorrected level ($P < 0.05$), and in ***bold-italic*** those that survive multiple testing correction within the particular analysis indicated (see text). [1] Thalamus volume was not available from the NIH dataset.

Table S9. Full linear model results for the cortical surface area AIs in adults.

Cortical surface area AI		beta-coefficient			Standard Error			t-value			p-value			Cohen's d (95% CI)
	N cases / controls	diag	sex	age	diag	sex	age	diag	sex	age	diag	sex	age	diag
bankssts	595/514	0.0012	-0.0034	0.0001	0.0050	0.0051	0.0003	0.23	-0.67	0.58	0.814	0.502	0.561	0.01 (-0.1,0.13)
caudalanteriorcingulate	610/538	0.0027	0.0094	-0.0001	0.0073	0.0075	0.0004	0.38	1.25	-0.17	0.706	0.211	0.865	0.02 (-0.09,0.14)
caudalmiddlefrontal	613/539	0.0054	0.0006	-0.0007	0.0044	0.0045	0.0002	1.23	0.14	-3.18	0.218	0.891	**0.002**	0.07 (-0.04,0.19)
cuneus	612/539	-0.0079	-0.0040	-0.0002	0.0044	0.0045	0.0002	-1.82	-0.88	-0.69	0.068	0.378	0.488	-0.11 (-0.23,0.01)
entorhinal	550/478	-0.0112	-0.0004	0.0006	0.0068	0.0071	0.0004	-1.66	-0.06	1.58	0.097	0.954	0.115	-0.1 (-0.23,0.02)
frontalpole	613/539	-0.0061	0.0053	0.0001	0.0054	0.0055	0.0003	-1.14	0.97	0.52	0.253	0.330	0.603	-0.07 (-0.18,0.05)
fusiform	567/493	-0.0048	0.0073	0.0002	0.0030	0.0031	0.0002	-1.59	2.36	1.18	0.112	**0.018**	0.238	-0.1 (-0.22,0.02)
inferiorparietal	610/538	-0.0004	0.0018	0.0001	0.0030	0.0031	0.0002	-0.13	0.59	0.48	0.894	0.552	0.628	-0.01 (-0.12,0.11)
inferiortemporal	563/493	0.0037	0.0006	0.0001	0.0032	0.0033	0.0002	1.16	0.19	0.39	0.247	0.850	0.695	0.07 (-0.05,0.19)
insula	605/532	-0.0012	0.0046	-0.0001	0.0025	0.0026	0.0001	-0.46	1.76	-0.56	0.643	0.078	0.573	-0.03 (-0.15,0.09)
isthmuscingulate	613/539	0.0030	-0.0084	0.0003	0.0044	0.0045	0.0002	0.69	-1.86	1.27	0.488	0.064	0.204	0.04 (-0.07,0.16)
lateraloccipital	610/539	0.0055	0.0007	0.0000	0.0028	0.0029	0.0001	1.97	0.23	0.27	**0.049**	0.822	0.786	0.12 (0,0.23)
lateralorbitofrontal	613/539	-0.0017	0.0029	0.0001	0.0021	0.0021	0.0001	-0.81	1.35	0.67	0.416	0.176	0.500	-0.05 (-0.17,0.07)
lingual	568/494	0.0023	0.0024	-0.0001	0.0034	0.0036	0.0002	0.67	0.67	-0.76	0.504	0.500	0.446	0.04 (-0.08,0.16)

medialorbitofrontal	611/539	-0.0012	-0.0046	0.0003	0.0033	0.0034	0.0002	-0.35	-1.33	1.71	0.723	0.184	0.087	-0.02 (-0.13,0.1)
middletemporal	555/477	-0.0004	-0.0015	0.0002	0.0027	0.0028	0.0001	-0.14	-0.52	1.12	0.887	0.602	0.261	-0.01 (-0.13,0.12)
paracentral	612/538	-0.0041	0.0053	-0.0001	0.0038	0.0039	0.0002	-1.09	1.35	-0.63	0.277	0.178	0.530	-0.06 (-0.18,0.05)
parahippocampal	568/491	-0.0060	0.0100	0.0001	0.0040	0.0041	0.0002	-1.50	2.45	0.54	0.134	**0.014**	0.587	-0.09 (-0.21,0.03)
parsopercularis	611/539	0.0026	0.0081	0.0000	0.0047	0.0048	0.0002	0.56	1.70	-0.11	0.578	0.090	0.910	0.03 (-0.08,0.15)
parsorbitalis	612/539	0.0021	0.0049	0.0004	0.0036	0.0037	0.0002	0.60	1.33	2.44	0.549	0.185	**0.015**	0.04 (-0.08,0.15)
parstriangularis	612/538	-0.0026	-0.0006	0.0002	0.0046	0.0047	0.0002	-0.56	-0.12	0.86	0.574	0.902	0.388	-0.03 (-0.15,0.08)
pericalcarine	612/539	0.0000	-0.0029	-0.0003	0.0039	0.0040	0.0002	0.00	-0.73	-1.46	0.997	0.469	0.145	0.0003 (-0.12,0.12)
postcentral	607/528	-0.0001	0.0004	-0.0001	0.0027	0.0028	0.0001	-0.03	0.14	-0.84	0.975	0.892	0.398	-0.002 (-0.12,0.11)
posteriorcingulate	613/539	0.0009	0.0055	0.0004	0.0046	0.0047	0.0002	0.21	1.15	1.54	0.836	0.249	0.124	0.01 (-0.1,0.13)
precentral	609/537	-0.0047	0.0023	0.0000	0.0024	0.0025	0.0001	-1.93	0.90	-0.18	0.054	0.366	0.857	-0.11 (-0.23,0)
precuneus	612/539	0.0011	0.0054	0.0001	0.0025	0.0025	0.0001	0.45	2.13	0.70	0.650	**0.034**	0.486	0.03 (-0.09,0.14)
rostralanteriorcingulate	610/535	-0.0055	0.0035	0.0005	0.0060	0.0062	0.0003	-0.92	0.57	1.65	0.360	0.569	0.100	-0.05 (-0.17,0.06)
rostralmiddlefrontal	613/538	-0.0012	0.0008	0.0001	0.0024	0.0024	0.0001	-0.53	0.34	0.90	0.597	0.732	0.371	-0.03 (-0.15,0.08)
superiorfrontal	610/535	-0.0032	-0.0008	0.0000	0.0021	0.0021	0.0001	-1.58	-0.38	-0.40	0.115	0.702	0.692	-0.09 (-0.21,0.02)
superiorparietal	610/539	-0.0010	-0.0001	0.0000	0.0025	0.0026	0.0001	-0.41	-0.06	0.30	0.683	0.954	0.761	-0.02 (-0.14,0.09)
superiortemporal	547/476	-0.0034	-0.0095	-0.0002	0.0026	0.0027	0.0001	-1.30	-3.51	-1.80	0.195	0.000	0.073	-0.08 (-0.2,0.04)

Table S9. Full linear model results for the cortical surface area AIs in adults [continued]

Cortical surface area AI	N cases / controls	beta-coefficient			Standard Error			t-value			p-value			Cohen's d (95% CI)
		diag	sex	age	diag	sex	age	diag	sex	age	diag	sex	age	diag
supramarginal	608/535	-0.0045	-0.0087	-0.0002	0.0037	0.0038	0.0002	-1.24	-2.31	-1.14	0.214	**0.021**	0.254	-0.07 (-0.19,0.04)
temporalpole	568/494	-0.0045	-0.0124	0.0001	0.0048	0.0050	0.0003	-0.95	-2.47	0.21	0.344	**0.014**	0.836	-0.06 (-0.18,0.06)
transversetemporal	568/494	0.0004	0.0041	0.0002	0.0051	0.0052	0.0003	0.07	0.78	0.91	0.941	0.437	0.361	0.005 (-0.12,0.13)
total surface area	613/539	-0.0008	0.0005	0.0000	0.0006	0.0007	0.0000	-1.20	0.72	0.27	0.232	0.474	0.789	-0.07 (-0.19,0.05)

P-values in **bold** are significant at the uncorrected level ($P < 0.05$), and in ***bold-italic*** those that survive multiple testing correction within the particular analysis indicated (see text).

Table S10. Full linear model results for the cortical thickness AIs in adults.

Cortical thickness AI		beta-coefficient			Standard Error			t-value			p-value			Cohen's d (95% CI)
	N cases / controls	diag	sex	age	diag	sex	age	diag	sex	age	diag	sex	age	diag
bankssts	595/514	-0.0011	0.0004	-0.0001	0.0024	0.0025	0.0001	-0.46	0.17	-0.51	0.644	0.865	0.608	-0.03 (-0.15,0.09)
caudalanteriorcingulate	610/538	0.0061	0.0042	0.0000	0.0033	0.0034	0.0002	1.89	1.25	-0.23	0.060	0.213	0.815	0.11 (0,0.23)
caudalmiddlefrontal	613/539	0.0005	-0.0019	-0.0001	0.0015	0.0016	0.0001	0.35	-1.21	-1.18	0.728	0.226	0.240	0.02 (-0.09,0.14)
cuneus	612/539	0.0034	-0.0005	0.0001	0.0018	0.0019	0.0001	1.89	-0.25	1.08	0.059	0.799	0.279	0.11 (0,0.23)
entorhinal	550/478	-0.0017	0.0009	-0.0002	0.0036	0.0039	0.0002	-0.46	0.24	-1.09	0.646	0.809	0.275	-0.03 (-0.15,0.09)
frontalpole	613/539	0.0044	-0.0071	-0.0002	0.0033	0.0034	0.0002	1.32	-2.08	-1.12	0.187	**0.038**	0.263	0.08 (-0.04,0.19)
fusiform	567/493	0.0004	0.0007	-0.0001	0.0014	0.0015	0.0001	0.26	0.44	-1.20	0.792	0.662	0.229	0.02 (-0.1,0.14)
inferiorparietal	610/538	0.0008	0.0024	0.0000	0.0012	0.0012	0.0001	0.66	1.94	-0.54	0.512	0.053	0.593	0.04 (-0.08,0.16)
inferiortemporal	563/493	-0.0003	0.0011	-0.0002	0.0015	0.0016	0.0001	-0.20	0.69	-1.94	0.845	0.490	0.053	-0.01 (-0.13,0.11)
insula	605/532	-0.0001	-0.0007	-0.0001	0.0014	0.0015	0.0001	-0.07	-0.47	-1.27	0.945	0.641	0.203	0 (-0.12,0.11)
isthmuscingulate	613/539	-0.0021	0.0000	0.0000	0.0022	0.0023	0.0001	-0.93	0.00	-0.30	0.351	0.996	0.766	-0.06 (-0.17,0.06)
lateraloccipital	610/539	0.0026	0.0010	0.0001	0.0012	0.0012	0.0001	2.22	0.82	1.11	**0.026**	0.415	0.269	0.13 (0.02,0.25)
lateralorbitofrontal	613/539	0.0023	-0.0035	-0.0001	0.0015	0.0016	0.0001	1.48	-2.18	-1.25	0.139	**0.030**	0.212	0.09 (-0.03,0.2)
lingual	568/494	-0.0008	-0.0008	-0.0001	0.0015	0.0016	0.0001	-0.53	-0.50	-1.34	0.593	0.619	0.182	-0.03 (-0.15,0.09)
medialorbitofrontal	611/539	0.0039	-0.0033	-0.0001	0.0020	0.0020	0.0001	2.01	-1.63	-0.58	**0.045**	0.103	0.560	0.12 (0,0.24)

Table S10. Full linear model results for the cortical thickness AIs in adults [continued]

Cortical thickness AI		beta-coefficient			Standard Error			t-value			p-value			Cohen's d (95% CI)
	N cases / controls	diag	sex	age	diag	sex	age	diag	sex	age	diag	sex	age	diag
middletemporal	555/477	-0.0038	-0.0029	0.0001	0.0015	0.0016	0.0001	-2.63	-1.87	0.84	**0.009**	0.061	0.401	-0.17 (-0.29,-0.04)
paracentral	612/538	-0.0005	0.0008	0.0000	0.0015	0.0016	0.0001	-0.29	0.47	0.34	0.768	0.637	0.731	-0.02 (-0.13,0.1)
parahippocampal	568/491	0.0026	0.0045	0.0000	0.0030	0.0031	0.0002	0.85	1.42	-0.02	0.394	0.156	0.984	0.05 (-0.07,0.17)
parsopercularis	611/539	-0.0002	0.0018	0.0000	0.0017	0.0018	0.0001	-0.14	0.98	-0.12	0.892	0.326	0.903	-0.01 (-0.12,0.11)
parsorbitalis	612/539	0.0021	-0.0008	0.0000	0.0023	0.0024	0.0001	0.90	-0.32	0.13	0.370	0.748	0.896	0.05 (-0.06,0.17)
parstriangularis	612/538	-0.0002	0.0012	0.0000	0.0019	0.0020	0.0001	-0.12	0.61	-0.29	0.901	0.545	0.769	-0.01 (-0.12,0.11)
pericalcarine	612/539	0.0062	-0.0002	-0.0002	0.0021	0.0022	0.0001	2.90	-0.11	-1.61	**0.004**	0.912	0.107	0.17 (0.06,0.29)
postcentral	607/528	-0.0033	-0.0016	0.0000	0.0013	0.0014	0.0001	-2.48	-1.13	-0.01	**0.013**	0.260	0.990	-0.15 (-0.27,-0.03)
posteriorcingulate	613/539	-0.0003	0.0002	-0.0002	0.0018	0.0019	0.0001	-0.17	0.13	-1.83	0.868	0.897	0.068	-0.01 (-0.13,0.11)
precentral	609/537	0.0018	-0.0029	-0.0001	0.0013	0.0013	0.0001	1.39	-2.21	-1.58	0.166	0.027	0.115	0.08 (-0.03,0.2)
precuneus	612/539	0.0005	-0.0018	-0.0001	0.0012	0.0012	0.0001	0.41	-1.43	-1.08	0.685	0.154	0.279	0.02 (-0.09,0.14)
rostralanteriorcingulate	610/535	0.0027	0.0021	-0.0004	0.0030	0.0031	0.0002	0.92	0.69	-2.17	0.355	0.493	**0.030**	0.06 (-0.06,0.17)
rostralmiddlefrontal	613/538	-0.0011	0.0001	-0.0001	0.0012	0.0013	0.0001	-0.96	0.11	-1.04	0.339	0.910	0.298	-0.06 (-0.17,0.06)
superiorfrontal	610/535	0.0004	-0.0005	-0.0001	0.0009	0.0009	0.0000	0.47	-0.50	-1.68	0.640	0.616	0.094	0.03 (-0.09,0.14)
superiorparietal	610/539	0.0002	-0.0009	0.0000	0.0010	0.0010	0.0001	0.19	-0.91	0.92	0.850	0.363	0.360	0.01 (-0.1,0.13)

superiorparietal	610/539	0.0002	-0.0009	0.0000	0.0010	0.0010	0.0001	0.19	-0.91	0.92	0.850	0.363	0.360	0.01 (-0.1,0.13)
superiortemporal	547/476	-0.0014	0.0013	0.0000	0.0015	0.0016	0.0001	-0.92	0.78	-0.29	0.358	0.435	0.770	-0.06 (-0.18,0.06)
supramarginal	608/535	-0.0001	0.0011	-0.0001	0.0014	0.0015	0.0001	-0.08	0.78	-0.67	0.934	0.437	0.505	0 (-0.12,0.11)
temporalpole	568/494	-0.0016	0.0019	0.0001	0.0033	0.0035	0.0002	-0.49	0.54	0.74	0.624	0.590	0.460	-0.03 (-0.15,0.09)
transversetemporal	568/494	-0.0027	0.0031	-0.0003	0.0029	0.0029	0.0001	-0.95	1.06	-1.79	0.344	0.289	0.073	-0.06 (-0.18,0.06)
total thickness	613/539	0.0001	-0.0003	-0.0001	0.0004	0.0004	0.0000	0.31	-0.63	-2.51	0.754	0.526	0.012	0.02 (-0.1,0.13)

P-values in **bold** are significant at the uncorrected level (P < 0.05), and in ***bold-italic*** those that survive multiple testing correction within the particular analysis indicated (see text).

Table S11. Full linear model results for the subcortical volume AIs in all age groups combined.

AI region		beta-coefficient			Standard Error			t-value			p-value			Cohen's d (95% CI)
	N cases/ controls	diag	sex	age	diag	sex	age	diag	sex	age	diag	sex	age	diag
Accumbens	1689/1566	-0.0026	0.0032	-0.0002	0.0027	0.0029	0.0002	-0.99	1.08	-1.24	0.321	0.280	0.217	-0.03 (-0.1,0.03)
Amygdala	1688/1565	-0.0010	0.0045	0.0000	0.0020	0.0022	0.0001	-0.52	2.06	0.11	0.606	0.039	0.911	-0.02 (-0.09,0.05)
Caudate Nucleus	1689/1566	0.0008	-0.0021	0.0002	0.0010	0.0011	0.0001	0.83	-1.92	2.68	0.407	0.055	**0.007**	0.03 (-0.04,0.1)
Globus Pallidus	1691/1562	-0.0049	-0.0021	-0.0006	0.0019	0.0021	0.0001	-2.57	-1.02	-4.10	**0.010**	0.308	**4.3×10^{-5}**	-0.09 (-0.16,-0.02)
Hippocampus	1688/1562	0.0007	0.0003	-0.0001	0.0014	0.0016	0.0001	0.50	0.22	-1.64	0.616	0.828	0.102	0.02 (-0.05,0.09)
Putamen	1690/1556	-0.0012	0.0008	0.0003	0.0011	0.0012	0.0001	-1.12	0.65	4.49	0.263	0.514	**7.2×10^{-6}**	-0.04 (-0.11,0.03)
Thalamus[1]	1574/1448	0.0015	-0.0014	0.0007	0.0010	0.0012	0.0001	1.45	-1.19	8.67	0.146	0.233	**7.0×10^{-18}**	0.05 (-0.02,0.12)

P-values in **bold** are significant at the uncorrected level ($P < 0.05$), and in ***bold-italic*** those that survive multiple testing correction within the particular analysis indicated (see text).

Table S12. Full linear model results for the cortical surface area AI in all age groups combined.

Surface area AI		**beta-coefficient**			**Standard Error**			**t-value**			**p-value**			**Cohen's d (95% CI)**
	N cases / controls	**diag**	**sex**	**age**	**diag**	**sex**	**age**	**diag**	**sex**	**age**	**diag**	**sex**	**age**	**diag**
bankssts	1825/1726	-0.0020	-0.0003	0.0002	0.0029	0.0031	0.0002	-0.70	-0.10	0.85	0.483	0.917	0.394	-0.02 (-0.09,0.04)
caudalanteriorcingulate	1928/1821	-0.0018	0.0070	-0.0004	0.0039	0.0042	0.0002	-0.47	1.66	-2.07	0.636	0.097	**0.038**	-0.02 (-0.08,0.05)
caudalmiddlefrontal	1933/1825	0.0032	-0.0017	-0.0005	0.0025	0.0027	0.0002	1.31	-0.61	-3.17	0.192	0.540	**0.002**	0.04 (-0.02,0.11)
cuneus	1926/1826	-0.0007	0.0001	0.0003	0.0023	0.0025	0.0002	-0.33	0.03	1.96	0.742	0.978	0.050	-0.01 (-0.07,0.05)
entorhinal	1817/1717	-0.0038	-0.0074	0.0007	0.0039	0.0043	0.0002	-0.96	-1.72	3.11	0.337	0.086	**0.002**	-0.03 (-0.1,0.03)
frontalpole	1933/1828	-0.0050	0.0062	0.0000	0.0030	0.0032	0.0001	-1.65	1.92	0.14	0.099	0.055	0.891	-0.05 (-0.12,0.01)
fusiform	1877/1776	-0.0025	0.0043	0.0000	0.0017	0.0018	0.0001	-1.45	2.38	0.34	0.147	**0.017**	0.733	-0.05 (-0.11,0.02)
inferiorparietal	1925/1818	0.0014	0.0072	-0.0001	0.0017	0.0019	0.0001	0.78	3.78	-0.97	0.435	**0.000**	0.332	0.03 (-0.04,0.09)
inferiortemporal	1843/1765	0.0021	-0.0017	0.0001	0.0018	0.0020	0.0001	1.15	-0.88	0.57	0.251	0.381	0.571	0.04 (-0.03,0.1)
insula	1918/1813	0.0016	0.0043	-0.0004	0.0015	0.0017	0.0001	1.07	2.62	-3.87	0.283	**0.009**	**0.0002**	0.04 (-0.03,0.1)
isthmuscingulate	1931/1821	-0.0008	-0.0091	-0.0001	0.0025	0.0027	0.0001	-0.32	-3.39	-0.72	0.752	**0.001**	0.472	-0.01 (-0.07,0.05)
lateraloccipital	1929/1828	0.0011	0.0009	0.0001	0.0016	0.0017	0.0001	0.71	0.49	1.27	0.478	0.623	0.204	0.02 (-0.04,0.09)
lateralorbitofrontal	1932/1828	-0.0024	0.0038	0.0001	0.0013	0.0014	0.0001	-1.86	2.66	1.08	0.064	**0.008**	0.280	-0.06 (-0.12,0)
lingual	1885/1780	-0.0002	0.0022	0.0000	0.0018	0.0019	0.0001	-0.10	1.15	0.37	0.917	0.249	0.708	0 (-0.07,0.06)
medialorbitofrontal	1922/1819	0.0041	-0.0042	0.0004	0.0019	0.0021	0.0001	2.19	-2.05	2.60	**0.029**	0.041	**0.009**	0.07 (0.01,0.14)

Table S12. Full linear model results for the cortical surface area AI in all age groups combined [continued]

Surface area AI		beta-coefficient			Standard Error			t-value			p-value			Cohen's d (95% CI)
	N cases / controls	diag	sex	age	diag	sex	age	diag	sex	age	diag	sex	age	diag
middletemporal	1786/1704	0.0013	0.0002	0.0002	0.0015	0.0017	0.0001	0.88	0.14	1.91	0.382	0.885	0.056	0.03 (-0.04,0.1)
paracentral	1931/1827	-0.0047	0.0044	-0.0001	0.0021	0.0023	0.0001	-2.19	1.90	-1.41	0.029	0.058	0.159	-0.07 (-0.14,-0.01)
parahippocampal	1879/1771	-0.0008	0.0112	-0.0001	0.0024	0.0026	0.0001	-0.35	4.25	-0.63	0.729	**2.1×10 -5**	0.526	-0.01 (-0.08,0.05)
parsopercularis	1922/1824	0.0026	0.0047	0.0000	0.0027	0.0029	0.0001	0.95	1.62	0.36	0.340	0.105	0.719	0.03 (-0.03,0.1)
parsorbitalis	1932/1826	0.0047	0.0040	0.0001	0.0020	0.0022	0.0001	2.32	1.85	1.31	0.021	0.065	0.189	0.08 (0.01,0.14)
parstriangularis	1925/1828	0.0030	-0.0004	-0.0001	0.0025	0.0027	0.0001	1.19	-0.15	-0.52	0.236	0.879	0.600	0.04 (-0.03,0.1)
pericalcarine	1929/1827	0.0001	-0.0043	0.0001	0.0020	0.0022	0.0001	0.07	-1.98	0.92	0.941	**0.048**	0.355	0.002 (-0.06,0.07)
postcentral	1909/1800	0.0013	-0.0021	0.0001	0.0015	0.0016	0.0001	0.86	-1.27	0.60	0.391	0.203	0.551	0.03 (-0.04,0.09)
posteriorcingulate	1930/1823	-0.0014	0.0046	0.0001	0.0025	0.0028	0.0001	-0.53	1.66	0.87	0.594	0.097	0.382	-0.02 (-0.08,0.05)
precentral	1913/1818	-0.0023	0.0009	0.0000	0.0013	0.0014	0.0001	-1.71	0.65	0.11	0.087	0.514	0.910	-0.06 (-0.12,0.01)
precuneus	1931/1825	0.0010	0.0056	0.0000	0.0013	0.0014	0.0001	0.73	3.90	-0.11	0.464	**0.0001**	0.912	0.02 (-0.04,0.09)
rostralanteriorcingulate	1918/1817	-0.0022	0.0029	-0.0001	0.0033	0.0036	0.0002	-0.67	0.81	-0.68	0.506	0.416	0.496	-0.02 (-0.09,0.04)
rostralmiddlefrontal	1932/1824	-0.0021	-0.0005	-0.0001	0.0013	0.0014	0.0001	-1.59	-0.38	-1.69	0.112	0.704	0.092	-0.05 (-0.12,0.01)
superiorfrontal	1926/1821	0.0007	-0.0014	0.0001	0.0011	0.0012	0.0001	0.59	-1.19	1.92	0.555	0.234	0.056	0.02 (-0.04,0.08)
superiorparietal	1925/1826	0.0016	0.0023	-0.0001	0.0014	0.0015	0.0001	1.10	1.49	-2.34	0.273	0.137	0.019	0.04 (-0.03,0.1)

superiortemporal	1743/1692	0.0007	-0.0103	0.0000	0.0014	0.0015	0.0001	0.50	-6.92	0.04	0.620	5.4×10 -12	0.972	0.02 (-0.05,0.08)
supramarginal	1910/1809	-0.0024	-0.0076	-0.0002	0.0020	0.0022	0.0001	-1.18	-3.38	-1.31	0.236	**0.001**	0.189	-0.04 (-0.1,0.03)
temporalpole	1879/1777	0.0001	-0.0096	0.0002	0.0027	0.0030	0.0002	0.04	-3.23	1.28	0.965	**0.001**	0.201	0.001 (-0.06,0.07)
transversetemporal	1882/1780	0.0002	0.0032	0.0003	0.0027	0.0029	0.0002	0.09	1.09	1.59	0.931	0.276	0.111	0.003 (-0.06,0.07)
Total surface area	1933/1829	0.0002	0.0005	0.0000	0.0003	0.0003	0.0000	0.61	1.59	0.04	0.540	0.112	0.965	0.02 (-0.04,0.08)

P-values in **bold** are significant at the uncorrected level (P<0.05), and in ***bold-italic*** those that survive multiple testing correction within the particular analysis indicated (see text).

Table S13. Full linear model results for the cortical thickness AIs in all age groups combined.

Thickness AI		beta-coefficient			Standard Error			t-value			p-value			Cohen's d (95% CI)
	N cases / controls	diag	sex	age	diag	sex	age	diag	sex	age	diag	sex	age	diag
bankssts	1824/1726	-0.0025	0.0021	-0.0001	0.0013	0.0014	0.0001	-1.90	1.43	-0.91	0.058	0.152	0.363	-0.06 (-0.13,0)
caudalanteriorcingulate	1927/1821	0.0028	0.0002	-0.0001	0.0018	0.0019	0.0001	1.61	0.10	-0.69	0.108	0.919	0.488	0.05 (-0.01,0.12)
caudalmiddlefrontal	1931/1826	0.0019	-0.0016	-0.0001	0.0008	0.0009	0.0001	2.22	-1.75	-1.65	**0.027**	0.081	0.099	0.07 (0.01,0.14)
cuneus	1926/1827	0.0006	-0.0001	0.0000	0.0011	0.0012	0.0001	0.58	-0.05	0.30	0.563	0.964	0.764	0.02 (-0.05,0.08)
entorhinal	1817/1718	-0.0023	-0.0057	-0.0003	0.0020	0.0022	0.0001	-1.12	-2.56	-2.48	0.261	**0.010**	0.013	-0.04 (-0.1,0.03)
frontalpole	1931/1828	0.0020	-0.0034	-0.0002	0.0021	0.0023	0.0001	0.95	-1.48	-1.41	0.344	0.140	0.160	0.03 (-0.03,0.1)
fusiform	1878/1777	-0.0001	0.0001	-0.0001	0.0007	0.0008	0.0001	-0.08	0.18	-1.44	0.938	0.858	0.151	-0.003 (-0.07,0.06)
inferiorparietal	1925/1821	0.0002	0.0003	0.0000	0.0007	0.0007	0.0000	0.24	0.45	-0.72	0.810	0.654	0.471	0.01 (-0.06,0.07)
inferiortemporal	1842/1764	-0.0004	0.0000	-0.0002	0.0009	0.0010	0.0001	-0.40	0.00	-2.36	0.691	0.999	**0.018**	-0.01 (-0.08,0.05)
insula	1918/1814	-0.0017	0.0004	-0.0001	0.0008	0.0009	0.0001	-1.97	0.42	-2.26	**0.049**	0.678	**0.024**	-0.06 (-0.13,0)
isthmuscingulate	1928/1821	-0.0001	-0.0001	0.0001	0.0013	0.0014	0.0001	-0.11	-0.09	1.66	0.909	0.927	0.098	-0.004 (-0.07,0.06)
lateraloccipital	1928/1829	0.0006	0.0005	0.0001	0.0007	0.0008	0.0001	0.82	0.58	1.57	0.410	0.563	0.118	0.03 (-0.04,0.09)
lateralorbitofrontal	1931/1828	0.0007	-0.0032	0.0000	0.0009	0.0010	0.0001	0.81	-3.19	-0.18	0.419	**0.001**	0.856	0.03 (-0.04,0.09)
lingual	1884/1779	-0.0009	-0.0013	-0.0001	0.0009	0.0010	0.0001	-1.06	-1.37	-1.63	0.290	**0.171**	0.104	-0.04 (-0.1,0.03)
medialorbitofrontal	1922/1820	0.0000	-0.0045	0.0001	0.0012	0.0013	0.0001	0.03	-3.43	0.67	0.974	**0.001**	0.500	0.001 (-0.06,0.07)

middletemporal	1786/1706	-0.0014	-0.0014	-0.0001	0.0008	0.0009	0.0001	-1.59	-1.46	-1.01	0.111	0.144	0.313	-0.05 (-0.12,0.01)
paracentral	1930/1827	-0.0005	0.0006	-0.0001	0.0009	0.0009	0.0001	-0.62	0.69	-1.66	0.534	0.493	0.098	-0.02 (-0.08,0.04)
parahippocampal	1879/1772	-0.0025	-0.0024	0.0000	0.0016	0.0017	0.0001	-1.57	-1.37	-0.52	0.117	0.170	0.605	-0.05 (-0.12,0.01)
parsopercularis	1921/1824	0.0007	0.0013	-0.0002	0.0010	0.0010	0.0001	0.76	1.22	-3.02	0.448	0.221	**0.003**	0.02 (-0.04,0.09)
parsorbitalis	1931/1826	0.0016	0.0009	0.0000	0.0015	0.0017	0.0001	1.03	0.54	0.14	0.304	0.587	0.890	0.03 (-0.03,0.1)
parstriangularis	1923/1828	0.0001	0.0004	-0.0002	0.0010	0.0011	0.0001	0.10	0.39	-2.29	0.919	0.694	**0.022**	0.003 (-0.06,0.07)
pericalcarine	1926/1826	0.0019	0.0001	-0.0004	0.0013	0.0014	0.0001	1.45	0.06	-4.70	0.148	0.955	**2.7×10^{-6}**	0.05 (-0.02,0.11)
postcentral	1906/1802	-0.0012	-0.0018	0.0000	0.0007	0.0008	0.0000	-1.59	-2.27	-0.28	0.112	**0.023**	0.779	-0.05 (-0.12,0.01)
posteriorcingulate	1928/1826	-0.0010	-0.0006	-0.0001	0.0010	0.0012	0.0001	-0.94	-0.50	-1.27	0.349	0.616	0.205	-0.03 (-0.09,0.03)
precentral	1911/1816	0.0013	-0.0014	-0.0001	0.0006	0.0007	0.0000	1.93	-2.01	-1.39	0.054	**0.045**	0.165	0.06 (0,0.13)
precuneus	1930/1825	0.0006	-0.0006	0.0000	0.0007	0.0007	0.0000	0.91	-0.90	-0.74	0.360	0.369	0.458	0.03 (-0.03,0.09)
rostralanteriorcingulate	1917/1815	0.0020	0.0015	-0.0004	0.0016	0.0018	0.0001	1.26	0.84	-3.41	0.209	0.401	**0.001**	0.04 (-0.02,0.11)
rostralmiddlefrontal	1931/1825	-0.0001	-0.0002	-0.0001	0.0007	0.0008	0.0001	-0.19	-0.29	-2.39	0.846	0.772	0.017	-0.01 (-0.07,0.06)
superiorfrontal	1925/1821	0.0005	-0.0004	-0.0001	0.0005	0.0006	0.0000	1.03	-0.72	-2.15	0.305	0.474	0.032	0.03 (-0.03,0.1)
superiorparietal	1924/1826	-0.0002	-0.0008	0.0000	0.0006	0.0006	0.0000	-0.29	-1.30	0.59	0.772	0.193	0.554	-0.01 (-0.07,0.05)
superiortemporal	1743/1695	0.0009	0.0016	0.0000	0.0008	0.0009	0.0001	1.09	1.80	-0.67	0.276	0.073	0.501	0.04 (-0.03,0.1)

Table S13. Full linear model results for the cortical thickness AIs in all age groups combined [continued]

Thickness AI		beta-coefficient			Standard Error			t-value			p-value			Cohen's d (95% CI)
	N cases / controls	diag	sex	age	diag	sex	age	diag	sex	age	diag	sex	age	diag
supramarginal	1909/1812	-0.0010	0.0008	0.0000	0.0008	0.0008	0.0000	-1.31	1.00	0.05	0.191	0.317	0.960	-0.04 (-0.11,0.02)
temporalpole	1877/1776	0.0006	0.0013	-0.0003	0.0019	0.0021	0.0001	0.29	0.60	-2.38	0.771	0.551	0.017	0.01 (-0.06,0.07)
transversetemporal	1881/1780	0.0000	0.0031	-0.0001	0.0016	0.0018	0.0001	-0.03	1.77	-1.30	0.980	0.076	0.194	-0.0008 (-0.07,0.06)
Total thickness	1932/1829	0.0001	-0.0004	-0.0001	0.0003	0.0003	0.0000	0.28	-1.52	-3.21	0.778	0.128	**0.001**	0.01 (-0.05,0.07)

P-values in **bold** are significant at the uncorrected level ($P < 0.05$), and in ***bold-italic*** those that survive multiple testing correction within the particular analysis indicated (see text).

Table S14 Directions of asymmetry changes in ADHD individuals versus controls for those AIs that had shown nominally significant (P<0.05) associations with diagnosis in any of the main analyses.

AI region	Mean AI ± SD in controls	Mean AI ± SD in ADHD	Cohen's *d* (95% CI) Left hemisphere	Cohen's *d* (95% CI) Right hemisphere	Controls	ADHD
Children:						
Medial orbitofrontal cortex surface area	-0.0116 ± 0.06	-0.0028 ± 0.06	-0.175 (-0.27,-0.08)	-0.32 (-0.41,-0.23)	rightward	decreased
Paracentral lobule surface area	-0.0587 ± 0.07	-0.0656 ± 0.07	-0.215 (-0.31,-0.12)	-0.115 (-0.21,-0.02)	rightward	increased
Total average surface area	-0.0025 ± 0.01	-0.0018 ± 0.01	-0.32 (-0.41,-0.23)	-0.339 (-0.43,-0.25)	rightward	decreased
Banks of superior temporal sulcus thickness	-0.0132 ± 0.04	-0.0172 ± 0.04	-0.109 (-0.2,-0.02)	0.002 (-0.09,0.09)	rightward	increased
Caudal middle frontal cortex thickness	0.0022 ± 0.03	0.0059 ± 0.03	-0.024 (-0.11,0.07)	-0.103 (-0.19,-0.01)	leftward	increased
Insula thickness	0.0036 ± 0.03	5e-04 ± 0.03	-0.146 (-0.24,-0.05)	-0.058 (-0.15,0.03)	leftward	decreased
Precentral gyrus thickness	0.0061 ± 0.02	0.0088 ± 0.02	-0.122 (-0.21,-0.03)	-0.192 (-0.28,-0.1)	leftward	increased
Adolescents:						
Pars orbitalis of inferior frontal gyrus surface area	-0.1115 ± 0.06	-0.1007 ± 0.06	-0.047 (-0.19,0.1)	-0.224 (-0.37,-0.08)	rightward	decreased
Cuneus thickness	-0.0022 ± 0.03	-0.0077 ± 0.03	-0.086 (-0.23,0.06)	0.026 (-0.12,0.17)	rightward	increased
Adults:						
Globus Pallidus	0.0085 ± 0.06	0.0031 ± 0.07	-0.039 (-0.16,0.08)	0.134 (0.01,0.26)	leftward	decreased
Lateral occipital cortex surface area	0.0125 ± 0.05	0.0182 ± 0.05	0.021 (-0.09,0.14)	-0.069 (-0.18,0.05)	leftward	increased
Lateral occipital cortex thickness	-0.0141 ± 0.02	-0.012 ± 0.02	0.166 (0.05,0.28)	0.07 (-0.05,0.19)	rightward	decreased
Medial orbitofrontal cortex thickness	0.0128 ± 0.04	0.0174 ± 0.04	-0.029 (-0.14,0.09)	-0.137 (-0.25,-0.02)	leftward	increased
Middle temporal gyrus thickness	-0.0035 ± 0.02	-0.007 ± 0.02	-0.04 (-0.16,0.08)	0.05 (-0.07,0.17)	rightward	increased
Pericalcarine cortex thickness	-0.0087 ± 0.04	-0.0025 ± 0.04	0.121 (0,0.24)	-0.035 (-0.15,0.08)	rightward	decreased
Postcentral gyrus thickness	0.0065 ± 0.02	0.0031 ± 0.02	-0.009 (-0.13,0.11)	0.096 (-0.02,0.21)	leftward	decreased
Total:						
Globus Pallidus	0.0256 ± 0.06	0.0199 ± 0.06	-0.128 (-0.2,-0.06)	-0.063 (-0.13,0.01)	leftward	decreased
Medial orbitofrontal cortex surface area	-0.0067 ± 0.06	-0.0013 ± 0.06	-0.115 (-0.18,-0.05)	-0.202 (-0.27,-0.14)	rightward	decreased
Paracentral lobule surface area	-0.0611 ± 0.07	-0.0663 ± 0.06	-0.123 (-0.19,-0.06)	-0.056 (-0.12,0.01)	rightward	increased
Pars orbitalis of inferior frontal gyrus surface area	-0.1072 ± 0.06	-0.1031 ± 0.06	-0.085 (-0.15,-0.02)	-0.165 (-0.23,-0.1)	rightward	decreased
Caudal middle frontal cortex thickness	0.0023 ± 0.03	0.0051 ± 0.03	-0.023 (-0.09,0.04)	-0.081 (-0.14,-0.02)	leftward	increased
Insula thickness	0.0028 ± 0.03	6e-04 ± 0.03	-0.081 (-0.14,-0.02)	-0.023 (-0.09,0.04)	leftward	decreased

The raw means and standard deviations are indicated, as well as the Cohen's d effect sizes for left and right hemispheric measures (i.e., when the left or right hemispheric measures were analyzed separately as dependent variables). Additionally, the average direction of asymmetry in controls (derived from the raw mean AI) and its change in ADHD is shown. Positive AI values indicate leftward asymmetry, negative AI values indicate rightward asymmetry.

Table S15. Sensitivity analyses for the effects of diagnosis in all age groups combined, for subcortical volume AIs.

	Main analysis		**Non-linear age**		**Winsorized**	
Subcortical volume AI	***P***	***d***	***P***	***d***	***P***	***d***
Accumbens	0.321	-0.035	0.324	-0.035	0.311	-0.036
Amygdala	0.606	-0.018	0.605	-0.018	0.682	-0.014
Caudate Nucleus	0.407	0.029	0.412	0.029	0.411	0.029
Globus Pallidus	**0.010**	-0.091	**0.011**	-0.090	**0.010**	-0.091
Hippocampus	0.616	0.018	0.617	0.018	0.524	0.022
Putamen	0.263	-0.040	0.260	-0.040	0.359	-0.032
Thalamus[1]	0.146	0.053	0.149	0.053	0.137	0.055

P-values (P) and Cohen's *d* values (d) for the effects of diagnosis are indicated. P-values in **bold** are significant at the uncorrected level ($P < 0.05$), and in ***bold-italic*** those that survive multiple testing correction within the particular analysis indicated (see text). [1]Thalamus volume was not available from the NIH dataset.

.

Table S16. Sensitivity analyses for the effects of diagnosis in all age groups combined, for cortical surface area AIs.

	Main analysis		**Non-linear age**		**Winsorized**	
Cortical surface area AI	***P***	***d***	***P***	***d***	***P***	***d***
banks of superior temporal sulcus	0.483	-0.024	0.473	-0.024	0.493	-0.023
caudal anterior cingulate cortex	0.636	-0.016	0.645	-0.015	0.630	-0.016
caudal middle frontal cortex	0.192	0.043	0.198	0.042	0.182	0.044
cuneus	0.742	-0.011	0.710	-0.012	0.771	-0.010
entorhinal cortex	0.337	-0.033	0.335	-0.033	0.372	-0.030
frontal pole	0.099	-0.054	0.100	-0.054	0.093	-0.055
fusiform gyrus	0.147	-0.048	0.152	-0.048	0.174	-0.045
inferior parietal cortex	0.435	0.026	0.416	0.027	0.437	0.026
inferior temporal gyrus	0.251	0.038	0.240	0.039	0.241	0.039
insula	0.283	0.035	0.262	0.037	0.315	0.033
isthmus cingulate cortex	0.752	-0.010	0.791	-0.009	0.853	-0.006
lateral occipital cortex	0.478	0.023	0.474	0.023	0.470	0.024
lateral orbitofrontal cortex	0.064	-0.061	0.063	-0.061	0.062	-0.061
lingual gyrus	0.917	-0.003	0.907	-0.004	0.900	0.004
medial orbitofrontal cortex	**0.029**	0.072	**0.029**	0.072	**0.027**	0.073
middle temporal gyrus	0.382	0.030	0.388	0.029	0.364	0.031
paracentral lobule	**0.029**	-0.072	**0.029**	-0.072	**0.029**	-0.072
parahippocampal gyrus	0.729	-0.012	0.735	-0.011	0.728	-0.012
pars opercularis of inferior frontal gyrus	0.340	0.031	0.336	0.032	0.338	0.032
pars orbitalis of inferior frontal gyrus	**0.021**	0.076	**0.019**	0.077	**0.019**	0.077
pars triangularis of inferior frontal gyrus	0.236	0.039	0.228	0.040	0.213	0.041
pericalcarine cortex	0.941	0.002	0.960	0.002	0.807	0.008
postcentral gyrus	0.391	0.028	0.400	0.028	0.365	0.030
posterior cingulate cortex	0.594	-0.017	0.598	-0.017	0.581	-0.018
precentral gyrus	0.087	-0.056	0.087	-0.056	0.101	-0.054
precuneus	0.464	0.024	0.458	0.024	0.389	0.028
rostral anterior cingulate cortex	0.506	-0.022	0.532	-0.021	0.500	-0.022
rostral middle frontal gyrus	0.112	-0.052	0.122	-0.051	0.123	-0.051
superior frontal gyrus	0.555	0.019	0.583	0.018	0.483	0.023
superiorparietal	0.273	0.036	0.255	0.037	0.227	0.040
superior temporal gyrus	0.620	0.017	0.642	0.016	0.496	0.023

Table S16. Sensitivity analyses for the effects of diagnosis in all age groups combined, for cortical surface area AIs [continued]

	Main analysis		**Non-linear age**		**Winsorized**	
Cortical surface area AI	***P***	***d***	***P***	***d***	***P***	***d***
supramarginal gyrus	0.236	-0.039	0.232	-0.039	0.268	-0.037
temporal pole	0.965	0.001	0.992	0.000	0.961	0.002
transverse temporal gyrus	0.931	0.003	0.946	0.002	0.830	0.007
total average surface area	0.540	0.020	0.532	0.020	0.147	0.048

P-values (P) and Cohen's *d* values (d) for the effects of diagnosis are indicated. P-values in **bold** are significant at the uncorrected level ($P < 0.05$), and in ***bold-italic*** those that survive multiple testing correction within the particular analysis indicated (see text).

Table S17. Sensitivity analyses for the effects of diagnosis in all age groups combined, for cortical thickness AIs.

	Main analysis		**Non-linear age**		**Winsorized**	
Cortical thickness AI	***P***	***d***	***P***	***d***	***P***	***d***
banks of superior temporal sulcus	0.058	-0.064	0.059	-0.064	0.060	-0.064
caudal anterior cingulate cortex	0.108	0.053	0.105	0.053	0.105	0.053
caudal middle frontal cortex	**0.027**	0.073	**0.026**	0.073	**0.027**	0.073
cuneus	0.563	0.019	0.555	0.019	0.566	0.019
entorhinal cortex	0.261	-0.038	0.263	-0.038	0.273	-0.037
frontal pole	0.344	0.031	0.345	0.031	0.337	0.032
fusiform gyrus	0.938	-0.003	0.932	-0.003	0.972	-0.001
inferior parietal cortex	0.810	0.008	0.801	0.008	0.813	0.008
inferior temporal gyrus	0.691	-0.013	0.683	-0.014	0.694	-0.013
insula	**0.049**	-0.065	0.050	-0.065	0.061	-0.062
isthmus cingulate cortex	0.909	-0.004	0.869	-0.005	0.914	-0.004
lateral occipital cortex	0.410	0.027	0.418	0.027	0.405	0.027
lateral orbitofrontal cortex	0.419	0.027	0.417	0.027	0.431	0.026
lingual gyrus	0.290	-0.035	0.280	-0.036	0.308	-0.034
medial orbitofrontal cortex	0.974	0.001	0.979	0.001	0.962	0.002
middle temporal gyrus	0.111	-0.054	0.113	-0.054	0.130	-0.052
paracentral lobule	0.534	-0.020	0.549	-0.020	0.534	-0.020
parahippocampal gyrus	0.117	-0.052	0.114	-0.053	0.116	-0.052
pars opercularis of inferior frontal gyrus	0.448	0.025	0.429	0.026	0.424	0.026
pars orbitalis of inferior frontal gyrus	0.304	0.034	0.293	0.034	0.306	0.034
pars triangularis of inferior frontal gyrus	0.919	0.003	0.898	0.004	0.835	0.007
pericalcarine cortex	0.148	0.048	0.140	0.048	0.145	0.048
postcentral gyrus	0.112	-0.053	0.110	-0.053	0.113	-0.052
posterior cingulate cortex	0.349	-0.031	0.344	-0.031	0.370	-0.029
precentral gyrus	0.054	0.064	0.053	0.064	0.050	0.065
precuneus	0.360	0.030	0.366	0.030	0.366	0.030
rostral anterior cingulate cortex	0.209	0.041	0.206	0.042	0.216	0.041
rostral middle frontal gyrus	0.846	-0.006	0.855	-0.006	0.834	-0.007
superior frontal gyrus	0.305	0.034	0.298	0.034	0.354	0.030
superiorparietal	0.772	-0.010	0.779	-0.009	0.783	-0.009
superior temporal gyrus	0.276	0.037	0.278	0.037	0.225	0.042

Table S17. Sensitivity analyses for the effects of diagnosis in all age groups combined, for cortical thickness AIs [continued]

	Main analysis		**Non-linear age**		**Winsorized**	
Cortical thickness AI	***P***	***d***	***P***	***d***	***P***	***d***
supramarginal gyrus	0.191	-0.043	0.192	-0.043	0.216	-0.041
temporal pole	0.771	0.010	0.748	0.011	0.751	0.011
transverse temporal gyrus	0.980	-0.001	0.963	-0.002	0.992	0.0003
total average thickness	0.778	0.009	0.776	0.009	0.815	0.008

P-values (P) and Cohen's *d* values (d) for the effects of diagnosis are indicated. P-values in **bold** are significant at the uncorrected level ($P < 0.05$), and in ***bold-italic*** those that survive multiple testing correction within the particular analysis indicated (see text).

Table S18. Associations of subcortical volume AIs with disorder severity in ADHD individuals, all age groups combined.

	Hyperactivity/impulsivity				Inattention			
Region	**N subjects**	**N datasets**	**p-value**	**t-value**	**N subjects**	**N datasets**	**p-value**	**t-value**
Accumbens	281	6	0.540	0.613	281	6	0.363	0.912
Amygdala	282	6	0.830	0.215	282	6	0.932	0.085
Caudate Nucleus	281	6	0.306	1.026	281	6	0.830	0.215
Globus Pallidus	283	6	0.964	0.045	283	6	0.668	0.429
Hippocampus	281	6	0.431	0.788	281	6	0.565	-0.577
Putamen	282	6	0.520	0.644	282	6	0.408	0.829
Thalamus[1]	280	6	0.125	-1.538	280	6	0.396	-0.851

P-values (P) and t-values (t) for the effects of ADHD severity, as measured by hyperactivity/impulsivity and inattention symptoms, are indicated. P-values in **bold** are significant at the uncorrected level ($P < 0.05$), and in ***bold-italic*** those that survive multiple testing correction within the particular analysis indicated (see text). [1]Thalamus volume was not available from the NIH dataset

Table S19. Associations of cortical surface area AIs with disorder severity in ADHD individuals, all age groups combined.

Region	Hyperactivity/impulsivity				Inattention			
	N subjects	N datasets	p-value	t-value	N subjects	N datasets	p-value	t-value
banks of superior temporal sulcus	286	7	0.971	-0.036	286	7	0.352	0.931
caudal anterior cingulate cortex	322	7	0.157	1.420	322	7	0.901	0.125
caudal middle frontal cortex	322	7	0.301	-1.037	322	7	0.552	-0.596
cuneus	322	7	0.112	1.595	322	7	0.865	-0.170
entorhinal cortex	320	7	**0.035**	2.123	320	7	0.066	1.846
frontal pole	322	7	0.687	-0.403	322	7	0.921	-0.099
fusiform gyrus	322	7	0.600	0.525	322	7	0.283	1.075
inferior parietal cortex	321	7	0.313	-1.010	321	7	0.757	-0.310
inferior temporal gyrus	321	7	0.255	-1.139	321	7	0.428	0.794
insula	322	7	0.905	-0.120	322	7	0.160	1.408
isthmus cingulate cortex	321	7	0.694	-0.393	321	7	0.087	-1.715
lateral occipital cortex	321	7	0.646	0.460	321	7	0.720	0.359
lateral orbitofrontal cortex	322	7	0.180	1.343	322	7	0.169	1.379
lingual gyrus	322	7	0.443	-0.768	322	7	0.838	0.205
medial orbitofrontal cortex	321	7	0.624	0.490	321	7	0.674	0.422
middle temporal gyrus	308	7	0.919	0.102	308	7	0.243	-1.169
paracentral lobule	322	7	0.585	-0.547	322	7	0.179	1.347
parahippocampal gyrus	322	7	0.646	-0.459	322	7	0.206	1.267
pars opercularis of inferior frontal gyrus	322	7	0.473	-0.719	322	7	0.165	-1.393

Table S19. Associations of cortical surface area AIs with disorder severity in ADHD individuals, all age groups combined [continued]

Region	Hyperactivity/impulsivity				Inattention			
	N subjects	N datasets	p-value	t-value	N subjects	N datasets	p-value	t-value
pars orbitalis of inferior frontal gyrus	322	7	0.284	1.072	322	7	0.286	1.069
pars triangularis of inferior frontal gyrus	322	7	0.816	-0.233	322	7	0.832	-0.212
pericalcarine cortex	322	7	0.992	0.010	322	7	0.754	-0.313
postcentral gyrus	321	7	0.773	-0.289	321	7	0.328	0.980
posterior cingulate cortex	321	7	0.516	-0.650	321	7	0.493	-0.686
precentral gyrus	322	7	0.937	-0.079	322	7	0.135	1.497
precuneus	321	7	0.548	0.602	321	7	0.382	-0.876
rostral anterior cingulate cortex	321	7	0.330	0.976	321	7	0.550	0.598
rostral middle frontal gyrus	322	7	0.848	-0.191	322	7	0.828	0.217
superior frontal gyrus	322	7	0.067	-1.839	322	7	0.125	1.540
superiorparietal	322	7	0.622	0.494	322	7	0.119	-1.561
superior temporal gyrus	291	7	0.182	1.337	291	7	0.507	0.664
supramarginal gyrus	314	7	0.411	0.823	314	7	0.806	-0.246
temporal pole	321	7	0.615	0.503	321	7	0.383	0.873
transverse temporal gyrus	322	7	0.637	-0.472	322	7	0.067	-1.836
total average surface area	322	7	0.812	-0.238	322	7	0.430	0.791

P-values (P) and t-values (t) for the effects of ADHD severity, as measured by hyperactivity/impulsivity and inattention symptoms, are indicated. P-values in **bold** are significant at the uncorrected level ($P < 0.05$), and in ***bold-italic*** those that survive multiple testing correction within the particular analysis indicated (see text).

Table S20. Associations of cortical thickness AIs with disorder severity in ADHD individuals, all age groups combined.

	Hyperactivity/impulsivity				Inattention			
Region	**N subjects**	**N datasets**	**p-value**	**t-value**	**N subjects**	**N datasets**	**p-value**	**t-value**
banks of superior temporal sulcus	286	7	0.784	-0.274	286	7	0.802	0.250
caudal anterior cingulate cortex	322	7	**0.008**	2.655	322	7	0.305	1.027
caudal middle frontal cortex	321	7	0.083	1.741	321	7	0.799	-0.255
cuneus	322	7	0.736	0.337	322	7	0.707	0.376
entorhinal cortex	321	7	0.684	-0.407	321	7	0.624	0.490
frontal pole	321	7	0.372	0.895	321	7	0.792	0.264
fusiform gyrus	322	7	0.447	-0.762	322	7	0.849	0.191
inferior parietal cortex	321	7	0.824	-0.222	321	7	0.301	-1.035
inferior temporal gyrus	321	7	0.793	0.263	321	7	0.393	0.855
insula	322	7	0.764	-0.301	322	7	0.082	-1.744
isthmus cingulate cortex	321	7	0.990	0.013	321	7	0.468	-0.726
lateral occipital cortex	321	7	0.645	0.462	321	7	0.633	-0.479
lateral orbitofrontal cortex	322	7	0.740	0.332	322	7	0.596	-0.531
lingual gyrus	322	7	0.584	-0.549	322	7	0.576	-0.559
medial orbitofrontal cortex	322	7	0.109	1.605	322	7	0.942	0.073
middle temporal gyrus	308	7	0.832	-0.212	308	7	0.810	-0.241
central lobule	322	7	0.110	1.603	322	7	0.744	0.327
parahippocampal gyrus	322	7	0.970	-0.037	322	7	0.907	-0.117

Table S20. Associations of cortical thickness AIs with disorder severity in ADHD individuals, all age groups combined [continued]

	Hyperactivity/impulsivity				Inattention			
Region	**N subjects**	**N datasets**	**p-value**	**t-value**	**N subjects**	**N datasets**	**p-value**	**t-value**
pars orbitalis of inferior frontal gyrus	322	7	0.081	-1.750	322	7	0.437	-0.778
pars triangularis of inferior frontal gyrus	320	7	0.311	1.016	320	7	0.083	1.740
pericalcarine cortex	322	7	**0.017**	2.410	322	7	0.311	1.015
postcentral gyrus	321	7	0.494	-0.684	321	7	0.137	1.490
posterior cingulate cortex	320	7	0.272	1.101	320	7	0.275	1.093
precentral gyrus	322	7	0.482	0.705	322	7	0.387	-0.867
precuneus	321	7	0.209	1.260	321	7	0.493	0.687
rostral anterior cingulate cortex	321	7	0.766	-0.298	321	7	0.481	0.705
rostral middle frontal gyrus	322	7	0.346	0.944	322	7	0.762	0.303
superior frontal gyrus	322	7	0.139	1.482	322	7	0.587	-0.543
superiorparietal	322	7	0.872	-0.162	322	7	0.450	-0.757
superior temporal gyrus	291	7	0.778	-0.283	291	7	0.352	0.932
supramarginal gyrus	314	7	0.457	-0.745	314	7	0.805	0.247
temporal pole	321	7	0.283	-1.076	321	7	0.975	-0.031
transverse temporal gyrus	322	7	0.311	1.014	322	7	0.173	1.367
total average thickness	322	7	0.113	1.591	322	7	0.486	0.697

P-values (P) and t-values (t) for the effects of ADHD severity, as measured by hyperactivity/impulsivity and inattention symptoms, are indicated. P-values in **bold** are significant at the uncorrected level ($P < 0.05$), and in ***bold-italic*** those that survive multiple testing correction within the particular analysis indicated (see text).

Table S21. Associations of subcortical volume AIs with psychostimulant medication use in ADHD individuals, all age groups combined.

	Lifetime medication use				Current medication use			
Region	**N subjects (no/yes)**	**N datasets**	**p-value**	**t-value**	**N subjects (no/yes)**	**N datasets**	**p-value**	**t-value**
Accumbens	183/268	7	0.406	-0.831	320/318	13	0.951	0.062
Amygdala	182/268	7	0.459	0.742	319/318	13	0.534	0.623
Caudate Nucleus	183/268	7	0.359	-0.919	320/318	13	0.793	0.263
Globus Pallidus	183/268	7	0.517	-0.648	320/318	13	0.089	1.701
Hippocampus	182/267	7	0.834	0.210	320/317	13	0.715	0.365
Putamen	183/268	7	0.306	-1.024	320/318	13	0.075	1.785
Thalamus[1]	182/267	7	0.406	0.832	290/235	12	0.985	-0.019

P-values (P) and t-values (t) for the effects of current and lifetime psychostimulant medication use are indicated. P-values in **bold** are significant at the uncorrected level (P < 0.05), and in ***bold-italic*** those that survive multiple testing correction within the particular analysis indicated (see text). [1]Thalamus volume was not available from the NIH dataset.

Table S22. Associations of cortical surface area AIs with psychostimulant medication use in ADHD individuals, all age groups combined.

	Lifetime medication use				Current medication use			
Region	**N subjects (no/yes)**	**N datasets**	**p-value**	**t-value**	**N subjects (no/yes)**	**N datasets**	**p-value**	**t-value**
banks of superior temporal sulcus	177/310	9	0.148	1.451	349/337	15	0.308	-1.019
caudal anterior cingulate cortex	188/335	9	0.134	1.501	377/359	15	0.293	1.053
caudal middle frontal cortex	188/337	9	0.837	0.206	377/361	15	0.169	-1.376
cuneus	186/337	9	0.604	0.520	374/358	15	0.745	0.325
entorhinal cortex	176/335	9	0.484	-0.701	366/337	14	0.753	-0.315
frontal pole	188/337	9	0.674	0.421	377/361	15	0.870	0.164
fusiform gyrus	187/335	9	0.978	-0.027	376/359	15	0.965	-0.044
inferior parietal cortex	186/335	9	0.989	-0.013	375/361	15	0.939	-0.077
inferior temporal gyrus	187/301	9	0.837	0.206	361/339	15	0.074	1.791
insula	179/336	9	**0.043**	-2.026	377/360	15	0.068	-1.825
isthmus cingulate cortex	188/337	9	0.268	1.110	377/360	15	0.276	-1.091
lateral occipital cortex	186/337	9	0.611	-0.509	377/361	15	0.982	-0.022
lateral orbitofrontal cortex	188/337	9	0.814	-0.235	376/361	15	0.605	-0.517
lingual gyrus	187/337	9	0.750	0.319	377/361	15	0.243	-1.168
medial orbitofrontal cortex	187/337	9	0.415	0.816	374/357	15	0.085	1.726
middle temporal gyrus	183/285	9	0.598	0.527	344/327	15	0.095	1.670
paracentral lobule	188/336	9	0.488	-0.694	377/361	15	0.584	0.547
parahippocampal gyrus	188/337	9	0.282	1.076	377/359	15	0.362	0.911

pars opercularis of inferior frontal gyrus	188/334	9	0.318	0.999	376/356	15	0.996	0.005
pars orbitalis of inferior frontal gyrus	187/337	9	0.543	0.609	377/361	15	0.061	1.874
pars triangularis of inferior frontal gyrus	188/337	9	0.098	1.657	377/358	15	0.943	-0.072
pericalcarine cortex	187/336	9	0.999	-0.001	376/359	15	0.949	-0.064
postcentral gyrus	180/337	9	0.294	-1.050	370/354	15	0.987	-0.016
posterior cingulate cortex	188/337	9	0.868	-0.166	377/360	15	0.765	-0.299
precentral gyrus	183/335	9	0.344	0.948	373/351	15	0.284	1.072
precuneus	187/337	9	0.972	-0.035	377/361	15	**0.025**	-2.251
rostral anterior cingulate cortex	187/334	9	**0.049**	1.972	372/357	15	0.257	1.135
rostral middle frontal gyrus	188/337	9	0.830	-0.214	377/360	15	0.084	-1.728
superior frontal gyrus	188/335	9	0.617	0.500	376/358	15	0.723	0.355
superiorparietal	185/336	9	0.282	1.078	375/360	15	0.187	-1.320
superior temporal gyrus	179/276	9	0.620	0.496	336/317	15	0.099	1.651
supramarginal gyrus	184/333	9	**0.038**	-2.082	374/353	15	0.507	-0.664
temporal pole	188/337	9	0.700	-0.386	377/360	15	0.401	-0.840
transverse temporal gyrus	188/337	9	0.182	-1.335	377/361	15	**0.020**	-2.336
total average surface area	188/337	9	0.919	0.102	377/361	15	0.140	-1.477

P-values (P) and t-values (t) for the effects of current and lifetime psychostimulant medication use are indicated. P-values in **bold** are significant at the uncorrected level ($P < 0.05$), and in ***bold-italic*** those that survive multiple testing correction within the particular analysis indicated (see text).

Table S23. Associations of cortical thickness AIs with psychostimulant medication use in ADHD individuals, all age groups combined.

	Lifetime medication use				Current medication use			
Region	**N subjects (no/yes)**	**N datasets**	**p-value**	**t-value**	**N subjects (no/yes)**	**N datasets**	**p-value**	**t-value**
banks of superior temporal sulcus	177/310	9	0.259	-1.131	349/337	15	0.479	0.709
caudal anterior cingulate cortex	188/335	9	0.894	-0.133	377/359	15	0.940	-0.076
caudal middle frontal cortex	188/337	9	0.237	-1.183	376/361	15	0.811	0.239
cuneus	187/337	9	0.871	0.162	375/358	15	0.773	0.288
entorhinal cortex	176/335	9	0.953	0.059	366/337	14	0.497	-0.679
frontal pole	188/337	9	0.801	0.252	377/361	15	0.620	0.496
fusiform gyrus	188/336	9	0.626	0.488	377/360	15	0.257	-1.134
inferior parietal cortex	186/335	9	0.380	0.879	375/361	15	0.020	-2.332
inferior temporal gyrus	187/301	9	0.699	-0.386	361/339	15	0.397	-0.847
insula	179/336	9	0.724	0.353	377/360	15	0.845	0.196
isthmus cingulate cortex	188/337	9	0.998	-0.002	377/360	15	0.138	1.483
lateral occipital cortex	186/337	9	0.422	0.803	377/361	15	0.315	1.005
lateral orbitofrontal cortex	188/337	9	0.281	1.078	376/361	15	0.942	0.073
lingual gyrus	187/337	9	0.928	-0.091	377/361	15	0.617	0.500
medial orbitofrontal cortex	187/337	9	0.282	-1.077	374/357	15	0.417	0.813
middle temporal gyrus	184/285	9	0.491	0.689	345/327	15	0.485	0.699
paracentral lobule	188/336	9	0.032	2.152	377/361	15	0.552	0.596
parahippocampal gyrus	188/336	9	0.384	0.872	377/358	15	0.726	-0.350

pars opercularis of inferior frontal gyrus	188/334	9	0.251	1.149	376/356	15	0.380	0.878
pars orbitalis of inferior frontal gyrus	187/337	9	0.753	-0.314	377/361	15	0.642	-0.466
pars triangularis of inferior frontal gyrus	188/337	9	0.937	-0.079	375/358	15	0.940	-0.075
pericalcarine cortex	187/336	9	0.868	0.167	376/359	15	0.086	1.721
postcentral gyrus	181/336	9	0.587	-0.543	371/353	15	0.988	0.015
posterior cingulate cortex	188/337	9	0.460	-0.740	376/360	15	0.868	0.166
precentral gyrus	183/335	9	0.258	-1.131	373/351	15	**0.031**	-2.157
precuneus	187/337	9	0.466	0.729	377/361	15	0.605	0.517
rostral anterior cingulate cortex	187/334	9	0.656	-0.445	372/357	15	0.511	-0.658
rostral middle frontal gyrus	188/337	9	0.840	-0.202	377/360	15	0.378	0.882
superior frontal gyrus	188/335	9	0.704	-0.380	376/358	15	0.911	-0.112
superiorparietal	185/336	9	0.761	-0.304	375/360	15	0.740	0.332
superior temporal gyrus	179/276	9	0.287	-1.066	336/317	15	0.171	1.369
supramarginal gyrus	184/333	9	0.789	0.268	374/353	15	0.502	-0.672
temporal pole	188/337	9	0.861	0.176	377/360	15	0.901	-0.124
transverse temporal gyrus	188/337	9	0.207	1.265	377/361	15	0.796	0.259
total average thickness	188/337	9	0.676	0.419	377/361	15	0.801	0.253

P-values (P) and t-values (t) for the effects of current and lifetime psychostimulant medication use are indicated. P-values in **bold** are significant at the uncorrected level ($P < 0.05$), and in ***bold-italic*** those that survive multiple testing correction within the particular analysis indicated (see text)

Table S24 Associations of subcortical volume AIs with comorbidities in ADHD individuals, all age groups combined.

AI region	Mood				ODD				Anxiety				SUD			
	N subj (no/yes)	ND	p-value	t-value	N subj (no/yes)	ND	p-value	t-value	N subj (no/yes)	ND	p-value	t-value	N subj (no/yes)	ND	p-value	t-value
Accumbens	268/164	8	0.58	0.55	81/39	6	0.251	1.154	388/66	9	0.664	0.435	252/67	6	0.492	0.688
Amygdala	266/164	8	0.75	0.32	81/39	6	0.316	-1.007	386/66	9	0.276	1.090	252/67	6	0.325	-0.985
Caudate Nucleus	268/164	8	0.48	0.71	81/39	6	0.249	-1.158	388/66	9	0.173	-1.364	252/65	6	0.105	-1.624
Globus Pallidus	268/164	8	0.67	0.42	81/39	6	0.605	0.519	388/66	9	0.978	-0.028	252/67	6	0.583	0.550
Hippocampus	267/164	8	0.94	0.08	81/39	6	0.307	-1.027	387/66	9	0.918	-0.103	252/67	6	0.416	-0.814
Putamen	268/164	8	0.11	-1.61	81/39	6	0.618	-0.501	388/66	9	0.669	-0.428	252/66	6	0.658	-0.444
Thalamus[1]	266/164	8	0.81	-0.24	81/39	6	0.419	-0.812	386/66	9	0.486	0.697	251/65	6	0.548	0.602

P-values (P) and t-values (t) for the effects of ADHD comorbidities are indicated. P-values in **bold** are significant at the uncorrected level ($P < 0.05$), and in ***bold-italic*** those that survive multiple testing correction within the particular analysis indicated (see text). ND= Number of datasets.

Table S25. Associations of cortical surface area AIs with comorbidities in ADHD individuals, all age groups combined.

AI region	Mood				ODD				Anxiety				SUD			
	N subj (no/yes)	ND	p-value	t-value	N subj (no/yes)	NS	p-value	t-value	N subj (no/yes)	ND	p-value	t-value	N subj (no/yes)	ND	p-value	t-value
bankssts	378/174	8	0.422	0.804	145/78	6	0.411	-0.824	493/81	9	0.786	0.272	331/75	6	0.111	1.599
caudalanteriorcingulate	383/178	8	0.154	-1.427	151/78	6	0.302	-1.034	501/82	9	0.301	-1.035	333/77	6	0.797	-0.257
caudalmiddlefrontal	384/179	8	0.814	-0.235	151/80	6	0.453	0.751	503/82	9	0.080	1.757	335/77	6	0.647	0.458
cuneus	383/178	8	0.086	1.718	149/79	6	0.849	0.191	501/82	9	0.212	1.249	335/76	6	0.857	0.181
entorhinal	344/157	8	0.128	1.525	127/76	6	0.523	-0.639	443/77	9	0.115	1.578	295/69	6	0.335	0.965
frontalpole	384/179	8	0.935	0.081	151/80	6	0.466	-0.730	503/82	9	0.256	1.138	335/77	6	0.793	0.262
fusiform	355/158	8	0.699	-0.386	148/78	6	0.792	-0.264	455/77	9	0.152	1.436	294/70	6	0.412	-0.821
inferiorparietal	383/176	8	0.742	-0.330	150/80	6	0.139	-1.484	500/81	9	0.928	-0.091	334/74	6	0.768	-0.296
inferiortemporal	357/159	8	**0.046**	2.003	150/77	6	0.074	1.795	458/77	9	0.484	0.700	295/70	6	0.883	0.147
insula	377/176	8	0.885	0.145	149/78	6	0.785	-0.273	422/81	9	0.714	-0.367	335/76	6	0.718	0.362
isthmuscingulate	384/179	8	0.262	-1.123	151/79	6	0.689	-0.400	503/82	9	0.686	0.404	335/77	6	0.316	1.003
lateraloccipital	383/177	8	0.863	-0.173	151/80	6	0.493	-0.687	500/82	9	0.296	-1.046	335/74	6	0.502	0.673
lateralorbitofrontal	383/179	8	0.185	1.328	150/80	6	0.627	0.486	502/82	9	0.968	0.041	334/77	6	0.713	0.368
lingual	356/159	8	0.438	-0.776	151/80	6	0.153	-1.434	458/79	9	0.227	-1.209	295/69	6	0.531	0.628
medialorbitofrontal	382/179	8	0.564	-0.577	149/79	6	0.779	-0.281	501/82	9	0.934	-0.083	335/76	6	0.127	1.531
middletemporal	353/155	8	0.714	0.367	147/74	6	0.746	-0.324	451/76	9	0.726	-0.350	290/68	6	0.989	0.014
paracentral	383/179	8	0.326	-0.982	151/80	6	0.070	1.823	503/81	9	0.990	-0.013	334/77	6	0.061	1.879
parahippocampal	357/158	8	0.319	-0.997	149/77	6	0.358	0.921	458/76	9	0.108	1.609	295/70	6	0.763	-0.301

Table S25. Associations of cortical surface area AIs with comorbidities in ADHD individuals, all age groups combined [continued]

AI region	Mood				ODD				Anxiety				SUD			
	N subj (no/yes)	ND	p-value	t-value	N subj (no/yes)	NS	p-value	t-value	N subj (no/yes)	ND	p-value	t-value	N subj (no/yes)	ND	p-value	t-value
parsopercularis	381/178	8	0.400	0.843	145/79	6	0.933	-0.084	499/82	9	0.566	0.575	331/77	6	0.673	0.400
parsorbitalis	383/179	8	0.845	-0.195	151/80	6	0.843	0.198	502/82	9	0.179	1.345	334/77	6	0.450	0.845
parstriangularis	384/179	8	0.826	-0.220	147/79	6	0.152	-1.436	503/82	9	0.852	-0.187	335/77	6	0.323	0.826
pericalcarine	382/179	8	0.194	1.299	148/80	6	0.481	-0.706	501/82	9	0.958	0.052	334/76	6	0.943	0.194
postcentral	379/176	8	0.683	0.408	149/75	6	0.970	-0.038	421/81	9	0.778	-0.282	334/77	6	**0.006**	0.683
posteriorcingulate	384/179	8	0.357	0.922	149/80	6	0.322	-0.992	503/82	9	0.166	-1.386	335/77	6	0.272	0.357
precentral	379/177	8	0.498	0.678	146/78	6	0.807	0.244	419/81	9	0.277	-1.088	331/77	6	0.945	0.498
precuneus	383/179	8	0.900	0.125	151/80	6	0.754	0.314	502/82	9	0.975	-0.032	335/76	6	0.833	0.900
rostralanteriorcingulate	381/177	8	**0.029**	-2.186	148/78	6	0.512	-0.657	500/80	9	0.388	-0.865	332/75	6	0.628	**0.029**
rostralmiddlefrontal	384/179	8	0.583	0.550	150/80	6	0.194	-1.303	503/82	9	0.879	0.152	335/77	6	0.053	0.583
superiorfrontal	383/178	8	0.988	-0.015	150/79	6	0.066	1.850	501/82	9	0.252	-1.147	333/77	6	0.447	0.988
superiorparietal	382/177	8	0.869	-0.165	150/79	6	0.585	0.547	500/81	9	0.784	0.274	334/74	6	0.442	0.869
superiortemporal	352/151	8	0.195	-1.297	147/76	6	0.521	-0.642	446/76	9	0.682	0.410	288/67	6	0.375	0.195
supramarginal	382/176	8	0.618	-0.499	151/79	6	0.280	-1.084	498/82	9	0.165	1.390	334/76	6	**0.033**	0.618
temporalpole	357/159	8	0.984	0.020	150/77	6	**0.037**	2.103	458/77	9	0.821	0.226	295/70	6	0.206	0.984
transversetemporal	357/159	8	0.633	-0.478	150/78	6	0.674	-0.422	458/77	9	0.658	0.443	295/70	6	0.230	0.633
total surface area	384/179	8	0.865	0.170	151/80	6	0.687	-0.404	503/82	9	0.601	0.524	335/77	6	0.372	0.865

P-values (P) and t-values (t) for the effects of ADHD comorbidities are indicated. P-values in **bold** are significant at the uncorrected level ($P < 0.05$), and in ***bold-italic*** those that survive multiple testing correction within the particular analysis indicated (see text). subj=subjects; ND=number of datasets.

Table S26. Associations of cortical thickness AIs with comorbidities in ADHD individuals, all age groups combined.

AI region	Mood				ODD				Anxiety				SUD			
	N subj (no/yes)	ND	p-value	t-value	N subj (no/yes)	ND	p-value	t-value	N subj (no/yes)	ND	p-value	t-value	N subj (no/yes)	ND	p-value	t-value
bankssts	378/174	8	0.633	-0.478	145/78	6	0.219	1.233	493/81	9	0.364	-0.909	331/75	6	0.407	0.83
caudalanteriorcingulate	383/178	8	0.584	-0.548	151/78	6	0.558	0.586	501/82	9	0.553	0.593	333/77	6	0.298	-1.042
caudalmiddlefrontal	383/179	8	0.295	1.049	151/79	6	0.192	1.309	502/82	9	0.395	0.852	335/77	6	0.92	-0.101
cuneus	383/179	8	0.979	0.026	149/79	6	0.745	0.326	502/82	9	**0.019**	2.358	335/76	6	**0.007**	2.7
entorhinal	345/157	8	**0.004**	-2.891	128/76	6	0.451	-0.755	444/77	9	0.794	-0.261	295/69	6	0.951	-0.061
frontalpole	384/179	8	0.77	-0.292	151/80	6	0.579	0.555	503/82	9	0.105	-1.622	335/77	6	0.164	1.395
fusiform	357/158	8	0.555	0.591	148/78	6	0.207	-1.265	457/77	9	0.241	1.173	294/70	6	0.692	0.396
inferiorparietal	383/176	8	0.196	1.295	150/80	6	0.963	0.046	500/81	9	0.705	-0.379	334/74	6	0.515	-0.652
inferiortemporal	357/159	8	0.902	0.123	150/77	6	0.671	-0.425	458/77	9	0.433	-0.785	295/70	6	0.538	-0.616
insula	377/176	8	0.487	-0.695	149/78	6	0.209	-1.26	422/81	9	0.961	-0.049	335/76	6	0.437	-0.779
isthmuscingulate	384/179	8	0.928	0.091	151/79	6	0.736	0.337	503/82	9	0.605	-0.517	335/77	6	0.749	-0.321
lateraloccipital	383/177	8	0.441	-0.771	151/80	6	0.66	-0.441	500/82	9	**0.016**	-2.413	335/74	6	0.309	1.019
lateralorbitofrontal	383/179	8	0.76	0.305	150/80	6	0.295	-1.05	502/82	9	0.482	0.703	334/77	6	0.234	1.193
lingual	356/159	8	0.296	1.047	151/80	6	0.17	-1.376	458/79	9	0.253	-1.145	295/69	6	0.734	0.34
medialorbitofrontal	383/179	8	0.503	0.67	150/79	6	**0.029**	-2.195	502/82	9	0.056	-1.914	335/76	6	0.522	-0.641
middletemporal	354/155	8	0.345	0.945	147/74	6	0.917	-0.104	452/76	9	0.406	-0.832	290/68	6	0.78	-0.28
paracentral	383/179	8	0.495	-0.683	151/80	6	0.794	0.261	503/81	9	0.060	1.882	334/77	6	**0.047**	1.991
parahippocampal	356/159	8	0.192	1.305	149/78	6	0.551	0.598	458/77	9	0.158	1.413	295/70	6	0.258	1.133
parsopercularis	381/178	8	0.805	-0.246	145/79	6	0.807	0.245	499/82	9	0.965	-0.044	331/77	6	0.131	1.512

Table S26. Associations of cortical thickness AIs with comorbidities in ADHD individuals, all age groups combined [continued]

AI region	Mood				ODD				Anxiety				SUD			
	N subj (no/yes)	ND	p-value	t-value	N subj (no/yes)	ND	p-value	t-value	N subj (no/yes)	ND	p-value	t-value	N subj (no/yes)	ND	p-value	t-value
bankssts	378/174	8	0.633	-0.478	145/78	6	0.219	1.233	493/81	9	0.364	-0.909	331/75	6	0.407	0.83
parsorbitalis	383/179	8	0.511	0.657	151/80	6	0.434	0.784	502/82	9	0.952	-0.060	334/77	6	0.309	1.018
parstriangularis	383/178	8	0.027	2.221	147/77	6	0.353	0.931	501/82	9	0.401	0.840	335/77	6	0.389	-0.862
pericalcarine	382/179	8	0.046	1.997	148/80	6	0.632	-0.48	501/82	9	0.990	0.013	334/76	6	0.429	-0.792
postcentral	379/176	8	0.409	-0.826	149/75	6	0.235	1.192	421/81	9	0.822	0.225	334/77	6	0.053	1.944
posteriorcingulate	383/179	8	0.076	1.777	149/79	6	0.1	1.651	502/82	9	0.929	0.090	335/77	6	0.601	0.524
precentral	379/177	8	0.742	0.33	146/78	6	0.563	0.579	419/81	9	0.953	0.059	331/77	6	0.465	-0.732
precuneus	383/179	8	0.024	-2.264	151/80	6	0.617	0.501	502/82	9	0.236	-1.185	335/76	6	0.09	1.702
rostralanteriorcingulate	381/177	8	0.05	1.962	148/78	6	0.626	-0.488	500/80	9	0.478	-0.710	332/75	6	0.117	-1.571
rostralmiddlefrontal	384/179	8	0.0002	3.702	150/80	6	0.365	0.908	503/82	9	0.780	0.280	335/77	6	0.373	0.891
superiorfrontal	383/178	8	0.985	-0.018	150/79	6	0.68	-0.414	501/82	9	0.838	0.204	333/77	6	0.586	-0.546
superiorparietal	382/177	8	0.578	0.557	150/79	6	0.083	1.742	500/81	9	0.730	-0.345	334/74	6	0.381	0.878
superior temporal gyrus	352/151	8	0.847	0.193	147/76	6	0.329	0.979	446/76	9	0.184	-1.332	288/67	6	0.625	0.489
supramarginal gyrus	382/176	8	0.679	0.415	151/79	6	0.232	1.198	498/82	9	0.117	1.568	334/76	6	0.402	-0.838
temporal pole	357/159	8	0.921	-0.099	150/77	6	0.263	-1.122	458/77	9	0.197	-1.292	295/70	6	0.715	-0.366
transverse temporal gyrus	357/159	8	0.036	2.101	150/78	6	0.748	0.322	458/77	9	0.976	-0.030	295/70	6	0.376	0.887
total average thickness	384/179	8	0.109	1.608	151/80	6	0.669	0.428	503/82	9	0.600	-0.524	335/77	6	0.642	0.465

P-values (P) and t-values (t) for the effects of ADHD comorbidities are indicated. P-values in **bold** are significant at the uncorrected level (P 0.05), and in ***bold-italic*** those that survive multiple testing correction within the particular analysis indicated (see text). subj=subjects; ND=number of datasets.

Table S27. Associations of subcortical volume AIs with IQ in all age groups combined.

	IQ cases			IQ controls		
AI region	**N subjects**	**p-value**	**t-value**	**N subjects**	**p-value**	**t-value**
Accumbens	1508	**0.031**	2.162	1386	0.772	-0.290
Amygdala	1507	0.277	1.087	1385	0.132	1.507
Caudate Nucleus	1508	0.410	0.824	1386	0.670	-0.427
Globus Pallidus	1510	0.619	0.497	1382	0.074	-1.790
Hippocampus	1507	**0.039**	-2.063	1382	0.661	0.439
Putamen	1509	0.282	1.076	1376	0.233	1.193
Thalamus[1]	1405	0.294	-1.051	1273	0.405	-0.834

P-values (P) and t-values (t) for the effects of IQ in cases and controls are indicated. P-values in **bold** are significant at the uncorrected level ($P < 0.05$), and in ***bold-italic*** those that survive multiple testing correction within the particular analysis indicated (see text). [1]Thalamus volume was not available from the NIH dataset.

Table S28. Associations of cortical surface area AIs with IQ in all age groups combined.

	IQ cases			IQ controls		
AI region	**N subjects**	**p-value**	**t-value**	**N subjects**	**p-value**	**t-value**
banks of superior temporal sulcus	1637	0.867	0.167	1541	0.448	0.759
caudal anterior cingulate cortex	1738	0.511	0.658	1631	0.877	0.155
caudal middle frontal cortex	1743	0.072	1.802	1635	0.072	-1.800
cuneus	1736	0.864	-0.172	1636	0.145	1.458
entorhinal cortex	1634	0.615	0.503	1530	0.532	0.625
frontal pole	1743	0.701	0.384	1638	0.622	0.493
fusiform gyrus	1688	0.914	-0.108	1587	0.264	-1.118
inferior parietal cortex	1735	0.304	1.028	1628	0.743	-0.327
inferior temporal gyrus	1654	0.869	-0.165	1575	0.465	-0.730
insula	1732	0.931	-0.087	1625	0.913	-0.110
isthmus cingulate cortex	1741	0.508	0.662	1631	0.402	0.839
lateral occipital cortex	1739	**0.030**	-2.174	1638	0.342	-0.950
lateral orbitofrontal cortex	1742	0.169	-1.376	1638	0.333	-0.969
lingual gyrus	1695	0.727	-0.349	1590	0.159	1.408
medial orbitofrontal cortex	1732	0.493	0.686	1629	0.621	-0.495
middle temporal gyrus	1597	0.253	1.144	1516	**0.012**	-2.521
paracentral lobule	1741	0.728	-0.348	1637	0.396	-0.849
parahippocampal gyrus	1690	0.239	-1.178	1581	0.095	-1.669
pars opercularis of inferior frontal gyrus	1733	0.520	-0.643	1634	0.381	0.876
pars orbitalis of inferior frontal gyrus	1742	0.579	-0.555	1636	0.770	0.293
pars triangularis of inferior frontal gyrus	1737	0.375	0.887	1638	0.062	1.864
pericalcarine cortex	1739	0.090	-1.696	1637	0.340	0.955
postcentral gyrus	1721	0.780	-0.280	1611	0.919	0.102
posterior cingulate cortex	1740	0.568	0.571	1633	0.966	0.043
precentral gyrus	1725	0.153	1.431	1628	0.857	0.180
precuneus	1741	0.571	-0.567	1635	0.851	-0.188
rostral anterior cingulate cortex	1728	0.817	-0.231	1627	0.806	0.245
rostral middle frontal gyrus	1742	0.389	-0.862	1634	0.955	-0.056
superior frontal gyrus	1736	0.058	-1.894	1631	0.604	-0.519
superiorparietal	1735	0.578	0.556	1636	0.990	0.013

Table S28. Associations of cortical surface area AIs with IQ in all age groups combined [continued]

	IQ cases			IQ controls		
AI region	**N subjects**	**p-value**	**t-value**	**N subjects**	**p-value**	**t-value**
superior temporal gyrus	1556	0.335	-0.965	1504	0.446	0.762
supramarginal gyrus	1720	0.244	1.164	1620	0.578	0.556
temporal pole	1690	0.856	0.182	1587	0.319	-0.996
transverse temporal gyrus	1693	0.791	-0.265	1590	0.431	0.787
total average surface area	1743	0.560	-0.583	1639	0.627	-0.485

P-values (P) and t-values (t) for the effects of IQ in cases and controls are indicated. P-values in **bold** are significant at the uncorrected level ($P < 0.05$), and in ***bold-italic*** those that survive multiple testing correct.

Table S29. Associations of cortical thickness AIs with IQ in all age groups combined.

AI region	IQ cases			IQ controls		
	N subjects	p-value	t-value	N subjects	p-value	t-value
banks of superior temporal sulcus	1636	0.605	0.517	1541	0.581	-0.553
caudal anterior cingulate cortex	1737	0.348	0.938	1631	0.056	1.914
caudal middle frontal cortex	1741	0.211	-1.251	1636	0.499	0.676
cuneus	1736	0.916	0.106	1637	0.915	-0.107
entorhinal cortex	1634	0.400	0.842	1531	0.319	0.996
frontal pole	1741	0.384	0.871	1638	0.145	1.459
fusiform gyrus	1689	0.545	0.606	1587	0.066	1.842
inferior parietal cortex	1735	0.730	-0.346	1632	0.377	-0.884
inferior temporal gyrus	1653	0.312	-1.012	1574	0.172	1.367
insula	1731	0.067	1.830	1626	0.992	0.010
isthmus cingulate cortex	1738	0.503	-0.670	1631	0.953	0.059
lateral occipital cortex	1738	0.157	1.416	1639	0.706	-0.377
lateral orbitofrontal cortex	1741	0.976	0.031	1638	0.770	-0.293
lingual gyrus	1694	0.624	0.490	1590	0.279	1.083
medial orbitofrontal cortex	1732	0.516	0.650	1630	0.248	1.156
middle temporal gyrus	1597	0.724	0.354	1518	0.101	1.643
paracentral lobule	1740	0.403	-0.837	1637	0.256	-1.137
parahippocampal gyrus	1690	0.964	-0.045	1582	0.676	0.418
pars opercularis of inferior frontal gyrus	1732	0.444	0.765	1634	0.131	1.511
pars orbitalis of inferior frontal gyrus	1741	0.571	0.567	1636	0.252	-1.145
pars triangularis of inferior frontal gyrus	1734	0.412	0.820	1638	0.481	0.705
pericalcarine cortex	1736	0.418	0.809	1636	0.716	-0.364
postcentral gyrus	1718	0.209	1.256	1613	0.625	0.488
posterior cingulate cortex	1738	0.935	-0.082	1636	0.720	-0.358
precentral gyrus	1723	0.331	-0.973	1626	0.586	-0.544
precuneus	1740	0.549	-0.599	1635	0.494	-0.685
rostral anterior cingulate cortex	1727	0.532	0.625	1625	**0.014**	2.466
rostral middle frontal gyrus	1741	0.979	0.027	1635	0.398	0.845
superior frontal gyrus	1735	0.887	0.142	1631	0.891	0.138
superiorparietal	1734	0.133	1.504	1636	0.353	0.930

Table S29. Associations of cortical thickness AIs with IQ in all age groups combined [continued]

AI region	IQ cases			IQ controls		
	N subjects	p-value	t-value	N subjects	p-value	t-value
superior temporal gyrus	1556	0.187	-1.322	1507	0.711	-0.370
supramarginal gyrus	1719	0.679	0.413	1623	**0.011**	-2.552
temporal pole	1689	0.224	-1.217	1586	0.246	1.161
transverse temporal gyrus	1692	0.508	0.663	1590	0.687	0.403
total average thickness	1742	0.371	0.895	1639	0.501	0.674

P-values (P) and t-values (t) for the effects of IQ in cases and controls are indicated. P-values in **bold** are significant at the uncorrected level ($P < 0.05$), and in ***bold-italic*** those that survive multiple testing correction within the particular analysis indicated (see text).

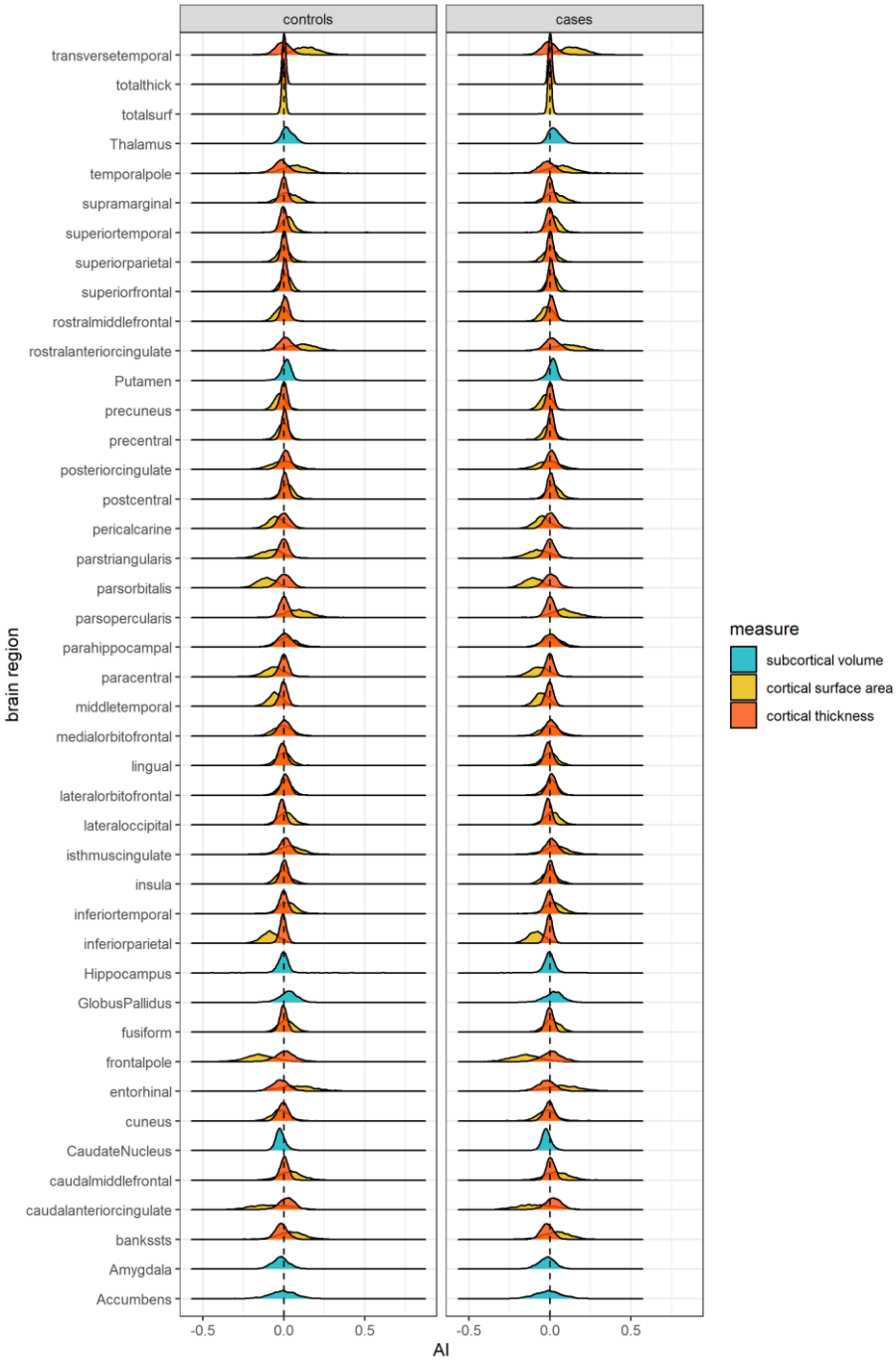


Figure S1. Joyplot of the distributions of AIs in the total study sample

(without winsorization), in ADHD cases (right) and controls (left). Shown for subcortical volumes (*cyan*), cortical surface areas (*orange*), and cortical thicknesses (*red*).

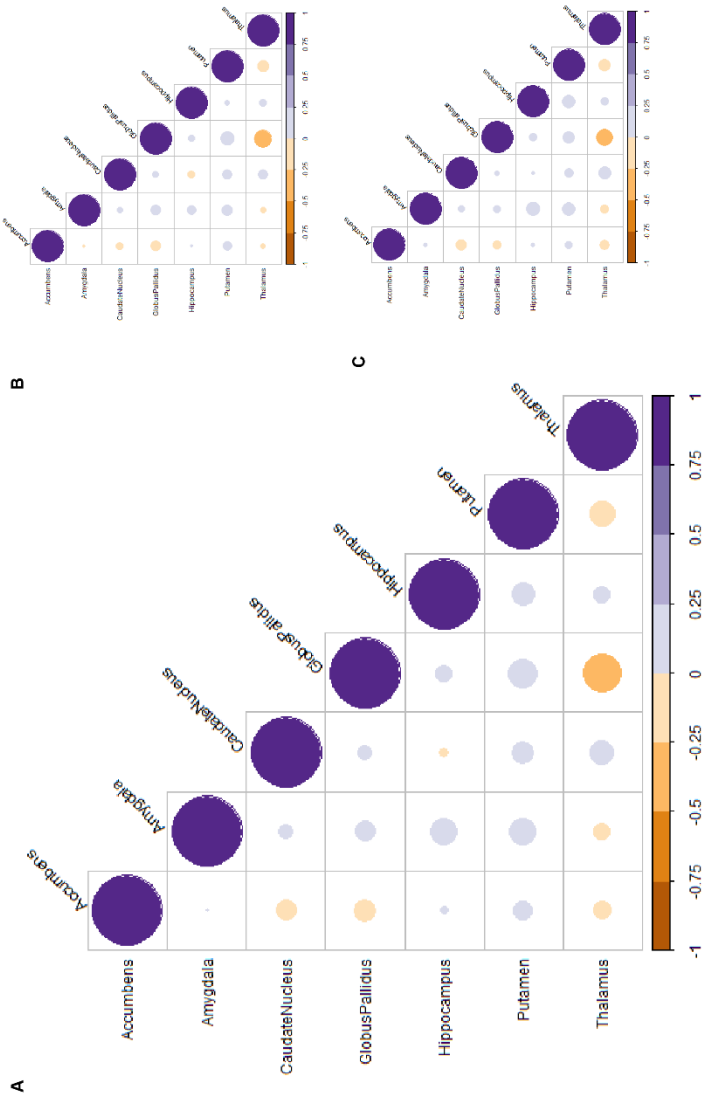


Figure S2. Correlations between AIs of subcortical volumes in (**A**) the total study sample, as well as in (**B**) cases and (**C**) controls. Correlations ranged from -0.30 (between globus pallidus and thalamus) to 0.17 (between globus pallidus and putamen). Negative correlations are in orange and positive correlations are in purple. Color intensities and circle sizes are proportional to the magnitudes of the correlation coefficients, where the area of the circles scales with the absolute value of the corresponding correlation coefficients

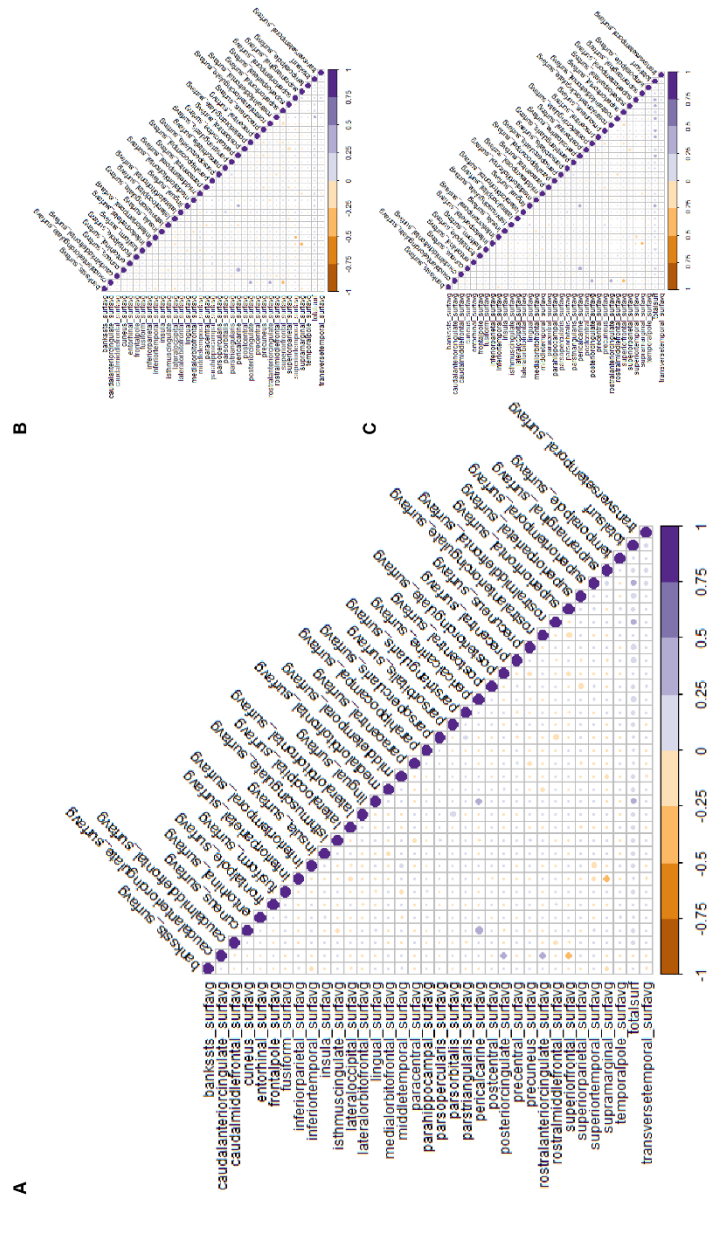


Figure S3. Correlations between AIs of cortical surface areas in (**A**) the total study sample, (**B**) cases, and (**C**) controls. Correlations ranged from -0.42 (between caudal anterior cingulate cortex and superior frontal gyrus) to 0.46 (between cuneus and pericalcarine cortex). Negative correlations are in orange and positive correlations are in purple. Color intensities and circle sizes are proportional to the magnitudes of the correlation coefficients, where the area of the circles scales with the absolute value of the corresponding correlation coefficients.

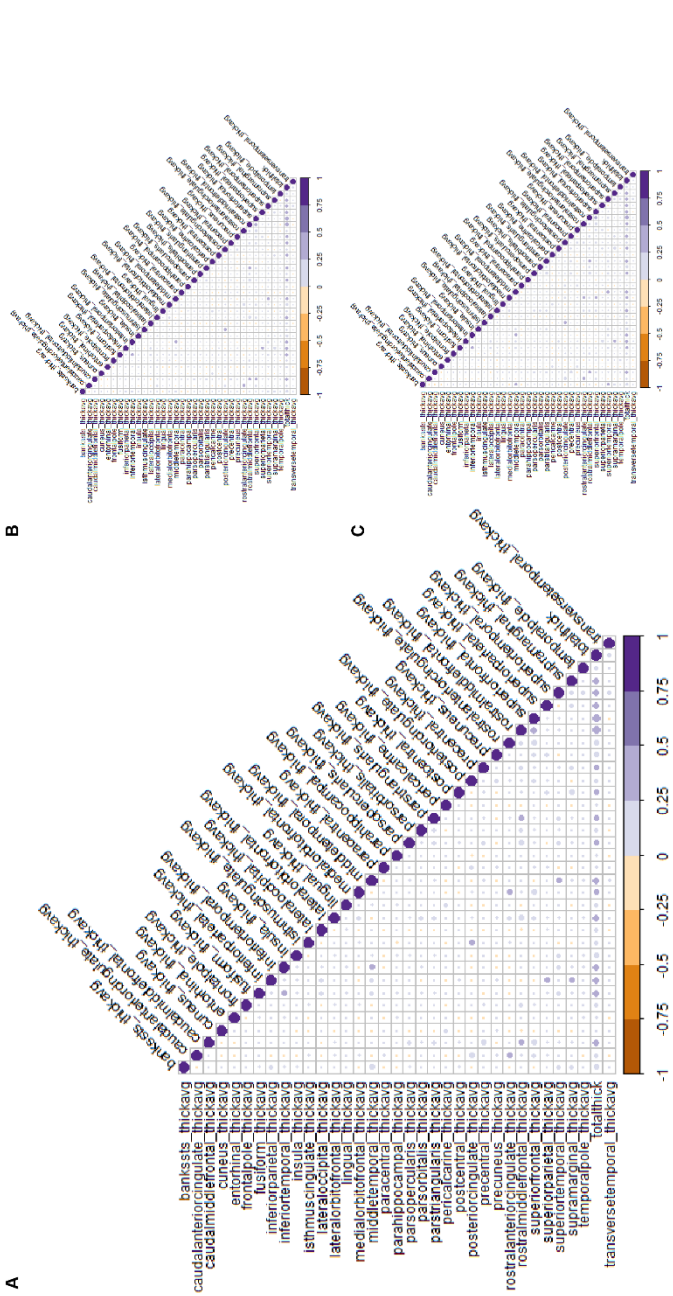


Figure S4. Correlations between AIs of cortical thickness in (**A**) the total study sample, (**B**) cases, and (**C**) controls. Correlations ranged from -0.11 (between lateral occipital cortex and rostral middle frontal cortex) to 0.49 (between rostral middle frontal cortex and total average thickness). Negative correlations are in orange and positive correlations are in purple. Color intensities and circle sizes are proportional to the magnitudes of the correlation coefficients, where the area of the circles scales with the absolute value of the corresponding correlation coefficients

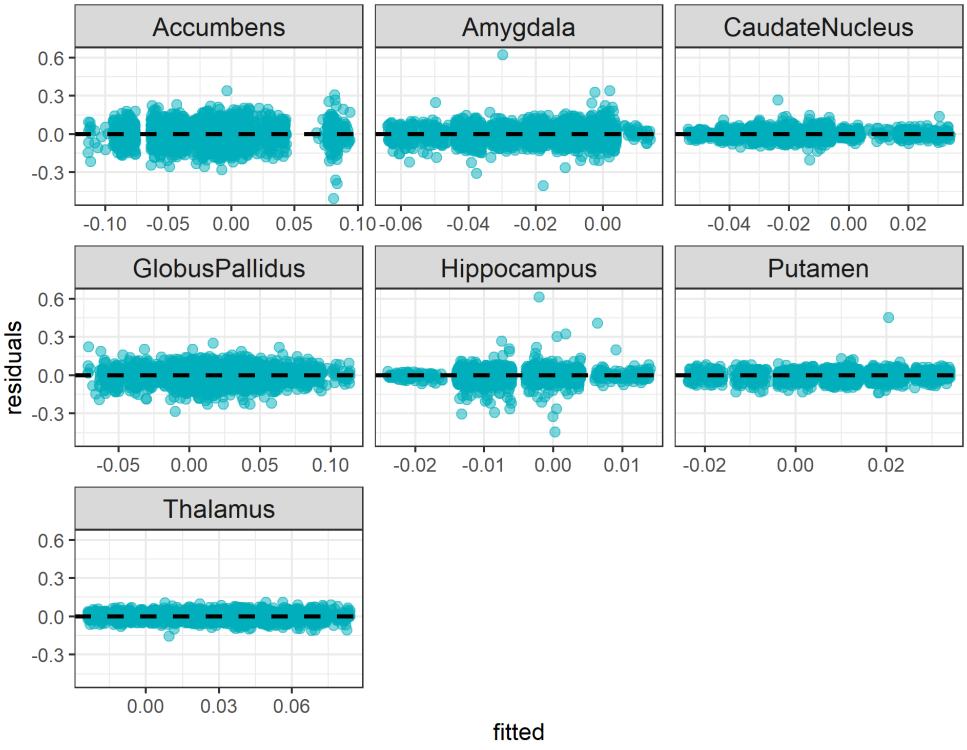


Figure S5. Residual plots of the linear mixed effects model analysis of subcortical volume AIs in the total study sample. The ggplot2 package in R was used to visualize residuals.

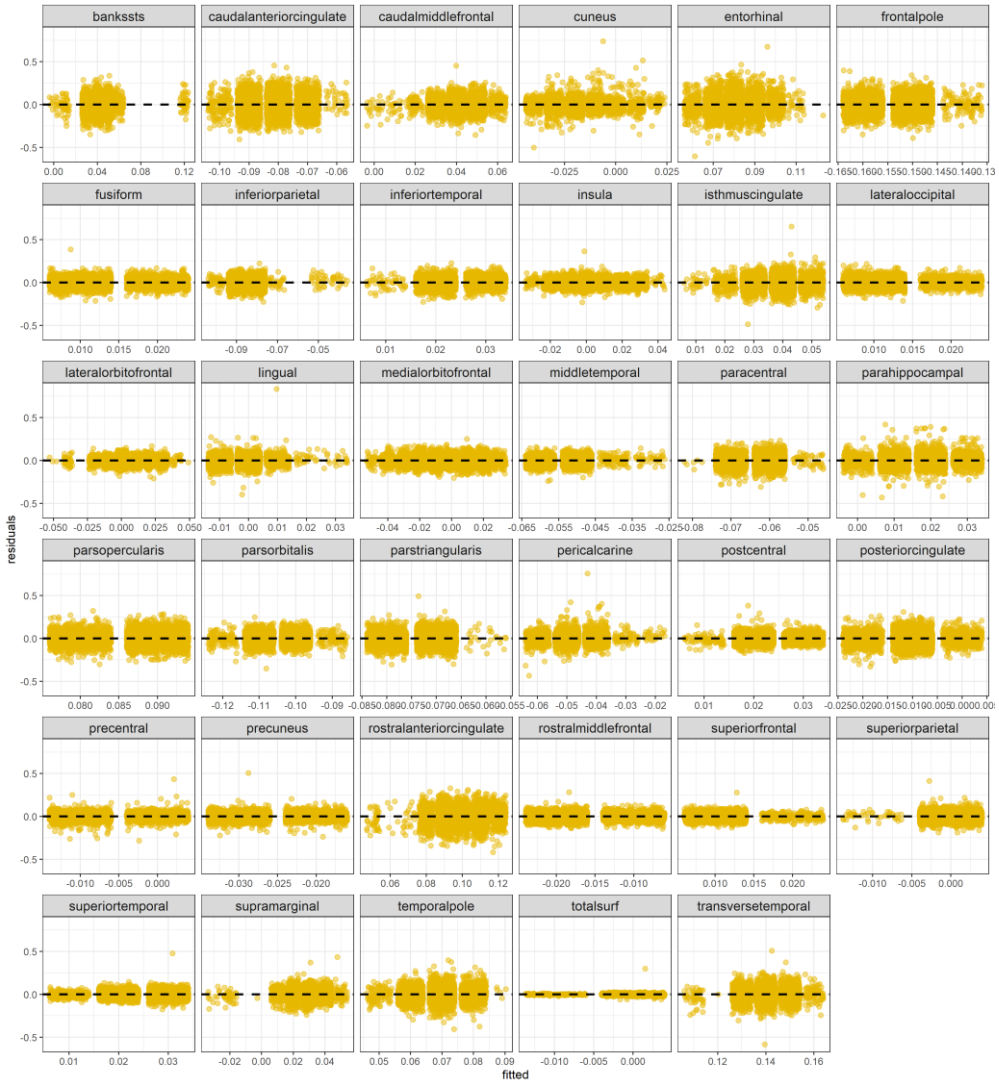


Figure S6. Residual plots of the linear mixed effects model analysis of cortical surface area AIs and the AI of the total average surface area (totalsurf) in the total study sample. The ggplot2 package in R was used to visualize residuals.

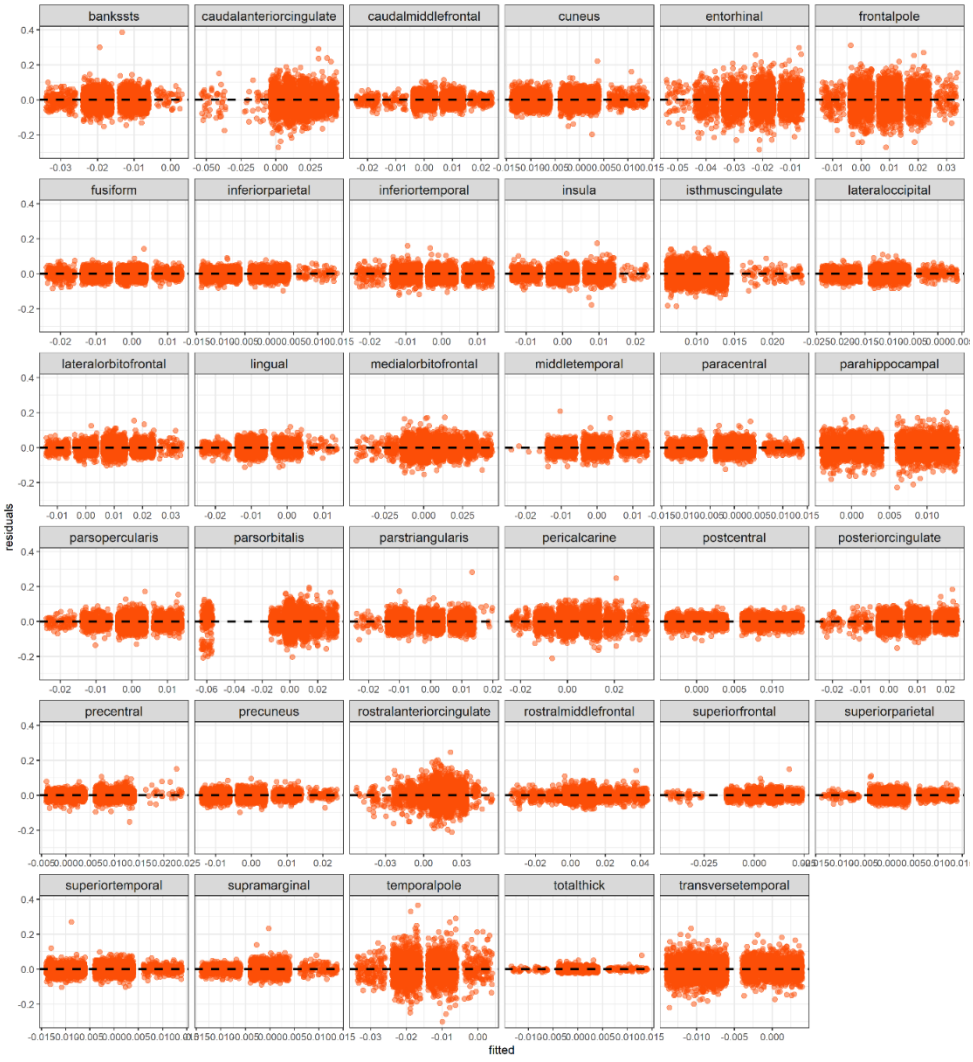


Figure S7. Residual plots of the linear mixed effects model analysis of cortical thickness AIs and the AI of the total average thickness (totalthick) in the total study sample. The ggplot2 package in R was used to visualize residuals.

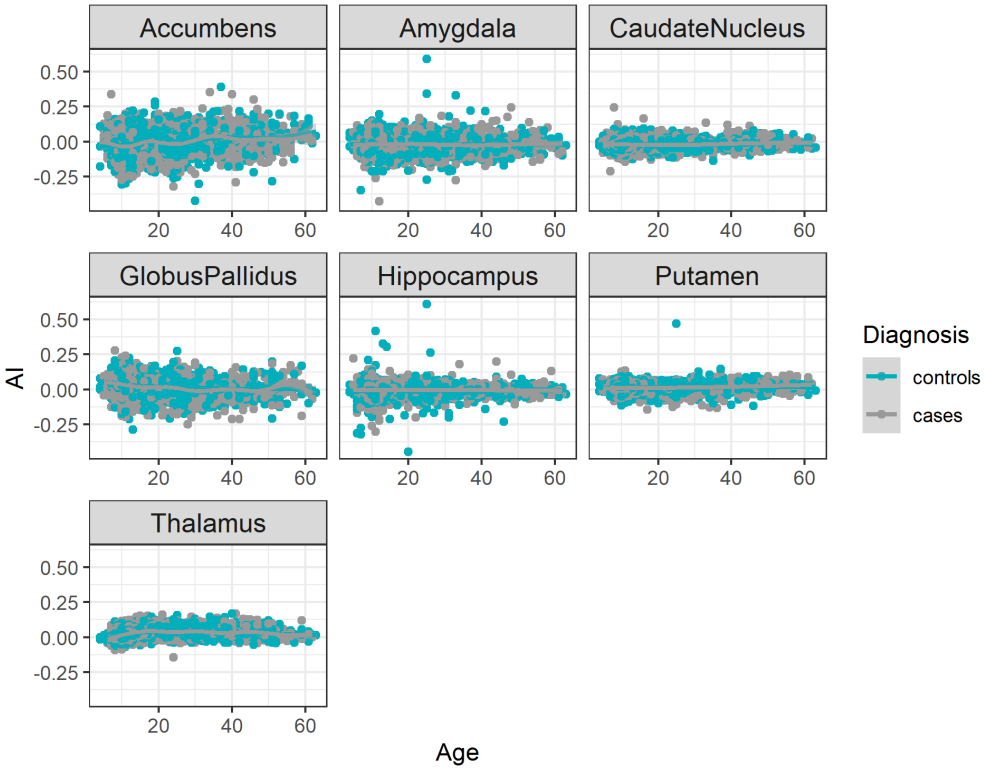


Figure S8. Scatter plots of the relationship between age and AIs of the subcortical volumes.

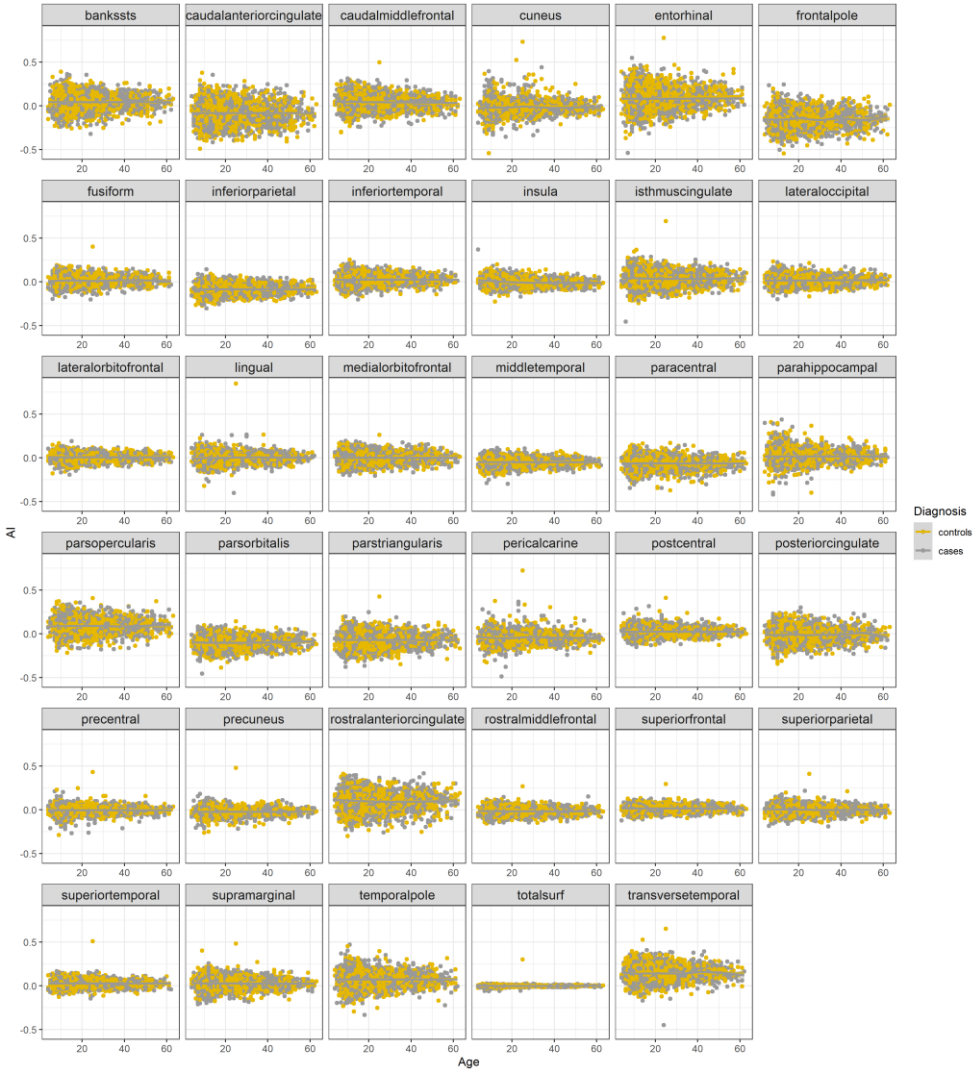


Figure S9. Scatter plots of the relationship between age and AIs of the cortical surface areas.

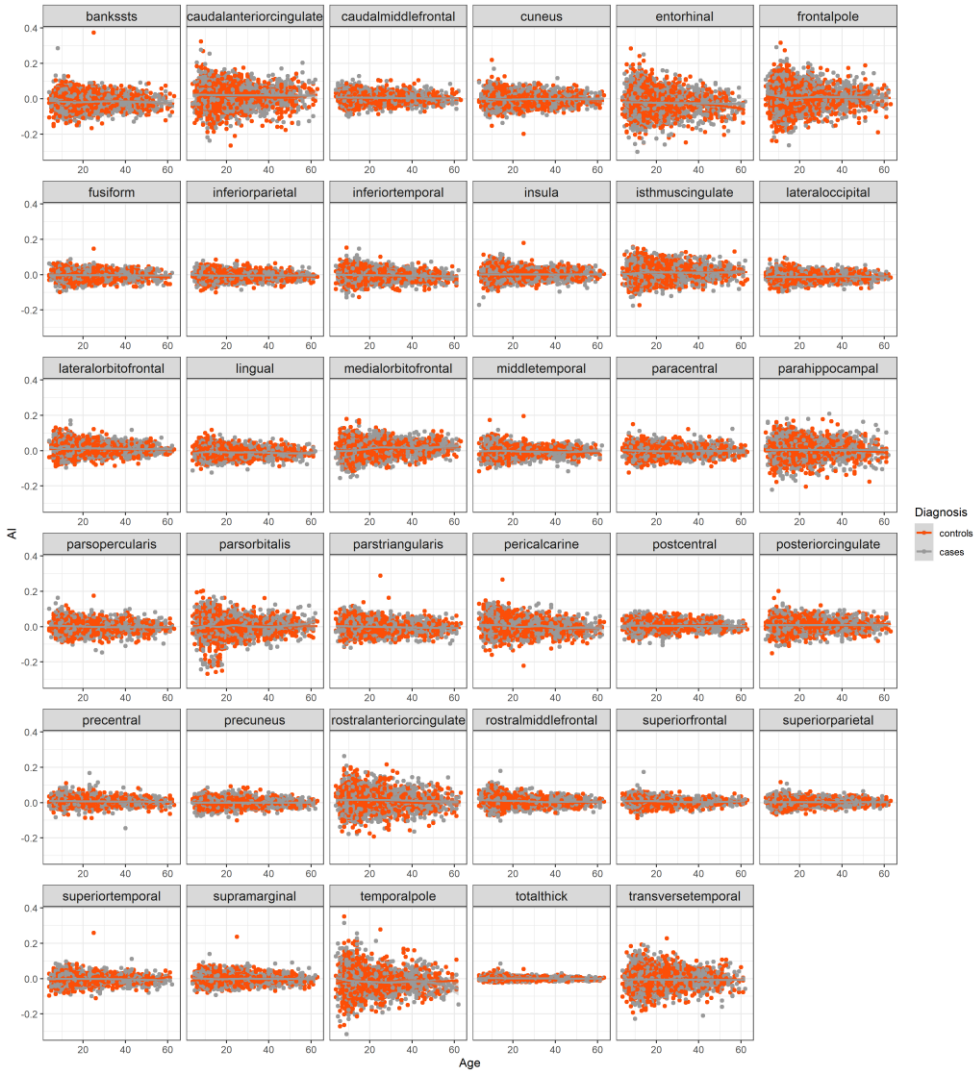
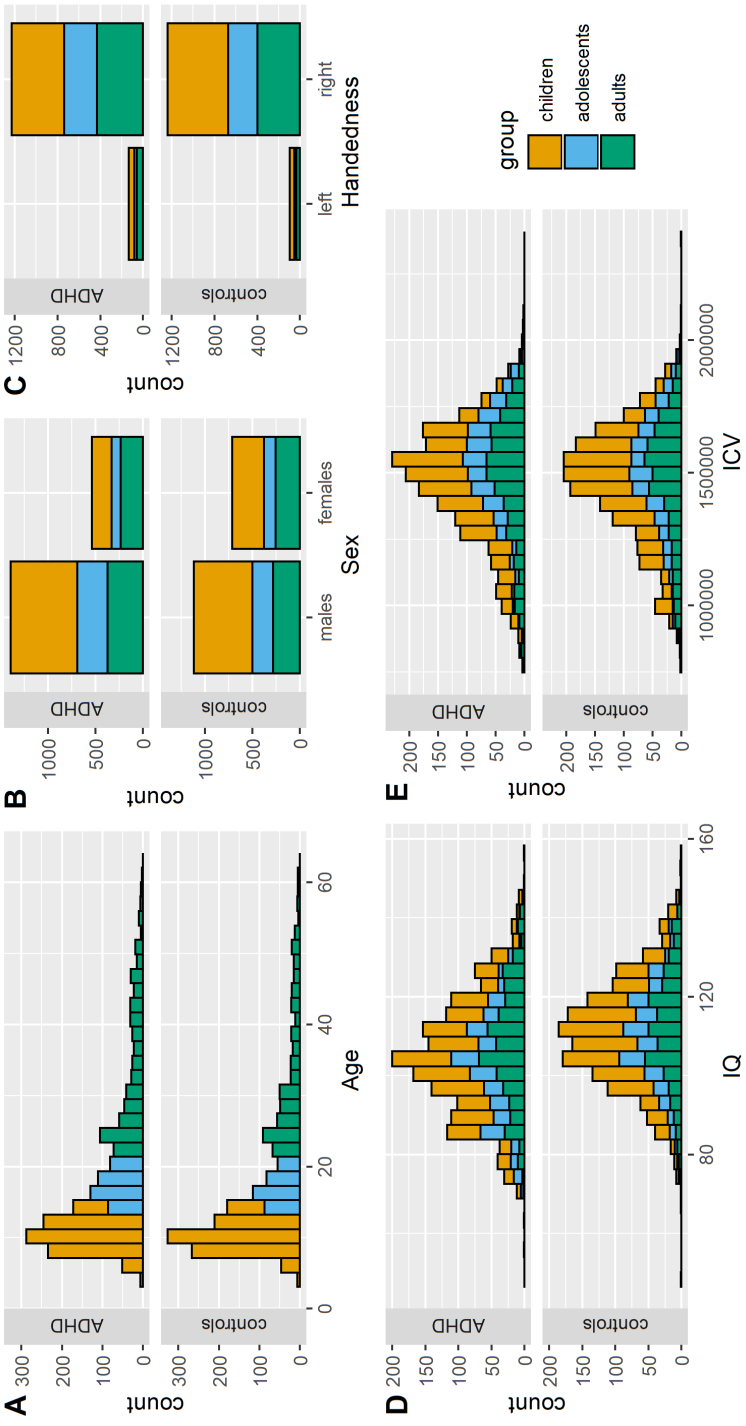


Figure S10. Scatter plots of the relationship between age and AIs of the cortical thickness.



igure S11. Distributions of (**A**) age, (**B**) sex (**C**) handedness, (**D**) IQ and (**E**) ICV in ADHD and controls, colored by children (*orange*), adolescents (*blue*) and adults (*green*)

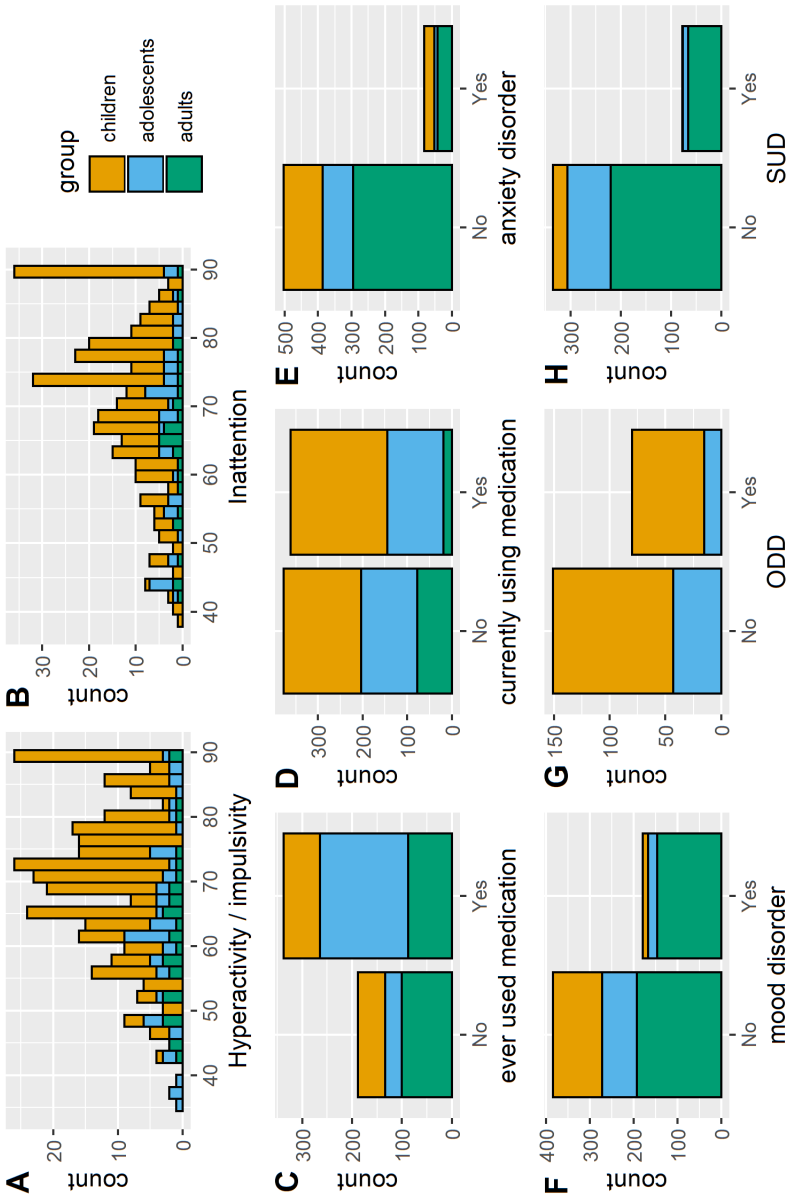


Figure S12. Distributions within ADHD cases of (**A**) hyperactivity/impulsivity symptom scores, (**B**) inattention symptom scores, (**C-D**) medication status (ever used stimulants (**C**) and currently using stimulants and (**D**)), and (**E-H**) comorbidity status ((**E**) anxiety disorder, (**F**) mood disorder, (**G**) obsessive compulsive disorder, and (**H**) substance use disorder), colored by children (*orange*), adolescents (*blue*) and adults (*green*)

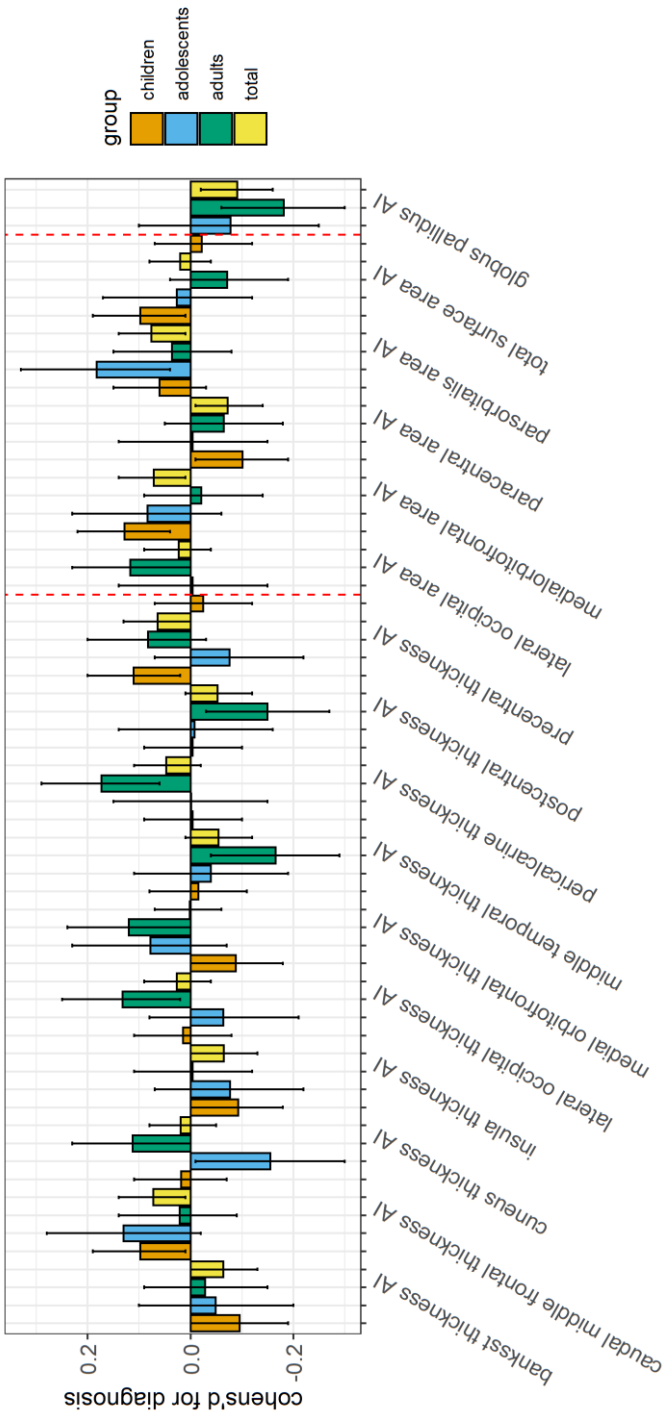


Figure 13. Bar plots of the Cohen's *d* effect sizes for diagnosis in the different age groups analyzed. Shown are only those AIs that showed a nominally significant effect of diagnosis in any of the analyses. All Cohen's *d* values above zero represent a mean shift towards greater leftward or reduced rightward asymmetry in ADHD compared to controls, and those below zero represent mean shifts towards greater rightward of reduced leftward asymmetry in ADHD compared to controls. The different age groups are shown in different colors: *orange* = children; *blue* = adolescents; *green* = adults; *yellow* = all age groups combined. The solid vertical lines reflect the error bars, indicating the 95% CI interval around Cohen's *d*, and the dotted vertical lines separate the different types of measure (i.e., thickness AIs, surface area AIs, subcortical volume AIs).

Supplementary References

1. Spitzer RL, Williams JB, Gibbon M, First MB. The Structured Clinical Interview for DSM-III-R (SCID). I: History, rationale, and description. Arch Gen Psychiatry. 1992;49(8):624-9.
2. Kaufman J, Birmaher B, Brent D, Rao U, Flynn C, Moreci P, et al. Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL): initial reliability and validity data. Journal of the American Academy of Child and Adolescent Psychiatry. 1997;36(7):980-8.
3. Conners CK. Rating scales in attention-deficit/hyperactivity disorder: use in assessment and treatment monitoring. J Clin Psychiatry. 1998;59 Suppl 7:24-30.
4. Oldfield RC. The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia. 1971;9(1):97-113.

CHAPTER

The genetics of situs inversus without primary ciliary dyskinesia

4

Abstract:

Situs inversus (SI), a left-right mirror reversal of the visceral organs, can occur with recessive Primary Ciliary Dyskinesia (PCD). However, most people with SI do not have PCD, and the etiology of their condition remains poorly studied. We sequenced the genomes of 15 people with SI, of which six had PCD, as well as 15 controls. Subjects with non-PCD SI in this sample had an elevated rate of left-handedness (five out of nine), which suggested possible developmental mechanisms linking brain and body laterality. The six SI subjects with PCD all had likely recessive mutations in genes already known to cause PCD. Two non-PCD SI cases also had recessive mutations in known PCD genes, suggesting reduced penetrance for PCD in some SI cases. One non-PCD SI case had recessive mutations in *PKD1L1*, and another in *CFAP52* (also known as *WDR16*). Both of these genes have previously been linked to SI without PCD. However, five of the nine non-PCD SI cases, including three of the left-handers in this dataset, had no obvious monogenic basis for their condition. Environmental influences, or possible random effects in early development, must be considered.

Key words: *Situs inversus*, primary ciliary dyskinesia, left-handedness, whole genome sequencing

---

This chapter is published as: Postema, M.C., Carrion-Castillo, A., Fisher, S.E. *et al*. The genetics of situs inversus without primary ciliary dyskinesia. *Sci Rep* **10,** 3677 (2020). https://doi.org/10.1038/s41598-020-60589-z

4.1 Introduction

Roughly 1:6,000-8,000 people have *situs inversus* (SI), a mirror reversal of the normal asymmetrical arrangement of the viscera (Arnold, Bixler, & Girod, 1983; Deng, Xia, & Deng, 2015). SI can occur in combination with Primary Ciliary Dyskinesia (PCD), a recessive genetic disorder which involves mutations that disrupt motile cilia (Deng et al., 2015). Cilia are hair-like organelles that protrude from the cell surface into the extracellular space (Mirvis, Stearns, & James Nelson, 2018). They are expressed in various tissues including the respiratory epithelium (Fliegauf, Benzing, & Omran, 2007), so that a disruption of ciliary motility can cause symptoms such as chronic bronchitis, inflamed or infected sinuses (Reiter & Leroux, 2017), which are often present in PCD.

Motile cilia are also expressed early in development, within an embryonic structure called the 'node', where they generate a leftward fluid flow that helps to create the left-right body axis (Fliegauf et al., 2007). The leftward direction of the nodal flow may be explained by a posterior tilt of the cilia together with their clockwise rotation, arising ultimately from molecular chirality of their component proteins (Cartwright, Piro, & Tuval, 2004; Hilfinger & Julicher, 2008; Okada, Takeda, Tanaka, Belmonte, & Hirokawa, 2005). When leftward nodal flow is absent due to recessive PCD-causing mutations, many of which affect component proteins of the ciliary cytoskeleton, it becomes a matter of chance whether the viscera will take up the normal or mirror-reversed positioning (Leigh et al., 2009). The same principle of randomization appears in mice, for example through recessive mutations in the *Iv* gene (Brueckner, D'Eustachio, & Horwich, 1989; Layton, 1976). Randomization can also explain why some human monozygotic twins are discordant for SI (Torgersen, 1950).

However, roughly three quarters of people with SI do not have PCD (Deng et al., 2015), and the causes of their SI remain largely unknown. A sibling recurrence risk of 100 times the population prevalence has been classically documented for SI (Cavalli-Sforza & Bodmer, 1971), although it is not clear whether SI individuals with PCD were distinguished from those without PCD when making this calculation, and we have been unable to obtain the primary

reference (Newcombe, 1964). A few genes have been reported to be involved in SI without PCD, and some of these can also cause partial disruptions of visceral laterality, known as heterotaxy or *situs ambiguus* (SA), including *ZIC3* [MIM:300265] (A. H. Li et al., 2019; Ware et al., 2004), *CCDC11* [MIM:614759] (Gur et al., 2017), *CFAP52* [MIM:609804, also known as *WDR16*] (Ta-Shma et al., 2015), *NME7* [MIM:613465] (Reish et al., 2016), and *PKD1L1* [MIM:609721] (Vetrini et al., 2016). The mechanisms by which these non-PCD genes influence visceral laterality are not well understood, but some code for cilia-associated proteins, rather than coding directly for cytoskeletal components of cilia.

A recent exome sequencing study of 323 people with laterality defects included 11 that had SI, while the rest had SA or cardiac malformations such as dextrocardia (A. H. Li et al., 2019). Of the 11 people with SI, only one had a clear candidate recessive mutation, in the known SA gene *MMP21*. PCD was not an inclusion criterion for that study, and none of the people with SI had recessive mutations in genes known to cause PCD. One additional participant with SI had a dominant mutation in *ROCK2*, and another had an X-linked mutation in *RAI2* – both of these genes are important for heart development, but it is not clear whether mutations in either were causative for SI (A. H. Li et al., 2019). That was the largest previous study, of which we are aware, that screened the exome for mutations in people with non-PCD SI.

Approximately 90% of people are right-handed and have left-hemisphere dominance for language, among other lateralized functions, but the developmental basis for brain asymmetry remains unknown(de Kovel et al., 2017; de Kovel, Lisgo, Fisher, & Francks, 2018; Francks, 2015). One possibility is that early embryonic mechanisms that give rise to asymmetries of the visceral organs also impact on brain asymmetries. Indeed, such a link has already been established in the *frequent-situs-inversus (fsi)* line of zebrafish (Barth et al., 2005). One study of humans reported a possible genetic link between a continuous measure of left-versus-right hand motor skill and genes involved in visceral laterality, based on analyzing genetic variants which are common in the population (Brandler et al., 2013). However, the sample size of under 3000 subjects was low for complex-trait genome-wide association analysis using common genetic variants. A much

larger study of over 300,000 subjects from the UK Biobank found no support for a link of left-handedness to genes involved in visceral asymmetry (de Kovel & Francks, 2019). The same large study identified an association of a common variant at the microtubule-associated gene *MAP2* with left-handedness, with a small effect (de Kovel & Francks, 2019).

Intriguingly, the general population rate of 10-15% left-handedness is not altered in SI with PCD (Afzelius & Stenram, 2006; McManus, Martin, Stubbings, Chung, & Mitchison, 2004). This implies a developmental dissociation between brain laterality for hand motor control and nodal-ciliary visceral patterning. However, in the only previous study of handedness in SI to specifically target non-PCD individuals, Vingerhoets *et al*. (Vingerhoets et al., 2018) recently reported that five of nine non-PCD SI subjects were left-handed. Assuming a population rate for left-handedness of 10%, an exact, one-tailed binomial test gives P = 0.0009 for this elevated rate of left-handedness in non-PCD SI. A further six SI subjects studied by Vingerhoets *et al.* (Vingerhoets et al., 2018) had PCD – but, as expected, they showed no elevated rate of left-handedness (1 out of 6).

The data of Vingerhoets *et al*. (Vingerhoets et al., 2018) suggest the existence of developmental mechanisms that link brain asymmetry and visceral laterality, independently of genes involved in PCD, although the small sample size cannot provide proof of this. In fact, a study from 1950 included 270 people with SI that were identified based on mass x-ray photography of the Norwegian population, and left-handedness was not over-represented compared to the general population(Torgersen, 1950). The ascertainment method in that study suggests that most SI individuals did not necessarily have PCD, although this was not investigated formally(Torgersen, 1950). If most SI individuals in the 1950 study did not have PCD, it would indicate that there is no elevation of left/handedness in non/PCD SI, and therefore that the elevated rate of left-handedness among the nine non-PCD SI participants of Vingerhoets *et al.* (Vingerhoets et al., 2018) was a chance finding in a small sample, or else an unusual feature of this particular set of subjects. Like the 1950 study, an even older study, from 1938, also reported a normal population rate of left-handedness in SI, although here ‘most of those included were examined on account of illness or a congenital malformation’ (Cockayne,

1938) . In the 1938 study, therefore, it seems possible that many of the SI individuals had PCD, although further information was not given (Cockayne, 1938). Left-handedness has a heritability of roughly 25% based on twin and family data (Medland et al., 2009), but only around 2% based on genome-wide genotype data for common polymorphisms, within the UK Biobank dataset (Neale, 2017). However, it has been proposed that left-handedness may sometimes occur due to genetic mutations that are rare in the population, but with substantial effects on brain laterality when present (Armour, Davison, & McManus, 2014; Kavaklioglu, Ajmal, Hameed, & Francks, 2016). Rare genetic effects are not well captured in genome-wide association or SNP-heritability studies, which are typically based on common variation in the genome (de Kovel & Francks, 2019).

Here we performed an exploratory whole genome sequencing study in the same set of 15 SI subjects studied by Vingerhoets *et al*.(Vingerhoets et al., 2018) (six with PCD and nine without), as well as 15 healthy controls matched for age, sex, education, and handedness (**Table 1**). The primary goal was to identify rare, highly penetrant mutations in the nine people with non-PCD SI, as relatively little is known about the genetic contributions to non-PCD SI. As these nine subjects showed an elevated rate of left-handedness, our study also had the potential to yield novel insights into the developmental origins of left-right patterning of both brain and body. The six SI subjects with PCD were included as positive controls, in which we expected to find recessive mutations in known PCD-causing genes.

4.2 Methods

4.2.1 Participants

Table 1 gives an overview of the participants and their characteristics. Fifteen people with radiologically documented SI (mean age = 33 years, 8 males), and 15 controls with normal *situs* matched for age, sex, education and handedness (mean age = 33 years, 8 males), were recruited from Ghent University Hospital and Middelheim Hospital Antwerp. All participants were tested and scanned at Ghent University Hospital following ethical approval by the “Commissie voor Medische Ethiek, Universitair Ziekenhuis Gent”,

study approval number 2012/073. All participants gave informed consent for DNA sample collection and genomic analysis in relation to body, brain and behavioural laterality. All methods were performed in accordance with the relevant guidelines and regulations.

Details about the recruitment, diagnosis and selection/exclusion procedure have been described previously (Vingerhoets et al., 2018). Radiological information (RX or CT) of thorax and complete abdomen was available for eight SI participants, and of thorax and upper abdomen in seven SI participants. The medical reports confirmed that all 15 participants presented with radiologically documented SI.

In five participants with SI, a formal diagnosis of PCD or Kartagener syndrome was found in their medical records. Kartagener syndrome refers to the clinical triad of SI, chronic sinusitis, and bronchiectasis(Leigh et al., 2009). A sixth SI-participant was identified on account of a radiological consultation regarding infertility. The participant also presented with chronic sinusitis and mild chronic bronchitis. Although no formal diagnosis of PCD was obtained in this case, the presence of chronic upper and lower respiratory infection and infertility in an individual with SI warrants suspicion of Kartagener syndrome. Consequently, we included the participant within the PCD group. In addition, all six members of the PCD group reported having a daily wet cough, and had PICADAR scores of between 8 and 12 (Behan et al., 2016), and thus predictive probabilities of having PCD of between 66% and 99%, based on this recently-developed, questionnaire-based tool (note that five of these subjects anyway had formal medical diagnoses).

A seventh SI subject (SI03) had no medical record pertaining to PCD but did report daily wet cough. This subject had a PICADAR score of 8. In the study of Vingerhoets *et al.* this SI subject was classified as non-PCD, before the PICADAR assessment was available (Vingerhoets et al., 2018). For the purposes of the present study, we also treated this subject as non-PCD given the lack of formal medical history of PCD and the ambiguous PICADAR score, as well as the fact that we found no likely recessive mutations in PCD-causing genes in this subject (see Results).

Eight other SI subjects had no medical record of PCD or PCD-like symptoms, and were classified as non-PCD SI cases. Six of these reported no daily wet cough, and two did not answer in this regard. PICADAR scores can only be calculated in the presence of daily wet cough, so that none of these eight cases received PICADAR scores. Three of these cases had been previously diagnosed with congenital heart disease that required surgical treatment, and their radiological files all referred to their cardiac condition. Congenital heart disease is a frequent comorbidity of SI, as the cardiac circulation appears particularly sensitive to perturbation in normal left-right positional information (Kosaki & Casey, 1998).

Handedness was assessed using the Edinburgh Handedness Inventory (EHI) (Oldfield, 1971). All participants had EHI scores > |0.5| (**Table 1**), so that none were ambidextrous (J. Robinson, 2013). Note that one non-PCD SI subject reported being forced to switch from left- to right-handedness in childhood, in which case five out of nine non-PCD SI cases were naturally left-handed. One of the six cases with PCD also reported forced left-to-right switching, otherwise the rest were right-handed (**Table 1**).

4.2.2 Whole Genome Sequencing (WGS) and Pre-processing

DNA was extracted from saliva samples using the Oragene kit (Oragene). WGS was performed by Novogene (Hong Kong) using Illumina high throughput sequencing (HiSeq-PE150), creating paired end reads with a length of 150 base pairs (bp). Raw reads, stored in BAM files, were aligned to the human reference genome (the extended "decoy" version of b37) using Burrows-Wheeler Aligner (BWA) software (H. Li & Durbin, 2009), and sorted and reordered using SAMtools (v1.3.1) (H. Li et al., 2009). PCR duplicates, which could arise during cluster amplification, were marked using Picard (v2.9.0). Genome Analysis Toolkit (GATK v3.7) (McKenna et al., 2010) software was used to realign reads around insertions/deletions (indels) and to recalibrate base quality scores per sample.

4.2.3 Variant Calling and Quality Control

Indels (insertion/deletions) and single nucleotide polymorphisms (SNPs) were called as recommended by the GATK Best Practices workflow. Variant Quality Score Recalibration (VQSR) was performed, and variants with a high probability of being false positive were flagged according to their sensitivity tranche (90, 99, 99.9 and 100). All SNPs and indels within a VQSR tranche of 99.9% or higher were discarded. Variants with a quality depth ≤ 9 or a call rate ≤ 0.8 were also removed. Vcftools v1.13 was used to create a summary table in the Variant Call Format (vcf). A total of 13,989,941 SNPs and indels were identified across the 30 subjects, with an average number of 5,186,055 (min= 5,053,188, max = 5,272,561) alternative alleles per subject (i.e. different to the reference genome, build37 decoy version).

4.2.4 Genomic-level evaluation

For overall descriptive analysis of the participant genomes, a subset of 40,387 variants distributed genome-wide was used that had known Minor Allele Frequencies (MAFs) > 0.1, and were the result of pruning to be in low linkage disequilibrium with one another. For this, the flag --indep-pairwise function in Plink (v.1.90b3w) (Purcell et al., 2007) was applied with a pairwise linkage disequilibrium r^2 greater than 0.2, based on a SNP-SNP correlation matrix of 1500 by 150 in window size. The resulting data were then used as the basis for inferring pairwise identity by descent (IBD) sharing between subjects (i.e. genetic relatedness), inbreeding, and inconsistencies with reported sex, using the Plink operations --genome, --ibc/het, and --check-sex.

A different subset of 61,467 independent variants distributed genome-wide was used for visualizing population structure through Multi-Dimensional Scaling (MDS) analysis with respect to the 1000 genomes reference dataset (v37) (Auton et al., 2015) (downloaded from: ftp://climb.genomics.cn/pub/10.5524/100001_101000/100116/1kg_phase1_all.tar.gz) using the Plink operation –mds-plot.

4.2.5 Annotation and filtering of single nucleotide variants affecting protein sequences

Variants were normalized using the software tool *vt normalize* (v0.5772-60f436c3) (Tan, Abecasis, & Kang, 2015), which ensures consistent representation of variants across the genome. Normalized SNVs were annotated using Annovar (Wang, Li, & Hakonarson, 2010) and Variant Effect Predictor (v88) (McLaren et al., 2016). Gemini (v 0.20.0) was used to select protein coding variants with 'MEDIUM' or 'HIGH' impact severity annotations, as well as non-coding variants with 'HIGH' impact severity annotations (in practice those altering splice donor or acceptor sites). Additional filtering was done in R and comprised the removal of synonymous variants, and of 'MEDIUM' variants with a PolyPhen (Adzhubei et al., 2010) prediction of "benign" unless annotated as 'deleterious' by Sift (Kumar, Henikoff, & Ng, 2009) , and those annotated as "tolerated" by Sift unless predicted to be 'probably damaging' by PolyPhen. Data were then processed and analyzed separately under recessive and dominant models:

4.2.5.1 Recessive model

For the recessive model we further excluded variants with a known population frequency ≥ 0.005 in any of the following databases: GNOMAD (Lek et al., 2016), ESP (Fu et al., 2013), 1000 Genomes (Auton et al., 2015) and ExAC databases (Lek et al., 2016). The remaining low-frequency variants were considered as putative mutations. Gene-level mutation counts per subject were then made, with a given gene being assigned as recessively mutated when it carried two copies of the same mutation (homozygous) or two different mutations (possible compound heterozygous). Integrative Genome Viewer (IGV v2.3.97) (J. T. Robinson et al., 2011) was used to visualize the possible compound heterozygous mutations, and genes carrying these were discarded as recessive candidates when both mutations were definitely present on the same allele (i.e. "in phase") on a given sequence read. Finally, genes recessively mutated according to these criteria in any of the fifteen unaffected control subjects were excluded as being potentially causative in cases, and also removed for the purposes of Gene Set Enrichment Analysis (GSEA; see below): this step had the advantage of removing false variants from potential technical artifacts, or variants which are common in the population but not annotated as such in on-line databases. In total, 17 genes were excluded based on overlap with the unaffected control subjects.

Genes on chromosome X were processed as part of the recessive pipeline, such that females would need to carry two mutations in a given gene, and males one mutation.

4.2.5.2 Dominant model

A maximum population frequency threshold of 5×10^{-5} was applied in this case, and genes carrying at least one rare variant according to this criterion were considered as potentially causative. Again, genes mutated according to this criterion in any of the fifteen unaffected control subjects (N = 47 genes) were excluded as being potentially dominantly causative in cases, and removed for the purposes of GSEA analysis (below).

4.2.6 Gene Set Enrichment Analysis

To test whether a list of mutated genes in a given set of subjects contained functionally related genes across subjects, we performed Gene Set Enrichment Analysis (GSEA) using the g:Profiler R package (version 0.6.1) (Reimand et al., 2016). A gene classification scheme derived from the Gene Ontology (GO) database (Ashburner et al., 2000, 2017) was used. This specified a total of 6380 functionally defined gene-sets, based on a background of 19,327 genes with functional annotations. Many genes are not represented in the GO, particularly mRNA transcripts of no known function or homology, so that the numbers of mutated genes in a given set of subjects was always higher than the subset used as input for GSEA.

Mutated genes on chromosome X were combined with recessively mutated autosomal genes for the purposes of GSEA. GSEAs were performed separately for the mutated gene lists in SI subjects with PCD (54 genes, of which 40 are in the GO), non-PCD SI subjects (60 genes, 38 in the GO), left handed non-PCD SI subjects (42 genes, 26 in the GO), and non-PCD SI cases that were unsolved under the recessive model (27 genes, 13 in the GO).

For the dominant model, GSEA was performed separately for mutated genes in the non-PCD SI subjects (330 genes, 271 in the GO), the subset of non-PCD SI subjects that remained unsolved under the recessive model (183 genes, 148 in the GO), and the left handed subset of non-PCD subjects (201

genes, 163 in the GO). PCD subjects were not tested under a dominant model, as PCD is known to be a recessive phenotype.

In order to confirm that unaffected controls showed no significant pathway enrichment among their mutated genes, a mirrored exclusion was performed whereby any genes mutated in cases were excluded from the control gene lists. This resulted in 56 genes exclusively mutated in controls for the recessive model (of which 34 genes are in the GO), and 533 genes exclusively mutated in controls under the dominant model (440 in the GO).

The identities of genes were first converted using the g:Convert tool (Reimand et al., 2016) to ensure recognition by the GO schema. The following settings were then used for GSEA: minimum set size = 15, maximum set size = 500, minimum intersect number = 2, hierarchical filtering = moderate. P-values were corrected for multiple testing across gene sets, based on the gSCS correction method in g:Profiler, separately for each input list of mutated genes corresponding to a given set of subjects. This method of multiple testing correction takes into account the hierarchical structure of the sets (Reimand et al., 2016). We applied a cut-off P value of adjusted 0.01.

4.2.7 Candidate Gene Lists

We queried the genetic data in R (v3.3), based on searching the Clinvar (Landrum et al., 2018) column of our annotated variant data for the terms: “situs inversus”, “heterotaxy | heterotaxia | situs ambiguus”, “PCD | ciliary dyskinesia | Kartagener”, “left-right”, and “asymmetry | laterality”.

A candidate gene list was also created which included genes from the literature that were suggested to be associated with human laterality phenotypes and/or ciliopathies, but not (yet) present in our initial list of candidate genes (**Table S1**). Specifically, a list of ciliopathy genes from a review of this topic (Reiter & Leroux, 2017) was searched for the same terms as mentioned above plus ‘congenital heart disease’, yielding 57 candidate genes. Additionally, a list of genes potentially associated with human laterality disorders, as compiled in a 2015 review (Deng et al., 2015), resulted in the addition of 38 candidate genes. Five more genes were included that

were considered as possibly causal of laterality defects in a recent exome sequencing study (A. H. Li et al., 2019) (see Introduction). Eleven genes containing variants that were previously reported to associate with measures of hand preference(de Kovel & Francks, 2019) or *planum temporale* asymmetry(Carrion-Castillo et al., 2020) were also added to the list of candidate genes. Finally, 47 more genes – arising from a search for the 'situs inversus' phenotype - were retrieved from the Mouse Genome Database (MGD) (Smith, Blake, Kadin, Richardson, & Bult, 2018) at the Mouse Genome Informatics website, the Jackson laboratory, Bar Harbor, Maine. (URL: http://www.informatics.jax.org) [Oct, 2017].

4.2.8 Structural Variants

For all participant genomes, structural Variants (SVs) (>50 kilobases in length) were called using a combination of two SV calling algorithms: CNVnator (v0.3.3) (Abyzov, Urban, Snyder, & Gerstein, 2011) and Lumpy (v 0.2.13) (Layer, Chiang, Quinlan, & Hall, 2014). These algorithms complement each other regarding the kinds of signals in WGS data that they can detect, as CNVnator is based on read-depth, whereas Lumpy is based on paired-end mapping (Pirooznia, Goes, & Zandi, 2015). For CNVnator the bin size was set to 100 base pairs for all genomes except for two, where it was 150 base pairs (we determined a roughly optimal bin size for each subject's genome, such that that the average read depth for that genome would be at least 4 standard deviations from zero). Lumpy was run using default parameters in "lumpyexpress".

SVs were then annotated using SV2 (Antaki, Brandler, & Sebat, 2017). Variants that were present in any of the 15 healthy controls were removed from consideration as potentially causative for SI. Only variants that passed the default SV2 filtering criteria for quality were included (Antaki et al., 2017).

4.3 Results

4.3.1 Protein-altering single nucleotide variants

4.3.1.1 Recessive mutations

Our variant calling, filtering and annotation pipeline produced between 5 and 15 recessively mutated genes per SI subject. We included six SI subjects with PCD as positive controls, in order to ensure that the variant calling and mutation definition criteria were well calibrated. As PCD is known to be a recessive phenotype for which at least 30 different genes have already been identified (Deng et al., 2015), we expected most, or all, of these six subjects to have identifiable mutations in known PCD-causing genes. As expected, each of these six cases had just one recessively mutated gene which was annotated 'Kartagener' or 'PCD' in the Clinvar database (Landrum et al., 2018), and was therefore the most likely monogenic cause for their condition (**Table 1**). These genes were *LRRC6* [MIM:614930], *DNAH11* [MIM:603339], *DNAAF1* [MIM:613190], *CCDC114* [MIM:615038], and *DNAH5* [MIM: 603335] (the latter mutated in two SI cases with PCD, consistent with *DNAH5* being the most common cause of PCD in European-ancestry populations (Hornef et al., 2006)) (**Table 1**). The PCD subject with a homozygous *LRRC6* mutation (subject SI06) was the only individual to show an elevated inbreeding coefficient and non-European ancestry (**Table S2, Figure S1**).

None of the fifteen unaffected control subjects had any recessively mutated genes annotated 'Kartagener', 'PCD', 'SA' or 'SI' in Clinvar.

Surprisingly, two of the nine non-PCD SI cases had recessive mutations in genes annotated 'Kartagener' in the Clinvar database (Landrum et al., 2018). These were subject SI12 (again involving mutations in *DNAH5*), and subject SI16 (in *CCDC151* [MIM:615956]) (**Table 1**). As these subjects had no medical records pertaining to PCD-like symptoms, and no daily wet cough, then the findings suggest reduced penetrance for PCD.

Two of the nine non-PCD SI subjects had recessive mutations in genes annotated in Clinvar with '*heterotaxy*' and/or '*situs inversus totalis'*, but not annotated as PCD-causing (**Table 1**). These genes were *PKD1L1* [MIM: 617205] in subject SI02, and *CFAP52* [MIM: 609804] in subject SI14. A homozygous missense mutation in *PKD1L1* was previously reported in an

Table 1. The most likely causal recessive mutations in the fifteen SI subjects.

Subj	SI group	Sex/ Age	EHI	NH	CHD	Daily wet cough	Type	Gene	Clinvar	rs ID	Start position	ref	alt	MAF	AAC	impact
SI02	non-PCD	M/50	0.9	R	0	?	hom	*PKD1L1*	SA | SI	rs528302390	47870810	T	TCA	0.00109	-	splice donor
SI03	non-PCD	F/26	-0.8	L	0	yes		unsolved								
SI04	non-PCD	M/23	-1	L	1	no		unsolved								
SI05	non-PCD	M/27	0.9	R	0	no		unsolved								
SI07	non-PCD	F/35	0.9	R	1	?		unsolved								
SI09	non-PCD	F/36	0.7	L[§]	0	no		unsolved								
SI12	non-PCD	F/40	0.9	R	0	no	chet	*DNAH5*	Kartagener	None	13900415	A	C	-	E/*	stop gained
							chet	*DNAH5*	Kartagener	None	13769691	G	GC	0.00002	A/X	frameshift
SI14	non-PCD	M/18	-0.8	L	1	no	chet	*CFAP52/WDR16*	SI	rs148905544	9489207	T	A	0.00066	H/Q	Missense variant
							chet	*CFAP52/WDR16*	SI	rs374035006	9536224	C	T	0.00012	R/*	Stop gained
SI16	non-PCD	M/21	-1	L	0	no	hom	*CCDC151*	Kartagener	rs765121016	11533429	T	TCTC	0.00011	E/-	inframe deletion
SI06	PCD	M/46	1	R	0	yes	hom	*LRRC6*	Kartagener	rs767624733	133687728	T	C	0.00017	-	splice donor
SI08	PCD	F/23	0.9	R	0	yes	hom	*DNAH11*	PCD	rs373706559	21659620	A	C	0.00006	Y/*	stop gained
SI11	PCD	F/32	0.9	R	0	yes	chet	*DNAAF1*	Kartagener	rs569633512	84203963	C	T	-	-	splice donor
							chet	*DNAAF1*	Kartagener	rs373103805	84193302	G	C	0.00023	N/K	missense
SI13	PCD	M/48	0.6	L[§]	0	yes	hom	*CCDC114*	Kartagener	rs779459076	48814907	CACG	C	0.00093	-/R	Inframe insert
SI15	PCD	F/31	0.7	R	0	yes	chet	*DNAH5*	Kartagener	None	13786289	A	T	-	K/*	stop gained
							chet	*DNAH5*	Kartagener	rs397515540	13753397	G	GA	0.00034	D/X	frameshift

Table 1. The most likely causal recessive mutations in the fifteen SI subjects [continued]

Subj	SI group	Sex/ Age	EHI	NH	CHD	Daily wet cough	Type	Gene	Clinvar	rs ID	Start position	ref	alt	MAF	AAC	impact
SI17	PCD	M/39	0.5	R	0	yes	chet	*DNAH5*	Kartagener	rs548521732	13839638	C	T	0.00021	-	splice acceptor
							chet	*DNAH5*	Kartagener	rs769458738	13753597	T	C	0.00004	R/H	missense

Subjects shown with two mutations have possible compound heterozygous mutations, otherwise they have homozygous mutations. The Genome Reference Consorium (GRC) build 37 decoy version was used as reference sequence. EHI: Edinburgh Handedness Inventory score; NH: natural handedness; CHD: Congenital Heart Disease. Type: type of genetic mutation, i.e., homozygous (hom) or compound heterozygous (chet); MAF: minor allele frequency in population databases, if known; AAC: amino acid change. [§]Self-identified natural lefthander made to convert to right-handedness.

individual with SI and congenital heart disease but no PCD, as well as recessive splicing mutations in two individuals with heterotaxy (Vetrini et al., 2016). Our subject SI02 had no diagnosis of congenital heart disease (CHD), and was right-handed (**Table 1**). As *PKD1L1* is a known recessive cause of non-PCD SI, we consider this gene to be the most likely cause of the non-PCD SI in subject SI02. As regards *CFAP52* (also known as *WDR16*), a homozygous deleterious deletion was previously reported to segregate with situs anomalies, including heterotaxy and situs inversus totalis, in a consanguineous family (Ta-Shma et al., 2015). SI14 carried possible compound heterozygous mutations in this gene (**Table 1**), which are therefore the most likely causes of non-PCD SI in this individual. SI14 is left-handed and has CHD (**Table 1**).

Subject SI03, who had no formal medical history of PCD, but had reported an intermediate PICADAR score, showed no recessive mutations in known PCD genes, which tends to support non-PCD status for this subject. This meant that five non-PCD SI cases did not have recessive mutations in genes known to cause human laterality disorders, as annotated in Clinvar, and therefore remained 'unsolved' under a recessive model (**Table 1**). Among these five non-PCD SI cases, three were left-handed, and were therefore of primary interest for the present study. Three of these same subjects also had CHD (**Table 1**).

We constructed a list of known or suspected laterality genes with reference to the literature and mouse laterality phenotypes (**Methods**; **Table S1**), but none of these genes had recessive mutations in the five unsolved non-PCD SI subjects.

We note possible compound heterozygous mutations in *PKD1* [MIM:601313], as a paralogue of *PKD1L1* [MIM:609721], in subject SI14 (**Table S3**). However, mutations in this gene would be expected to cause autosomal dominant Polycystic Kidney Disease (Ranjzad et al., 2018; Xu et al., 2018), and since there was no such diagnostic record for this subject, one or both of these specific mutations probably has limited functional impact and is therefore unlikely to be a monogenic cause for SI either. As noted above, SI14 anyway had likely causative mutations in *CFAP52*.

KIF13B [MIM:607350] was putatively recessively mutated in subjects SI12 (likely solved through *DNAH5*) and SI14 (likely solved through *CFAP52*) (**Table S3**). Although no literature has linked *KIF13B* to PCD or laterality phenotypes, *KIF3A* [MIM:604683], another kinesin encoding gene, is known to be involved in LR determination (Hirokawa, Tanaka, & Okada, 2009). Moreover, *KIF3B* [MIM: 603754] is known to affect motility of nodal cilia and LR determination (Nonaka et al., 1998). Together with dyneins, kinesins allow ciliary proteins to enter the organelle via the transition zone by transporting them as cargo (Goetz & Anderson, 2010; Verhey, Dishinger, & Kee, 2011), and thereby play an important role in ciliary construction and maintenance (Goetz & Anderson, 2010; Verhey et al., 2011). The mutations in *KIF13B* might therefore conceivably have affected phenotypic outcomes in subjects SI12 and SI14, but we cannot confidently assign roles to *KIF13B* on the basis of this evidence.

4.3.1.2 Dominant mutations

For the five non-PCD SI cases who remained unsolved under a recessive model, we also considered a dominant model using a maximum known mutation frequency of 5×10^{-5}, and again cross-referenced the mutated genes against Clinvar and our candidate gene list (**Table S1**), but no likely causative genes emerged (**Table S3**) (see Discussion).

Subject SI05 showed a heterozygous mutation in *LRRC6* [MIM:614930] (**Table S3**), which was included among our candidate genes. However, since recessive – but not dominant - mutations in this gene have been associated with PCD (Horani et al., 2013), it is unlikely that this mutation is causal for non-PCD SI in this subject.

We also note a heterozygous mutation in *WDR62* [MIM:613583] in the unsolved case SI09 (**Table S3**). Although mice with mutations in this gene have shown dextrocardia and right pulmonary isomerism (MGI:5437081) (Smith et al., 2018), humans with recessive *WDR62* mutations do not show altered laterality. Instead, they have shown infantile spasm, microcephaly and intellectual disability (Nardello et al., 2018). It is therefore unlikely that *WDR62* is a dominant cause of altered laterality in humans.

The same subject SI09 also had a heterozygous mutation in *PLXND1* [MIM:604282] (**Table S3**), a gene which appears among search results for the phenotype 'situs inversus' within the Mouse Genome Database (Smith et al., 2018). However, while *PLXND1* is annotated as a cause of aortic arch and atrial abnormalities in this database, it is not annotated with situs inversus or heterotaxia, so that the basis for the search result is unclear. We did not find evidence for this gene's involvement in visceral laterality in a further literature search.

Subject SI03 had a heterozygous mutation in *SPEF2* [MIM:610172] (**Table S3**). Loss of *SPEF2* function in mice results in spermatogenesis defects and primary ciliary dyskinesia (Sironen et al., 2011), but this is through recessive mutations, so that we do not consider the heterozygous mutation in SI03 as a strong causal candidate for non-PCD SI.

4.3.2 Gene set enrichment analysis

We first performed gene set enrichment analysis using the positive control set of six SI subjects with PCD. As noted above, the six PCD subjects had likely recessive monogenic causes in five different genes (two subjects had mutations in the same gene, *DNAH5*). As expected, from the total combined list of recessively or X-linked mutated genes in these six subjects, of which 40 genes were included in the GO schema, we obtained significant results for various cilia-related pathways, such as 'axoneme' ($p = 0.004$), 'outer dynein arm assembly' ($p = 3.85\times10^{-5}$), 'dynein complex' ($p = 4.89\times10^{-5}$), and 'microtubule based movement' ($p = 1.41 \times10^{-4}$) (**Table 2**) (all P values adjusted for multiple testing across gene sets, see Methods). When the single most likely causative gene for each PCD subject was removed from the gene list, which left 36 recessively or X-linked mutated genes in these subjects that are present in the GO schema, there were no longer any significant enrichment terms, which further supports that the monogenic causes had been correctly

Table 2. Gene set enrichment analysis under a recessive mutation model.

Group	p-value[1]	term size	query size [2]	overlap size	term ID	term name	intersection	samples	Pval Fisher[3]
Non-PCD SI (n=9) [4]	NS		38						
Non-PCD SI LH (n=5) [4]	NS		26						
Unsolved cases (n=5) [4]	NS		22						
SI with PCD (n=6)	3.85×10^{-5}	17	40	4	GO:0036158	outer dynein arm assembly	DNAH5, CCDC114, LRRC6, DNAAF1	SI06, SI11, SI12, SI13, SI15, SI17	7.71×10^{-13}
SI with PCD (n=6)	1.41×10^{-4}	296	40	8	GO:0007018	microtubule-based movement	DNAH5, CCDC114, DNAH11, WDR60, LRRC6, DNAAF1, FMN2, DNAH12	SI06, SI08, SI11, SI12, SI13, SI15, SI17	9.74×10^{-15}
SI with PCD (n=6)	4.89×10^{-5}	48	40	5	GO:0030286	dynein complex	DNAH5, CCDC114, DNAH11, WDR60, DNAH12	SI08, SI11, SI12, SI13, SI15, SI17	1.68×10^{-13}
SI with PCD (n=6)	0.004	113	40	5	GO:0005930	axoneme	DNAH5, CCDC114, DNAH11, DNAAF1, RP1L1	SI03, SI08, SI11, SI12, SI13, SI15, SI17	1.38×10^{-11}
SI with PCD (n=6)	0.002	427	40	8	GO:0099568	cytoplasmic region	DNAH5, CCDC114, DNAH11, SHROOM2, DNAAF1, FMN2, RP1L1, PCLO	SI03, SI06, SI08, SI11, SI12, SI13, SI15, SI17	1.88×10^{-13}
Unaffected controls (n=15)	NS		34						

LH: left-handed.[1] P-values are corrected for multiple testing using the gSCS method. NS: no significant gene sets. [2] The number of mutated genes present in the GO schema. [3]Uncorrected P-values were calculated using Fisher's exact test. [4] Results were the same after excluding subject SI03.

identified in these subjects. The gene set enrichment analysis in the PCD subjects confirmed that, despite a relatively small number of subjects (i.e. six), the analysis was well powered to identify affected biological processes, even when most individual subjects had monogenic causes in different genes, and each subject had other, non-causative mutated genes.

With this in mind, we then performed gene set enrichment analysis in the non-PCD SI subjects, who were of primary interest for the present study. However, no significant enrichments were found when testing the list of recessively or X-linked mutated genes in the nine subjects with non-PCD SI, of which 38 genes were included in the GO schema. There was also no significant functional enrichment when testing the recessively mutated or X-linked genes in the five left-handed subjects with non-PCD SI, of which 26 genes were in the GO schema, and nor when testing the list of genes in the five unsolved non-PCD SI subjects, of which 12 genes were present in the GO schema (**Table 2**). In addition, gene-set enrichment analysis with dominantly mutated genes as input did not produce significant results in the non-PCD SI cases, nor the left-handed or unsolved subsets.

As expected, the lists of recessively/X-linked and dominantly mutated genes in the fifteen unaffected control subjects did not produce any significant gene set enrichment terms (**Table 2**).

4.3.4 Structural variant analysis

We additionally screened the genomes of all subjects for structural variants (such as larger-scale deletions and duplications), using two complementary algorithms (see Methods). SI subjects had SVs affecting an average of 9.7 genes per subject (min = 2 SVs, max = 16 SVs), and controls had SVs affecting an average of 9.6 genes per subject (min = 5 SVs, max = 16 SVs). None of the SI subjects, regardless of PCD status, had SVs affecting genes that were annotated SI, PCD, Kartagener, *situs ambiguus* (SA), or Heterotaxy (HTX) in Clinvar, nor affecting genes from our list of candidate laterality genes.

4.4 Discussion

In this study we aimed to identify rare, highly penetrant genetic mutations that cause SI without PCD, by analysing whole genome sequence data from nine SI subjects without medical histories of PCD, five of whom were left-handed. We additionally included six SI subjects with PCD as positive controls, and fifteen unaffected subjects as negative technical controls.

Likely monogenic causes were identified in all positive controls, i.e. each of the six PCD subjects had one recessively mutated gene (usually a different gene in each subject) that is already known to cause PCD when mutated. The six PCD subjects also confirmed that significant pathway enrichment could be detected on the basis of their mutated gene lists, as multiple gene sets related to ciliary functions were detected. The fifteen unaffected control subjects showed no recessive mutations in genes known to cause PCD or laterality phenotypes, as expected.

Two of the nine non-PCD SI subjects had likely recessive monogenic causes in known PCD genes. This may indicate reduced penetrance of these mutations for PCD-like symptoms, such as lung symptoms, although they can apparently affect *situs* in early development. One non-PCD SI subject, who was right-handed, had a homozygous mutation in a gene already known to cause SI without PCD when mutated, i.e. *PKD1L1* (Vetrini et al., 2016). This gene encodes a component of a calcium channel which is associated with non-motile cilia (Masyuk, Gradilone, & LaRusso, 2014). Another non-PCD SI subject, who was left-handed and had CHD, had compound heterozygous mutations in *CFAP52*, which is also known as a cause of either heterotaxy or SI without PCD, when mutated (Ta-Shma et al., 2015). CFAP52 protein has been proposed to play a role in cilia-related signal transduction processes(Hirschner et al., 2007).

However, the five remaining non-PCD SI subjects had no obvious monogenic basis for their condition, i.e., they did not have likely causative mutations in genes known to cause human laterality disorders as annotated in the Clinvar database, or within a list of known or suspected laterality genes based on literature searching and mouse phenotypes. Our findings are similar to the findings from Li *et al.(A. H. Li et al., 2019)*, whose exome sequencing study

of laterality defects included 11 subjects with SI, of which 8 had no likely candidate mutations (see Introduction).

Furthermore, gene set enrichment analysis of the mutated genes in non-PCD SI subjects did not identify significant biological pathways, in either the whole set, or the left-handed subset, or the subset that was 'unsolved' by recessive monogenic causes. In other words, the biology of their non-PCD SI could not be linked via the mutations that they carried. Finally, we also considered larger genomic rearrangements known as Copy Number Variants (CNVs), but no obvious candidate genes emerged. A monogenic model is still possible for the majority of the non-PCD SI cases, but would have to involve genetic heterogeneity across a set of genes which are not currently linked in terms of their known biology in the Gene Ontology, at least to an extent that was discernible in this dataset, as it was for the PCD subjects.

Among the five 'unsolved' non-PCD SI subjects, three were left-handed, and therefore comprised the majority of left-handed non-PCD SI subjects in this study. This meant that we did not identify a genetic-developmental pathway that links handedness and visceral asymmetry in this study. The question of whether, and how, handedness is linked developmentally to visceral laterality in humans remains unanswered. As noted in the Introduction, the existing literature remains ambiguous, and on balance does not support overt links (Afzelius & Stenram, 2006; Brandler et al., 2013; Cockayne, 1938; de Kovel & Francks, 2019; de Kovel et al., 2018; McManus et al., 2004; Torgersen, 1950; Vingerhoets et al., 2018). It is possible that continuous measures of lateralized hand performance, such as pegboard measures or hand grip strength, might reveal stronger links with genes involved in visceral laterality (Brandler et al., 2013). However, we had no such measures available in the present study.

Some studies based on small samples have suggested that aspects of brain structural asymmetry can be altered in SI (Ihara et al., 2010; Kennedy et al., 1999; Schuler et al., 2017; Tubbs, Wellons, Salter, Blount, & Oakes, 2003). Indeed, this was previously assessed in the same SI subjects and controls as analyzed in the present study (Mannaert, Verhelst, Gerrits, Bogaert, & Vingerhoets, 2019; Vingerhoets et al., 2018) - briefly, *planum temporale* and

arcuate fasciculus asymmetries were not altered in SI, but the overall cerebral torque showed evidence of left-right reversal in non-PCD SI. At that time there was no genetic information available for these individuals. In light of the genetic data in the present study, the group of nine non-PCD SI individuals has now fragmented into two with likely low-penetrance PCD, two with likely causative mutations in non-PCD-causing genes, and five genetically unsolved. We do not attempt here to do group-based statistics with quantitative measures of brain structural asymmetry in these very small groups. However, this should be a line of research as larger datasets become available.

In the present study we focused on rare, protein-coding mutations, but genetic contributions to non-PCD SI might also involve non-coding variation, or rare combinations of multiple common variants. The noncoding genome comprises 98% of the genome, but interpreting the variation within these regions is challenging. Several attempts have been made to rank potentially causative variants across the genome based on scores that integrate different types of information, including conservation of DNA sequence, regulatory information (Dunham et al., 2012), and population genomic data (Dong et al., 2015; Ioannidis et al., 2016; Jagadeesh et al., 2016; Kircher et al., 2014; Quang, Chen, & Xie, 2015; Ritchie, Dunham, Zeggini, & Flicek, 2014). However, these ranking approaches are currently not very concordant with each other (Eilbeck, Quinlan, & Yandell, 2017). For the present study we did not address these possibilities, which must await larger sample sizes and an improved understanding of the role of rare, non-coding mutations in phenotypic variation.

In utero environmental effects such as prenatal drug exposure might also affect left-right determination (van Veenendaal, Kusters, Oostra, Bergman, & Cobben, 2016). As regards handedness, this is known to associate with various early life factors including birthweight and breastfeeding, although not to a degree which is remotely predictive at the individual level (de Kovel, Carrion-Castillo, & Francks, 2019). As noted in the introduction, left-handedness has a heritability of roughly 25% in family and twin studies, and lower in SNP-based heritability studies. Regardless, the primary causes of this trait remain unknown.

In this study there was a degree of diagnostic uncertainty as regards the PCD status of some SI subjects. However, it was not the purpose of the present study to achieve a clinical diagnosis of the presence or absence of PCD, nor to confirm already-known PCD genes. In this context we did not, for example, sequence the mutations in the PCD subjects by another technique for validation, nor confirm allelic phase in the compound heterozygotes. Rather, we were concerned to identify potentially causative mutations in the nine SI subjects without medical histories of PCD, who also showed an elevated rate of left-handedness, with the potential to yield important basic insights into body and brain laterality. If we had found clear candidate mutations in genes not previously linked to PCD or SI, then further validation and functional characterization of those mutations would have been appropriate, but this did not arise.

Regardless of the PCD status of any individual SI subject in this study, the overall pattern of results was clearly different between the PCD and non-PCD groups, where all six positive control subjects in the PCD group had mutations in known PCD-causing genes, while five out of nine in the non-PCD group had no obvious monogenic mutations. Also, the pathway analyses in various different subsets of the non-PCD SI group showed a consistent lack of significant enrichment, whereas clear signals emerged from the PCD group, which further supports an overall distinction of the groups. Nonetheless, further detailed investigation of subjects SI12 and SI16 with PCD diagnostic tools might reveal lung and other ciliary symptoms to an extent, as these two subjects had likely causative mutations for their SI in known PCD-causing genes(Shapiro et al., 2018).

Although ciliary-induced nodal flow plays a crucial role in symmetry breaking, some organisms, such as chicks and fruit flies (*Drosophilia melanogaster*), do not have motile nodal cilia, yet they do show asymmetrical organs (Vandenberg & Levin, 2013). For example, left-right asymmetry in chicks involves rearranging the relative orientations of cells expressing critical genes at the node, in a non-ciliary-dependent manner (Gros, Feistel, Viebahn, Blum, & Tabin, 2009). Furthermore, left-right asymmetry of the *Drosophilia* gut is established by intracellular cytoskeletal organization that may give rise to cellular shape chirality, by means of unidirectional tilting of

radial fibers, and anti-clockwise swirling of transverse fibers (Tee et al., 2015). Similarly, the actin-related *Lsdia1* gene affects shell coiling direction in the water snail (Abe & Kuroda, 2019). Whether such mechanisms also influence left-right organ asymmetry in mammals is unclear. In the present study we saw no mutations in the orthologues of two genes implicated in cellular chirality in *Drosophilia*, *MYO1D* or *MYO1C* (Lebreton et al., 2018), nor in the orthologues of two genes thought to affect laterality in chicks and frogs, *SLC6A4* and *SLC18A2* (Fukumoto, Blakely, & Levin, 2005). We were not able to identify a human homologue of *Lsdia1*.

Future studies in larger human cohorts may help to identify more genetic contributions to non-PCD SI, and clarify the extent to which developmental mechanisms might be shared with brain asymmetry.

4.5 References

Abe, M., & Kuroda, R. (2019). The development of CRISPR for a mollusc establishes the formin Lsdia1 as the long-sought gene for snail dextral/sinistral coiling. *Development, 146*(9). doi:10.1242/dev.175976

Abyzov, A., Urban, A. E., Snyder, M., & Gerstein, M. (2011). CNVnator: an approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. *Genome Res, 21*(6), 974-984. doi:10.1101/gr.114876.110

Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., . . . Sunyaev, S. R. (2010). A method and server for predicting damaging missense mutations. *Nat Methods, 7*(4), 248-249. doi:10.1038/nmeth0410-248

Afzelius, B. A., & Stenram, U. (2006). Prevalence and genetics of immotile-cilia syndrome and left-handedness. *Int J Dev Biol, 50*(6), 571-573. doi:10.1387/ijdb.052132ba

Antaki, D., Brandler, W. M., & Sebat, J. (2017). SV2: Accurate Structural Variation Genotyping and De Novo Mutation Detection from Whole Genomes. *Bioinformatics*. doi:10.1093/bioinformatics/btx813

Armour, J. A., Davison, A., & McManus, I. C. (2014). Genome-wide association study of handedness excludes simple genetic models. *Heredity (Edinb), 112*(3), 221-225. doi:10.1038/hdy.2013.93

Arnold, G. L., Bixler, D., & Girod, D. (1983). Probable autosomal recessive inheritance of polysplenia, situs inversus and cardiac defects in an Amish family. *Am J Med Genet, 16*(1), 35-42. doi:10.1002/ajmg.1320160107

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., . . . Sherlock, G. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet, 25*(1), 25-29. doi:10.1038/75556

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., . . . Sherlock, G. (2017). Expansion of the Gene Ontology knowledgebase and resources. *Nucleic Acids Res, 45*(D1), D331-d338. doi:10.1093/nar/gkw1108

Auton, A., Brooks, L. D., Durbin, R. M., Garrison, E. P., Kang, H. M., Korbel, J. O., . . . Abecasis, G. R. (2015). A global reference for human genetic variation. *Nature, 526*(7571), 68-74. doi:10.1038/nature15393

Barth, K. A., Miklosi, A., Watkins, J., Bianco, I. H., Wilson, S. W., & Andrew, R. J. (2005). fsi zebrafish show concordant reversal of laterality of viscera, neuroanatomy, and a subset of behavioral responses. *Curr Biol, 15*(9), 844-850. doi:10.1016/j.cub.2005.03.047

Behan, L., Dimitrov, B. D., Kuehni, C. E., Hogg, C., Carroll, M., Evans, H. J., . . . Lucas, J. S. (2016). PICADAR: a diagnostic predictive tool for primary ciliary dyskinesia. *Eur Respir J, 47*(4), 1103-1112. doi:10.1183/13993003.01551-2015

Brandler, W. M., Morris, A. P., Evans, D. M., Scerri, T. S., Kemp, J. P., Timpson, N. J., . . . Paracchini, S. (2013). Common variants in left/right asymmetry genes and pathways are associated with relative hand skill. *PLoS Genet, 9*(9), e1003751. doi:10.1371/journal.pgen.1003751

Brueckner, M., D'Eustachio, P., & Horwich, A. L. (1989). Linkage mapping of a mouse gene, iv, that controls left-right asymmetry of the heart and viscera. *Proc Natl Acad Sci U S A, 86*(13), 5035-5038. doi:10.1073/pnas.86.13.5035

Carrion-Castillo, A., Pepe, A., Kong, X. Z., Fisher, S. E., Mazoyer, B., Tzourio-Mazoyer, N., . . . Francks, C. (2020). Genetic effects on planum temporale asymmetry and their limited relevance to neurodevelopmental disorders, intelligence or educational attainment. *Cortex, 124*, 137-153. doi:10.1016/j.cortex.2019.11.006

Cartwright, J. H., Piro, O., & Tuval, I. (2004). Fluid-dynamical basis of the embryonic development of left-right asymmetry in vertebrates. *Proc Natl Acad Sci U S A, 101*(19), 7234-7239. doi:10.1073/pnas.0402001101

Cavalli-Sforza, L. L., & Bodmer, W. F. (1971). *The genetics of human populations*. Oxford, England: W. H. Freeman.

Cockayne. (1938). the genetics of transposition of the viscera.

de Kovel, C. G. F., Carrion-Castillo, A., & Francks, C. (2019). A large-scale population study of early life factors influencing left-handedness. *Sci Rep, 9*(1), 584. doi:10.1038/s41598-018-37423-8

de Kovel, C. G. F., & Francks, C. (2019). The molecular genetics of hand preference revisited. *Sci Rep, 9*(1), 5986. doi:10.1038/s41598-019-42515-0

de Kovel, C. G. F., Lisgo, S., Karlebach, G., Ju, J., Cheng, G., Fisher, S. E., & Francks, C. (2017). Left-Right Asymmetry of Maturation Rates in Human Embryonic Neural Development. *Biol Psychiatry, 82*(3), 204-212. doi:10.1016/j.biopsych.2017.01.016

de Kovel, C. G. F., Lisgo, S. N., Fisher, S. E., & Francks, C. (2018). Subtle left-right asymmetry of gene expression profiles in embryonic and foetal human brains. *Sci Rep, 8*(1), 12606. doi:10.1038/s41598-018-29496-2

Deng, H., Xia, H., & Deng, S. (2015). Genetic basis of human left-right asymmetry disorders. *Expert Rev Mol Med, 16*, e19. doi:10.1017/erm.2014.22

Dong, C., Wei, P., Jian, X., Gibbs, R., Boerwinkle, E., Wang, K., & Liu, X. (2015). Comparison and integration of deleteriousness prediction methods for nonsynonymous SNVs in whole exome sequencing studies. *Hum. Mol. Genet., 24*(8), 2125-2137.

Dunham, I., Kundaje, A., Aldred, S. F., Collins, P. J., Davis, C. A., Doyle, F., . . . Birney, E. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature, 489*(7414), 57-74.

Eilbeck, K., Quinlan, A., & Yandell, M. (2017). Settling the score: variant prioritization and Mendelian disease. *Nat. Rev. Genet.*

Fliegauf, M., Benzing, T., & Omran, H. (2007). When cilia go bad: cilia defects and ciliopathies. *Nat Rev Mol Cell Biol, 8*(11), 880-893. doi:10.1038/nrm2278

Francks, C. (2015). Exploring human brain lateralization with molecular genetics and genomics. *Ann N Y Acad Sci, 1359*, 1-13. doi:10.1111/nyas.12770

Fu, W., O'Connor, T. D., Jun, G., Kang, H. M., Abecasis, G., Leal, S. M., . . . Akey, J. M. (2013). Analysis of 6,515 exomes reveals the recent origin of most human protein-coding variants. *Nature, 493*(7431), 216-220. doi:10.1038/nature11690

Fukumoto, T., Blakely, R., & Levin, M. (2005). Serotonin transporter function is an early step in left-right patterning in chick and frog embryos. *Dev Neurosci, 27*(6), 349-363. doi:10.1159/000088451

Goetz, S. C., & Anderson, K. V. (2010). The primary cilium: a signalling centre during vertebrate development. *Nat Rev Genet, 11*(5), 331-344. doi:10.1038/nrg2774

Gros, J., Feistel, K., Viebahn, C., Blum, M., & Tabin, C. J. (2009). Cell movements at Hensen's node establish left/right asymmetric gene expression in the chick. *Science, 324*(5929), 941-944. doi:10.1126/science.1172478

Gur, M., Cohen, E. B., Genin, O., Fainsod, A., Perles, Z., & Cinnamon, Y. (2017). Roles of the cilium-associated gene CCDC11 in left-right patterning and in laterality disorders in humans. *Int J Dev Biol, 61*(3-4-5), 267-276. doi:10.1387/ijdb.160442yc

Hilfinger, A., & Julicher, F. (2008). The chirality of ciliary beats. *Phys Biol, 5*(1), 016003. doi:10.1088/1478-3975/5/1/016003

Hirokawa, N., Tanaka, Y., & Okada, Y. (2009). Left-right determination: involvement of molecular motor KIF3, cilia, and nodal flow. *Cold Spring Harb Perspect Biol, 1*(1), a000802. doi:10.1101/cshperspect.a000802

Hirschner, W., Pogoda, H. M., Kramer, C., Thiess, U., Hamprecht, B., Wiesmuller, K. H., . . . Verleysdonk, S. (2007). Biosynthesis of Wdr16, a marker protein for kinocilia-bearing cells, starts at the time of kinocilia formation in rat, and wdr16 gene knockdown causes hydrocephalus in zebrafish. *J Neurochem, 101*(1), 274-288. doi:10.1111/j.1471-4159.2007.04500.x

Horani, A., Ferkol, T. W., Shoseyov, D., Wasserman, M. G., Oren, Y. S., Kerem, B., . . . Kerem, E. (2013). LRRC6 mutation causes primary ciliary dyskinesia with dynein arm defects. *PLoS One, 8*(3), e59436. doi:10.1371/journal.pone.0059436

Hornef, N., Olbrich, H., Horvath, J., Zariwala, M. A., Fliegauf, M., Loges, N. T., . . . Omran, H. (2006). DNAH5 mutations are a common cause of primary ciliary dyskinesia with outer dynein arm defects. *Am J Respir Crit Care Med, 174*(2), 120-126. doi:10.1164/rccm.200601-084OC

Ihara, A., Hirata, M., Fujimaki, N., Goto, T., Umekawa, Y., Fujita, N., . . . Murata, T. (2010). Neuroimaging study on brain asymmetries in situs inversus totalis. *J Neurol Sci, 288*(1-2), 72-78. doi:10.1016/j.jns.2009.10.002

Ioannidis, N. M., Rothstein, J. H., Pejaver, V., Middha, S., McDonnell, S. K., Baheti, S., . . . Sieh, W. (2016). REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. *Am. J. Hum. Genet., 99*(4), 877-885.

Jagadeesh, K. A., Wenger, A. M., Berger, M. J., Guturu, H., Stenson, P. D., Cooper, D. N., . . . Bejerano, G. (2016). M-CAP eliminates a majority of variants of uncertain significance in clinical exomes at high sensitivity. *Nat. Genet., 48*(12), 1581-1586.

Kavaklioglu, T., Ajmal, M., Hameed, A., & Francks, C. (2016). Whole exome sequencing for handedness in a large and highly consanguineous family. *Neuropsychologia, 93*(Pt B), 342-349. doi:10.1016/j.neuropsychologia.2015.11.010

Kennedy, D. N., O'Craven, K. M., Ticho, B. S., Goldstein, A. M., Makris, N., & Henson, J. W. (1999). Structural and functional brain asymmetries in human situs inversus totalis. *Neurology, 53*(6), 1260-1265. doi:10.1212/wnl.53.6.1260

Kircher, M., Witten, D. M., Jain, P., O'Roak, B. J., Cooper, G. M., & Shendure, J. (2014). A general framework for estimating the relative pathogenicity of human genetic variants. *Nat. Genet., 46*(3), 310-315.

Kosaki, K., & Casey, B. (1998). Genetics of human left-right axis malformations. *Semin Cell Dev Biol, 9*(1), 89-99. doi:10.1006/scdb.1997.0187

Kumar, P., Henikoff, S., & Ng, P. C. (2009). Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc, 4*(7), 1073-1081. doi:10.1038/nprot.2009.86

Landrum, M. J., Lee, J. M., Benson, M., Brown, G. R., Chao, C., Chitipiralla, S., . . . Maglott, D. R. (2018). ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res, 46*(D1), D1062-d1067. doi:10.1093/nar/gkx1153

Layer, R. M., Chiang, C., Quinlan, A. R., & Hall, I. M. (2014). LUMPY: a probabilistic framework for structural variant discovery. *Genome Biol, 15*(6), R84. doi:10.1186/gb-2014-15-6-r84

Layton, W. M., Jr. (1976). Random determination of a developmental process: reversal of normal visceral asymmetry in the mouse. *J Hered, 67*(6), 336-338. doi:10.1093/oxfordjournals.jhered.a108749

Lebreton, G., Geminard, C., Lapraz, F., Pyrpassopoulos, S., Cerezo, D., Speder, P., . . . Noselli, S. (2018). Molecular to organismal chirality is induced by the conserved myosin 1D. *Science, 362*(6417), 949-952. doi:10.1126/science.aat8642

Leigh, M. W., Pittman, J. E., Carson, J. L., Ferkol, T. W., Dell, S. D., Davis, S. D., . . . Zariwala, M. A. (2009). Clinical and genetic aspects of primary ciliary dyskinesia/Kartagener syndrome. *Genet Med, 11*(7), 473-487. doi:10.1097/GIM.0b013e3181a53562

Lek, M., Karczewski, K. J., Minikel, E. V., Samocha, K. E., Banks, E., Fennell, T., . . . MacArthur, D. G. (2016). Analysis of protein-coding genetic variation in 60,706 humans. *Nature, 536*(7616), 285-291. doi:10.1038/nature19057

Li, A. H., Hanchard, N. A., Azamian, M., D'Alessandro, L. C. A., Coban-Akdemir, Z., Lopez, K. N., . . . Belmont, J. W. (2019). Genetic architecture of

laterality defects revealed by whole exome sequencing. *Eur J Hum Genet*. doi:10.1038/s41431-018-0307-z

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics, 25*(14), 1754-1760. doi:10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., . . . Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics, 25*(16), 2078-2079. doi:10.1093/bioinformatics/btp352

Mannaert, L., Verhelst, H., Gerrits, R., Bogaert, S., & Vingerhoets, G. (2019). White matter asymmetries in human situs inversus totalis. *Brain Struct Funct, 224*(7), 2559-2565. doi:10.1007/s00429-019-01904-x

Masyuk, A. I., Gradilone, S. A., & LaRusso, N. F. (2014). Calcium signaling in cilia and ciliary-mediated intracellular calcium signaling: are they independent or coordinated molecular events? *Hepatology, 60*(5), 1783-1785. doi:10.1002/hep.27331

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., . . . DePristo, M. A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res, 20*(9), 1297-1303. doi:10.1101/gr.107524.110

McLaren, W., Gil, L., Hunt, S. E., Riat, H. S., Ritchie, G. R., Thormann, A., . . . Cunningham, F. (2016). The Ensembl Variant Effect Predictor. *Genome Biol, 17*(1), 122. doi:10.1186/s13059-016-0974-4

McManus, I. C., Martin, N., Stubbings, G. F., Chung, E. M., & Mitchison, H. M. (2004). Handedness and situs inversus in primary ciliary dyskinesia. *Proc Biol Sci, 271*(1557), 2579-2582. doi:10.1098/rspb.2004.2881

Medland, S. E., Duffy, D. L., Wright, M. J., Geffen, G. M., Hay, D. A., Levy, F., . . . Boomsma, D. I. (2009). Genetic influences on handedness: data from 25,732 Australian and Dutch twin families. *Neuropsychologia, 47*(2), 330-337. doi:10.1016/j.neuropsychologia.2008.09.005

Mirvis, M., Stearns, T., & James Nelson, W. (2018). Cilium structure, assembly, and disassembly regulated by the cytoskeleton. *Biochem J, 475*(14), 2329-2353. doi:10.1042/bcj20170453

Nardello, R., Fontana, A., Antona, V., Beninati, A., Mangano, G. D., Stallone, M. C., & Mangano, S. (2018). A novel mutation of WDR62 gene associated with severe phenotype including infantile spasm, microcephaly, and intellectual disability. *Brain Dev, 40*(1), 58-64. doi:10.1016/j.braindev.2017.07.003

Neale, B. (2017). Heritability of >2,000 traits and disorders in the UK Biobank.

Newcombe, H. B. (1964). In M. Fishbein, ed. Papers and Discussions of the Second International Conference on Congenital Malformations. New York: International Medical Congress.

Nonaka, S., Tanaka, Y., Okada, Y., Takeda, S., Harada, A., Kanai, Y., . . . Hirokawa, N. (1998). Randomization of left-right asymmetry due to loss of nodal cilia generating leftward flow of extraembryonic fluid in mice lacking KIF3B motor protein. *Cell, 95*(6), 829-837.

Okada, Y., Takeda, S., Tanaka, Y., Belmonte, J. I., & Hirokawa, N. (2005). Mechanism of nodal flow: a conserved symmetry breaking event in left-right axis determination. *Cell, 121*(4), 633-644. doi:10.1016/j.cell.2005.04.008

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia, 9*(1), 97-113.

Pirooznia, M., Goes, F. S., & Zandi, P. P. (2015). Whole-genome CNV analysis: advances in computational approaches. *Front Genet, 6*, 138. doi:10.3389/fgene.2015.00138

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., . . . Sham, P. C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet, 81*(3), 559-575. doi:10.1086/519795

Quang, D., Chen, Y., & Xie, X. (2015). DANN: a deep learning approach for annotating the pathogenicity of genetic variants. *Bioinformatics, 31*(5), 761-763.

Ranjzad, F., Aghdami, N., Tara, A., Mohseni, M., Moghadasali, R., & Basiri, A. (2018). Identification of Three Novel Frameshift Mutations in the PKD1 Gene in Iranian Families with Autosomal Dominant Polycystic Kidney Disease Using Efficient Targeted Next-Generation Sequencing. *Kidney Blood Press Res, 43*(2), 471-478. doi:10.1159/000488471

Reimand, J., Arak, T., Adler, P., Kolberg, L., Reisberg, S., Peterson, H., & Vilo, J. (2016). g:Profiler-a web server for functional interpretation of gene lists (2016 update). *Nucleic Acids Res, 44*(W1), W83-89. doi:10.1093/nar/gkw199

Reish, O., Aspit, L., Zouella, A., Roth, Y., Polak-Charcon, S., Baboushkin, T., . . . Parvari, R. (2016). A Homozygous Nme7 Mutation Is Associated with Situs Inversus Totalis. *Hum Mutat, 37*(8), 727-731. doi:10.1002/humu.22998

Reiter, J. F., & Leroux, M. R. (2017). Genes and molecular pathways underpinning ciliopathies. *Nat Rev Mol Cell Biol, 18*(9), 533-547. doi:10.1038/nrm.2017.60

Ritchie, G. R., Dunham, I., Zeggini, E., & Flicek, P. (2014). Functional annotation of noncoding sequence variants. *Nat. Methods, 11*(3), 294-296.

Robinson, J. (2013). Edinburgh Handedness Inventory. In: Volkmar F.R. (eds) Encyclopedia of Autism Spectrum Disorders, Springer, New York, NY.

.

Robinson, J. T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., & Mesirov, J. P. (2011). Integrative genomics viewer. *Nat Biotechnol, 29*(1), 24-26. doi:10.1038/nbt.1754

Schuler, A. L., Kasprian, G., Schwartz, E., Seidl, R., Diogo, M. C., Mitter, C., . . . Bartha-Doering, L. (2017). Mens inversus in corpore inverso? Language lateralization in a boy with situs inversus totalis. *Brain Lang, 174*, 9-15. doi:10.1016/j.bandl.2017.06.005

Shapiro, A. J., Davis, S. D., Polineni, D., Manion, M., Rosenfeld, M., Dell, S. D., . . . Lavergne, V. (2018). Diagnosis of Primary Ciliary Dyskinesia. An Official American Thoracic Society Clinical Practice Guideline. *Am J Respir Crit Care Med, 197*(12), e24-e39. doi:10.1164/rccm.201805-0819ST

Sironen, A., Kotaja, N., Mulhern, H., Wyatt, T. A., Sisson, J. H., Pavlik, J. A., . . . Lee, L. (2011). Loss of SPEF2 function in mice results in spermatogenesis defects and primary ciliary dyskinesia. *Biol Reprod, 85*(4), 690-701. doi:10.1095/biolreprod.111.091132

Smith, C. L., Blake, J. A., Kadin, J. A., Richardson, J. E., & Bult, C. J. (2018). Mouse Genome Database (MGD)-2018: knowledgebase for the laboratory mouse. *Nucleic Acids Res, 46*(D1), D836-d842. doi:10.1093/nar/gkx1006

Ta-Shma, A., Perles, Z., Yaacov, B., Werner, M., Frumkin, A., Rein, A. J., & Elpeleg, O. (2015). A human laterality disorder associated with a homozygous WDR16 deletion. *Eur J Hum Genet, 23*(9), 1262-1265. doi:10.1038/ejhg.2014.265

Tan, A., Abecasis, G. R., & Kang, H. M. (2015). Unified representation of genetic variants. *Bioinformatics, 31*(13), 2202-2204. doi:10.1093/bioinformatics/btv112

Tee, Y. H., Shemesh, T., Thiagarajan, V., Hariadi, R. F., Anderson, K. L., Page, C., . . . Bershadsky, A. D. (2015). Cellular chirality arising from the self-organization of the actin cytoskeleton. *Nat Cell Biol, 17*(4), 445-457. doi:10.1038/ncb3137

Torgersen. (1950). Situs inversus, Asymmetry, and Twinning.

Tubbs, R. S., Wellons, J. C., 3rd, Salter, G., Blount, J. P., & Oakes, W. J. (2003). Intracranial anatomic asymmetry in situs inversus totalis. *Anat Embryol (Berl), 206*(3), 199-202. doi:10.1007/s00429-002-0286-1

van Veenendaal, N. R., Kusters, C. D., Oostra, R. J., Bergman, J. E., & Cobben, J. M. (2016). When the right (Drug) should be left: Prenatal drug exposure and heterotaxy syndrome. *Birth Defects Res A Clin Mol Teratol, 106*(7), 573-579. doi:10.1002/bdra.23497

Vandenberg, L. N., & Levin, M. (2013). A unified model for left-right asymmetry? Comparison and synthesis of molecular models of embryonic laterality. *Dev Biol, 379*(1), 1-15. doi:10.1016/j.ydbio.2013.03.021

Verhey, K. J., Dishinger, J., & Kee, H. L. (2011). Kinesin motors and primary cilia. *Biochem Soc Trans, 39*(5), 1120-1125. doi:10.1042/bst0391120

Vetrini, F., D'Alessandro, L. C., Akdemir, Z. C., Braxton, A., Azamian, M. S., Eldomery, M. K., . . . Yang, Y. (2016). Bi-allelic Mutations in PKD1L1 Are Associated with Laterality Defects in Humans. *Am J Hum Genet, 99*(4), 886-893. doi:10.1016/j.ajhg.2016.07.011

Vingerhoets, G., Li, X., Hou, L., Bogaert, S., Verhelst, H., Gerrits, R., . . . Roberts, N. (2018). Brain structural and functional asymmetry in human situs inversus totalis. *Brain Struct Funct*. doi:10.1007/s00429-017-1598-5

Wang, K., Li, M., & Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res, 38*(16), e164. doi:10.1093/nar/gkq603

Ware, S. M., Peng, J., Zhu, L., Fernbach, S., Colicos, S., Casey, B., . . . Belmont, J. W. (2004). Identification and functional analysis of ZIC3 mutations in heterotaxy and related congenital heart defects. *Am J Hum Genet, 74*(1), 93-105. doi:10.1086/380998

Xu, D., Ma, Y., Gu, X., Bian, R., Lu, Y., Xing, X., & Mei, C. (2018). Novel Mutations in the PKD1 and PKD2 Genes of Chinese Patients with Autosomal Dominant Polycystic Kidney Disease. *Kidney Blood Press Res, 43*(2), 297-309. doi:10.1159/000487899

4.6 Supplementary Information

Table of contents

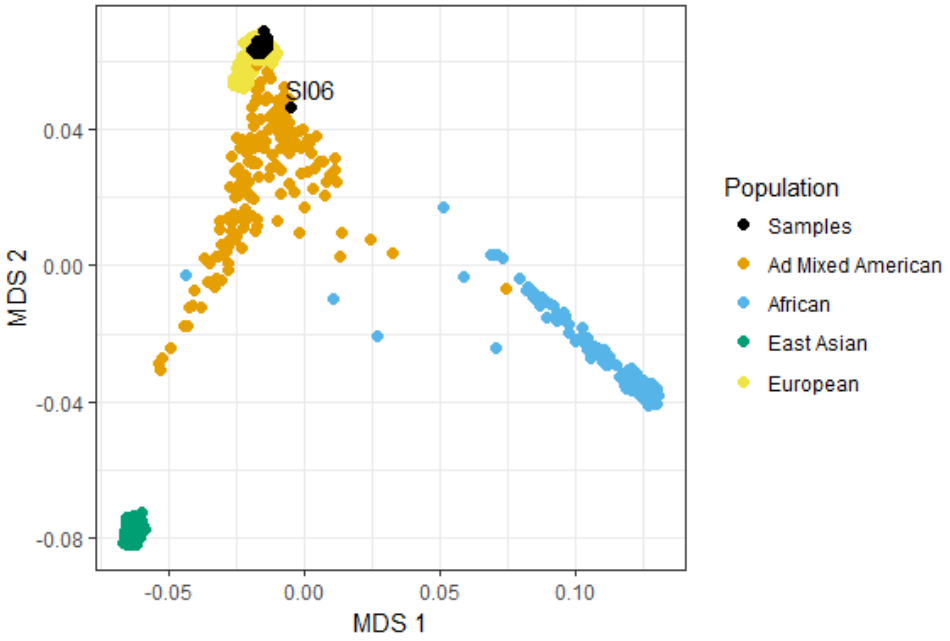


Figure S1. Multidimensional Scaling (MDS) to capture overall genomic diversity among the 30 study samples (black dots), in relation to the 1000 Genomes populations of known geographic ancestries. Subject SI06 is clearly distinct from the European-descent population (yellow dots).

Table S1. List of candidate genes, and the sources that led to their inclusion.

Approved Symbol	Cyto Location	Ensembl Gene ID	Reference(s)
ACVR2B	3p22-p21.3	ENSG00000114739	Deng, Xia, & Deng (2015), Mouse Genome Database (2018)
AK7	14q32.2	ENSG00000140057	Reiter & Leroux (2017)
AMN	14q32	ENSG00000166126	Mouse Genome Database (2018)
ANKS3	16p13.3	ENSG00000168096	Reiter & Leroux (2017)
ANKS6	9q22.33	ENSG00000165138	Reiter & Leroux (2017), Deng, Xia, & Deng (2015), Mouse Genome Database (2018)
AP1B1	22q12	ENSG00000100280	Mouse Genome Database (2018)
ARL17A	NA	NA	de Kovel & Francks (2019)
ARL17B	NA	NA	de Kovel & Francks (2019)
ARL2BP	16q13	ENSG00000102931	Deng, Xia, & Deng (2015)
ARMC4	10p12.1	ENSG00000169126	Reiter & Leroux (2017), Deng, Xia, & Deng (2015), Mouse Genome Database (2018)
ATMIN	16q23.2	ENSG00000166454	Mouse Genome Database (2018)
BBS1	11q13	ENSG00000174483	Mouse Genome Database (2018)
BBS2	16q21	ENSG00000125124	Deng, Xia, & Deng (2015)
BBS4	15q22.3-q23	ENSG00000140463	Mouse Genome Database (2018)
BICC1	10q21.2	ENSG00000122870	Mouse Genome Database (2018)
BOK	NA	NA	Carrion-Castillo et al. (2019)
C1orf127	NA	NA	Mouse Genome Database (2018)
C1ORF88	NA	NA	Deng, Xia, & Deng (2015)
C21orf59	21q22.11	ENSG00000159079	Reiter & Leroux (2017), Deng, Xia, & Deng (2015)
C2orf74	NA	NA	Mouse Genome Database (2018)
C9orf116	9q34.3	ENSG00000160345	Mouse Genome Database (2018)
CC2D2A	4p15.3	ENSG00000048342	Mouse Genome Database (2018)
CCDC103	17q21.31	ENSG00000214447	Reiter & Leroux (2017), Deng, Xia, & Deng (2015)

Approved Symbol	Cyto Location	Ensembl Gene ID	Reference(s)
CCDC114	19q13.3	ENSG00000105479	Reiter & Leroux (2017), Deng, Xia, & Deng (2015)
CCDC151	19p13.2	ENSG00000198003	Reiter & Leroux (2017), Deng, Xia, & Deng (2015), Mouse Genome Database (2018)
CCDC39	3q26.33	ENSG00000145075	Reiter & Leroux (2017), Deng, Xia, & Deng (2015), Mouse Genome Database (2018)
CCDC40	17q25.3	ENSG00000141519	Reiter & Leroux (2017), Deng, Xia, & Deng (2015)
CCDC65	12q13.12	ENSG00000139537	Reiter & Leroux (2017), Deng, Xia, & Deng (2015)
CCNO	5q11.2	ENSG00000152669	Reiter & Leroux (2017), Deng, Xia, & Deng (2015)
CEP290	12q21.3	ENSG00000198707	Mouse Genome Database (2018)
CERS1	19p12	ENSG00000223802	Mouse Genome Database (2018)
CFAP52	17p13.1	ENSG00000166596	Reiter & Leroux (2017)
CFAP53	18q21.1	ENSG00000172361	Reiter & Leroux (2017), Deng, Xia, & Deng (2015)
CFAP54	NA	NA	Reiter & Leroux (2017)
CFC1	2q21.1	ENSG00000136698	Deng, Xia, & Deng (2015), Mouse Genome Database (2018)
CFC1B	NA	NA	Mouse Genome Database (2018)
CITED2	6q23.3	ENSG00000164442	Deng, Xia, & Deng (2015), Mouse Genome Database (2018)
CRELD1	3p25.3	ENSG00000163703	Deng, Xia, & Deng (2015)
DAND5	19p13.2-p13.13	ENSG00000179284	Mouse Genome Database (2018)
DCTN5	16p12.2	ENSG00000166847	Mouse Genome Database (2018)
DLL1	NA	NA	Mouse Genome Database (2018)
DNAAF1	16q24.1	ENSG00000154099	Reiter & Leroux (2017), Deng, Xia, & Deng (2015), Deng, Xia, & Deng (2015)
DNAAF2	14q21.3	ENSG00000165506	Reiter & Leroux (2017), Deng, Xia, & Deng (2015), Mouse Genome Database (2018)

Approved Symbol	Cyto Location	Ensembl Gene ID	Reference(s)
DNAAF3	19q13.4	ENSG00000167646	Reiter & Leroux (2017), Deng, Xia, & Deng (2015), Mouse Genome Database (2018)
DNAAF4	15q21	ENSG00000256061	Reiter & Leroux (2017), Deng, Xia, & Deng (2015)
DNAAF5	7p22.3	ENSG00000164818	Reiter & Leroux (2017), Deng, Xia, & Deng (2015)
DNAH11	7p21	ENSG00000105877	Reiter & Leroux (2017), Deng, Xia, & Deng (2015), Mouse Genome Database (2018)
DNAH5	5p15-p14	ENSG00000039139	Reiter & Leroux (2017), Deng, Xia, & Deng (2015), Mouse Genome Database (2018)
DNAH6	2p11.2	ENSG00000115423	Reiter & Leroux (2017)
DNAH9	17p12	ENSG00000007174	Reiter & Leroux (2017)
DNAI1	9p13.3	ENSG00000122735	Reiter & Leroux (2017), Deng, Xia, & Deng (2015), Mouse Genome Database (2018)
DNAI2	17q25	ENSG00000171595	Reiter & Leroux (2017), Deng, Xia, & Deng (2015), Mouse Genome Database (2018)
DNAJB13	11q13.3	ENSG00000187726	Reiter & Leroux (2017)
DNAL1	14q24.3	ENSG00000119661	Reiter & Leroux (2017), Deng, Xia, & Deng (2015)
DPCD	10q24.32	ENSG00000166171	Reiter & Leroux (2017)
DRC1	2p23.3	ENSG00000157856	Reiter & Leroux (2017), Deng, Xia, & Deng (2015), Mouse Genome Database (2018)
DRC3	NA	NA	Reiter & Leroux (2017)
DTYMK	2q37.3	ENSG00000168393	Carrion-Castillo et al. (2019)
EPB41L5	2q14	ENSG00000115109	Deng, Xia, & Deng (2015)
FGF10	5p13-p12	ENSG00000070193	Mouse Genome Database (2018)
FOXH1	8q24.3	ENSG00000160973	Deng, Xia, & Deng (2015), Mouse Genome Database (2018)
FOXJ1	17q22-q25	ENSG00000129654	Mouse Genome Database (2018)
FOXL2	3q23	ENSG00000183770	Mouse Genome Database (2018)

Approved Symbol	Cyto Location	Ensembl Gene ID	Reference(s)
GALNT11	7q36.1	ENSG00000178234	Reiter & Leroux (2017), Deng, Xia, & Deng (2015)
GAS8	16q24.3	ENSG00000141013	Reiter & Leroux (2017), Mouse Genome Database (2018)
GATA4	8p23.1-p22	ENSG00000136574	Deng, Xia, & Deng (2015)
GDF1	19p12	ENSG00000223802	Deng, Xia, & Deng (2015), Mouse Genome Database (2018)
HES7	17p13.2	ENSG00000179111	Deng, Xia, & Deng (2015)
HM13	20q11.21	ENSG00000101294	Mouse Genome Database (2018)
HYDIN	16q22.2	ENSG00000157423	Reiter & Leroux (2017), Deng, Xia, & Deng (2015)
IFT122	3q21	ENSG00000163913	Mouse Genome Database (2018)
IFT140	16p13.3	ENSG00000187535	Mouse Genome Database (2018)
IFT27	22q12.3	ENSG00000100360	Mouse Genome Database (2018)
IFT74	9p21.2	ENSG00000096872	Mouse Genome Database (2018)
IFT88	13q12.1	ENSG00000032742	Reiter & Leroux (2017)
IHH	2q33-q35	ENSG00000163501	Mouse Genome Database (2018)
INVS	9q31	ENSG00000119509	Reiter & Leroux (2017), Deng, Xia, & Deng (2015), Mouse Genome Database (2018)
ITIH5	10p15	ENSG00000123243	Carrion-Castillo et al. (2019)
KANSL1	17q21.31	ENSG00000120071	de Kovel & Francks (2019)
KIF3A	5q31	ENSG00000131437	Reiter & Leroux (2017), Mouse Genome Database (2018)
LEFTY1	1q42.1	ENSG00000243709	Mouse Genome Database (2018)
LEFTY2	1q42.1	ENSG00000143768	Deng, Xia, & Deng (2015), Mouse Genome Database (2018)
LRD	NA	NA	Deng, Xia, & Deng (2015)
LRRC37A	17q21.31-q21.32	ENSG00000176681	de Kovel & Francks (2019)
LRRC37A2	17q21.31-q21.32	ENSG00000238083	de Kovel & Francks (2019)
LRRC6	8q24.22	ENSG00000129295	Reiter & Leroux (2017), Deng, Xia, & Deng (2015)

Approved Symbol	Cyto Location	Ensembl Gene ID	Reference(s)
LZTFL1	3p21.3	ENSG00000163818	Deng, Xia, & Deng (2015)
MAP2	2q34-q35	ENSG00000078018	de Kovel & Francks (2019)
MBD4	3q21-q22	ENSG00000129071	Mouse Genome Database (2018)
MCIDAS	5q11.2	ENSG00000234602	Reiter & Leroux (2017)
MED13L	12q24	ENSG00000123066	Deng, Xia, & Deng (2015)
MEGF8	19q12	ENSG00000105429	Deng, Xia, & Deng (2015), Mouse Genome Database (2018)
MGAT1	5q35	ENSG00000131446	Mouse Genome Database (2018)
MGRN1	16p13.3	ENSG00000102858	Mouse Genome Database (2018)
MKS1	17q23	ENSG00000011143	Mouse Genome Database (2018)
MMP21	10q26.13	ENSG00000154485	Mouse Genome Database (2018)
MNS1	15q21.3	ENSG00000138587	Reiter & Leroux (2017), Mouse Genome Database (2018)
NEK2	1q32.2-q41	ENSG00000117650	Deng, Xia, & Deng (2015)
NEK8	17q11.1	ENSG00000160602	Deng, Xia, & Deng (2015), Mouse Genome Database (2018)
NKX2-5	5q34	ENSG00000183072	Deng, Xia, & Deng (2015)
NME7	1q24	ENSG00000143156	Reiter & Leroux (2017), Mouse Genome Database (2018)
NME8	7p14.1	ENSG00000086288	Reiter & Leroux (2017), Deng, Xia, & Deng (2015)
NODAL	10q22.1	ENSG00000156574	Deng, Xia, & Deng (2015), Mouse Genome Database (2018)
NOTCH1	9q34.3	ENSG00000148400	Deng, Xia, & Deng (2015)
NOTCH2	1p13-p11	ENSG00000134250	Deng, Xia, & Deng (2015)
NOTO	NA	NA	Reiter & Leroux (2017)
NPHP3	3q22	ENSG00000113971	Deng, Xia, & Deng (2015), Mouse Genome Database (2018)
NPHP3-ACAD11	NA	NA	Mouse Genome Database (2018)
NPHP4	1p36	ENSG00000131697	Deng, Xia, & Deng (2015)

Approved Symbol	Cyto Location	Ensembl Gene ID	Reference(s)
NSF	17q21-q22	ENSG00000073969	de Kovel & Francks (2019)
NUP188	9q34.11	ENSG00000095319	Deng, Xia, & Deng (2015)
OFD1	Xp22.3-p22.2	ENSG00000046651	Reiter & Leroux (2017)
PAX8	2q12-q14	ENSG00000125618	Mouse Genome Database (2018)
PCSK5	9q21.3	ENSG00000099139	Mouse Genome Database (2018)
PCSK6	15q26	ENSG00000140479	Mouse Genome Database (2018)
PIH1D3	Xq22.3	ENSG00000080572	Reiter & Leroux (2017)
PITX2	4q25-q26	ENSG00000164093	Deng, Xia, & Deng (2015)
PKD1L1	7p13-p12	ENSG00000158683	Reiter & Leroux (2017), Mouse Genome Database (2018)
PKD2	4q21-q23	ENSG00000118762	Deng, Xia, & Deng (2015), Mouse Genome Database (2018)
PLXND1	3q22	ENSG00000004399	Mouse Genome Database (2018)
POLB	8p11.2	ENSG00000070501	Mouse Genome Database (2018)
PSKH1	16q22.1	ENSG00000159792	Mouse Genome Database (2018)
RFX3	9p24.2	ENSG00000080298	Reiter & Leroux (2017), Mouse Genome Database (2018)
ROCK2	2p24	ENSG00000134318	Deng, Xia, & Deng (2015)
RPGR	Xp11.4	ENSG00000156313	Reiter & Leroux (2017)
RPGRIP1L	16q12.2	ENSG00000103494	Mouse Genome Database (2018)
RSPH1	21q22.3	ENSG00000160188	Reiter & Leroux (2017), Deng, Xia, & Deng (2015)
RSPH3	6q25.3	ENSG00000130363	Reiter & Leroux (2017)
RSPH4A	6q22.1	ENSG00000111834	Reiter & Leroux (2017), Deng, Xia, & Deng (2015)
RSPH9	6p21.1	ENSG00000172426	Reiter & Leroux (2017), Deng, Xia, & Deng (2015)
SESN1	6q21	ENSG00000080546	Deng, Xia, & Deng (2015)
SH3PXD2A	NA	NA	Mouse Genome Database (2018)
SHH	7q36	ENSG00000164690	Mouse Genome Database (2018)
SHROOM3	4q21.1	ENSG00000138771	Deng, Xia, & Deng (2015)

Approved Symbol	Cyto Location	Ensembl Gene ID	Reference(s)
SLIT2	4p15.2	ENSG00000145147	Mouse Genome Database (2018)
SMAD2	18q21	ENSG00000175387	Deng, Xia, & Deng (2015)
SPAG1	8q22	ENSG00000104450	Reiter & Leroux (2017), Deng, Xia, & Deng (2015)
SPEF2	5p13.2	ENSG00000152582	Reiter & Leroux (2017)
SSH	NA	NA	Deng, Xia, & Deng (2015)
STH	17q21.1	ENSG00000256762	de Kovel & Francks (2019)
TBC1D32	6q22.31	ENSG00000146350	Reiter & Leroux (2017), Mouse Genome Database (2018)
TBXT	NA	NA	Mouse Genome Database (2018)
TGFBR2	3p22	ENSG00000163513	Deng, Xia, & Deng (2015)
TGIF1	18p11.3	ENSG00000177426	Mouse Genome Database (2018)
TMEM67	8q21.13-q22.1	ENSG00000164953	Mouse Genome Database (2018)
TRAPPC10	21q22.3	ENSG00000160218	Mouse Genome Database (2018)
TTC25	17q21.2	ENSG00000204815	Reiter & Leroux (2017)
TTC8	14q32.1	ENSG00000165533	Deng, Xia, & Deng (2015)
UVRAG	11q13	ENSG00000198382	Deng, Xia, & Deng (2015)
WDR62	19q13.12	ENSG00000075702	Mouse Genome Database (2018)
ZIC3	Xq26.2	ENSG00000156925	Deng, Xia, & Deng (2015), Mouse Genome Database (2018)
ZMYND10	3p21.3	ENSG00000004838	Reiter & Leroux (2017), Deng, Xia, & Deng (2015)

Table S2. Inbreeding coefficients per subject.

Subject	Group	O(HOM)	E(HOM)	NOMISS	Fhat1	Fhat2	Fhat3
SI02	non-PCD SI	21098	21680	40351	-0.034	-0.034	-0.034
SI03	non-PCD SI	20997	21680	40352	-0.035	-0.038	-0.036
SI04	non-PCD SI	21099	21680	40350	-0.027	-0.035	-0.031
SI05	non-PCD SI	21161	21680	40337	-0.026	-0.030	-0.028
SI07	non-PCD SI	20935	21690	40354	-0.047	-0.039	-0.043
SI09	non-PCD SI	21090	21690	40363	-0.037	-0.033	-0.035
SI12	non-PCD SI	21274	21690	40359	-0.032	-0.023	-0.028
SI14	non-PCD SI	20952	21690	40359	-0.041	-0.040	-0.040
SI16	non-PCD SI	21282	21680	40352	-0.018	-0.023	-0.021
SI06	**SI with PCD**	**22503**	**21680**	**40348**	**0.048**	**0.041**	**0.044**
SI08	SI with PCD	21039	21690	40357	-0.033	-0.035	-0.034
SI11	SI with PCD	21059	21680	40346	-0.036	-0.034	-0.035
SI13	SI with PCD	21011	21690	40360	-0.036	-0.036	-0.036
SI15	SI with PCD	21000	21690	40355	-0.032	-0.038	-0.035
SI17	SI with PCD	20888	21690	40363	-0.042	-0.047	-0.045
CO02	Unaffected control	21221	21690	40358	-0.031	-0.025	-0.028
CO03bis	Unaffected control	21028	21680	40341	-0.036	-0.036	-0.036
CO04	Unaffected control	21192	21680	40351	-0.026	-0.030	-0.028
CO15	Unaffected control	21177	21680	40348	-0.037	-0.025	-0.031
CO06bis	Unaffected control	21151	21680	40346	-0.035	-0.029	-0.032
CO07	Unaffected control	20989	21680	40349	-0.038	-0.038	-0.038
CO08	Unaffected control	21128	21680	40354	-0.034	-0.030	-0.032
CO09bis	Unaffected control	20923	21680	40350	-0.039	-0.045	-0.042
CO11	Unaffected control	21095	21680	40344	-0.038	-0.029	-0.033
CO12	Unaffected control	21095	21690	40356	-0.031	-0.034	-0.032
CO13bis	Unaffected control	21067	21690	40358	-0.039	-0.033	-0.036
CO14	Unaffected control	20983	21690	40362	-0.041	-0.039	-0.040
CO15	Unaffected control	20965	21690	40360	-0.036	-0.042	-0.039
CO16bis	Unaffected control	21007	21690	40356	-0.038	-0.038	-0.038
CO17	Unaffected control	21170	21680	40345	-0.023	-0.031	-0.027

O(HOM): observed number of homozygotes; E(HOM): expected number of homozygotes; NOMISS: Number of non-missing genotype calls; Fhat1: variance-standardized relationship minus 1; Fhat2: Excess homozygosity-based inbreeding estimate; Fhat3: estimate based on correlation between uniting gametes. Subject SI06 (bold font) is the only subject with positive inbreeding coefficients, indicating an excess of homozygosity

Table S3. Notable mutations in SI cases, which we nonetheless do not consider causative for SI (see the Results section of the main text for explanation).

Subj	SI group	Sex/ Age	EHI	NH	CHD	Daily wet cough	Type	Gene	Clinvar annotation for the gene	rs ID	Start position	ref	alt	MAF	AAC	impact
SI03	non-PCD	F/26	-0.8	L	0	yes	dom	PKD1	3-4 toe syndactyly | abnormality of the kidney | cerebral aneurysm | hepatic cysts | hereditary cancer-predisposing syndrome | hypertension | inborn genetic diseases | lymphangiomyomatosis | moderate sensorineural hearing impairment | multicystic kidney dysplasia | multiple renal cysts | pancreatic cysts | polycystic kidney disease\x2c adult type | polycystic kidney dysplasia | proteinuria | renal cyst | renovascular hypertension | stage 5 chronic kidney disease | tuberous sclerosis 2 | tuberous sclerosis and lymphangiomyomatosis | tuberous sclerosis syndrome	None	2163259	T	C	-1	M/V	missense variant
SI03	non-PCD	F/26	-0.8	L	0	yes	dom	SPEF2	None	None	35776366	A	T	-1	T/S	missense variant
SI05	non-PCD	M/27	0.9	R	0	no	dom	LRRC6	Kartagener	None	133687517	A	T	-1	NA	splice donor variant

SI09	non-PCD	F/36	0.7	L[§]	0	no	dom	WDR62	abnormality of neuronal migration | microcephaly\x2c cortical malformations\x2c and intellectual disability | primary microcephaly\x2c recessive | primary microcephaly 2 with or without cortical malformations | primary autosomal recessive microcephaly 2	None	36592512	CT	C	-1	L/X	frameshift variant
SI09	non-PCD	F/36	0.7	L[§]	0	no	dom	PLXND1	None	None	129286636	AGAC	A	9.01E-06	V/-	inframe deletion
SI12, SI14	non-PCD	M/18	-0.8	L	1	no	chet	KIF13B	None	None	28974427	C	A	-1	V/L	missense variant
SI12, SI14	non-PCD	M/18	-0.8	L	1	no	chet	KIF13B	None	rs753108980	28974427	C	T	0.000188	V/M	missense variant
SI03	non-PCD	F/26	-0.8	L	0	yes	dom	PKD1	3-4 toe syndactyly | abnormality of the kidney | cerebral aneurysm | hepatic cysts | hereditary cancer-predisposing syndrome | hypertension | inborn genetic diseases | sensorineural hearing impairment | lymphangiomyomatosis | moderate multicystic kidney dysplasia | multiple renal cysts | pancreatic cysts | polycystic kidney disease\x2c adult type |	None	2163259	T	C	-1	M/V	missense variant

Table S3. Notable mutations in SI cases, which we nonetheless do not consider causative for SI (see the Results section of the main text for explanation) [continued]

Subj	SI group	Sex/ Age	EHI	NH	CHD	Daily wet cough	Type	Gene	Clinvar annotation for the gene	rs ID	Start position	ref	alt	MAF	AAC	impact
									polycystic kidney disease\x2c adult type | polycystic kidney dysplasia | proteinuria | renal cyst | renovascular hypertension | stage 5 chronic kidney disease | tuberous sclerosis 2 | tuberous sclerosis and lymphangiomyomatosis | tuberous sclerosis syndrome							
SI03	non-PCD	F/26	-0.8	L	0	yes	dom	SPEF2	None	None	35776366	A	T	-1	T/S	missense variant
SI05	non-PCD	M/27	0.9	R	0	no	dom	LRRC6	Kartagener	None	133687517	A	T	-1	NA	splice donor variant
SI09	non-PCD	F/36	0.7	L[§]	0	no	dom	WDR62	abnormality of neuronal migration | microcephaly\x2c cortical malformations\x2c and intellectual disability | primary microcephaly\x2c recessive | primary microcephaly 2 with or without cortical malformations | primary autosomal recessive microcephaly 2	None	36594252	CT	C	-1	L/X	frameshift variant

SI09	non-PCD	F/36	0.7	L[§]	0	no	dom	PLXND1	None	None	129286636	AGAC	A	9.01E-06	V/-	inframe deletion
SI12, SI14	non-PCD	M/18	-0.8	L	1	no	chet	KIF13B	None	None	28974427	C	A	-1	V/L	missense variant
SI12, SI14	non-PCD	M/18	-0.8	L	1	no	chet	KIF13B	None	rs753108980	28974427	C	T	0.000188	V/M	missense variant
SI14	non-PCD	M/18	-0.8	L	1	no	chet	PKD1	See above	rs19970048 5	2154530	G	T	0.001591	T/N	missense variant
SI14	non-PCD	M/18	-0.8	L	1	no	chet	PKD1	See above	rs14273358 8	2153266	C	T	0.000245	G/S	missense variant

The Genome Reference Consorium (GRC) build 37 decoy version was used as reference sequence. EHI: Edinburgh Handedness Inventory score; NH: natural handedness; CHD: Congenital Heart Disease. Type: type of genetic mutation, i.e., heterozygous (dom), homozygous (hom) or compound heterozygous (chet); MAF: minor allele frequency in population databases, if known; AAC: amino acid change. [§]Self-identified natural lefthander made to convert to right-handedness.

CHAPTER

Genetic influences on left-right asymmetry of intrinsic functional connectivity in the human brain

5

Abstract:

Various cognitive functions are lateralized in the brain, but the genetic contributions to functional laterality remain largely unknown. Indices of intrinsic connectivity asymmetry, based on resting state functional magnetic resonance imaging (rs-fMRI), have previously been shown to relate to functional lateralization for language production. We used genetic and rs-fMRI data from 30,660 individuals from the UK Biobank to explore genetic contributions to variability in two indices of functional connectivity asymmetry. The rs-fMRI-based brain atlas of intrinsic connectivity of homotopic areas (AICHA) was first applied to define network nodes. Then, at the whole hemisphere level, a hemispheric intrinsic connectivity asymmetry similarity index (HICAs_pop) was calculated, which measured the similarity of each participant's left-right connectivity matrix to the average of the dataset. A graph-based measure of degree centrality asymmetry (Rs_DC_asymm) was also calculated per participant, targeted specifically at core brain regions involved in language. Single nucleotide polymorphism (SNP)-based heritability, genome-wide association scanning (GWAS), transcriptome-wide association scanning (TWAS), and other bioinformatic analyses were performed for SNPs with minor allele frequencies above 1%. The effects of rarer, coding mutations were explored using optimal unified sequence kernel association tests (SKAT-O), based on exome sequence data in 16,949 individuals. Phenome-wide association analysis was also performed for each measure of functional connectivity asymmetry. Both measures differed significantly between right- and left-handed individuals (HICAs_pop: $p= 1.6\times10^{-93}$, t=-21.161; Rs_DC_asymm:

p= 3.7×10^{-16}, t=-8.187). HICAs_pop showed 2.4% heritability (SE=0.012, p=0.020) and Rs_DC_asymm showed 3.5% heritability (SE=0.013, p=0.002). No individual loci, genes or gene sets were significantly associated with HICAs_pop or RS_DC_asymm in analysis of common or rare genetic variation, after correction for multiple testing. Both functional connectivity asymmetry measures showed associations with various other brain, cognitive and behavioural traits, and HICAs_pop was genetically correlated with educational attainment (rg=0.26, SE=0.090, p=0.0032). To conclude, hemispheric asymmetry of functional connectivity may subtly contribute to variation in cognitive performance and handedness. There is a very low but significantly heritable contribution to variation in functional connectivity asymmetry. These findings suggest that gene mapping efforts for functional connectivity asymmetry may be successful, but sample sizes will likely need to increase substantially beyond the 30,000 participants available, to reliably detect individual contributions of specific loci and genes. Different approaches for quantifying hemispheric functional connectivity asymmetry should also be pursued, which may provide more strongly heritable indices.

Key words:

Functional connectivity asymmetry, resting-state fMRI, genome-wide association scan, exome sequencing, phenome-wide association scan, hemispheric language dominance

5.1 Introduction

Functional magnetic resonance imaging (fMRI) during the performance of cognitive or behavioural tasks is a widely used tool to capture functional activity of the brain (Glover, 2011). During the resting-state (i.e., in the absence of a specific cognitive task demand), the brain also shows neural activity that is captured by low-frequency fluctuations of the blood oxygen level dependent (BOLD) signal (Lv et al., 2018). The BOLD-signal is strengthened by increased levels of oxygenated hemoglobin, which is transported via the vasculature as a result of increased metabolic demands of firing neurons (Glover, 2011).

Temporal correlations in the BOLD-signal between regions during the resting-state suggest that those regions are functionally connected, and that they work together in a neural network (Cole, Bassett, Power, Braver, & Petersen, 2014; Labache et al., 2019; Power et al., 2011; Smith et al., 2009). Indeed, some of the networks defined in data-driven analysis of rs-fMRI images have close correspondences to task-activated functional networks (Cole et al., 2014; Labache et al., 2019; Power et al., 2011). Resting-state fMRI data therefore offers the potential to study functional networks, even in the absence of task-specific fMRI data.

The human brain is asymmetric in both structure and function (Duboc, Dufourcq, Blader, & Roussigne, 2015; Toga & Thompson, 2003). Well known examples of functional brain asymmetries, at the population level, are left-hemispheric dominance for language and hand preference, and right-hemispheric dominance for visuospatial attention and face processing (Duboc et al., 2015). *In utero* behavioural (Hepper, 2013; Parma, Brasselet, Zoia, Bulgheroni, & Castiello, 2017) and neuroimaging studies (Abu-Rustum, Ziade, & Abu-Rustum, 2013; Gervain, Macagno, Cogoi, Pena, & Mehler, 2008; Kasprian et al., 2011) have shown that population-level brain laterality is present already before birth, which indicates a genetically regulated program of asymmetrical brain development (Francks, 2015). This has been further supported by studies reporting asymmetrical gene or microRNA expression in the central nervous systems of post mortem human embryos and

foetuses (C. G. F. de Kovel, Lisgo, Fisher, & Francks, 2018; Miao et al., 2020; Ocklenburg et al., 2017; Sun et al., 2005).

Recently, a multivariate GWAS in more than 30,000 individuals from the UK Biobank implicated 27 independent genetic variants in affecting brain anatomical asymmetries (Sha et al., 2020). The findings implicated microtubule-related genes and embryonic expression as being particularly relevant to brain anatomical asymmetry. This is intriguing because the cytoskeleton is known to be involved in cellular chirality and left-right axis formation in other organs of other species (Inaki, Liu, & Matsuno, 2016; McDowell, Rajadurai, & Levin, 2016; Okumura et al., 2008). Consistent with this, a recent GWAS of handedness in more than 1 million people (Cuellar-Partida et al., 2020) detected more than 40 genetic loci, and also implicated some of the same microtubule-related genes as Sha *et al.* (Sha et al., 2020). Inter-individual variation in hemispheric language dominance is associated with handedness (Knecht et al., 2000; Packheiser et al., 2020), although the association seems to be mostly driven by roughly 1% of individuals who have fully rightward hemispheric language dominance, most of whom are also left-handed (Mazoyer et al., 2014). The same rare subset of individuals has also been shown by Joliot *et al.* (Joliot et al., 2016) to have atypical hemispheric asymmetry of functional connectivity during the resting-state.

However, genetic studies of functional hemispheric language dominance have not been performed previously on a large scale. Twin- or family-based studies have indicated heritabilities from zero to roughly 30% (Bishop & Bates, 2019; Ocklenburg et al., 2016; Somers et al., 2015). The studies varied in their methods for assessing language dominance, as well as their sample sizes and types of population recruited. Specifically, a study in 103 families demonstrated non-significant heritability ($h^2=0.003$, $p=0.98$) for auditory language dominance as measured by the dichotic listening task, in which different syllables were played simultaneously into the right and left ears (participants tend to perceive the stimuli preferentially from the ear connected to the contralateral dominant hemisphere) (Ocklenburg et al., 2016). In contrast, another family-based study in 368 participants from 37 extended families reported a significant 31% heritability for atypical language

lateralization (Somers et al., 2015). They used functional transcranial Doppler sonography to assess language lateralization (which detects relatively increased blood flow in the middle temporal artery supplying the dominant hemisphere during language tasks), and the families came from an isolated population with relatively high numbers of left-handed individuals (Somers et al., 2015). A study of 141 mono- and dizygotic twin pairs (Bishop & Bates, 2019), also using functional transcranial Doppler sonography to assess language dominance, found a heritability of zero for language dominance, but the limited sample size meant that the 95% confidence interval for the heritability estimate ranged up to 15%.

Thus the heritability of hemispheric language dominance may be only a little lower than the roughly 25% value reported for left-handedness (Medland, Duffy, Wright, Geffen, & Martin, 2006), as well as certain anatomical brain asymmetries with twin-based heritabilities up to a similar level (Guadalupe et al., 2016; X. Z. Kong et al., 2018). This suggests that GWAS studies may be successful for functional language dominance, if performed on a sufficiently large scale, since such studies have already identified significant loci and genetic pathways for both handedness and anatomical brain asymmetry (see above).

An exploratory gene mapping study of functional language dominance was carried out by Carrion Castillo *et al* (Carrion-Castillo et al., 2019), based on 33 individuals with rightward hemispheric language dominance, and 34 individuals with typical leftward dominance as technical controls. This study was clearly far too small for a GWAS approach, because GWAS is based on detecting small genetic effects of common polymorphisms, and requires large samples (see Chapter 1). However, the approach taken by Carrion Castillo *et al*. was to sequence the genomes for rare DNA variants in protein coding regions of the genome, in order to test the hypothesis that some rare genetic contributions to altered functional brain asymmetry may involve large individual effects (high 'penetrance' in genetics terminology, similar to Mendelian disorders). The results tentatively implicated genes involved in the actin cytoskeleton in functional language dominance (Carrion-Castillo et al., 2019).

A very large population study known as the 'UK Biobank' has acquired a wide variety of phenotypic data, including genetic, health-related, cognitive and sociodemographic measures, from roughly 500,000 individuals aged 40-69 and living in the United Kingdom (UK) at the time of recruitment (Sudlow et al., 2015). At the time of the present study, brain imaging data had also been collected from around 40,000 of these individuals, of which post-quality control imaging and genetic data were available from just over 30,000 (see details below). While task-based functional imaging data on language dominance were not collected by the UK Biobank, resting-state fMRI data are available.

Here, we made use of the UK Biobank data to examine the genetic bases of human intrinsic connectivity asymmetry during the resting-state, as an indirect approach to the genetics of functional language dominance. We investigated common genetic variation (i.e. variants with population frequencies of 1% or higher) in a total of 30,660 individuals, as well as rare genetic variation (less than 1% population frequency) in 16,949 individuals with both imaging and whole exome sequencing data.

We calculated two measures of functional connectivity asymmetry; one hemisphere-wide index which we term HICAs_pop (i.e., a measure effectively the same as the 'intra-hemispheric intrinsic connectivity asymmetry similarity' (HICAs) index described by (Joliot et al., 2016) (see Methods)), and another based only on regions that constitute core components of the language network (which we term 'Rs_DC_asymm' and was previously proposed by Labache *et al*. (Labache et al., 2020). Importantly, both metrics have previously shown evidence for association with task based functional hemispheric language dominance, in studies where resting-state and task fMRI were both available. We will briefly summarize these studies and findings here:

The HICAs index is based on the atlas of intrinsic connectivity of homotopic areas (AICHA) brain atlas (Joliot et al., 2015), which was constructed from resting-state fMRI data to define 192 regions per

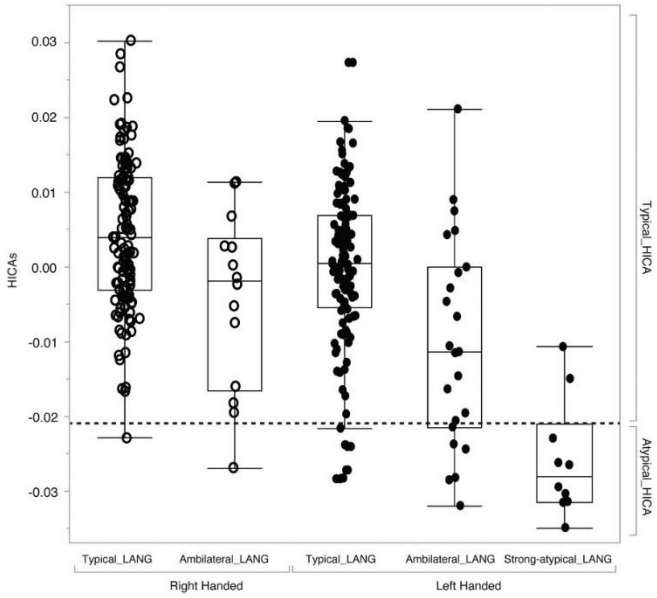


Fig.5.1. Box-plots of HICAs versus hemispheric lateralization for language production (Typical_LANG, Ambilateral_LANG, Strong-atypical_LANG), split by handedness. The dashed line shows the threshold used to distinguish individuals with typical HICAs from those with atypical HICAs. Figure reproduced from (Joliot, Tzourio-Mazoyer, & Mazoyer, 2016)

hemisphere with two essential properties: the voxels within each region have relatively homogeneous time-series with each other, and each region has a homotopic region in the contralateral hemisphere, with which it shares strong functional connectivity. Joliot *et al.* calculated a HICAs index for each of 290 individuals by calculating their separate left and right functional connectivity matrices (with AICHA regions as the matrix nodes), subtracting the right from the left hemispheric matrix separately per participant, and finally calculating the similarity of each individual's left-right difference matrix to that of all other individuals (using correlation analysis) (Joliot et al., 2016).

Joliot *et al.* observed that, from the 22 people who they defined as having atypical HICAs, 20 were left-handed, out of which 14 also showed ambilateral or strong atypical (rightward) language lateralization as assessed using task fMRI (**Figure 5.1**) (Joliot et al., 2016). In addition, all of the 10 individuals exhibiting strong atypical language lateralization were left-handed, out of which 8 also showed atypical HICAs. Thus the HICAs index, despite being based on resting-state data, showed a substantial relation to functional language dominance based on task fMRI.

The HICAs index is calculated with respect to intra-hemispheric connectivity within each entire hemisphere. However, a specific subset of 18 AICHA atlas regions, collectively known as the sentence core network (SENT_CORE), has previously been defined to contain regions that are essential in sentence-level processing (**Figure 5.2**) (Labache et al., 2019). This definition was on the basis of task fMRI data from 138 right-handed participants for three different modalities: reading, listening and covert sentence generation, when contrasting sentence-level tasks with word list tasks (i.e. the latter without sentence grammar). SENT_CORE regions were defined as showing activation across all three modalities, as well as average leftward laterality of activation, and also clustering together into a distinct network based on resting-state fMRI connectivity (Labache et al., 2019). In a follow-up study, Labache *et al*. (Labache et al., 2020) calculated a participant-specific, resting-state-based asymmetry index specifically for the SENT_CORE regions, i.e. the resting-state degree connectivity asymmetry (Rs_DC_asymm) index. First, the degree centrality in each hemisphere of each participant was calculated as the average sum of the positive time-series correlations between each region and all others of the SENT_CORE network. Then, the left minus right difference of the averaged Rs_DC values was calculated per individual, to yield Rs_DC_asymm as a measure of the asymmetry in intra-hemispheric connectivity strengths for the SENT_CORE network of regions.

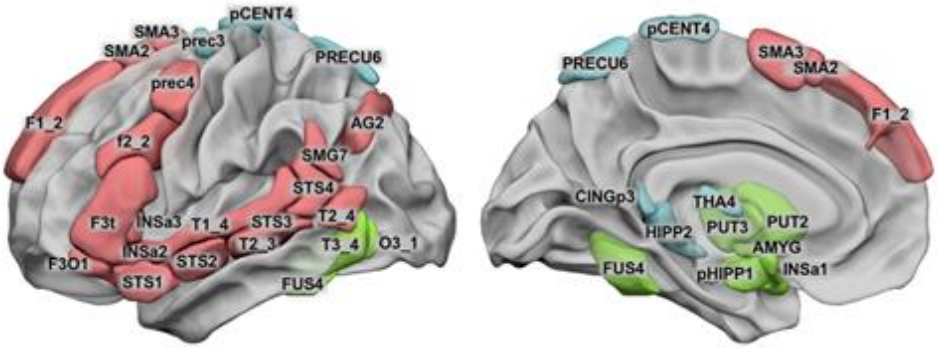


Fig.5.2. Lateral and medial view of the SENT_CORE network, indicated in red. The 18 homotopic AICHA atlas-based regions that comprise the SENT_CORE were conjointly activated and leftward asymmetrical during three sentence minus word-list tasks that were performed by 144 right-handers from the BIL&GIN dataset. Figure reproduced from (Labache et al., 2019).

In contrast to HICAs, hemispheric asymmetry of Rs_DC was not associated with handedness (Labache et al., 2020), but was significantly different between people with leftward language dominance and those with strongly atypical, rightward language dominance, as revealed by a significant 'language organization' by 'side' interaction for Rs_DC (Labache et al., 2020). The atypical group had higher Rs_DC than the typical group in the right hemisphere specifically. Thus Rs_DC_asymm can be taken as another resting-state-based correlate of functional language dominance, this time focused on 18 brain regions that are particularly involved in sentence-level processing.

5.2 Methods

5.2.1 Participants

Genetic and resting-state fMRI data were obtained from the UK Biobank as part of research application 16066, with Clyde Francks as principal applicant. Resting-state fMRI data were downloaded from 40,595 participants (February 2020 release, field 20227). The availability of genome-wide SNP genotype data and exome sequence data meant that common and rare variant imaging genetic analyses were conducted on 30,660 and 16,949 of these individuals respectively, after quality control procedures (see below).

Informed consent was obtained from all UK Biobank participants, and ethical procedures were controlled by a dedicated Ethics and Guidance Council (http://www.ukbiobank.ac.uk/ethics/).

5.2.2 UK Biobank genotyping

5.2.2.1. Genome-wide imputed data

Details about the UK Biobank procedure for generating the genome-wide imputed genotype data can be found elsewhere (Bycroft et al., 2018). Briefly, two different genotyping arrays were used on DNA extracted from blood samples: 49,950 participants with the UK BiLEVE Axiom Array, and

438,427 participants with the UK Biobank Axiom Array. A total of 812,428 unique genetic variants from both of these genotyping arrays were called for all 489,212 participants. Imputation was based on the Haplotype Reference Consortium (HRC) reference panel, as well as on the merged UK10K and 1000 Genomes phase 3 reference panels, and the HRC imputation was used whenever a variant was present in both panels. Data were available from 487,442 participants for a total of 93,095,623 autosomal SNPs, short insertion-deletions (indels) and larger structural variants, and were aligned against the human reference genome assembly GRCh37 (hg19) (Bycroft et al., 2018). For simplicity we will refer to all of these variants as SNPs throughout the following text.

5.2.2.2 Whole-exome sequencing data

Here we used the whole-exome sequencing (WES) data available from the UK Biobank as of October 2020 (https://biobank.ndph.ox.ac.uk/showcase/label.cgi?id=170). WES data from 200,636 individuals (mostly without brain image data) were processed using the original quality functional equivalent (OQFE) pipeline (https://hub.docker.com/r/dnanexus/oqfe).

In brief, whole exome sequences were captured from DNA samples using a slightly modified version of the IDT xGen Exome Research Panel v1.0, supplemented with additional probes. A total of 38,997,831 bases of the human genome were targeted, comprising 19,396 genes (on all chromosomes) (Van Hout et al., 2020). The Illumina NovaSeq 600 platform with either S2 or S4 flow cells (used for the first 50,000 or the additional 150,000 participants, respectively) was used for paired-end sequencing of multiplexed samples, using 75 base pair paired-end reads with two 10 base pair index reads. Raw sequencing data were uploaded to the DNAnexus platform for automated analysis, including conversion into FASTQ files, alignment to the GRCh38 reference genome with BWA-mem (Li & Durbin, 2009), and marking duplicate reads using Picard (http://broadinstitute.github.io/picard/). Samtools (Li et al., 2009) was used to convert BAM files to CRAM files, and various metric statistics were captured per sample using Picard, bcftools (http://samtools.github.io/bcftools/) and

FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc). The QQFE variant calling protocol was applied (Szustakowski et al., 2020), and variants were called with DeepVariant (v.0.10.0) (Lin et al., 2018), including variants in 100 base pair flanking regions up- and downstream of each target region. All 200,000 samples were joint genotyped, using GLnexus (v1.2.6) joint genotyping tool with default 'DeepVariantWES' parameters (Lin et al., 2018; Yun et al., 2021), to create a single, unfiltered project-level Variant Call Format (VCF) file. Further details on the UK Biobank's procedure for generating the WES data can be found elsewhere (Szustakowski et al., 2020; Van Hout et al., 2020) .

5.2.3 Genetic quality control

5.2.3.1. Genome-wide imputed genotype data

Individuals were removed when they showed a mismatch between self-reported and genetically inferred sex, or when they were classified as having putative sex chromosome aneuploidies. Also, those individuals that were outliers based on heterozygosity (ancestry-corrected heterozygosity > 0.1903 (Bycroft et al., 2018)) or genotype missingness (missing rate > 0.05) were removed, as well as those not being from white British ancestry (as defined by (Bycroft et al., 2018), 'in.white.British.ancestry.subset'). Finally, one individual from each related pair (kinship coefficient > 0.0442, as provided by the UK Biobank) was randomly removed from the study sample.

Variants were removed if they had an imputation quality info score below 0.7 (as provided by the UK Biobank with the imputed data), a minor allele frequency (MAF) below 0.01, a MAF difference of 0.2 or larger (based on the absolute difference in MAF between the T1 imaging subset and the total UK Biobank sample, and the subset and the HRC reference panel), or a Hardy-Weinberg Equilibrium (HWE) p-value lower than 1×10^{-7}. Additionally, multi-allelic SNPs were removed, as most software do not support these. A total of 9,855,079 SNPs passed quality control and were subsequently analyzed in the GWAS (see section 2.7.2).

Genetic quality control was applied after phenotype outlier removal (section 2.7), and rendered a final number of 30,660 individuals, which were used as input for subsequent analyses.

5.2.3.2 Whole-exome sequencing data

Of the 30,660 individuals that passed quality control procedures described in the previous section, 16,949 had whole-exome sequence data available (7963 males, 8986 females). Vcftools (v0.1.17) (Danecek et al., 2011) was used to extract the exome data for these individuals from target genomic regions defined in the project level UK Biobank file *xgen_plus_spikein.GRCh38.bed* (available from UK Biobank data showcase resource 3803), rendering 7,918,861 variant sites on which hard filtering was applied. Genotype-level hard filtering comprised setting genotypes with a low approximate read depth (DP < 7 for SNVs and DP < 10 for INDELs) and a low genotype quality (GQ < 20) to no call. Site-level hard filtering comprised removing sites with low average genotype quality across genotypes (average GQ < 35), a high site-level missingness rate (> 0.12), a minor allele count (MAC) of zero, and/or a low allele balance (AB < 0.15 for SNVs and < 0.20 for INDELs, calculated using GATK, v4.1.9.0). These filtering criteria covered all the filtering steps that were used in the WES UK Biobank study on the first 50,000 participants (Van Hout et al., 2020), plus some additional steps as described in a manuscript on quality control for WES studies (Carson et al., 2014). Finally, 48,639 multi-allelic sites (i.e., sites with > 2 alleles at the same genomic position) were excluded and VCFs were converted to PLINK binary format files using Plink (v2.00a2LM). Prior to association testing, two different sets of variants were selected based on their functional annotations from snpEff (Cingolani et al., 2012), and additional SIFT4G (Vaser, Adusumalli, Leng, Sikic, & Ng, 2016) and PolyPhen-2 (Adzhubei et al., 2010) annotations derived from dbNSFP (X. Liu, Li, Mou, Dong, & Tu, 2020). The first set (i.e., the strict filter set) comprised 507,636 variants that were most likely to disrupt protein function and were selected by filtering for variants with HIGH impact, as well as variants with MODERATE impact, 'Protein coding' Biotype, and at least a 'damaging' or 'probably damaging' annotation according to SIFT4G or PolyPhen-2 in dbNSFP, respectively. The second set (i.e., the loose filter set), comprising 1,352,297 variants, included a broader set of

variants with possibly more subtle effects on gene function. This set was selected by filtering for 'Protein coding' Biotype variants with HIGH impact, as well as for variants with MODERATE or MODIFIER impact where SIFT4G and PolyPhen-2 annotations were not 'tolerated' or 'benign', respectively. Plink (v1.90b6.9) (Chang et al., 2015)) was used to filter for variants with MAF <0.01 ('max-maf' function), rendering a final number of 541,181 variants in the strict filter set and 1,110,954 variants in the loose filter set.

5.2.4 UK Biobank brain imaging

Details of UK Biobank brain imaging procedures have been described elsewhere (Miller et al., 2016). In brief, brain imaging was performed in six different modalities, including T1 structural imaging and rs-fMRI, at three different centers in the UK. All imaging centers used identical scanners with fixed platforms (3T Siemens Skyra; software platform VD13) to maximize compatibility. For rs-fMRI, participants were instructed to relax and remain still, with their eyes fixated on a crosshair. The resting-state scans lasted 6 minutes and consisted of 490 time points (TR = 0.735s, TE=39ms; resolution: 2.4x2.4x2.4mm^3; field-of-view: 88x88x64 matrix; GE-EPI with x8 multi-slice acceleration, no iPAT, flip angle 52 degree, fat saturation).

5.2.5 Image processing and quality control

The following pre-processing steps were applied by the UK Biobank: motion correction using MCFLIRT (Jenkinson, Bannister, Brady, & Smith, 2002); grand-mean intensity normalization of the entire 4D dataset by a single multiplicative factor; high-pass filtering to remove temporal drift (Gaussian-weighted least-squares straight line fitting, with sigma=50.0s); EPI unwarping (utilizing the field maps); and gradient distortion correction. Structured artefacts were removed by Independent Component Analysis (ICA) and FMRIB's ICA-based X-noiseifier (FIX) (Beckmann & Smith, 2004; Griffanti et al., 2014; Salimi-Khorshidi et al., 2014). Low quality images, as defined by a combination of manual and automated checking

(https://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/brain_mri.pdf), were labeled 'unusable', and were removed from further analyses.

The preprocessed resting-state fMRI images were downloaded from the UK Biobank and spatially normalized into the AICHA atlas space (Joliot et al., 2015). The AICHA brain atlas parcellates the brain into 192 regions per hemisphere, each of which has a functionally homotopic region in the contralateral hemisphere. Seven homotopic regional pairs were removed from further analyses, as they showed signal loss due to magnetic susceptibility artefacts (i.e., artefacts induced by a local magnetic field inhomogeneity, e.g. due to the proximity of air-filled cavities).These were: G_Frontal_Sup_Orb-1, G_Frontal_Inf_Orb-2, S_Olfactory-1, G_Temporal_Inf-1, G_Temporal_Inf-2, G_Temporal_Pole_Mid-3, and G_subcallosal-1, located at the orbital and inferior temporal regions of the brain, which are prone to magnetic susceptibility artefacts (Krupa & Bekiesińska-Figatowska, 2015). This exclusion left 185 regions per hemisphere.

For each of the 185 regions, a continuous time course was extracted using *invwarp* and *applywarp* from FSL (v.5.0.10) (Jenkinson, Beckmann, Behrens, Woolrich, & Smith, 2012) and *mri_segstats* from FreeSurfer (v. stable6) (Fischl, 2012), by averaging the time courses of all voxels within a given region. Temporal correlation coefficients between the extracted time course from a given region and those from the other 184 regions in the same hemisphere were calculated using the *corrcoef* function from Python (v2.7) package numpy (v1.13.1), to determine the strength of the connectivity between each pair of regions of the functional network at rest. Correlation coefficients were transformed to Gaussian-distributed z-scores via Fisher's transformation to improve normality, resulting in a symmetric z-value matrix (i.e., functional connectivity matrix) per hemisphere for each participant. Functional connectivity matrices were successfully calculated for 37,840 of the 40,595 individuals with resting-state fMRI data (i.e. excluding those labeled as 'unsuable' by the UK Biobank).

R software (v3.5.3) was used to extract several measures from the UK Biobank's motion correction output, that is, mean relative motion averaged across space (same as UK Biobank field f25741), mean absolute motion

averaged across space, maximum relative motion, maximum absolute motion, standard deviation of relative motion, standard deviation of absolute motion, range of relative motion, and range of absolute motion. Distributions were plotted based on the 37,840 individuals for which functional connectivity matrices were calculated, using the *ggplot2* package in R. Individuals with values exceeding 6 times the inter quartile range (IQR) for any of these measures were listed as outliers and subsequently removed from further analysis. Likewise, outliers at six times the IQR were removed based on other imaging-related measures, which were extracted from the UK Biobank data using the R package *ukbtools*: discrepancy between rs-fMRI brain image and T1 brain image (UK Biobank field 25739), inverted temporal signal to noise ratio (SNR) in pre-processed rs-fMRI (UK Biobank field 25743), inverted SNR in artefact-cleaned pre-processed rs-fMRI (UK Biobank field 25744), scanner lateral X brain position (UK Biobank field 25756), scanner transverse Y brain position (UK Biobank field 25757), and scanner longitudinal Z brain position (UK Biobank field 25758). The distribution plots for each measure can be found in the supplement (**Figures S2-S15**).

For each individual, their right hemispheric functional connectivity matrix was then subtracted from their left (left – right), rendering a single functional connectivity asymmetry matrix (sized 185 x 185) per individual. The lower triangle and diagonal were removed from each matrix, leaving 185*184/2 values in the matrix for each individual. For each participant we calculated the mean from all 185*184/2 values within their matrix, and removed outliers at six times the IQR, as a quality control step.

On the basis of all metrics for defining outliers, 709 outliers were removed, leaving 37,131 individuals (i.e., before genetic quality control; section 2.3).

5.2.6 Hemispheric functional connectivity asymmetry

R statistical software (v 3.5.3) was used to calculate two indices of hemispheric functional connectivity asymmetry per individual: HICAs_pop and Rs_DC_asymm (see the Introduction of this chapter for the rationale).

5.2.6.1. HICAs_pop

The Pearson correlation was calculated between each individual's left-right asymmetry matrix and the population average asymmetry matrix (the latter calculated as the mean asymmetry matrix across the 30,660 individuals that passed imaging- and genetic QC). The correlations per individual were subsequently z-transformed using Fisher transformation (fisherz; R package *psych*). Note that this procedure is slightly different from the original HICAs calculation reported by Joliot *et al.* (Joliot et al., 2016), where each individual's asymmetry matrix was correlated with the asymmetry matrices of all other individuals in the dataset, in a one-by-one manner (**Figure 5.3**).

We favoured HICAs_pop over HICAs because of a significantly faster computation time, but the two measures are effectively the same, with a Pearson's correlation of 0.999 (**Figure S1**).

5.2.6.2. Rs_DC_asymm

Rs_DC_asymm was calculated from the SENT_CORE regions specifically (see Introduction), such that the left- and right hemispheric functional connectivity matrices were each of size 18x18, comprising the following

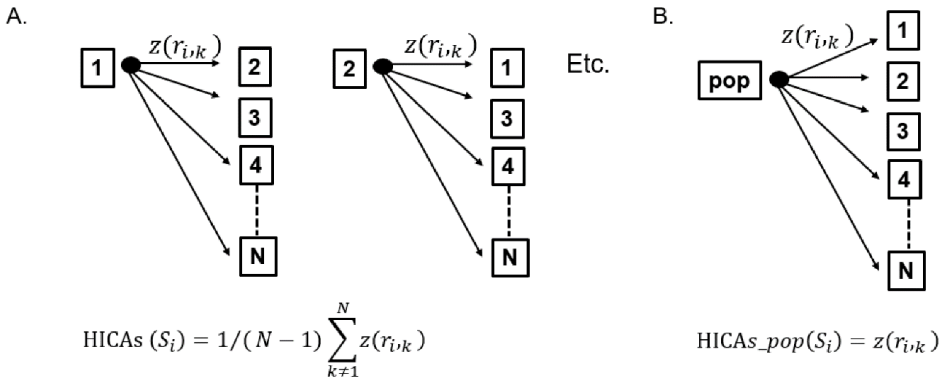


Figure 5.3. Schematic overview of the calculation of HICAs (**A**) and HICAs_pop (**B**). Each square represents a vector variable from an individual's functional connectivity asymmetry matrix (individual number indicated within the square, N=total number of individuals). HICAs is calculated by comparing each individuals's functional connectivity asymmetry matrix with that of all the other individuals in the sample. HICAs_pop is calculated by comparing a population average asymmetry matrix (pop) with all individuals in the sample.

AICHA-atlas regions: G_Frontal_Sup-2, S_Inf_Frontal-2, G_Frontal_Inf_Tri-1, G_Frontal_Inf_Orb_1, S_Precentral-4, G_Insula-anterior-2, G_Insula-anterior-3, G_Supramarginal-7, G_Angular-2, G_Temproral_Sup-4, S_Sup_Temporal-1, S_Sup_Temporal-2, S_Sup_Temporal-3, S_Sup-Temporal-4, G_Temporal_Mid-3, G_Temporal_Mid-4, G_Supp_Motor_Area-2, G_Supp_Motor_Area-3. Separately for the left- and right hemispheric functional connectivity matrices, the sum of all positive correlations was calculated (i.e., degree centrality), and the mean across regions was derived for each hemisphere. The mean degree centrality of the right hemisphere matrix was then subtracted from the left (left-right) to obtain a single Rs_DC_asymm value per individual.

The distributions of the raw HICAs_pop and RS_DC_asymm measures are in **Figures S16 and S17**. Additionally, rank-normalization was performed (with R v3.5.3; RankNorm, R package *RNomni* (McCaw, Lane, Saxena, Redline, & Lin, 2019)) because of a visible skew in the distribution of the raw HICAs_pop, which might have violated assumptions of normality in genetic analyses if left uncorrected (the normalized distributions are shown in **Figures S18, S19**).

5.2.7 Handedness

Handedness was assessed by means of a touchscreen question 'Are you right or left handed?', with four answering options: 'right-handed', 'left-handed', 'use both right and left hands equally', and 'prefer not to answer'. Data were used from the first visit of each individual to a UK Biobank assessment centre (field 1707_0_0). Some individuals also had handedness assessed during their imaging visit (field 1707_2_0) (which was a subsequent visit), and in case first- and imaging visit answers were inconsistent, data from the imaging visit were used. In total, our data comprised 27,265 right-handers, 2,926 left-handers, 465 ambidextrous individuals, and 4 who had indicated 'prefer not to answer'. The R package *ggplot2* was used to produce violin-boxplots of rank-normalized HICAs_pop and Rs_DC_asymm, split by handedness (**Figure 5.4**).

5.2.8 Consideration of brain injury or disease

The available ICD9 and ICD10 data (International Classification of Diseases) of the UK Biobank were searched for diagnostic codes related to viral disease of the nervous system, brain neoplasms, mental-, behavioural- and neurodevelopmental disorders, diseases of the nervous system, cerebrovascular diseases, brain-related perinatal injuries, congenital malformations of the nervous system, and brain injuries, such that a list with brain-related diagnostic codes was created (**Tables S1-S2**). Out of the 30,660 individuals passing both imaging and genetic quality control, 530 had received a brain-related diagnosis based on ICD9 (field 41203) or ICD10 (field 41202) main diagnostic criteria. Two Sample t-tests between diagnosed and non-diagnosed individuals showed no significant group difference in HICAs_pop ($t = -1.331$, $df = 30658$, $p = 0.1831$) nor Rs_DC_asymm ($t = -0.087$, $df = 30658$, $p = 0.9306$), after correcting for covariates (see below). Therefore, the 530 individuals that showed a brain-related ICD9 or ICD10 main diagnosis were retained in all analyses.

5.2.9 Linear modelling to adjust for confound variables

Linear modelling, performed in R (v3.5.3), was performed on each of the two rank-normalized functional connectivity asymmetry phenotypes, to derive residuals that were adjusted for a number of potential confound variables. The following variables, most of which were extracted using the R package *ukbtools*, were included as covariates in the linear models: age at time of imaging (calculated from year and month of birth (UK Biobank fields 34 and 52) and date of attending assessment centre (UK Biobank field 53)), non-linear age, i.e. $(age\text{-}sample\_mean\_age)^2$, sex (UK Biobank field 31), the first 15 principal components capturing genome-wide genetic ancestry, genotyping array, UK biobank assessment centre (UK Biobank field 54), scanner lateral X brain position (UK Biobank field 25756), scanner transverse Y brain position (UK Biobank field 25757), scanner longitudinal Z brain position (UK Biobank field 25758), mean rs-fMRI head motion averaged across space and time points (UK Biobank field 25741), and inverted temporal SNR in pre-processed rs-fMRI (UK Biobank field 25743).

Genotyping array was also included as a covariate, but only for those analyses based on the SNP data. The exome batch (first 50k or additional 150k) was included as a binary covariate for analysis of the exome data to account for the different flow cells used. The residuals were then extracted from the model outputs to be used for further analyses. Residuals-vs-fitted plots are in the supplement (**Figures S20, S21**).

Handedness was not included as a covariate in the linear models, because we were interested in all genetic effects on functional connectivity asymmetry, regardless of potentially pleiotropic effects on handedness. Adjusting for phenotypic covariates that are themselves partly heritable can lead to biased or spurious associations with the target trait in genetic analysis (Aschard, Vilhjálmsson, Joshi, Price, & Kraft, 2015).

5.2.10 Analyses based on the phenotypic data

5.2.10.1 Candidate phenotypic associations

Differences in HICAs_pop and Rs_DC_asymm between handedness groups were calculated by means of the Two Sample t-tests when variances were equal between groups (as first determined by an F-test to compare two variances), or by means of Welch's Two Sample t-test when variances were unequal between groups. The *GGally* package in R (v3.5.3) was used to visualize phenotypic correlations between HICAs_pop and Rs_DC_asymm, as well as their correlations with three structural laterality phenotypes that our group has recently studied in the UK Biobank, namely horizontal and vertical brain skews (related to brain torque) (X.-Z. Kong et al., 2020), and *planum temporale* grey matter volume asymmetry (Carrion-Castillo et al., 2020). The phenotypic correlations were plotted separately for left- and right handers (**Figure S22**).

5.2.10.2 Phenome-wide association study

The UK Biobank has collected a large amount of phenotypic data, including data on sociodemographic and lifestyle factors, psychological factors, derived imaging traits, and cognitive functioning and health. A selection of

phenotypes of potential relevance to the analysis of brain MRI and genetic data were available as part of research application 16066. To examine whether HICAs_pop or Rs_DC_asymm were associated with any of these phenotypes within the UK Biobank, a phenome-wide association Scan (PheWAS) analysis was conducted using PHEnome Scan Analysis Tool (PHESANT) (Millard, Davies, Gaunt, Davey Smith, & Tilling, 2017), which enables comprehensive phenome scans to be performed across continuous, integer and categorical variables in the UK Biobank. PHESANT makes use of a rule-based algorithm that determines how each variable is coded and hence which type of association test is appropriate to use. Dependent variables with continuous, binary, ordered categorical, and unordered categorical data types are tested using linear, logistic, ordered logistic, and multinominal logistic regression, respectively. Prior to testing, an inverse normal rank transformation is applied to variables of the continuous data type.

All analyses were adjusted for covariates as in the linear models for HICAs_pop and Rs_DC_asymm (see section 2.8). Phenotypes that had been included as covariates in these linear models were excluded from the PheWAS analysis, namely, UK Biobank assessment centre (field 54), scanner lateral x brain position (field 25756), scanner transverse y brain position (field 25757), scanner longitudinal z brain position (field 25758), mean rs-fMRI head motion averaged across space and time points (field 25741), and inverted temporal signal-to-noise ratio in preprocessed rs-fMRI (field 25743). In total, 1193 phenotypes were tested in the PheWAS, where some phenotypes with many categorical levels were tested multiple times, such that 3107 tests were performed, separately for HICAs_pop and for RS_DC_asymm. Bonferroni correction was used to correct for multiple testing, with a significance threshold determined by dividing 0.05 by the number of tests performed (i.e., n=3107 for HICAs_pop and separately for Rs_DC_asym).

5.2.11 Analyses based on the genome-wide SNP data

5.2.11.1 SNP heritability and genetic correlation within the UK Biobank data

For the purposes of heritability and genetic correlation analysis, a genetic relationship matrix (GRM), based on 33,582 'white British' (see above) individuals from the UK Biobank with T1 brain imaging data, was built from 8,297,399 non-multi-allelic SNPs with a MAF ≥ 0.01, a MAF difference ≤ 0.2 between the T1 imaging subset and the total UK Biobank dataset, or between the T1 imaging subset and the HRC reference panel, a genotype missingness rate ≤ 0.02, a Hardy Weinberg Equilibrium (HWE) p-value ≥ 1×10^{-6}, and an imputation INFO score ≥ 0.9, using GCTA (version 1.93.0 beta) (Yang, Lee, Goddard, & Visscher, 2011). Individuals with a genotyping rate < 98% were furthermore removed from the GRM (S. Turner et al., 2011), and also one random individual from any pair with a kinship coefficient higher than 0.025 (Yang et al., 2010), which left 28,350 individuals for this particular analysis. The genome-based restricted maximum likelihood (GREML) method in GCTA (version 1.26.0)(Yang et al., 2011) was used to estimate SNP-based heritabilities from all autosomes and chromosome X, and the bivariate GREML method was used to estimate the genetic correlation between HICAs_pop and Rs_DC_asym.

5.2.11.2 Genome-wide association analysis

GWAS were conducted using an additive linear genetic model on the phenotype residuals, using BGENIE (v1.2) (Bycroft et al., 2018). The R package '*qqman*' was used to create QQ-plots and Manhattan plots (S. D. Turner, 2014) in R (version 3.6.0 and version 3.5.3, respectively).

5.2.11.3 Gene-based and gene-set association analyses

MAGMA (v.1.08) (de Leeuw, Mooij, Heskes, & Posthuma, 2015) was used to conduct gene-based and gene-set analyses, in order to examine the joint effects of multiple SNPs or genes, respectively. SNPs were annotated to genes based on the 1000 Genomes reference build 37 (last updated 19-09-2018) (Auton et al., 2015), using a gene annotation window of 50 kb up- and

downstream of a gene, to include nearby cis regulatory regions. Gene-based analyses were performed using 1000 Genomes phase 3 European population reference data (last updated 19-09-2018) for linkage disequilibrium information (Auton et al., 2015), and the SNP-wise mean model (default). Competitive gene-set analyses were run across 5345 Gene Ontology (GO) (Ashburner et al., 2000; Consortium, 2019) gene sets that comprised a minimum of 15 and a maximum of 500 genes (downloaded from http://software.broadinstitute.org/gsea/msigdb/collections.jsp). Thirty-seven out of these 5345 GO gene sets contained the term 'MICROTUBULE', and were subsequently used for a candidate gene set analysis (see the Introduction of this chapter for the rationale as regards microtubules).

Candidate gene sets potentially related to handedness were also tested. For this purpose, 40 genes (after excluding *AL133166.1*, for which no valid Entrez Gene ID was found) located near variants associated with left-handedness in a large GWAS (i.e., 1,534,836 right-handers, 194,198 left-handers, and 37,637 ambidextrous individuals) were included in a gene-set that we called 'PARTIDA', named after the first author of the paper (Cuellar-Partida et al., 2020). Additionally, twelve candidate gene-sets for visceral laterality were taken from (C. G. F. de Kovel & Francks, 2019), of which 11 comprised human homologue genes for mouse laterality phenotypes (derived from the Mouse Genome Informatics resource), and 1 gene-set contained human genes for primary ciliary dyskinesia (PCD) (MIM #244400).

5.2.11.4. Transcriptome-wide association analysis

Transcriptome-wide association scanning (TWAS) was conducted using MetaXcan (Barbeira et al., 2018). This method uses GWAS summary statistics for a trait of interest (such as HICAs_pop) in combination with pre-computed gene expression predictions from prior training data in which post mortem mRNA expression and SNP genotypes had been measured, to test for association between the trait of interest and predicted gene expression levels. The training sets used here were GteX Brain Cortex (119 samples, 4,332 genes), GteX Brain Frontal Cortex BA9 (104 samples, 3,607 genes), and GteX Whole Blood (315 samples, 6,297 genes) (http://predictdb.org/). Multiple testing correction over individual genes was performed using the

False Discovery Rate (FDR) method with a threshold of 0.05, separately for each analysis.

5.2.11.5 Genetic correlations with cognitive, behavioural and disorder traits

Our GWAS summary statistics for HICAs_pop and Rs_DC_asymm were used to test for genetic correlations with regard to publically available GWAS summary statistics for schizophrenia (Ripke et al., 2014), autism spectrum disorder (Grove et al., 2019), attention-deficit hyperactivity disorder (Demontis et al., 2019), intelligence (Savage et al., 2018), educational attainment (J. J. Lee et al., 2018), and handedness (Carolien G.F. de Kovel & Francks, 2018). These genetic correlation analyses were performed using LDSC (v1.0.0) (Bulik-Sullivan et al., 2015).

5.2.12 Analyses based on the whole-exome sequencing data

5.2.12.1 Gene-based association analysis using rare variants

In order to test for associations between HICAs_pop or Rs_DC_asymm with rare genetic variants, optimal unified sequence kernel association tests (SKAT-O) (S. Lee et al., 2012) were performed separately per phenotype and per variant filter set (i.e., strict or loose filtering, see section 2.3.2). SKAT-O is designed to find the optimal balance, for each individual gene, between a classical burden test and the sequence kernel association test (SKAT). Burden tests have high statistical power to detect association in genes when most variants are causal, and in the same direction with respect to the phenotype. On the other hand, SKAT has larger power to detect association in genes where most variants are non-causal, or where the effects of causal variants can be in opposite directions with respect to the phenotype. The SKAT-O test adapts, for each gene, to find the optimal balance while retaining the correct type 1 error rate (S. Lee et al., 2012). The analyses were run in R (v3.5.3) on the Plink binary format files using the SKAT package (S. Lee et al., 2012). Bonferroni multiple testing correction was applied for the 18,835 and 19,109 genes, based on the strict and loose filtering respectively, tested across all chromosomes (note that chromosome X was analyzed using a separate

function for this chromosome within the SKAT package (Ma, Boehnke, & Lee, 2015)).

5.3 Results

5.3.1 Phenotypic associations with functional connectivity asymmetry

5.3.1.1 Phenotypic associations with handedness and other brain asymmetry traits

Distributions of the rank-normalized HICAs_pop and Rs_DC_asymm measures (after adjustment for potential confound variables) are plotted for each of the three handedness groups (i.e., left-handed, right-handed and ambidextrous) in **Figure 5.4**. For HICAs_pop, significant group differences were found between right- and left-handers (t=-21.16, Welch's $p=1.6\times10^{-93}$), between right-handers and ambidextrous individuals (t=-4.09, $p=4.3\times10^{-5}$), and between left-handers and ambidextrous individuals (t=-5.32, Welch's $p=1.4\times10^{-7}$). For Rs_DC_asymm, significant group differences were observed between right- and left-handers (t=-8.19, Welch's $p=3.7\times10^{-16}$) and

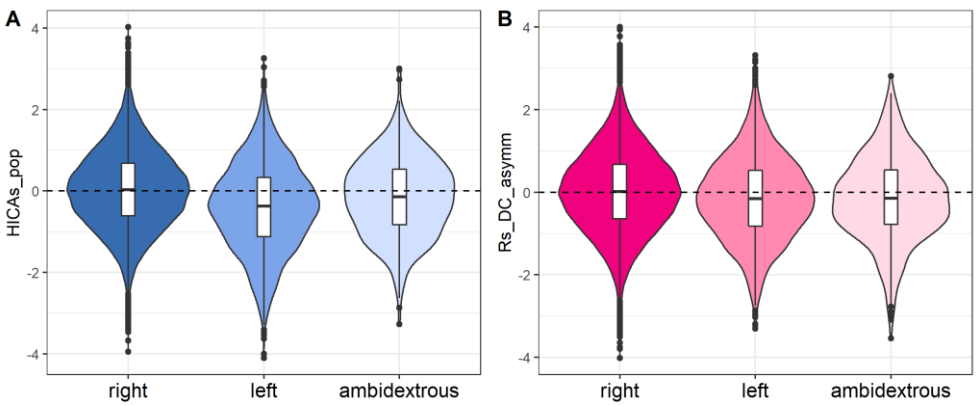


Figure 5.4. Violin-boxplots of rank-normalized (**A**) HICAs_pop, and (**B**) Rs_DC_asymm after adjustment for potential confound variables, split by handedness. The horizontal dotted line indicates zero, the point of symmetry. Lines of the box-plots represents the median, hinges (first and third quartiles), and whiskers (until 1.5 times the inter quartile range). Individuals exceeding the whiskers are plotted as points.

between right-handers and ambidextrous individuals (t=-3.4, p=0.0006), but not between left-handers and ambidextrous individuals (t=-0.057, p=0.95).

Phenotypic correlations were low between HICAs_pop, Rs_DC_asymm, and two UK Biobank-derived structural laterality phenotypes that have been recently studied, that is, horizontal and vertical brain skew (X.-Z. Kong et al., 2020) and *planum temporale* asymmetry (Carrion-Castillo et al., 2020), with the largest correlation between HICAs_pop and Rs_DC_asymm (r=0.19) (**Figure S22**).

5.3.1.2 Phenome-wide association study results

HICAs_pop showed 341 significant (Bonferroni corrected P < 0.05) associations with phenotypes in the UK Biobank (**Table S3, Figure 6**). Out of these, 310 associations were observed within the category '*Brain MRI'*, 14 within '*Imaging*', 5 within '*Anthropometry*', 3 within '*Trail making*', 2 within '*Symbol digit substitution*', 2 within '*Psychosocial factors*', 2 within '*Hand grip strength*', and one within each of '*Early life factors*', '*Tower rearranging*', and '*Reaction time*' . Notably, the phenotype within the '*Early life factors*' category was 'handedness (chirality/laterality)'.

Rs_DC_asymm showed 38 phenome-wide significant associations (**Table S4, Figure 5.6**), of which five were in common with HICAs_pop. These were (1) handedness (chirality/laterality) ('*Early life factors*'), (2) volume of grey matter in precentral gyrus (left) ('*Imaging*'), (3) volume of grey matter in middle frontal gyrus (right) ('*Imaging*'), (4) weighted-mean MO in tract posterior thalamic radiation (right) ('*Brain MRI*'), and (5) 90th percentile in z-statistic for faces-shapes contrast within a group-defined amygdala activation mask ('*Brain MRI*'). Most of the other associations were found within the categories '*Brain MRI*' (N= 26 in total) and '*Imaging*' (N=7 in total), but Rs_DC_asymm also showed significant associations with 'home location at assessment – north co-ordinate (rounded)', and with 'maximum digits remembered correctly' within '*Numeric memory*'.

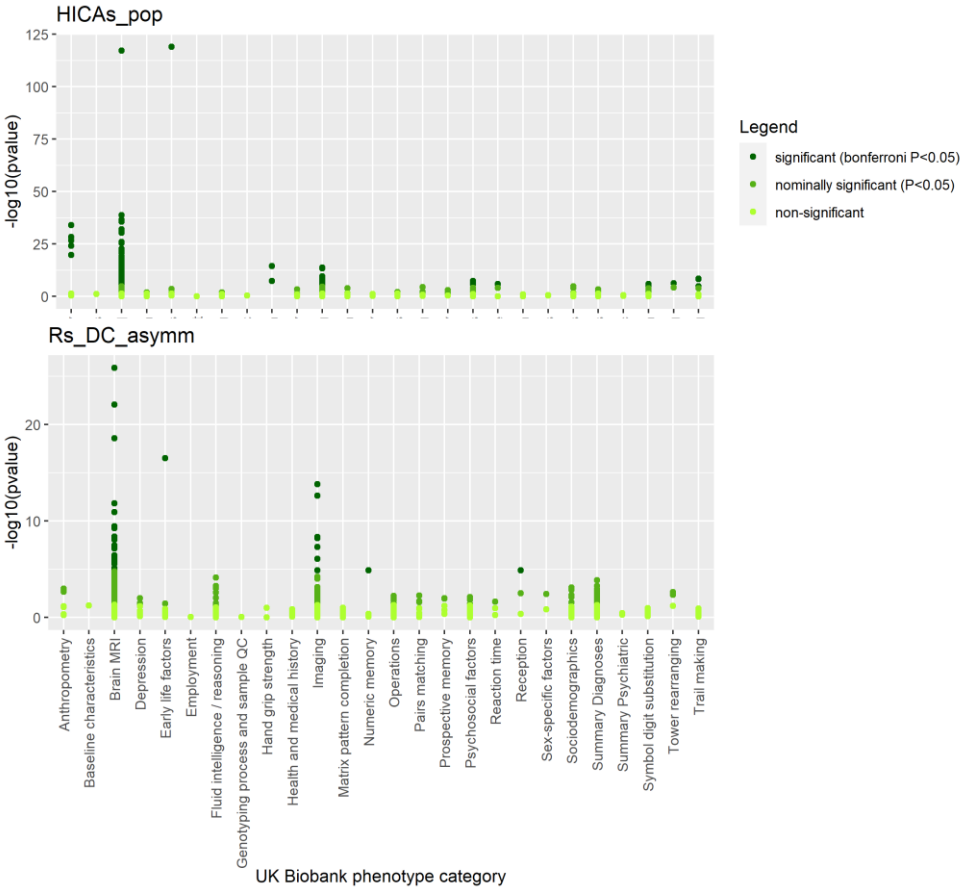


Figure 5.6. Dot plot of the associations between HICAs_pop (*top*) and Rs_DC_asymm (*bottom*) with phenotypes within different UK Biobank phenotype categories. For visualization purposes, the p-values on the y-axis were -log10-scaled. Bonferroni correction was based on N=3107 tests.

5.3.2 Analyses based on the whole-genome SNP data

5.3.2.1 SNP heritability and genetic correlation within the UK Biobank data

The influence of common genetic variation on HICAs_pop and Rs_DC_asymm was examined in 28,350 individuals from the UK Biobank that survived stricter genetic quality control for the GRM (i.e., those with genotype call rate >98% and a kinship coefficient < 0.025). Low but

significant SNP-based heritability estimates were observed for HICAs_pop (h^2=0.024, SE=0.012, p=0.020, N=28350) and Rs_DC_asymm (h^2=0.035, SE=0.013, p=0.002, N=28350). The genetic correlation between HICAs_pop and Rs_DC_asymm, as calculated based on individual-level data from the 28,359 individuals using GCTA (see Methods). However, this estimate had a substantial standard error and was not significant (SE=0.29, p=0.1).

5.3.2.2 Genome-wide association analyses

GWAS did not identify any genome-wide significant associations for HICAs_pop or Rs_DC_asymm (**Tables S5-S6**). The most significant association with HICAs_pop was for the SNP rs35898523 on chromosome 9 (MAF=0.07, β=0.08, SE=0.02, t=5.27, p=1.38×10^{-7}) (**Table S5**). The most significant association with Rs_DC_asymm was for rs116534877 on chromosome 3 (MAF=0.02, β=-0.16, SE=0.03, t=-5.38, p=7.38×10^{-8}) (**Fig.5.5, Table S6**). Among SNPs that have previously been associated with brain structural asymmetry at a genome-wide significant level (Sha et al., 2020), rs6658111 showed a nominally significant association (P<0.05) with HICAs_pop (t=2.62, p = 0.01, MAF = 0.36), and rs7781 with Rs_DC_asymm (t=2.22, p=0.026, MAF = 0.24), but these two associations would not survive Bonferroni correction for multiple testing over the 27 SNPs queried in this look-up (**Tables S7 and S8)**.

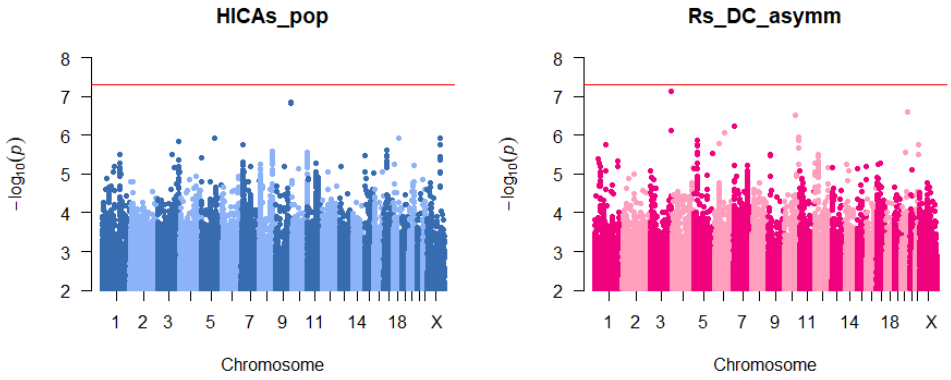


Figure 5.5 Manhattan plots of the GWAS results for HICAs_pop (*left*) and Rs_DC_asymm (*right*). The horizontal red line indicates the genome-wide significance threshold (p=5×10^{-8}), but none of the SNPs passed that threshold.

5.3.2.3 Gene-based and gene-set association analyses based on the GWAS results

Gene-based analyses of GWAS results showed no significant associations with HICAs_pop or Rs_DC_asymm after multiple testing correction. The most significant gene for HICAs_pop was *FAM167A* (Z=4.38, p=0.00001) (**Table S9**), and for Rs_DC_asymm this was with *EDA* (Z=3.86, p=0.0001) (**Table S10**), although these associations remain tentative in the context of multiple testing correction.

Likewise, no FDR significant results were found in the gene-set analyses, where all GO-terms (sized 15 to 500) were tested for associations with HICAs_pop (**Table S11**) and Rs_DC_asymm (**Table S12**). The most significant association for HICAs_pop was with 'GO camera type eye morphogenesis' (N=110 genes, β=0.30, p=0.0002), and for Rs_DC_asymm this was with 'GO vitamin transmembrane transporter activity' (N=24 genes, β=0.67, p=0.001).

We also examined associations with candidate gene-sets for microtubule-related biology, handedness and visceral laterality, but found only tentative associations with microtubule-related gene-sets, which did not survive multiple testing correction for the 37 microtubule candidate gene sets tested in total (Bonferroni threshold = 0.05/37=0.001). 'GO microtubule plus end binding' was tentatively associated with HICAs_pop (Ngenes=17, beta=0.39, unadjusted p=0.02) (**Table S13**), and 'GO microtubule organizing center localization' (Ngenes=28, beta=0.26, unadjusted p=0.04) was tentatively associated with Rs_DC_asymm (**Table S14**).

5.3.2.4 Transcriptome-wide association analyses

TWAS demonstrated no transcriptome-wide significant associations, either for HICAs_pop, or for Rs_DC_asymm (**Figure S23**). For HICAs_pop, the most significant association (i.e. with the highest Z-score) was observed with *ZNF462* (Z=3.94, p=0.0001) when analyzed using expression data from GteX Brain Frontal Cortex BA9 (**Tables S15-S17**). For Rs_DC_asymm, the most significant association was with *CUL7* (Z=4.28, p=0.00002) for GteX Whole Blood (**Tables S18-S20**). Neither of these associations, however, survived multiple testing correction for the number of genes in the corresponding expression datasets.

Table 5.1. Genetic correlations estimated from GWAS summary statistics for Rs_DC_asymm and HICAs_pop.

P1*	P2	rg	se	z	p	H2 obs	H2 obs se	H2 int	H2 int se	Gcov int	Gcov int se
Rs_DC_asymm	Schizophrenia	-0.136	0.060	-2.276	**0.023**	0.242	0.010	1.042	0.012	0.014	0.005
Rs_DC_asymm	ASD	-0.173	0.114	-1.510	0.131	0.194	0.017	1.009	0.010	0.012	0.006
Rs_DC_asymm	ADHD	-0.143	0.094	-1.527	0.127	0.237	0.015	1.032	0.011	0.010	0.005
Rs_DC_asymm	Intelligence	-0.125	0.074	-1.692	0.091	0.196	0.009	1.032	0.025	-0.011	0.011
Rs_DC_asymm	Educational attainment	0.018	0.048	0.375	0.708	0.112	0.003	0.983	0.016	-0.004	0.007
Rs_DC_asymm	handedness (left vs right)	0.035	0.136	0.260	0.795	0.013	0.002	1.032	0.007	-0.014	0.005
HICAs_pop	Schizophrenia	-0.028	0.084	-0.337	0.736	0.242	0.010	1.042	0.012	0.007	0.007
HICAs_pop	ASD	0.038	0.218	0.174	0.862	0.194	0.017	1.009	0.010	-0.005	0.008
HICAs_pop	ADHD	-0.152	0.103	-1.478	0.139	0.237	0.015	1.032	0.011	-0.003	0.006
HICAs_pop	Intelligence	0.102	0.112	0.908	0.364	0.196	0.009	1.032	0.025	0.005	0.011
HICAs_pop	Educational attainment	0.265	0.090	2.946	**0.003***	0.112	0.003	0.983	0.016	0.003	0.006
HICAs_pop	handedness (left vs right)	-0.058	0.151	-0.381	0.703	0.013	0.002	1.032	0.007	-0.033	0.004

P1=phenotype 1, P2= phenotype 2, rg= genetic correlation between phenotype 1 en 2, se=standard error of the genetic correlation, z= the z-statistic, p= p-value, h2 obs = heritability on the observed scale, h2 obs se = standard error of the heritability, h2 int = intercept of the heritability, h2 int se = standard error of the heritability intercept, Gcov int = intercept of the genetic covariance. Gcov in se = standard error of the genetic covariance intercept. P values in **bold font** are nominally significant at the alpha 0.05 level. *Significant after Bonferroni multiple testing correction over all twelve tests in this table.

5.3.2.2 Genetic correlations with cognitive, behavioural and disorder traits

GWAS summary statistics for HICAs_pop and Rs_DC_asymm were used as input for genetic correlation analysis with schizophrenia, ASD, ADHD, IQ, educational attainment, and handedness (using LD score regression and published GWAS summary statistics for these traits; see Methods). HICAs_pop was genetically correlated with educational attainment (rg=0.26, SE=0.090, Z=2.95, p=0.0032), which would remain significant after Bonferroni correction for six tests times two functional connectivity asymmetry measures. Rs_DC_asymm showed evidence for a negative genetic correlation with schizophrenia (rg=-0.136, SE=0.060, Z=-2.28, p=0.023) (**Table 5.1**), which would not survive multiple testing correction.

5.3.2 Analyses based on the whole exome sequencing data

5.3.2.1 Gene-based association analysis based on rare variants

Two differently filtered sets of rare variants (i.e., a strict filter set and a loose filter set; see section 2.3.2) were separately investigated for gene-based associations with HICAs_pop (**Table 5.2**) and Rs_DC_asymm (**Table 5.3**) in 16,949 individuals from the UK Biobank using SKAT-O. No significant gene-based associations were found after correcting across the 18,835 genes tested for the strict variant filter or the 19,109 genes tested for the loose variant filter. The most significant association was between Rs_DC_asymm and Activity-Dependent Neuroprotector Homeobox Protein 2 (*ADNP2*) (loose QC set, N=106 markers tested; $P_{uncorr.}$=0.00001, P_{bonf}=0.169) (**Table 5.3**). QQ-plots for the SKAT-O test results can be found in the supplement (**Figures** S**24-S27**).

Table 5.2. Top 10 results of the SKAT-O analyses for HICAs_pop based on the strict filter variant set (*top*) and the loose filter variant set (*bottom*).

Set ID	Gene	Chr	Number of variants	P.value	Bonferroni corrected
strict filter					
ENSG00000170365	SMAD1	4	13	0.0001	0.627
ENSG00000171634	BPTF	17	88	0.0001	0.627
ENSG00000148346	LCN2	9	12	0.0003	0.627
ENSG00000109610	SOD3	4	11	0.0003	0.627
ENSG00000100226	GTPBP1	22	21	0.0007	0.627

ENSG00000101222	SPEF1	20	7	0.0005	0.627
ENSG00000110002	VWA5A	11	42	0.0003	0.627
ENSG00000182223	ZAR1	4	19	0.0007	0.627
ENSG00000158813	EDA	X	7	0.0007	0.627
ENSG00000129003	VPS13C	15	133	0.0010	0.627
loose filter					
ENSG00000262406	MMP12	11	27	0.00002	0.408
ENSG00000105323	HNRNPUL1	19	123	0.00011	0.617
ENSG00000091409	ITGA6	2	122	0.00015	0.617
ENSG00000140374	ETFA	15	27	0.00035	0.617
ENSG00000148346	LCN2	9	15	0.00031	0.617
ENSG00000242247	ARFGAP3	22	103	0.00058	0.617
ENSG00000120662	MTRF1	13	46	0.00084	0.617
ENSG00000109956	B3GAT1	11	38	0.00023	0.617
ENSG00000172340	SUCLG2	3	135	0.00033	0.617
ENSG00000104818	CGB2	19	71	0.00027	0.617

Table 5.3. Top 10 results of the SKAT-O analyses for Rs_DC_asymm based on the strict variant filter (*top*) and the loose variant filter (*bottom*).

Set ID	Gene	Chr	Number of variants	P.value	Bonferroni corrected
strict filter					
ENSG00000161999	JMJD8	16	22	0.00004	0.356
ENSG00000109686	SH3D19	4	48	0.00006	0.356
ENSG00000197406	DIO3	14	8	0.00011	0.356
ENSG00000232196	MTRNR2L4	16	1	0.00011	0.356
ENSG00000120129	DUSP1	5	10	0.00017	0.455
ENSG00000261787	TCF24	8	8	0.00024	0.501
ENSG00000181751	MACIR	5	7	0.00021	0.488
ENSG00000143036	SLC44A3	1	29	0.00009	0.356
ENSG00000185090	MANEAL	1	23	0.00011	0.356
ENSG00000005022	SLC25A5	X	7	0.00041	0.623
loose filter					
ENSG00000101544	ADNP2	18	106	0.00001	0.169
ENSG00000197406	DIO3	14	21	0.00004	0.418
ENSG00000104714	ERICH1	8	69	0.00010	0.624
ENSG00000181751	MACIR	5	9	0.00018	0.696
ENSG00000261787	TCF24	8	15	0.00028	0.696
ENSG00000120129	DUSP1	5	13	0.00023	0.696
ENSG00000184140	OR4F6	15	23	0.00037	0.696
ENSG00000184471	C1QTNF8	16	33	0.00031	0.696
ENSG00000005022	SLC25A5	X	8	0.00041	0.696
ENSG00000114416	FXR1	3	74	0.00031	0.696

5.4 Discussion

Here we explored the genetic contributions to two functional connectivity asymmetry phenotypes that were extracted from UK Biobank resting-state fMRI data: HICAs_pop and Rs_DC_asymm. Both of these measures have previously been linked to functional hemispheric language dominance in other datasets (see Introduction) (Joliot et al., 2016; Labache et al., 2019). We analyzed imputed genotypes in 30,660 individuals for common genetic variants (population frequencies >= 1%), and whole-exome sequences from 16,949 individuals to test rarer variants (population frequencies < 1%) for potentially larger effects on gene function. This study had the potential to identify genes important for the establishment of functional brain asymmetry, and especially for hemispheric language dominance.

The proportion of phenotypic variation that could be explained by common variants (i.e., SNP-based heritability) was roughly 3% for HICAs_pop and 4% for Rs_DC_asymm. While these heritabilities are low, they were significant in this large dataset. Nonetheless, association analysis with common genetic variation did not identify any individually significant variants, genes or gene sets. This suggests that the common-variant contributions to these phenotypes are highly polygenic, comprised of multiple very small effects, and that even larger sample sizes will be needed in the future to detect significant individual effects. Also with regard to the rare exonic variation, no individual genes were found to be significantly associated with either trait. The rare-variant association with the lowest P-value was for Rs_DC_asymm and *ADNP2*, a gene important for myelin formation (Malishkevich, Leyk, Goldbaum, Richter-Landsberg, & Gozes, 2015) which is possibly implicated in schizophrenia, Alzheimer's disease, and autism spectrum disorder (ASD) (Sragovich, Merenlender-Wagner, & Gozes, 2017). As all of these disorders have been linked to altered brain laterality (H. Liu et al., 2018; Ocklenburg, Güntürkün, Hugdahl, & Hirnstein, 2015), *ADNP2* is worth highlighting as a putative candidate gene for future studies. However, such studies will require larger sample sizes to reliably identify individually significant genes.

The low heritabilities of the functional connectivity asymmetry metrics in this study might superficially indicate large environmental contributions. Indeed, various environmental factors have been proposed to influence brain structural or functional asymmetries, such as fetal position in the uterus, maternal hormone effects *in utero*, functional adaptation such as switching of hand preference in response to societal demands, maternal smoking or breastfeeding (C. G. F. de Kovel, Carrion-Castillo, & Francks, 2019; Toga & Thompson, 2003), although no substantial effects have been reported to our knowledge. In fact, random developmental variation may in fact account for most, or all, of the non-heritable variance in brain asymmetry measures (Bishop & Bates, 2019; C. G. F. de Kovel et al., 2019; Mitchell, 2018; Sha et al., 2020). In addition, in this exploratory study we had no data on intra-subject repeatability of the functional connectivity asymmetry metrics, or their stability over developmental time or aging. Thus the low heritabilities may be partly due to measurement instability, although there were robust associations with other stable traits such as handedness in our study (see below), which suggests that functional connectivity asymmetry is also substantially stable within individuals.

The UK Biobank protocol did not include functional brain imaging during a language task, and we therefore took an exploratory, indirect approach to hemispheric language dominance based on resting-state fMRI indices found to associate with this trait in the BIL&GIN dataset (see Introduction) (Joliot et al., 2016; Labache et al., 2019). HICAs_pop was a measure of total hemispheric functional connectivity asymmetry relative to the population average, whereas Rs_DC_asymm was specifically derived from functional connectivity within core regions of the language network. However, as the associations of HICAs_pop and Rs_DC_asymm with hemispheric language dominance were moderate in previous analyses in other data (see Introduction), then the heritable variance which we detected in these measures may not necessarily be relevant to their correlations with hemispheric language dominance. In other words, the heritable variance in these indices may be independent of functional language dominance, and reflect some other aspect of functional laterality. Direct, large-scale genetic analysis of hemispheric language dominance requires the future collection, or

aggregation, of language task functional data from tens of thousands of individuals, but no such resource yet exists.

It also remains possible that other approaches to analyzing resting-state fMRI data will yield measures that are more strongly associated with hemispheric language dominance, and/or are more strongly heritable, which could better support genetic mapping studies. For example, we used the AICHA functional atlas (Joliot et al., 2015), which was specifically built from resting-state fMRI data to define homotopic regions in the two hemispheres, making it well suited for studies of laterality. Nonetheless, other atlases may yield functional connectivity asymmetry indices with higher correlations with hemispheric language dominance, or higher heritability. We also adopted a common approach in the field by setting negative inter-regional correlations of intrinsic activity to zero, but there is increasing interest in whether including such negative correlations can have biological meaning and increase precision (Zhan et al., 2017). Another alternative approach may be multivariate genetic analysis, rather than reducing functional connectivity asymmetry matrices to single measures prior to genetic analysis. A multivariate approach was employed recently in a genetic study of 42 regional measures of brain structural asymmetry (Sha et al., 2020), where testing of each SNP simultaneously with all 42 regional measures gave rise to multiple significant loci after genome-wide correction for multiple testing, even though the heritabilities of the individual measures were not greatly higher than those in the present study.

Another important factor to consider in the analysis of resting-state fMRI data is whether to correct for the global signal, i.e. the average time-series of the BOLD-signal across all voxels, which can be removed from the time-series of each individual voxel through linear regression (Murphy & Fox, 2017), prior to calculating functional connectivity between regions. The global signal can partly reflect unwanted noise, such as the cardiac and respiratory cycles, but also is likely to partially reflect a neural activity component (Murphy & Fox, 2017). In the present study we did not perform global signal regression. This was partly because we were focused on hemispheric differences, and much of the global signal is likely to be removed

when calculating difference-based metrics (i.e. noise affecting both hemispheres in the same way is likely to be cancelled when subtracting one side's matrix from the other). Anti-correlations of intrinsic activity between regions can also be artificially induced by global signal regression (Murphy, Birn, Handwerker, Jones, & Bandettini, 2009). In addition, it is unclear whether a complex form of collider bias may arise (Day, Loh, Scott, Ong, & Perry, 2016), as the global signal may itself be causally influenced by both genetic variation and region-specific variations in the BOLD signal. Future genetic studies of functional connectivity asymmetry may further consider this aspect, for example by contrasting results obtained with and without global signal regression.

Both of our functional connectivity asymmetry metrics were associated with handedness. For HICAs_pop, this is consistent with the previous report by Joliot *et al.*, who developed this index from a smaller, independent sample (see Introduction) (Joliot et al., 2016). For Rs_DC_asymm, our association with handedness contrasts with a lack of association reported by Labache *et al.*, (Labache et al., 2019), but this may be a statistical power issue, as our study was two orders of magnitude larger. Another previous study of around 9,000 participants from the UK Biobank reported increased cross-hemispheric functional connectivity in left-handers compared to right-handers for independent components derived from resting-state fMRI, which corresponded roughly to the left- and right hemispheric language networks (Wiberg et al., 2019). Together with our data and other observations reported by Labache *et al.*, this indicates that left-handers tend to have stronger cross-hemispheric connectivity of language regions compared to right-handers (Labache et al., 2019), as well as reduced hemispheric asymmetry of connectivity between core language regions, and more atypical brain-wide organization of functional connectivity asymmetry.

There was no significant genetic correlation of either metric of functional connectivity asymmetry with handedness. This may have arisen due to the low heritabilities of handedness and the two functional connectivity asymmetry metrics, which limited the power of this kind of analysis based on linkage disequilibrium score regression (see **Table 1** where the standard

errors of the genetic correlations are substantial, even despite the large sample sizes involved). We detected one significant genetic correlation, i.e., a positive correlation between HICAs_pop and educational attainment. While genetic correlation cannot be taken as direct evidence for causation, it suggests that some genetic variants contribute to lower educational attainment through subtly disrupting the development of typical functional brain laterality. This possibility might be investigated in future studies designed to assess causative mediation, perhaps through Mendelian Randomization approaches (Bowden & Holmes, 2019; Davey Smith & Hemani, 2014).

In the phenome-wide analysis, there was evidence that the functional connectivity asymmetry metrics were positively associated with measures of cognitive performance. These findings are in line with the concept that typical laterality represents an optimal organization for supporting multiple domains of cognition. In fact, typical or fully reversed functional laterality may be optimal for performance, whereas a lack of hemispheric language dominance has been associated with subtly reduced performance across multiple cognitive domains (Mellet et al., 2014). In the current study we did not attempt to distinguish ambilateral from fully reversed patterns of functional connectivity asymmetry, but future studies may explore in this direction. The HICAs_pop metric reflects the similarity of an individual's functional connectivity matrix to the population average matrix, and cannot be interpreted as a straightforward laterality index as used elsewhere in this thesis. Similarity to the population average may be influenced in different individuals by alterations of connectivity that affect different subsets of network nodes. Therefore another future direction will be to assess whether there are subgroups of individuals with distinct forms of non-typical functional connectivity asymmetry

To conclude, this was the first large-scale genetic study of functional connectivity asymmetry metrics linked to hemispheric language dominance. Weak heritabilities were found, but larger sample sizes or alternative methodological approaches will likely be required to identify individual genetic effects on these phenotypes. In addition, we observed significant phenotypic and genetic correlations that were consistent with typical

functional connectivity asymmetry being an optimal organization for cognitive function.

5.5 References

Abu-Rustum, R. S., Ziade, M. F., & Abu-Rustum, S. E. (2013). Reference values for the right and left fetal choroid plexus at 11 to 13 weeks: an early sign of "developmental" laterality? *J Ultrasound Med, 32*(9), 1623-1629. doi:10.7863/ultra.32.9.1623

Aschard, H., Vilhjálmsson, B. J., Joshi, A. D., Price, A. L., & Kraft, P. (2015). Adjusting for heritable covariates can bias effect estimates in genome-wide association studies. *Am J Hum Genet, 96*(2), 329-339. doi:10.1016/j.ajhg.2014.12.021

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., . . . Sherlock, G. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet, 25*(1), 25-29. doi:10.1038/75556

Auton, A., Brooks, L. D., Durbin, R. M., Garrison, E. P., Kang, H. M., Korbel, J. O., . . . Abecasis, G. R. (2015). A global reference for human genetic variation. *Nature, 526*(7571), 68-74. doi:10.1038/nature15393

Barbeira, A. N., Dickinson, S. P., Bonazzola, R., Zheng, J., Wheeler, H. E., Torres, J. M., . . . Im, H. K. (2018). Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics. *Nat Commun, 9*(1), 1825. doi:10.1038/s41467-018-03621-1

Beckmann, C. F., & Smith, S. M. (2004). Probabilistic independent component analysis for functional magnetic resonance imaging. *IEEE Trans Med Imaging, 23*(2), 137-152. doi:10.1109/tmi.2003.822821

Bishop, D. V. M., & Bates, T. C. (2019). Heritability of language laterality assessed by functional transcranial Doppler ultrasound: a twin study. *Wellcome Open Res, 4*, 161. doi:10.12688/wellcomeopenres.15524.2

Bowden, J., & Holmes, M. V. (2019). Meta-analysis and Mendelian randomization: A review. *Res Synth Methods, 10*(4), 486-496. doi:10.1002/jrsm.1346

Bulik-Sullivan, B. K., Loh, P. R., Finucane, H. K., Ripke, S., Yang, J., Patterson, N., . . . Neale, B. M. (2015). LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet, 47*(3), 291-295. doi:10.1038/ng.3211

Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L. T., Sharp, K., . . . Marchini, J. (2018). The UK Biobank resource with deep phenotyping and genomic data. *Nature, 562*(7726), 203-209. doi:10.1038/s41586-018-0579-z

Carrion-Castillo, A., Pepe, A., Kong, X. Z., Fisher, S. E., Mazoyer, B., Tzourio-Mazoyer, N., . . . Francks, C. (2020). Genetic effects on planum temporale

asymmetry and their limited relevance to neurodevelopmental disorders, intelligence or educational attainment. *Cortex, 124*, 137-153. doi:10.1016/j.cortex.2019.11.006

Carrion-Castillo, A., Van der Haegen, L., Tzourio-Mazoyer, N., Kavaklioglu, T., Badillo, S., Chavent, M., . . . Francks, C. (2019). Genome sequencing for rightward hemispheric language dominance. *Genes Brain Behav, 18*(5), e12572. doi:10.1111/gbb.12572

Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience, 4*, 7. doi:10.1186/s13742-015-0047-8

Cole, M. W., Bassett, D. S., Power, J. D., Braver, T. S., & Petersen, S. E. (2014). Intrinsic and task-evoked network architectures of the human brain. *Neuron, 83*(1), 238-251. doi:10.1016/j.neuron.2014.05.014

Consortium, T. G. O. (2019). The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res, 47*(D1), D330-d338. doi:10.1093/nar/gky1055

Cuellar-Partida, G., Tung, J. Y., Eriksson, N., Albrecht, E., Aliev, F., Andreassen, O. A., . . . Medland, S. E. (2020). Genome-wide association study identifies 48 common genetic variants associated with handedness. *Nat Hum Behav*. doi:10.1038/s41562-020-00956-y

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., . . . Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics, 27*(15), 2156-2158. doi:10.1093/bioinformatics/btr330

Davey Smith, G., & Hemani, G. (2014). Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet, 23*(R1), R89-98. doi:10.1093/hmg/ddu328

Day, F. R., Loh, P. R., Scott, R. A., Ong, K. K., & Perry, J. R. (2016). A Robust Example of Collider Bias in a Genetic Association Study. *Am J Hum Genet, 98*(2), 392-393. doi:10.1016/j.ajhg.2015.12.019

de Kovel, C. G. F., Carrion-Castillo, A., & Francks, C. (2019). A large-scale population study of early life factors influencing left-handedness. *Sci Rep, 9*(1), 584. doi:10.1038/s41598-018-37423-8

de Kovel, C. G. F., & Francks, C. (2018). The molecular genetics of hand preference revisited. *bioRxiv*, 447177. doi:10.1101/447177

de Kovel, C. G. F., & Francks, C. (2019). The molecular genetics of hand preference revisited. *Sci Rep, 9*(1), 5986. doi:10.1038/s41598-019-42515-0

de Kovel, C. G. F., Lisgo, S. N., Fisher, S. E., & Francks, C. (2018). Subtle left-right asymmetry of gene expression profiles in embryonic and foetal human brains. *Sci Rep, 8*(1), 12606. doi:10.1038/s41598-018-29496-2

de Leeuw, C. A., Mooij, J. M., Heskes, T., & Posthuma, D. (2015). MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput Biol, 11*(4), e1004219. doi:10.1371/journal.pcbi.1004219

Demontis, D., Walters, R. K., Martin, J., Mattheisen, M., Als, T. D., Agerbo, E., . . . Neale, B. M. (2019). Discovery of the first genome-wide significant risk

loci for attention deficit/hyperactivity disorder. *Nat Genet, 51*(1), 63-75. doi:10.1038/s41588-018-0269-7

Duboc, V., Dufourcq, P., Blader, P., & Roussigne, M. (2015). Asymmetry of the Brain: Development and Implications. *Annu Rev Genet, 49*, 647-672. doi:10.1146/annurev-genet-112414-055322

Fischl, B. (2012). FreeSurfer. *Neuroimage, 62*(2), 774-781. doi:10.1016/j.neuroimage.2012.01.021

Francks, C. (2015). Exploring human brain lateralization with molecular genetics and genomics. *Ann N Y Acad Sci, 1359*, 1-13. doi:10.1111/nyas.12770

Gervain, J., Macagno, F., Cogoi, S., Pena, M., & Mehler, J. (2008). The neonate brain detects speech structure. *Proc Natl Acad Sci U S A, 105*(37), 14222-14227. doi:10.1073/pnas.0806530105

Glover, G. H. (2011). Overview of functional magnetic resonance imaging. *Neurosurg Clin N Am, 22*(2), 133-139, vii. doi:10.1016/j.nec.2010.11.001

Griffanti, L., Salimi-Khorshidi, G., Beckmann, C. F., Auerbach, E. J., Douaud, G., Sexton, C. E., . . . Smith, S. M. (2014). ICA-based artefact removal and accelerated fMRI acquisition for improved resting state network imaging. *Neuroimage, 95*, 232-247. doi:10.1016/j.neuroimage.2014.03.034

Grove, J., Ripke, S., Als, T. D., Mattheisen, M., Walters, R. K., Won, H., . . . Borglum, A. D. (2019). Identification of common genetic risk variants for autism spectrum disorder. *Nat Genet, 51*(3), 431-444. doi:10.1038/s41588-019-0344-8

Guadalupe, T., Mathias, S. R., vanErp, T. G., Whelan, C. D., Zwiers, M. P., Abe, Y., . . . Francks, C. (2016). Human subcortical brain asymmetries in 15,847 people worldwide reveal effects of age and sex. *Brain Imaging Behav*. doi:10.1007/s11682-016-9629-z

Hepper, P. G. (2013). The developmental origins of laterality: fetal handedness. *Dev Psychobiol, 55*(6), 588-595. doi:10.1002/dev.21119

Inaki, M., Liu, J., & Matsuno, K. (2016). Cell chirality: its origin and roles in left-right asymmetric development. *Philos Trans R Soc Lond B Biol Sci, 371*(1710). doi:10.1098/rstb.2015.0403

Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage, 17*(2), 825-841. doi:10.1016/s1053-8119(02)91132-8

Jenkinson, M., Beckmann, C. F., Behrens, T. E., Woolrich, M. W., & Smith, S. M. (2012). FSL. *Neuroimage, 62*(2), 782-790. doi:10.1016/j.neuroimage.2011.09.015

Joliot, M., Jobard, G., Naveau, M., Delcroix, N., Petit, L., Zago, L., . . . Tzourio-Mazoyer, N. (2015). AICHA: An atlas of intrinsic connectivity of homotopic areas. *J Neurosci Methods, 254*, 46-59. doi:10.1016/j.jneumeth.2015.07.013

Joliot, M., Tzourio-Mazoyer, N., & Mazoyer, B. (2016). Intra-hemispheric intrinsic connectivity asymmetry and its relationships with handedness and

language Lateralization. *Neuropsychologia, 93*(Pt B), 437-447. doi:10.1016/j.neuropsychologia.2016.03.013

Kasprian, G., Langs, G., Brugger, P. C., Bittner, M., Weber, M., Arantes, M., & Prayer, D. (2011). The prenatal origin of hemispheric asymmetry: an in utero neuroimaging study. *Cereb Cortex, 21*(5), 1076-1083. doi:10.1093/cercor/bhq179

Knecht, S., Drager, B., Deppe, M., Bobe, L., Lohmann, H., Floel, A., . . . Henningsen, H. (2000). Handedness and hemispheric language dominance in healthy humans. *Brain, 123 Pt 12*, 2512-2518.

Kong, X.-Z., Postema, M., Carrión Castillo, A., Pepe, A., Crivello, F., Joliot, M., . . . Francks, C. (2020). Large-scale Phenomic and Genomic Analysis of Brain Asymmetrical Skew. *bioRxiv*, 756395. doi:10.1101/756395

Kong, X. Z., Mathias, S. R., Guadalupe, T., Glahn, D. C., Franke, B., Crivello, F., . . . Francks, C. (2018). Mapping cortical brain asymmetry in 17,141 healthy individuals worldwide via the ENIGMA Consortium. *Proc Natl Acad Sci U S A*. doi:10.1073/pnas.1718418115

Krupa, K., & Bekiesińska-Figatowska, M. (2015). Artifacts in magnetic resonance imaging. *Pol J Radiol, 80*, 93-106. doi:10.12659/pjr.892628

Labache, L., Joliot, M., Saracco, J., Jobard, G., Hesling, I., Zago, L., . . . Tzourio-Mazoyer, N. (2019). A SENtence Supramodal Areas AtlaS (SENSAAS) based on multiple task-induced activation mapping and graph analysis of intrinsic connectivity in 144 healthy right-handers. *Brain Struct Funct, 224*(2), 859-882. doi:10.1007/s00429-018-1810-2

Labache, L., Mazoyer, B., Joliot, M., Crivello, F., Hesling, I., & Tzourio-Mazoyer, N. (2020). Typical and atypical language brain organization based on intrinsic connectivity and multitask functional asymmetries. *eLife, 9*. doi:10.7554/eLife.58722

Lee, J. J., Wedow, R., Okbay, A., Kong, E., Maghzian, O., Zacher, M., . . . Cesarini, D. (2018). Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat Genet, 50*(8), 1112-1121. doi:10.1038/s41588-018-0147-3

Lee, S., Emond, M. J., Bamshad, M. J., Barnes, K. C., Rieder, M. J., Nickerson, D. A., . . . Lin, X. (2012). Optimal unified approach for rare-variant association testing with application to small-sample case-control whole-exome sequencing studies. *Am J Hum Genet, 91*(2), 224-237. doi:10.1016/j.ajhg.2012.06.007

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics, 25*(14), 1754-1760. doi:10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., . . . Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics, 25*(16), 2078-2079. doi:10.1093/bioinformatics/btp352

Lin, M. F., Rodeh, O., Penn, J., Bai, X., Reid, J. G., Krasheninina, O., & Salerno, W. J. (2018). GLnexus: joint variant calling for large cohort sequencing. *bioRxiv*, 343970. doi:10.1101/343970

Liu, H., Zhang, L., Xi, Q., Zhao, X., Wang, F., Wang, X., . . . Lin, Q. (2018). Changes in Brain Lateralization in Patients with Mild Cognitive Impairment and Alzheimer's Disease: A Resting-State Functional Magnetic Resonance Study from Alzheimer's Disease Neuroimaging Initiative. *Front Neurol, 9*, 3. doi:10.3389/fneur.2018.00003

Liu, X., Li, C., Mou, C., Dong, Y., & Tu, Y. (2020). dbNSFP v4: a comprehensive database of transcript-specific functional predictions and annotations for human nonsynonymous and splice-site SNVs. *Genome Med, 12*(1), 103. doi:10.1186/s13073-020-00803-9

Lv, H., Wang, Z., Tong, E., Williams, L. M., Zaharchuk, G., Zeineh, M., . . . Wintermark, M. (2018). Resting-State Functional MRI: Everything That Nonexperts Have Always Wanted to Know. *AJNR Am J Neuroradiol, 39*(8), 1390-1399. doi:10.3174/ajnr.A5527

Ma, C., Boehnke, M., & Lee, S. (2015). Evaluating the Calibration and Power of Three Gene-Based Association Tests of Rare Variants for the X Chromosome. *Genet Epidemiol, 39*(7), 499-508. doi:10.1002/gepi.21935

Malishkevich, A., Leyk, J., Goldbaum, O., Richter-Landsberg, C., & Gozes, I. (2015). ADNP/ADNP2 expression in oligodendrocytes: implication for myelin-related neurodevelopment. *J Mol Neurosci, 57*(2), 304-313. doi:10.1007/s12031-015-0640-4

Mazoyer, B., Zago, L., Jobard, G., Crivello, F., Joliot, M., Perchey, G., . . . Tzourio-Mazoyer, N. (2014). Gaussian mixture modeling of hemispheric lateralization for language in a large sample of healthy individuals balanced for handedness. *PLoS One, 9*(6), e101165. doi:10.1371/journal.pone.0101165

McCaw, Z. R., Lane, J. M., Saxena, R., Redline, S., & Lin, X. (2019). Operating characteristics of the rank-based inverse normal transformation for quantitative trait analysis in genome-wide association studies. *Biometrics*. doi:10.1111/biom.13214

McDowell, G., Rajadurai, S., & Levin, M. (2016). From cytoskeletal dynamics to organ asymmetry: a nonlinear, regulative pathway underlies left-right patterning. *Philos Trans R Soc Lond B Biol Sci, 371*(1710). doi:10.1098/rstb.2015.0409

Medland, S. E., Duffy, D. L., Wright, M. J., Geffen, G. M., & Martin, N. G. (2006). Handedness in twins: joint analysis of data from 35 samples. *Twin Res Hum Genet, 9*(1), 46-53. doi:10.1375/183242706776402885

Mellet, E., Zago, L., Jobard, G., Crivello, F., Petit, L., Joliot, M., . . . Tzourio-Mazoyer, N. (2014). Weak language lateralization affects both verbal and spatial skills: an fMRI study in 297 subjects. *Neuropsychologia, 65*, 56-62. doi:10.1016/j.neuropsychologia.2014.10.010

Miao, N., Lai, X., Zeng, Z., Cai, W., Chen, W., & Sun, T. (2020). Differential expression of microRNAs in the human fetal left and right cerebral cortex. *Mol Biol Rep*. doi:10.1007/s11033-020-05708-9

Millard, L. A. C., Davies, N. M., Gaunt, T. R., Davey Smith, G., & Tilling, K. (2017). Software Application Profile: PHESANT: a tool for performing automated phenome scans in UK Biobank. *Int J Epidemiol*. doi:10.1093/ije/dyx204

Miller, K. L., Alfaro-Almagro, F., Bangerter, N. K., Thomas, D. L., Yacoub, E., Xu, J., . . . Smith, S. M. (2016). Multimodal population brain imaging in the UK Biobank prospective epidemiological study. *Nat Neurosci, 19*(11), 1523-1536. doi:10.1038/nn.4393

Mitchell, K. (2018). Innate: How the Wiring of Our Brains Shapes Who We Are. *Princeton University Press*. doi:doi:10.2307/j.ctvc77m71

Murphy, K., Birn, R. M., Handwerker, D. A., Jones, T. B., & Bandettini, P. A. (2009). The impact of global signal regression on resting state correlations: are anti-correlated networks introduced? *Neuroimage, 44*(3), 893-905. doi:10.1016/j.neuroimage.2008.09.036

Murphy, K., & Fox, M. D. (2017). Towards a consensus regarding global signal regression for resting state functional connectivity MRI. *Neuroimage, 154*, 169-173. doi:10.1016/j.neuroimage.2016.11.052

Ocklenburg, S., Güntürkün, O., Hugdahl, K., & Hirnstein, M. (2015). Laterality and mental disorders in the postgenomic age--A closer look at schizophrenia and language lateralization. *Neurosci Biobehav Rev, 59*, 100-110. doi:10.1016/j.neubiorev.2015.08.019

Ocklenburg, S., Schmitz, J., Moinfar, Z., Moser, D., Klose, R., Lor, S., . . . Gunturkun, O. (2017). Epigenetic regulation of lateralized fetal spinal gene expression underlies hemispheric asymmetries. *eLife, 6*. doi:10.7554/eLife.22784

Ocklenburg, S., Strockens, F., Bless, J. J., Hugdahl, K., Westerhausen, R., & Manns, M. (2016). Investigating heritability of laterality and cognitive control in speech perception. *Brain Cogn, 109*, 34-39. doi:10.1016/j.bandc.2016.09.003

Okumura, T., Utsuno, H., Kuroda, J., Gittenberger, E., Asami, T., & Matsuno, K. (2008). The development and evolution of left-right asymmetry in invertebrates: lessons from Drosophila and snails. *Dev Dyn, 237*(12), 3497-3515. doi:10.1002/dvdy.21788

Packheiser, J., Schmitz, J., Arning, L., Beste, C., Güntürkün, O., & Ocklenburg, S. (2020). A large-scale estimate on the relationship between language and motor lateralization. *Sci Rep, 10*(1), 13027. doi:10.1038/s41598-020-70057-3

Parma, V., Brasselet, R., Zoia, S., Bulgheroni, M., & Castiello, U. (2017). The origin of human handedness and its role in pre-birth motor control. *Sci Rep, 7*(1), 16804. doi:10.1038/s41598-017-16827-y

Power, J. D., Cohen, A. L., Nelson, S. M., Wig, G. S., Barnes, K. A., Church, J. A., . . . Petersen, S. E. (2011). Functional network organization of the human brain. *Neuron, 72*(4), 665-678. doi:10.1016/j.neuron.2011.09.006

Ripke, S., Neale, B. M., Corvin, A., Walters, J. T., Farh, K. H., Holmans, P. A., . . . O'Donovan, M. C. (2014). Biological insights from 108 schizophrenia-associated genetic loci. *Nature, 511*(7510), 421-427. doi:10.1038/nature13595

Salimi-Khorshidi, G., Douaud, G., Beckmann, C. F., Glasser, M. F., Griffanti, L., & Smith, S. M. (2014). Automatic denoising of functional MRI data: combining independent component analysis and hierarchical fusion of classifiers. *Neuroimage, 90*, 449-468. doi:10.1016/j.neuroimage.2013.11.046

Savage, J. E., Jansen, P. R., Stringer, S., Watanabe, K., Bryois, J., de Leeuw, C. A., . . . Posthuma, D. (2018). Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nat Genet, 50*(7), 912-919. doi:10.1038/s41588-018-0152-6

Sha, Z., Schijven, D., Carrion-Castillo, A., Joliot, M., Mazoyer, B., Fisher, S. E., . . . Francks, C. (2020). The genetic architecture of structural left-right asymmetry of the human brain. *bioRxiv*, 2020.2006.2030.179721. doi:10.1101/2020.06.30.179721

Smith, S. M., Fox, P. T., Miller, K. L., Glahn, D. C., Fox, P. M., Mackay, C. E., . . . Beckmann, C. F. (2009). Correspondence of the brain's functional architecture during activation and rest. *Proc Natl Acad Sci U S A, 106*(31), 13040-13045. doi:10.1073/pnas.0905267106

Somers, M., Ophoff, R. A., Aukes, M. F., Cantor, R. M., Boks, M. P., Dauwan, M., . . . Sommer, I. E. (2015). Linkage analysis in a Dutch population isolate shows no major gene for left-handedness or atypical language lateralization. *J Neurosci, 35*(23), 8730-8736. doi:10.1523/jneurosci.3287-14.2015

Sragovich, S., Merenlender-Wagner, A., & Gozes, I. (2017). ADNP Plays a Key Role in Autophagy: From Autism to Schizophrenia and Alzheimer's Disease. *Bioessays, 39*(11). doi:10.1002/bies.201700054

Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., . . . Collins, R. (2015). UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med, 12*(3), e1001779. doi:10.1371/journal.pmed.1001779

Sun, T., Patoine, C., Abu-Khalil, A., Visvader, J., Sum, E., Cherry, T. J., . . . Walsh, C. A. (2005). Early asymmetry of gene transcription in embryonic human left and right cerebral cortex. *Science, 308*(5729), 1794-1798. doi:10.1126/science.1110324

Szustakowski, J. D., Balasubramanian, S., Sasson, A., Khalid, S., Bronson, P. G., Kvikstad, E., . . . Reid, J. G. (2020). Advancing Human Genetics Research and Drug Discovery through Exome Sequencing of the UK Biobank. *medRxiv*, 2020.2011.2002.20222232. doi:10.1101/2020.11.02.20222232

Toga, A. W., & Thompson, P. M. (2003). Mapping brain asymmetry. *Nat Rev Neurosci, 4*(1), 37-48. doi:10.1038/nrn1009

Turner, S., Armstrong, L. L., Bradford, Y., Carlson, C. S., Crawford, D. C., Crenshaw, A. T., . . . Ritchie, M. D. (2011). Quality control procedures for genome-wide association studies. *Curr Protoc Hum Genet, Chapter 1*, Unit1.19. doi:10.1002/0471142905.hg0119s68

Turner, S. D. (2014). qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. *bioRxiv*.

Van Hout, C. V., Tachmazidou, I., Backman, J. D., Hoffman, J. D., Liu, D., Pandey, A. K., . . . Baras, A. (2020). Exome sequencing and characterization of 49,960 individuals in the UK Biobank. *Nature, 586*(7831), 749-756. doi:10.1038/s41586-020-2853-0

Vaser, R., Adusumalli, S., Leng, S. N., Sikic, M., & Ng, P. C. (2016). SIFT missense predictions for genomes. *Nat Protoc, 11*(1), 1-9. doi:10.1038/nprot.2015.123

Wiberg, A., Ng, M., Al Omran, Y., Alfaro-Almagro, F., McCarthy, P., Marchini, J., . . . Furniss, D. (2019). Handedness, language areas and neuropsychiatric diseases: insights from brain imaging and genetics. *Brain*. doi:10.1093/brain/awz257

Yang, J., Benyamin, B., McEvoy, B. P., Gordon, S., Henders, A. K., Nyholt, D. R., . . . Visscher, P. M. (2010). Common SNPs explain a large proportion of the heritability for human height. *Nat Genet, 42*(7), 565-569. doi:10.1038/ng.608

Yang, J., Lee, S. H., Goddard, M. E., & Visscher, P. M. (2011). GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet, 88*(1), 76-82. doi:10.1016/j.ajhg.2010.11.011

Yun, T., Li, H., Chang, P. C., Lin, M. F., Carroll, A., & McLean, C. Y. (2021). Accurate, scalable cohort variant calls using DeepVariant and GLnexus. *Bioinformatics*. doi:10.1093/bioinformatics/btaa1081

Zhan, L., Jenkins, L. M., Wolfson, O. E., GadElkarim, J. J., Nocito, K., Thompson, P. M., . . . Leow, A. D. (2017). The significance of negative correlations in brain connectivity. *J Comp Neurol, 525*(15), 3251-3265. doi:10.1002/cne.24274

5.6 Supplementary Information

Table of contents

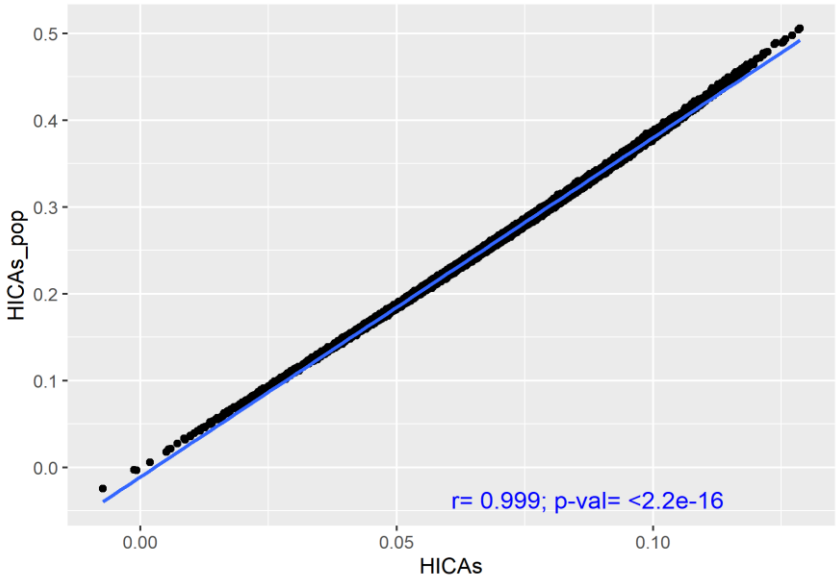


Figure S1. Correlation plot between HICAs and HICAs_pop. The two measures are highly correlated (Pearson's R=0.999, P < 2.2.e-16).

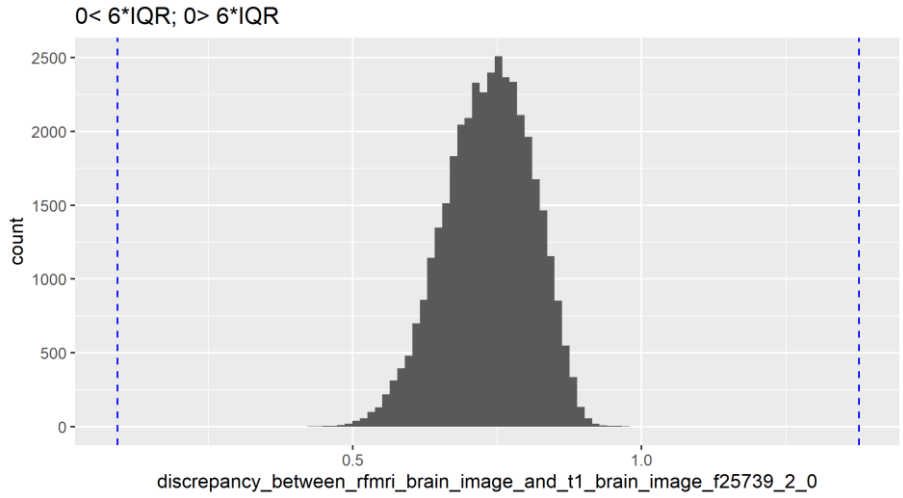


Figure S2. Frequency histogram of the discrepancy between resting-state fMRI brain image and T1 brain image (UK Biobank field 25739) in 37,840 individuals. Dotted blue lines indicate the outlier removal threshold of 6 times the inter quartile range (IQR).

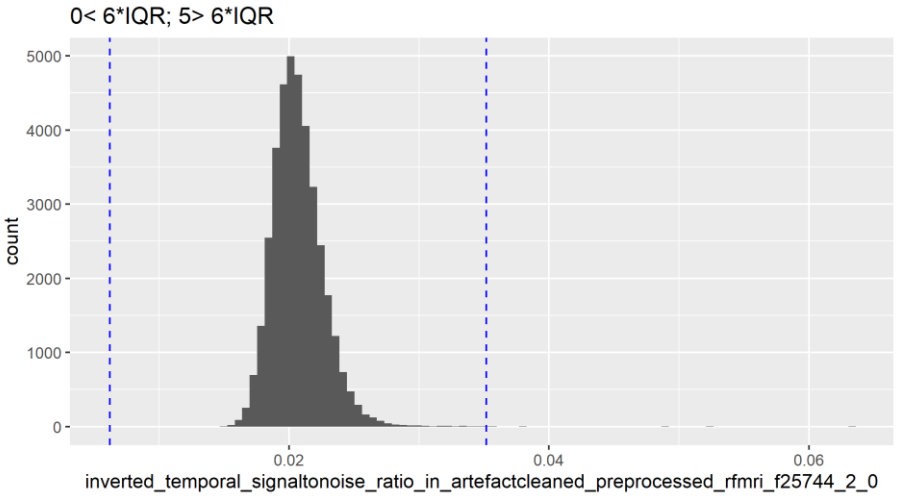


Figure S3. Frequency histogram of the inverted temporal signal-to-noise ratio in artefact-cleaned preprocessed resting-state fMRI (UK Biobank field 25744) in 37,840 individuals. Dotted blue lines indicate the outlier removal threshold of 6 times the inter quartile range (IQR).

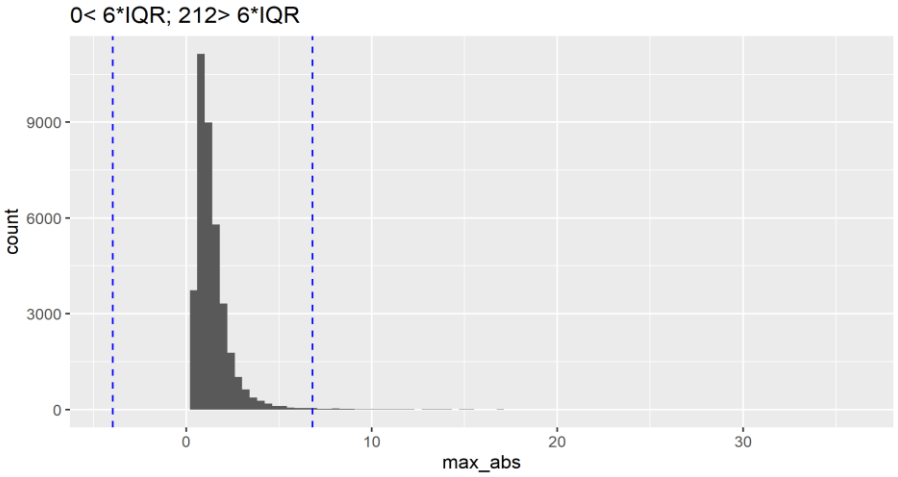


Figure S4. Frequency histogram of the maximum absolute resting-state fMRI head motion averaged across space and time points, as calculated from 'prefiltered_func_data_mcf_abs.rms', in 37,840 individuals. Dotted blue lines indicate the outlier removal threshold of 6 times the inter quartile range (IQR).

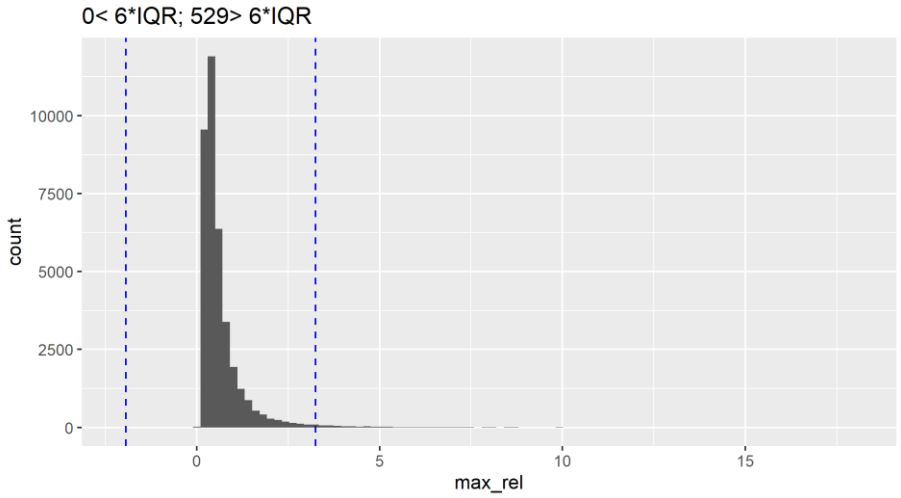


Figure S5. Frequency histogram of the maximum relative resting-state fMRI head motion averaged across space and time points, as calculated from 'prefiltered_func_data_mcf_rel.rms', in 37,840 individuals. Dotted blue lines indicate the outlier removal threshold of 6 times the inter quartile range (IQR).

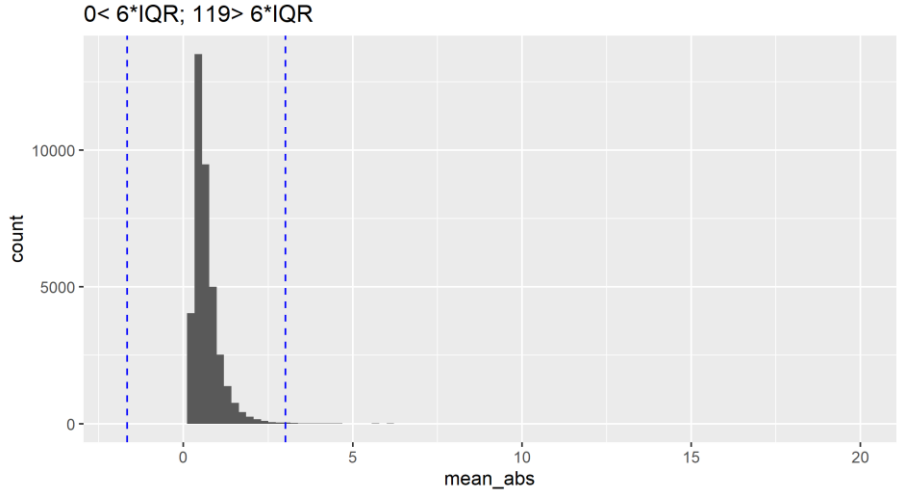


Figure S6. Frequency histogram of the mean absolute resting-state fMRI head motion averaged across space and time points, as calculated from 'prefiltered_func_data_mcf_abs.rms', in 37,840 individuals. Dotted blue lines indicate the outlier removal threshold of 6 times the inter quartile range (IQR).

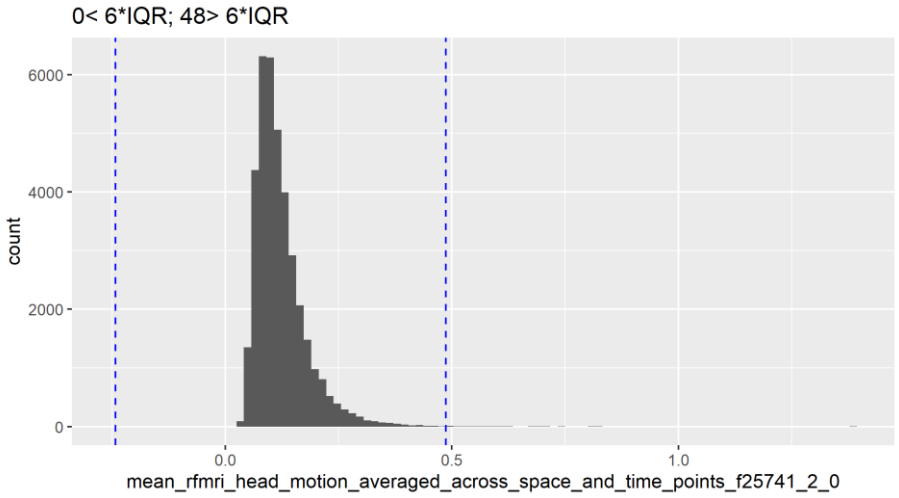


Figure S7. Frequency histogram of the mean relative resting-state fMRI head motion averaged across space and time points (UK Biobank field 25741), in 37,840 individuals. Dotted blue lines indicate the outlier removal threshold of 6 times the inter quartile range (IQR)

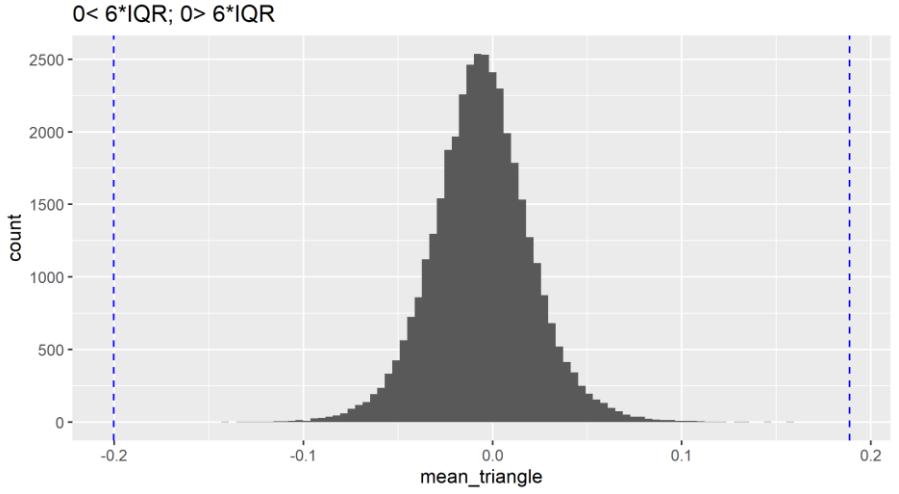


Figure S8. Frequency histogram of 'mean triangle', i.e., the per-individual average value of the functional connectivity asymmetry matrix in 37,840 individuals. Dotted blue lines indicate the outlier removal threshold of 6 times the inter quartile range (IQR).

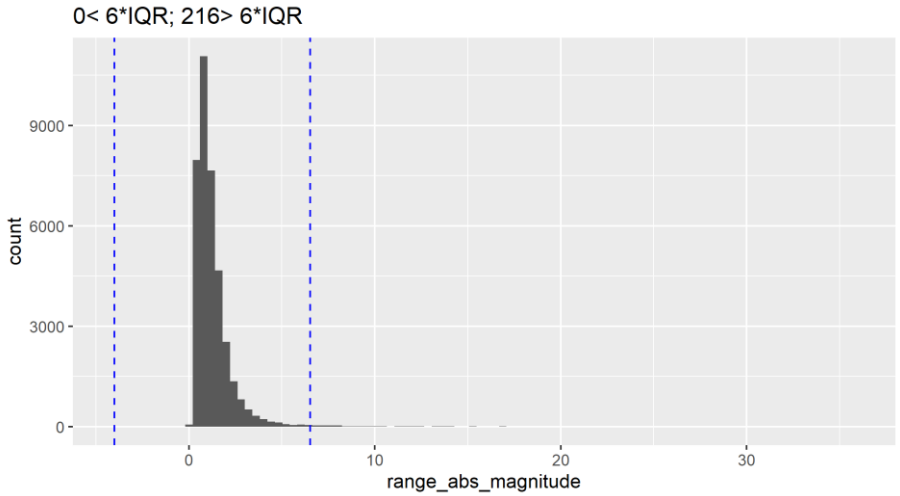


Figure S9. Frequency histogram of the magnitude (difference) between the upper and lower range of absolute resting-state fMRI head motion averaged across space and time points in 37,840 individuals. Dotted blue lines indicate the outlier removal threshold of 6 times the inter quartile range (IQR).

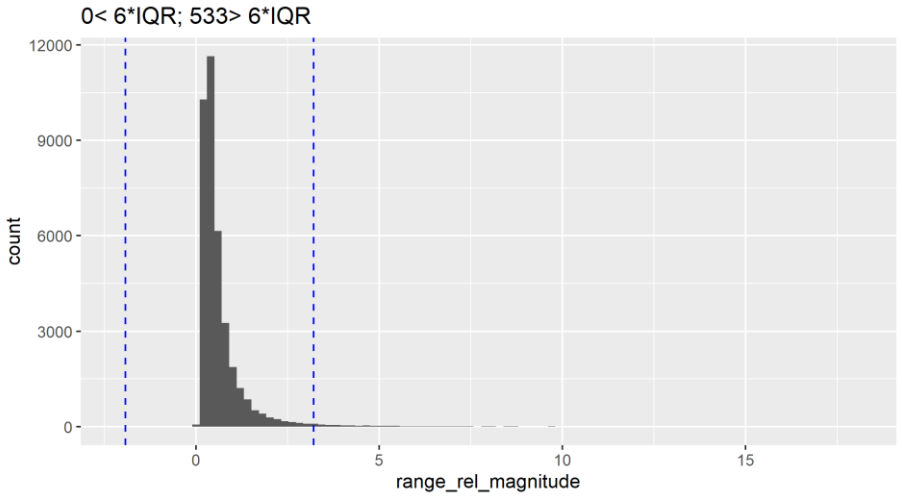


Figure S10. Frequency histogram of the magnitude (difference) between the upper and lower range of relative resting-state fMRI head motion averaged across space and time points in 37,840 individuals. Dotted blue lines indicate the outlier removal threshold of 6 times the inter quartile range (IQR).

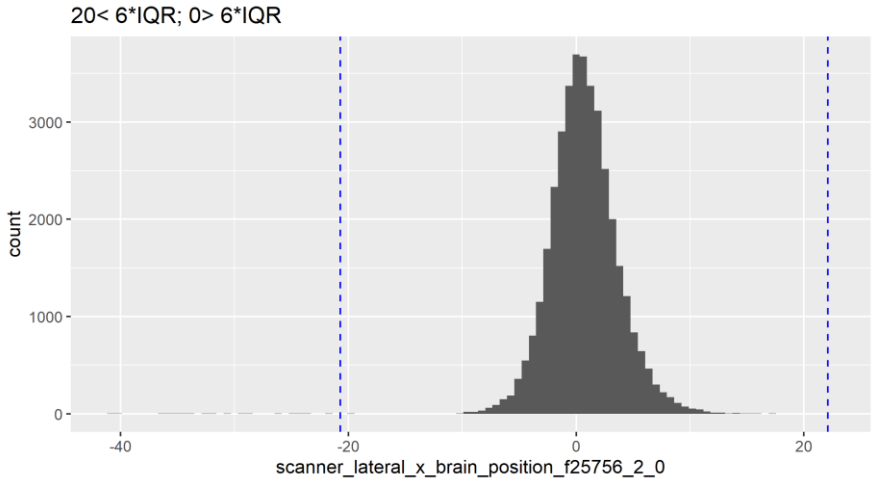


Figure S11. Frequency histogram of scanner lateral x brain position (UK Biobank field 25756) in 37,840 individuals. Dotted blue lines indicate the outlier removal threshold of 6 times the inter quartile range (IQR).

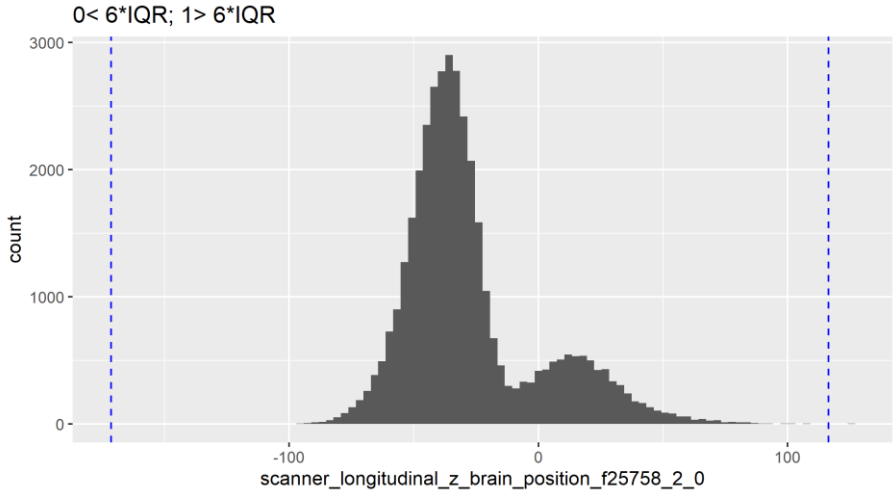


Figure S12. Frequency histogram of scanner longitudinal z brain position (UK Biobank field 25758) in 37,840 individuals. Dotted blue lines indicate the outlier removal threshold of 6 times the inter quartile range (IQR).

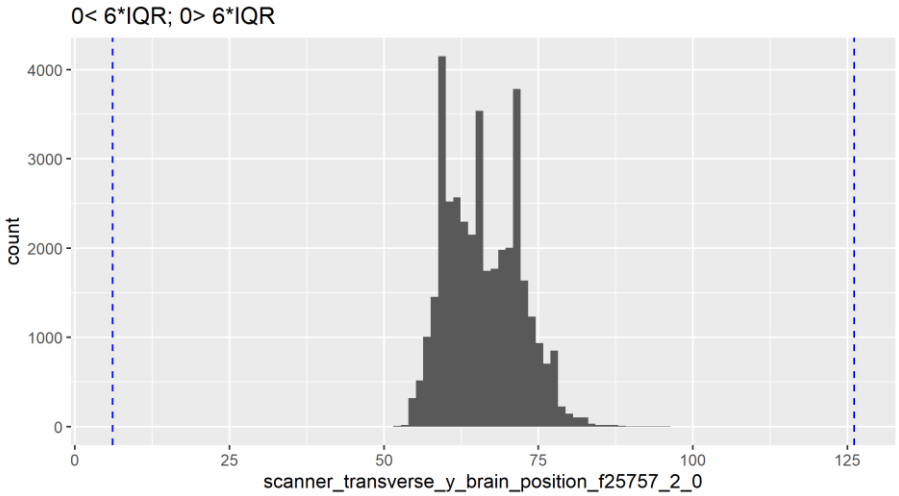


Figure S13. Frequency histogram of scanner transverse y brain position (UK Biobank field 25757) in 37,840 individuals. Dotted blue lines indicate the outlier removal threshold of 6 times the inter quartile range (IQR).

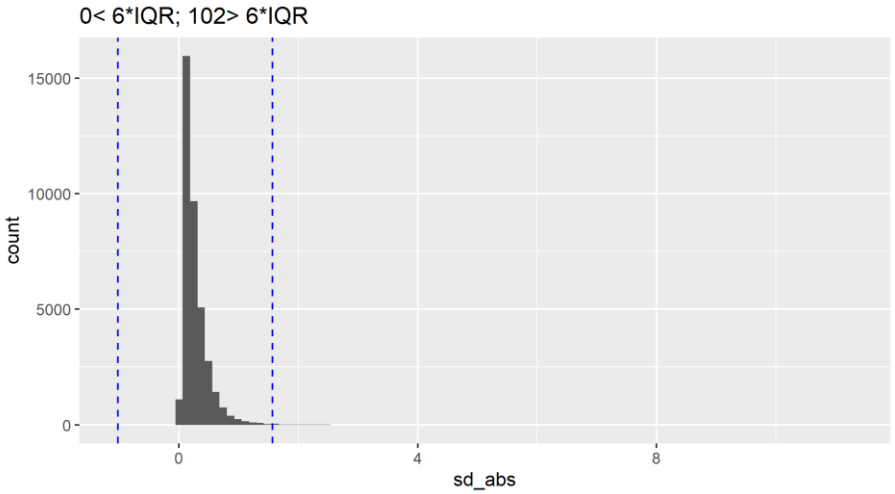


Figure S14. Frequency histogram of the standard deviation of absolute resting-state fMRI head motion averaged across space and time points in 37,840 individuals. Dotted blue lines indicate the outlier removal threshold of 6 times the inter quartile range (IQR).

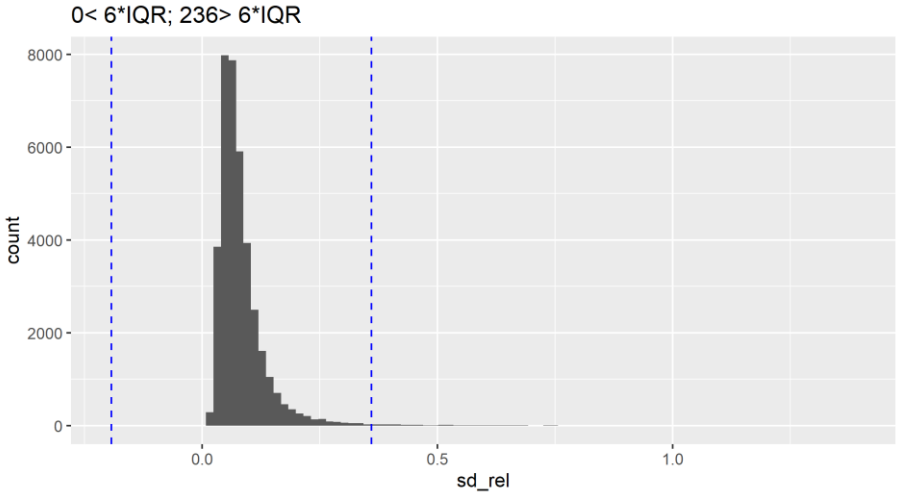


Figure S15. Frequency histogram of the standard deviation of relative resting-state fMRI head motion averaged across space and time points in 37,840 individuals. Dotted blue lines indicate the outlier removal threshold of 6 times the inter quartile range (IQR).

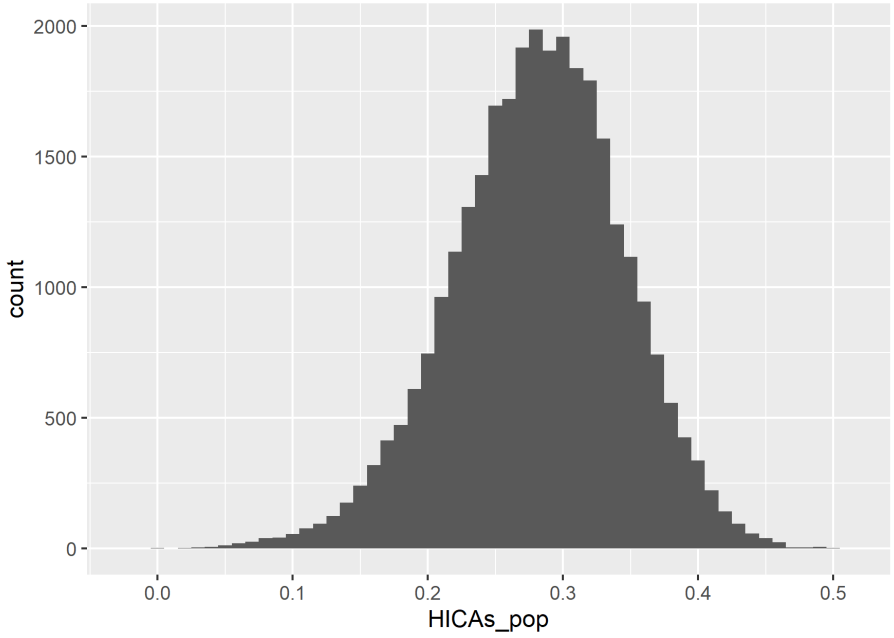


Figure S16. Frequency histogram of HICAs_pop in 30,660 individuals.

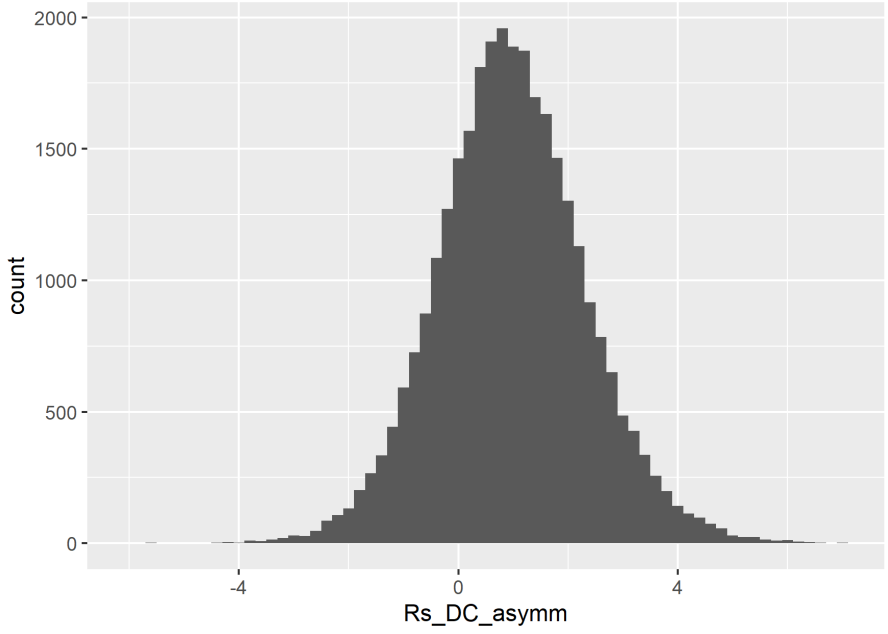


Figure S17. Frequency histogram of Rs_DC_asymm in 30,660 individuals

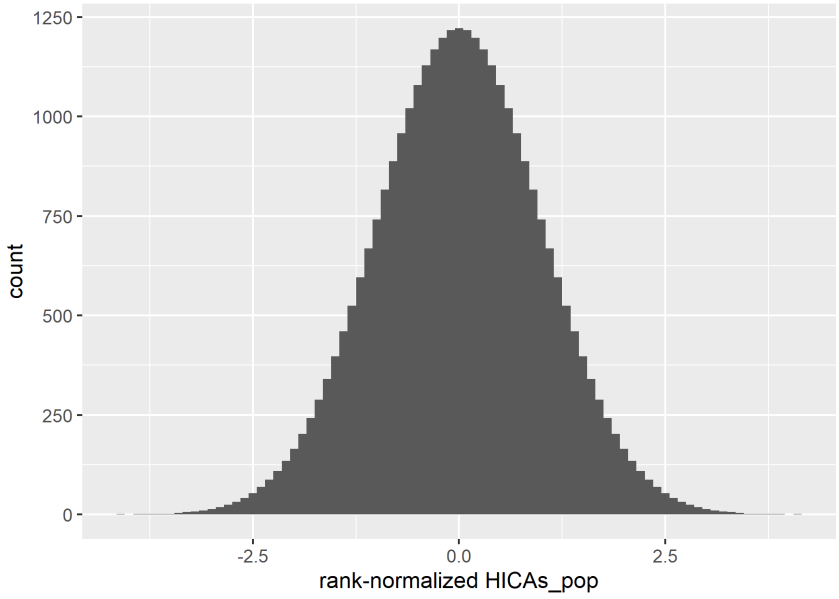


Figure S18. Frequency histogram of rank-normalized HICAs_pop in 30,660 individuals.

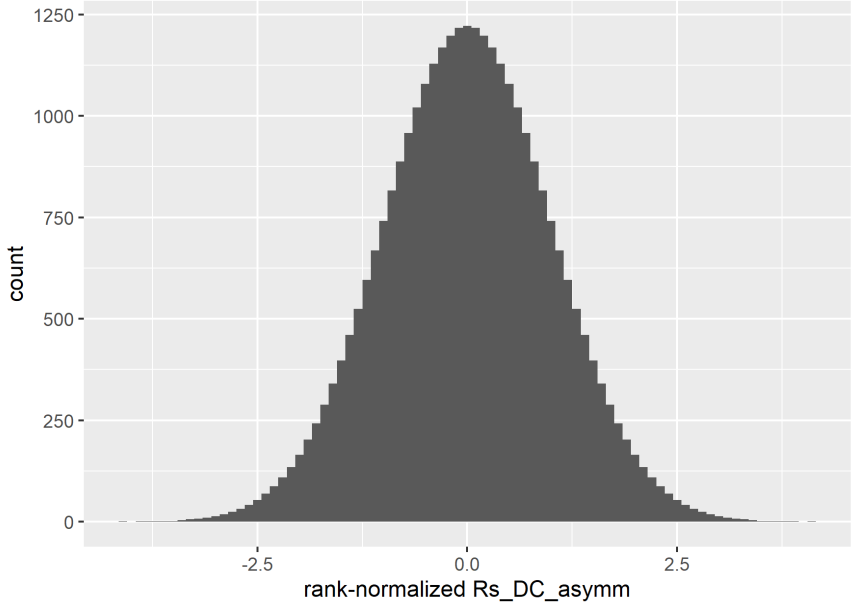


Figure S19. Frequency histogram of rank-normalized Rs_DC_asymm in 30,660 individuals.

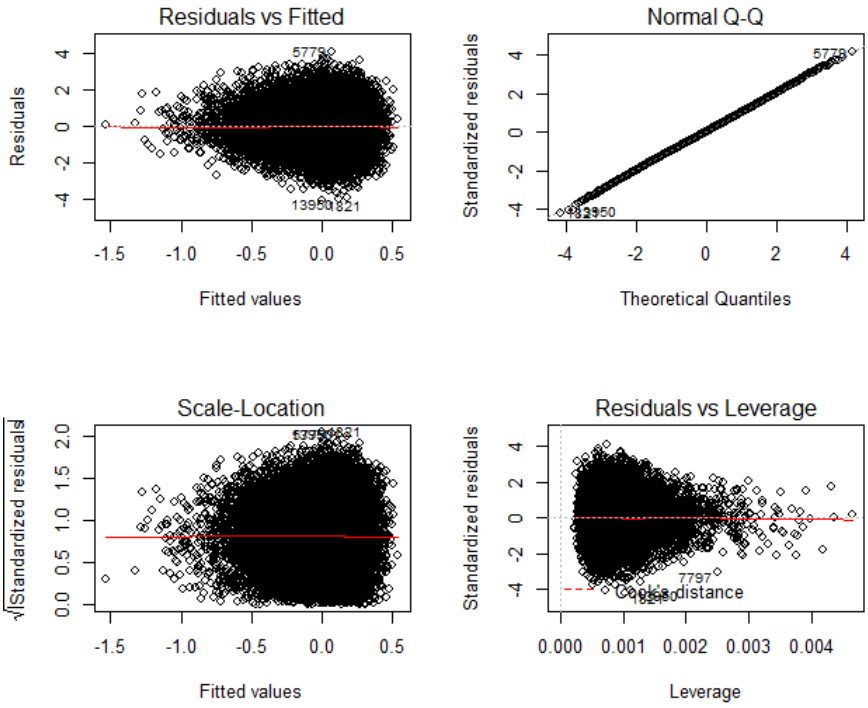


Figure S20. Diagnostic plots for the linear regression analyses on rank-normalized HICAs_pop in 30,660 individuals, including genotyping array as a covariate.

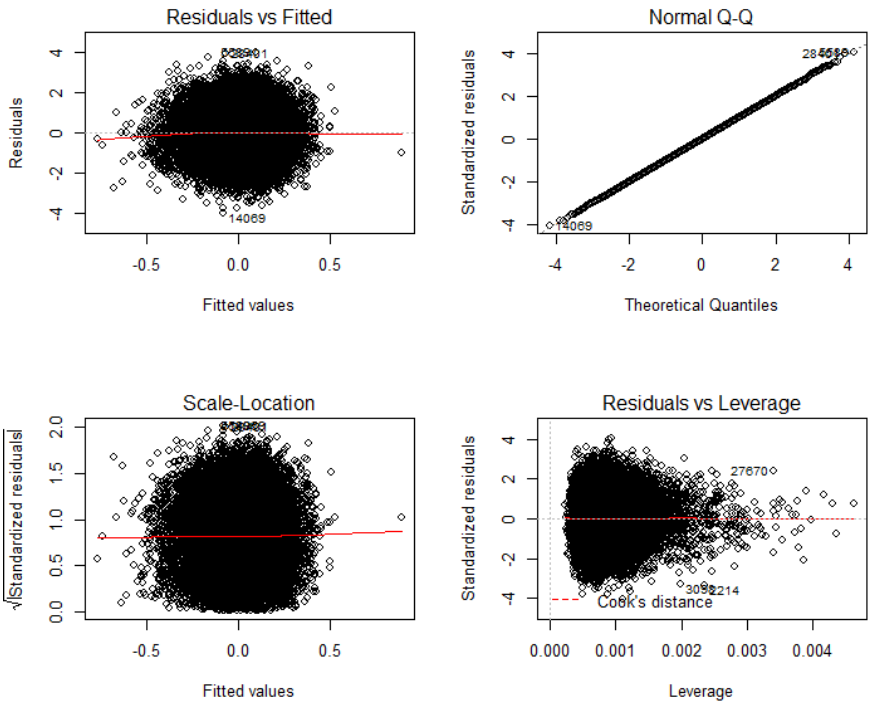


Figure S21. Diagnostic plots for the linear regression analyses on rank-normalized Rs_DC_asymm in 30,660 individuals, including genotyping array as a covariate.

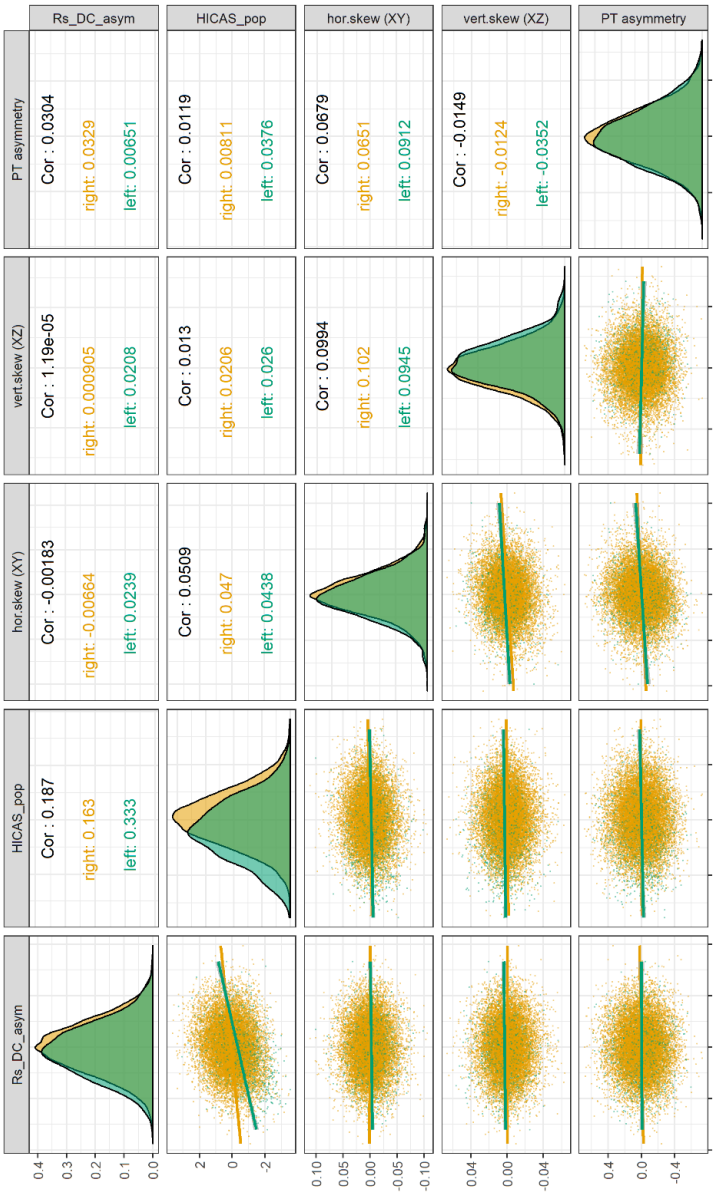


Figure S22. Phenotypic correlations between Rs_DC_asymm, HICAs_pop, horizontal brain skew, vertical brain skew, and planum temporale (PT) asymmetry, split by handedness. *Orange*=right-handers, *green*=left-handers.

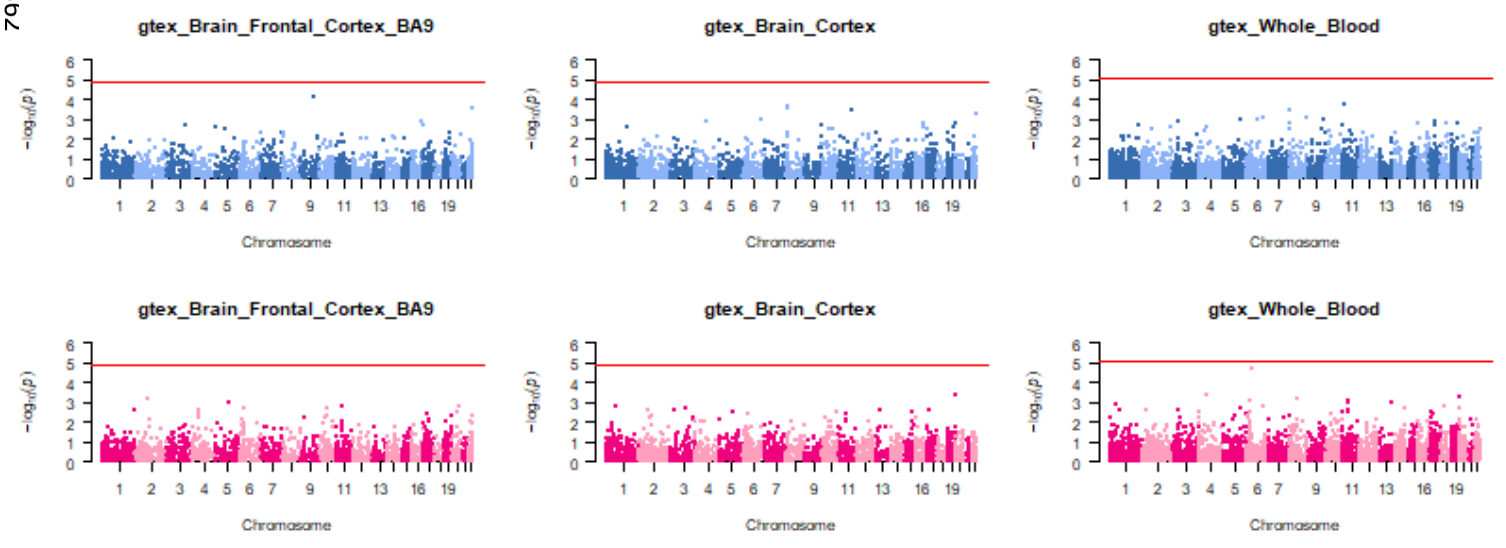


Figure S23. Gene-based manhattan plots for HICAs_pop (top) and Rs_DC_asymm (bottom) based on the TWAS with GteX Brain Frontal Cortex BA9 (left), GteX Brain Cortex (middle) and GteXWhole Blood (right). The horizontal red line indicates the Bonferroni threshold (i.e., 0.05/number of genes tested).

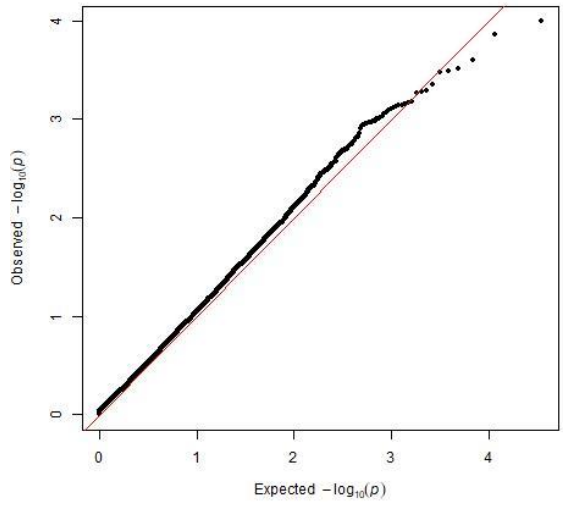


Figure S24. QQ-plot of the p-values from the SKAT-O for HICAs_pop, based on the strict filter variant set.

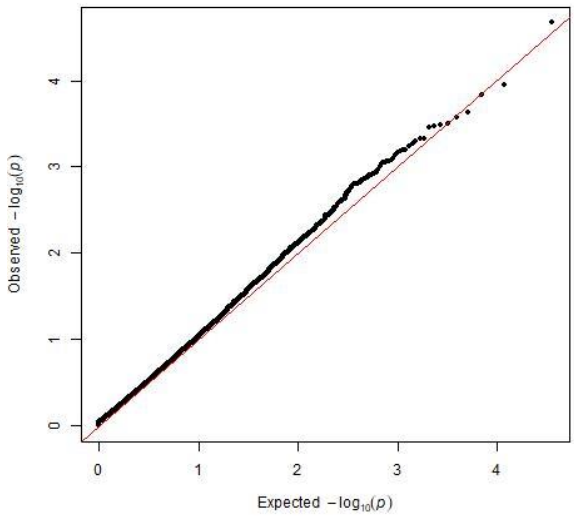


Figure S25. QQ-plot of the p-values from the SKAT-O for HICAs_pop, based on the loose filter variant set

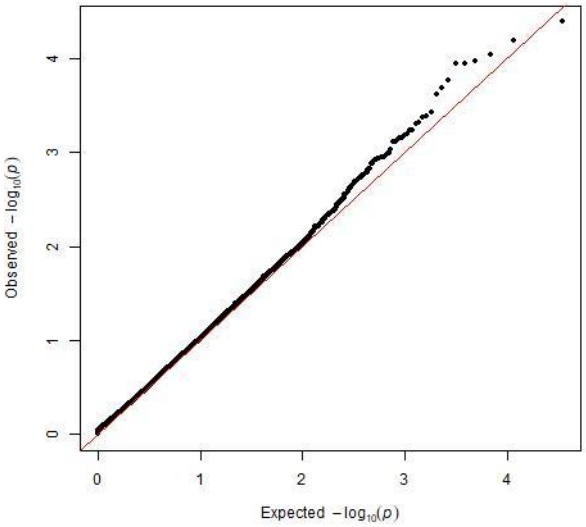


Figure S26. QQ-plot of the p-values from the SKAT-O for Rs_DC_asymm, based on the strict filter variant set**.**

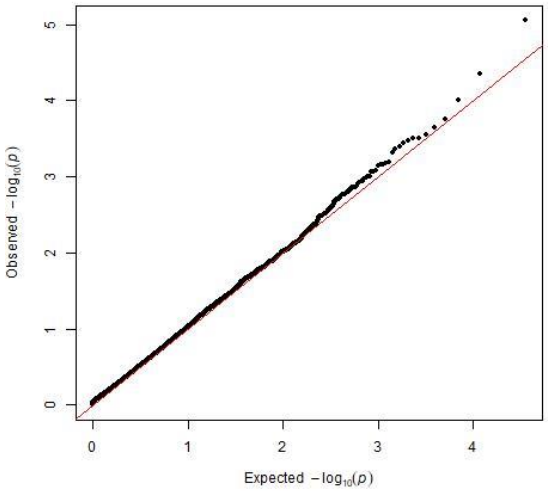


Figure S27. QQ-plot of the p-values from the SKAT-O for Rs_DC_asymm, based on the loose filter variant set**.**

Table S1. List of brain-related diagnostic codes from ICD9.

ICD9 coding	meaning	Node id	Parent id
45	045 Acute poliomyelitis	722	25
450	0450 Acute paralytic poliomylitis specified as bulbar	723	722
451	0451 Acute poliomyelitis with other paralysis	724	722
452	0452 Acute nonparalytic poliomyelitis	725	722
459	0459 Acute poliomyelitis, unspecified	726	722
46	046 Slow virus infection of central nervous system	727	25
460	0460 Kuru	728	727
461	0461 Jakob-creutzfeldt disease	729	727
462	0462 Subacute sclerosing panencephalitis	730	727
463	0463 Progressive multifocal leucoencephalopathy	731	727
468	0468 Other slow virus infection of central nervous system	732	727
469	0469 Unspecified slow virus infection of central nervous system	733	727
47	047 Meningitis due to enterovirus	734	25
470	0470 Meningitis due to coxsackie virus	735	734
471	0471 Meningitis due to echo virus	736	734
478	0478 Meningitis due to other virus	737	734
479	0479 Meningitis due to unspecified virus	738	734
48	048 Other enterovirus diseases of central nervous system	739	25
489	0489 Other enterovirus diseases of central nervous system	740	739
49	049 Other non-arthropod-borne viral diseases of central nervous system	741	25
490	0490 Lymphocytic choriomeningitis	742	741
491	0491 Meningitis due to adenovirus	743	741
498	0498 Other non-arthropod-borne viral diseases of central nervous system	744	741
499	0499 Unspecified non-arthropod-borne viral diseases of central nervous system	745	741
191	191 Malignant neoplasm of brain	1522	42
1910	1910 Malignant neoplasm of cerebrum, except lobes and ventricles	1523	1522
1911	1911 Malignant neoplasm of brain, frontal lobe	1524	1522
1912	1912 Malignant neoplasm of brain, temporal lobe	1525	1522
1913	1913 Malignant neoplasm of brain, parietal lobe	1526	1522
1914	1914 Malignant neoplasm of brain, occipital lobe	1527	1522
1915	1915 Malignant neoplasm of brain, ventricle	1528	1522
1916	1916 Malignant neoplasm of cerebellum	1529	1522
1917	1917 Malignant neoplasm of brain stem	1530	1522
1918	1918 Malignant neoplasm of brain, other	1531	1522
1919	1919 Malignant neoplasm of brain, unspecified	1532	1522

ICD9 coding	meaning	Node id	Parent id
225	225 Benign neoplasm of brain and other parts of nervous system	1759	45
2250	2250 Benign neoplasm of brain	1760	1759
2251	2251 Benign neoplasm of cranial nerves	1761	1759
2252	2252 Benign neoplasm of cerebral meninges	1762	1759
2253	2253 Benign neoplasm of spinal cord	1763	1759
2254	2254 Benign neoplasm of spinal meninges	1764	1759
2258	2258 Benign neoplasm, other spec. part - brain, nervous system	1765	1759
2259	2259 Benign neoplasm, unspec. part - brain, nervous system	1766	1759
2904	2904 Arteriosclerotic dementia	2531	2526
320	320 Bacterial meningitis	2756	59
3200	3200 Haemophilus meningitis	2757	2756
3201	3201 Pneumococcal meningitis	2758	2756
3202	3202 Streptococcal meningitis	2759	2756
3203	3203 Staphylococcal meningitis	2760	2756
3204	3204 Tuberculous meningitis	2761	2756
3205	3205 Meningococcal meningitis	2762	2756
3207	3207 Meningitis in other bacterial diseases classified elsewhere	2763	2756
3208	3208 Meningitis due to other specified bacteria	2764	2756
3209	3209 Meningitis due to unspecified bacterium	2765	2756
321	321 Meningitis due to other organisms	2766	59
3210	3210 Fungal meningitis	2767	2766
3211	3211 Meningitis due to coxsackie virus	2768	2766
3212	3212 Meningitis due to echo virus	2769	2766
3213	3213 Meningitis due to herpes zoster virus	2770	2766
3214	3214 Meningitis due to herpes simplex virus	2771	2766
3215	3215 Meningitis due to mumps virus	2772	2766
3216	3216 Meningitis due to lymphocytic choriomeningitis virus	2773	2766
3217	3217 Meningitis due to other and unspecified viruses	2774	2766
3218	3218 Meningitis due to other organism	2775	2766
322	322 Meningitis of unspecified cause	2776	59
3220	3220 Nonpyogenic meningitis	2777	2776
3221	3221 Eosinophilic meningitis	2778	2776
3222	3222 Chronic meningitis	2779	2776
3229	3229 Meningitis, unspecified	2780	2776
323	323 Encephalitis, myelitis and encephalomyelitis	2781	59
3230	3230 Kuru	2782	2781
3231	3231 Subacute sclerosing panencephalitis	2783	2781
3232	3232 Poliomyelitis	2784	2781

ICD9 coding	meaning	Node id	Parent id
3233	3233 Arthropod - borne viral encephalitis	2785	2781
3234	3234 Other encephalitis due to infection	2786	2781
3235	3235 Encephalitis following immunization procedures	2787	2781
3236	3236 Postinfectious encephalitis	2788	2781
3237	3237 Toxic encephalitis	2789	2781
3238	3238 Encephalitis, myelitis and encephalomyelitis, other specified cause	2790	2781
3239	3239 Encephalitis, myelitis and encephalomyelitis, unspecified cause	2791	2781
324	324 Intracranial and intraspinal abscess	2792	59
3240	3240 Intracranial abscess	2793	2792
3241	3241 Intraspinal abscess	2794	2792
3249	3249 Intracranial and intraspinal abscess of unspecified site	2795	2792
325	325 Phlebitis and thrombophlebitis of intracranial venous sinuses	2796	59
3259	3259 Phlebitis and thrombophlebitis of intracranial venous sinuses	2797	2796
326	326 Late effects of intracranial abscess or pyogenic infection	2798	59
3269	3269 Late effects of intracranial abscess or pyogenic infection	2799	2798
330	330 Cerebral degenerations usually manifest in childhood	2800	60
3300	3300 Leucodystrophy	2801	2800
3301	3301 Cerebral lipidoses	2802	2800
3302	3302 Cerebral degeneration in generalized lipidoses	2803	2800
3303	3303 Cerebral degen. childhood in other diseases classif. elsewhere	2804	2800
3308	3308 Other cerebral degenerations in childhood	2805	2800
3309	3309 Cerebral degeneration usually manifest in childhood, unspecified	2806	2800
331	331 Other cerebral degenerations	2807	60
3310	3310 Alzheimer's disease	2808	2807
3311	3311 Pick's disease	2809	2807
3312	3312 Senile degeneration of brain	2810	2807
3313	3313 Communicating hydrocephalus	2811	2807
3314	3314 Obstructive hydrocephalus	2812	2807
3315	3315 Jakob-creutzfeldt disease	2813	2807
3316	3316 Progressive multifocal leucoencephalopathy	2814	2807
3317	3317 Cerebral degeneration in other diseases classified elsewhere	2815	2807
3318	3318 Other specified cerebral degeneration	2816	2807
3319	3319 Cerebral degeneration, unspecified	2817	2807
332	332 Parkinson's disease	2818	60

ICD9 coding	meaning	Node id	Parent id
3320	3320 Paralysis agitans	2819	2818
3321	3321 Secondary parkinsonism	2820	2818
333	333 Other extrapyramidal disease and abnormal movement disorders	2821	60
3330	3330 Other degenerative diseases of the basal ganglia	2822	2821
3331	3331 Essential and other specified forms of tremor	2823	2821
3332	3332 Myoclonus	2824	2821
33320	33320 Myoclonus (familial essential)	2825	2824
33321	33321 Myoclonus (myoclonic encephalopathy)	2826	2824
33322	33322 Myoclonus (progressive myoclonic epilepsy)	2827	2824
33329	33329 Myoclonus (other and unspecified)	2828	2824
3333	3333 Tics of organic origin	2829	2821
3334	3334 Huntington's chorea	2830	2821
3335	3335 Other choreas	2831	2821
3336	3336 Idiopathic torsion dystonia	2832	2821
3337	3337 Symptomatic torsion dystonia	2833	2821
33370	33370 Symptomatic torsion dystonia (athetoid cerebral palsy or vogt's disease)	2834	2833
33371	33371 Symptomatic torsion dystonia (double athetosis)	2835	2833
33379	33379 Symptomatic torsion dystonia (other)	2836	2833
3338	3338 Fragments of torsion dystonia	2837	2821
3339	3339 Other extrapyramidal disease and abnormal movement disorders	2838	2821
334	334 Spinocerebellar disease	2839	60
3340	3340 Friedreich's ataxia	2840	2839
3341	3341 Hereditary spastic paraplegia	2841	2839
3342	3342 Primary cerebellar degeneration	2842	2839
3343	3343 Other cerebellar ataxia	2843	2839
3344	3344 Cerebellar ataxia in disease classified elsewhere	2844	2839
3348	3348 Other specified spinocerebellar disease	2845	2839
3349	3349 Spinocerebellar disease, unspecified	2846	2839
335	335 Anterior horn cell disease	2847	60
3350	3350 Werdnig-hoffman disease	2848	2847
3351	3351 Spinal muscular atrophy	2849	2847
3352	3352 Motor neurone disease	2850	2847
3358	3358 Other specified anterior horn cell disease	2851	2847
3359	3359 Anterior horn cell disease, unspecified	2852	2847
340	340 Multiple sclerosis	2864	63
3409	3409 Multiple sclerosis	2865	2864
341	341 Other demyelinating diseases of central nervous system	2866	63

ICD9 coding	meaning	Node id	Parent id
3410	3410 Neuromyelitis optica	2867	2866
3411	3411 Schilder's disease	2868	2866
3418	3418 Other specified demyelinating diseases of the central nervous system	2869	2866
3419	3419 Demyelinating diseases of the central nervous system, unspecified	2870	2866
342	342 Hemiplegia	2871	63
3420	3420 Flaccid hemiplegia	2872	2871
3421	3421 Spastic hemiplegia	2873	2871
3429	3429 Hemiplegia, unspecified	2874	2871
343	343 Infantile cerebral palsy	2875	63
3430	3430 Infantile cerebral palsy, diplegic	2876	2875
34300	34300 Infantile cerebral palsy, diplegic (unqualified)	2877	2876
34301	34301 Infantile cerebral palsy, diplegic (paraplegic)	2878	2876
3431	3431 Infantile cerebral palsy, hemiplegic	2879	2875
3432	3432 Infantile cerebral palsy, quadriplegic	2880	2875
3433	3433 Infantile cerebral palsy, monoplegic	2881	2875
3434	3434 Infantile hemiplegia	2882	2875
3438	3438 Other specified infantile cerebral palsy	2883	2875
34380	34380 Other specified infantile cerebral palsy (ataxic)	2884	2883
34381	34381 Other specified infantile cerebral palsy (flaccid)	2885	2883
34389	34389 Other specified infantile cerebral palsy	2886	2883
3439	3439 Infantile cerebral palsy, unspecified	2887	2875
344	344 Other paralytic syndromes	2888	63
3440	3440 Quadriplegia	2889	2888
3441	3441 Paraplegia	2890	2888
3442	3442 Diplegia of upper limbs	2891	2888
3443	3443 Monoplegia of lower limb	2892	2888
3444	3444 Monoplegia of upper limb	2893	2888
3445	3445 Unspecified monoplegia	2894	2888
3446	3446 Cauda equina syndrome	2895	2888
3448	3448 Other specified paralytic syndromes	2896	2888
34480	34480 Other specified paralytic syndromes (todd's paralysis)	2897	2896
34489	34489 Other specified paralytic syndromes	2898	2896
3449	3449 Unspecified paralytic syndromes	2899	2888
345	345 Epilepsy	2900	63
3450	3450 Generalized nonconvulsive epilepsy	2901	2900
34500	34500 Generalized nonconvulsive epilepsy (petit mal and epileptic absences)	2902	2901
34501	34501 Generalized nonconvulsive epilepsy (atonic seizure)	2903	2901

ICD9 coding	meaning	Node id	Parent id
34502	34502 Generalized nonconvulsive epilepsy (akinetic seizure)	2904	2901
34509	34509 Generalized nonconvulsive epilepsy (other and unspecified)	2905	2901
3451	3451 Generalized convulsive epilepsy	2906	2900
34510	34510 Generalized convulsive epilepsy (grand mal and major epilepsy)	2907	2906
34511	34511 Generalized convulsive epilepsy (epileptic seizure, clonic)	2908	2906
34512	34512 Generalized convulsive epilepsy (epileptic seizure, myoclonic)	2909	2906
34513	34513 Generalized convulsive epilepsy (epileptic seizure, tonic)	2910	2906
34519	34519 Generalized convulsive epilepsy (other and unspecified)	2911	2906
3452	3452 Petit mal status	2912	2900
3453	3453 Grand mal status	2913	2900
3454	3454 Partial epilepsy, with impairment of consciousness	2914	2900
3455	3455 Partial epilepsy, without mention of impairment of consciousness	2915	2900
34550	34550 Partial epilepsy, without mention of impairment of consciousness (focal)	2916	2915
34551	34551 Unilateral epilepsy, without mention of impairment of consciousness	2917	2915
34552	34552 Partial epilepsy with autonomic symptoms without mention unconsciousness	2918	2915
34553	34553 Partial epilepsy with spec/somatosensory symptoms without unconsciousness	2919	2915
34559	34559 Other/unspec. partial epilepsy without mention impairment consciousness	2920	2915
3456	3456 Infantile spasms	2921	2900
3457	3457 Epilepsia partialis continua	2922	2900
3458	3458 Other specified epilepsy	2923	2900
3459	3459 Epilepsy, unspecified	2924	2900
348	348 Other conditions of brain	2933	63
3480	3480 Cerebral cysts	2934	2933
34800	34800 Cerebral cysts (arachnoid)	2935	2934
34809	34809 Cerebral cysts (other and unspecified)	2936	2934
3481	3481 Anoxic brain damage	2937	2933
3482	3482 Benign intracranial hypertension	2938	2933
3483	3483 Encephalopathy, unspecified	2939	2933
3484	3484 Compression of brain	2940	2933
3485	3485 Cerebral oedema	2941	2933
3488	3488 Other specified conditions of brain	2942	2933
3489	3489 Unspecified condition of brain	2943	2933

ICD9 coding	meaning	Node id	Parent id
349	349 Other and unspecified disorders of the nervous system	2944	63
3490	3490 Reaction to spinal or lumbar puncture	2945	2944
3491	3491 Nervous system complications from surgically implanted device	2946	2944
3492	3492 Disorders of meninges, not elsewhere classified	2947	2944
3498	3498 Other specified disorders of the nervous system	2948	2944
3499	3499 Unspecified disorders of the nervous system	2949	2944
430	430 Subarachnoid haemorrhage	3492	73
4309	4309 Subarachnoid haemorrhage	3493	3492
431	431 Intracerebral haemorrhage	3494	73
4319	4319 Intracerebral haemorrhage	3495	3494
432	432 Other and unspecified intracranial haemorrhage	3496	73
4320	4320 Nontraumatic extradural haemorrhage	3497	3496
4321	4321 Subdural haemorrhage	3498	3496
4329	4329 Unspecified intracranial haemorrhage	3499	3496
434	434 Occlusion of cerebral arteries	3507	73
4340	4340 Cerebral thrombosis	3508	3507
4341	4341 Cerebral embolism	3509	3507
4349	4349 Occlusion of cerebral arteries, unspecified	3510	3507
435	435 Transient cerebral ischaemia	3511	73
4359	4359 Transient cerebral ischaemia	3512	3511
436	436 Acute but ill-defined cerebrovascular disease	3513	73
4369	4369 Acute but ill-defined cerebrovascular disease	3514	3513
437	437 Other and ill-defined cerebrovascular disease	3515	73
4370	4370 Cerebral atherosclerosis	3516	3515
4371	4371 Other generalized ischaemic cerebrovascular disease	3517	3515
4372	4372 Hypertensive encephalopathy	3518	3515
4373	4373 Cerebral aneurysm, nonruptured	3519	3515
4374	4374 Cerebral arteritis	3520	3515
4375	4375 Moyamoya disease	3521	3515
4376	4376 Nonpyogenic thrombosis of intracranial venous sinus	3522	3515
4378	4378 Other ill-defined cerebrovascular disease	3523	3515
4379	4379 Ill-defined cerebrovascular disease, unspecified	3524	3515
438	438 Late effects of cerebrovascular disease	3525	73
4389	4389 Late effects of cerebrovascular disease	3526	3525
7670	7670 Birth trauma: subdural and cerebral haemorrhage	7297	7296
76700	76700 Birth trauma: subdural haemorrhage (trauma or intrapartum anoxia/hypoxia)	7298	7297

ICD9 coding	meaning	Node id	Parent id
76701	76701 Birth trauma: cerebral haemorrhage (trauma or intrapartum anoxia/hypoxia)	7299	7297
76702	76702 Birth trauma: subdural and cerebral haemorrhage (tentorial tear)	7300	7297
76703	76703 Birth injury to brain, not otherwise specified	7301	7297
76709	76709 Birth trauma: subdural and cerebral haemorrhage (other and unspecified)	7302	7297
7674	7674 Birth trauma: injury to spine and spinal cord	7318	7296
7675	7675 Birth trauma: facial nerve injury	7319	7296
7676	7676 Birth trauma: injury to brachial plexus	7320	7296
76760	76760 Birth trauma: injury to brachial plexus (erb's palsy)	7321	7320
76761	76761 Birth trauma: injury to brachial plexus (klumpke's palsy)	7322	7320
76762	76762 Birth trauma: injury to brachial plexus (phrenic nerve palsy)	7323	7320
76769	76769 Birth trauma: injury to brachial plexus (other and unspecified)	7324	7320
7677	7677 Birth trauma: other cranial and peripheral nerve injuries	7325	7296
76770	76770 Birth trauma: other cranial and peripheral nerve injuries (cranial nerve)	7326	7325
76771	76771 Birth trauma: other peripheral nerve injuries	7327	7325
76779	76779 Birth trauma: other cranial and peripheral nerve injuries	7328	7325
7721	7721 Fetal and neonatal intraventricular haemorrhage	7444	7435
7722	7722 Fetal and neonatal subarachnoid haemorrhage	7445	7435
740	740 Anencephalus and similar anomalies	6261	107
7400	7400 Anencephalus	6262	6261
74000	74000 Anencephalus (absence of brain)	6263	6262
74001	74001 Anencephalus (acrania)	6264	6262
74002	74002 Anencephalus (anencephaly)	6265	6262
74003	74003 Anencephalus (hemianencephaly)	6266	6262
74008	74008 Anencephalus (other)	6267	6262
7401	7401 Craniorachischisis	6268	6261
7402	7402 Iniencephaly	6269	6261
74020	74020 Iniencephaly (closed)	6270	6269
74021	74021 Iniencephaly (open)	6271	6269
74029	74029 Iniencephaly (unspecified)	6272	6269
741	741 Spina bifida	6273	107
7410	7410 Spina bifida with hydrocephalus	6274	6273
74100	74100 Spina bifida with hydrocephalus (aperta, any site)	6275	6274
74101	74101 Spina bifida cystica with arnold chiari malformation and hydrocephalus	6276	6274

ICD9 coding	meaning	Node id	Parent id
74102	74102 Spina bifida cystica with stenosed aqueduct of sylvius	6277	6274
74103	74103 Spina bifida with hydrocephalus (cystica, cervical)	6278	6274
74104	74104 Spina bifida with hydrocephalus (cystica, thoracic)	6279	6274
74105	74105 Spina bifida with hydrocephalus (cystica, lumbar)	6280	6274
74106	74106 Spina bifida with hydrocephalus (cystica, sacral)	6281	6274
74107	74107 Spina bifida with hydrocephalus (late onset, any site)	6282	6274
74108	74108 Spina bifida with hydrocephalus (other)	6283	6274
74109	74109 Spina bifida with hydrocephalus (unspecified)	6284	6274
7419	7419 Spina bifida without mention of hydrocephalus	6285	6273
74190	74190 Spina bifida without mention of hydrocephalus (aperta)	6286	6285
74191	74191 Spina bifida without mention of hydrocephalus (cystica, cervical)	6287	6285
74192	74192 Spina bifida without mention of hydrocephalus (cystica, thoracic)	6288	6285
74193	74193 Spina bifida without mention of hydrocephalus (cystica, lumbar)	6289	6285
74194	74194 Spina bifida without mention of hydrocephalus (cystica, sacral)	6290	6285
74198	74198 Spina bifida without mention of hydrocephalus (other)	6291	6285
74199	74199 Spina bifida without mention of hydrocephalus (unspecified)	6292	6285
742	742 Other congenital anomalies of nervous system	6293	107
7420	7420 Encephalocele	6294	6293
74200	74200 Encephalocele (occipital)	6295	6294
74208	74208 Encephalocele (other specified site)	6296	6294
74209	74209 Encephalocele (unspecified)	6297	6294
7421	7421 Microcephalus	6298	6293
7422	7422 Reduction deformities of brain	6299	6293
74220	74220 Reduction deformities of brain (anomalies of cerebrum)	6300	6299
74221	74221 Reduction deformities of brain (anomalies of corpus callosum)	6301	6299
74222	74222 Reduction deformities of brain (anomalies of hypothalamus)	6302	6299
74223	74223 Reduction deformities of brain (anomalies of cerebellum)	6303	6299
74224	74224 Reduction deformities of brain (agyria and lissencephaly)	6304	6299
74225	74225 Reduction deformities of brain (microgyria)	6305	6299

ICD9 coding	meaning	Node id	Parent id
74226	74226 Reduction deformities of brain (holoprosencephaly)	6306	6299
74228	74228 Reduction deformities of brain (other)	6307	6299
74229	74229 Reduction deformities of brain (unspecified)	6308	6299
7423	7423 Congenital hydrocephalus	6309	6293
74230	74230 Congenital hydrocephalus (anomalies of aqueduct of sylvius)	6310	6309
74231	74231 Congenital hydrocephalus (atresia of foramina of magendie and luschka)	6311	6309
74232	74232 Congenital hydrocephalus (hydraencephaly)	6312	6309
74238	74238 Congenital hydrocephalus (other)	6313	6309
74239	74239 Congenital hydrocephalus (unspecified)	6314	6309
7424	7424 Other specified anomalies of brain	6315	6293
74240	74240 Other specified anomalies of brain (enlarged brain)	6316	6315
74241	74241 Other specified anomalies of brain (porencephaly)	6317	6315
74242	74242 Other specified anomalies of brain (multiple cerebral cysts)	6318	6315
74248	74248 Other specified anomalies of brain	6319	6315
7425	7425 Other specified anomalies of spinal cord	6320	6293
74250	74250 Other specified anomalies of spinal cord (amyelia)	6321	6320
74251	74251 Other specified anomalies of spinal cord (hypoplasia and dysplasia)	6322	6320
74252	74252 Other specified anomalies of spinal cord (diastematomyelia)	6323	6320
74253	74253 Other specified anomalies of spinal cord (other cauda equina anomalies)	6324	6320
74254	74254 Other specified anomalies of spinal cord (hydromyelia)	6325	6320
74258	74258 Other specified anomalies of spinal cord	6326	6320
7428	7428 Other specified anomalies of nervous system	6327	6293
74280	74280 Other specified anomalies of nervous system (jaw-winking syndrome)	6328	6327
74281	74281 Other specified anomalies of nervous system (familial dysautonomia)	6329	6327
74288	74288 Other specified anomalies of nervous system	6330	6327
7429	7429 Unspecified anomalies of brain, spinal cord and nervous system	6331	6293
74290	74290 Unspecified anomalies of brain	6332	6331
74291	74291 Unspecified anomalies of spinal cord	6333	6331
74299	74299 Unspecified anomalies of nervous system, unspecified	6334	6331
851	851 Cerebral laceration and contusion	8072	129
8510	8510 Cerebral laceration/contusion without open intracranial wound	8073	8072

ICD9 coding	meaning	Node id	Parent id
8511	8511 Cerebral laceration and contusion with open intracranial wound	8074	8072
852	852 Subarachnoid, subdural and extradural haemorrhage, following injury	8075	129
8520	8520 Subarach./subdural/extradural haem. foll.inj.without open intracr.wound	8076	8075
8521	8521 Subarach./subdural/extradural haem.foll.inj.with open intracranial wound	8077	8075
853	853 Other and unspecified intracranial haemorrhage following injury	8078	129
8530	8530 Other/unspec. intracranial haem.foll.without open intracranial wound	8079	8078
8531	8531 Other/unspec. intracranial haem.foll.injury with open intracranial wound	8080	8078
854	854 Intracranial injury of other and unspecified nature	8081	129
8540	8540 Other/unspec. intracranial injury without open intracranial wound	8082	8081
8541	8541 Other/unspec. intracranial injury with open intracranial wound	8083	8081

Table S2. List of brain-related diagnostic codes from ICD10.

ICD10 coding	meaning	Node id	Parent id
A80	A80 Acute poliomyelitis	663	31
A800	A80.0 Acute paralytic poliomyelitis, vaccine-associated	664	663
A801	A80.1 Acute paralytic poliomyelitis, wild virus, imported	665	663
A802	A80.2 Acute paralytic poliomyelitis, wild virus, indigenous	666	663
A803	A80.3 Acute paralytic poliomyelitis, other and unspecified	667	663
A804	A80.4 Acute nonparalytic poliomyelitis	668	663
A809	A80.9 Acute poliomyelitis, unspecified	669	663
A81	A81 Atypical virus infections of central nervous system	670	31
A810	A81.0 Creutzfeldt-Jakob disease	671	670
A811	A81.1 Subacute sclerosing panencephalitis	672	670
A812	A81.2 Progressive multifocal leukoencephalopathy	673	670
A818	A81.8 Other atypical virus infections of central nervous system	674	670
A819	A81.9 Atypical virus infection of central nervous system, unspecified	675	670
A82	A82 Rabies	676	31
A820	A82.0 Sylvatic rabies	677	676
A821	A82.1 Urban rabies	678	676

ICD10 coding	meaning	Node id	Parent id
A829	A82.9 Rabies, unspecified	679	676
A83	A83 Mosquito-borne viral encephalitis	680	31
A830	A83.0 Japanese encephalitis	681	680
A831	A83.1 Western equine encephalitis	682	680
A832	A83.2 Eastern equine encephalitis	683	680
A833	A83.3 St. Louis encephalitis	684	680
A834	A83.4 Australian encephalitis	685	680
A835	A83.5 California encephalitis	686	680
A836	A83.6 Rocio virus disease	687	680
A838	A83.8 Other mosquito-borne viral encephalitis	688	680
A839	A83.9 Mosquito-borne viral encephalitis, unspecified	689	680
A84	A84 Tick-borne viral encephalitis	690	31
A840	A84.0 Far Eastern tick-borne encephalitis [Russian spring-summer encephalitis]	691	690
A841	A84.1 Central European tick-borne encephalitis	692	690
A848	A84.8 Other tick-borne viral encephalitis	693	690
A849	A84.9 Tick-borne viral encephalitis, unspecified	694	690
A85	A85 Other viral encephalitis, not elsewhere classified	695	31
A850	A85.0 Enteroviral encephalitis	696	695
A851	A85.1 Adenoviral encephalitis	697	695
A852	A85.2 Arthropod-borne viral encephalitis, unspecified	698	695
A858	A85.8 Other specified viral encephalitis	699	695
A86	A86 Unspecified viral encephalitis	700	31
A87	A87 Viral meningitis	701	31
A870	A87.0 Enteroviral meningitis	702	701
A871	A87.1 Adenoviral meningitis	703	701
A872	A87.2 Lymphocytic choriomeningitis	704	701
A878	A87.8 Other viral meningitis	705	701
A879	A87.9 Viral meningitis, unspecified	706	701
A88	A88 Other viral infections of central nervous system, not elsewhere classified	707	31
A880	A88.0 Enteroviral exanthematous fever [Boston exanthem]	708	707
A881	A88.1 Epidemic vertigo	709	707
A888	A88.8 Other specified viral infections of central nervous system	710	707
A89	A89 Unspecified viral infection of central nervous system	711	31
C70	C70 Malignant neoplasm of meninges	1563	54
C700	C70.0 Cerebral meninges	1564	1563
C701	C70.1 Spinal meninges	1565	1563
C709	C70.9 Meninges, unspecified	1566	1563

ICD10 coding	meaning	Node id	Parent id
C71	C71 Malignant neoplasm of brain	1567	54
C710	C71.0 Cerebrum, except lobes and ventricles	1568	1567
C711	C71.1 Frontal lobe	1569	1567
C712	C71.2 Temporal lobe	1570	1567
C713	C71.3 Parietal lobe	1571	1567
C714	C71.4 Occipital lobe	1572	1567
C715	C71.5 Cerebral ventricle	1573	1567
C716	C71.6 Cerebellum	1574	1567
C717	C71.7 Brain stem	1575	1567
C718	C71.8 Overlapping lesion of brain	1576	1567
C719	C71.9 Brain, unspecified	1577	1567
C72	C72 Malignant neoplasm of spinal cord, cranial nerves and other parts of central nervous system	1578	54
C720	C72.0 Spinal cord	1579	1578
C721	C72.1 Cauda equina	1580	1578
C722	C72.2 Olfactory nerve	1581	1578
C723	C72.3 Optic nerve	1582	1578
C724	C72.4 Acoustic nerve	1583	1578
C725	C72.5 Other and unspecified cranial nerves	1584	1578
C728	C72.8 Overlapping lesion of brain and other parts of central nervous system	1585	1578
C729	C72.9 Central nervous system, unspecified	1586	1578
D33	D33 Benign neoplasm of brain and other parts of central nervous system	1987	60
D330	D33.0 Brain, supratentorial	1988	1987
D331	D33.1 Brain, infratentorial	1989	1987
D332	D33.2 Brain, unspecified	1990	1987
D333	D33.3 Cranial nerves	1991	1987
D334	D33.4 Spinal cord	1992	1987
D337	D33.7 Other specified parts of central nervous system	1993	1987
D339	D33.9 Central nervous system, unspecified	1994	1987
F01	F01 Vascular dementia	2714	76
F010	F01.0 Vascular dementia of acute onset	2715	2714
F011	F01.1 Multi-infarct dementia	2716	2714
F012	F01.2 Subcortical vascular dementia	2717	2714
F013	F01.3 Mixed cortical and subcortical vascular dementia	2718	2714
F018	F01.8 Other vascular dementia	2719	2714
F019	F01.9 Vascular dementia, unspecified	2720	2714
G00	G00 Bacterial meningitis, not elsewhere classified	3176	87
G000	G00.0 Haemophilus meningitis	3177	3176

ICD10 coding	meaning	Node id	Parent id
G001	G00.1 Pneumococcal meningitis	3178	3176
G002	G00.2 Streptococcal meningitis	3179	3176
G003	G00.3 Staphylococcal meningitis	3180	3176
G008	G00.8 Other bacterial meningitis	3181	3176
G009	G00.9 Bacterial meningitis, unspecified	3182	3176
G01	G01 Meningitis in bacterial diseases classified elsewhere	3183	87
G02	G02 Meningitis in other infectious and parasitic diseases classified elsewhere	3184	87
G020	G02.0 Meningitis in viral diseases classified elsewhere	3185	3184
G021	G02.1 Meningitis in mycoses	3186	3184
G028	G02.8 Meningitis in other specified infectious and parasitic diseases classified elsewhere	3187	3184
G03	G03 Meningitis due to other and unspecified causes	3188	87
G030	G03.0 Nonpyogenic meningitis	3189	3188
G031	G03.1 Chronic meningitis	3190	3188
G032	G03.2 Benign recurrent meningitis [Mollaret]	3191	3188
G038	G03.8 Meningitis due to other specified causes	3192	3188
G039	G03.9 Meningitis, unspecified	3193	3188
G04	G04 Encephalitis, myelitis and encephalomyelitis	3194	87
G040	G04.0 Acute disseminated encephalitis	3195	3194
G041	G04.1 Tropical spastic paraplegia	3196	3194
G042	G04.2 Bacterial meningoencephalitis and meningomyelitis, not elsewhere classified	3197	3194
G048	G04.8 Other encephalitis, myelitis and encephalomyelitis	3198	3194
G049	G04.9 Encephalitis, myelitis and encephalomyelitis, unspecified	3199	3194
G05	G05 Encephalitis, myelitis and encephalomyelitis in diseases classified elsewhere	3200	87
G050	G05.0 Encephalitis, myelitis and encephalomyelitis in bacterial diseases classified elsewhere	3201	3200
G051	G05.1 Encephalitis, myelitis and encephalomyelitis in viral diseases classified elsewhere	3202	3200
G052	G05.2 Encephalitis, myelitis and encephalomyelitis in other infectious and parasitic diseases classified elsewhere	3203	3200
G058	G05.8 Encephalitis, myelitis and encephalomyelitis in other diseases classified elsewhere	3204	3200
G06	G06 Intracranial and intraspinal abscess and granuloma	3205	87
G060	G06.0 Intracranial abscess and granuloma	3206	3205
G061	G06.1 Intraspinal abscess and granuloma	3207	3205

ICD10 coding	meaning	Node id	Parent id
G062	G06.2 Extradural and subdural abscess, unspecified	3208	3205
G07	G07 Intracranial and intraspinal abscess and granuloma in diseases classified elsewhere	3209	87
G08	G08 Intracranial and intraspinal phlebitis and thrombophlebitis	3210	87
G09	G09 Sequelae of inflammatory diseases of central nervous system	3211	87
G10	G10 Huntington's disease	3212	88
G11	G11 Hereditary ataxia	3213	88
G110	G11.0 Congenital nonprogressive ataxia	3214	3213
G111	G11.1 Early-onset cerebellar ataxia	3215	3213
G112	G11.2 Late-onset cerebellar ataxia	3216	3213
G113	G11.3 Cerebellar ataxia with defective DNA repair	3217	3213
G114	G11.4 Hereditary spastic paraplegia	3218	3213
G118	G11.8 Other hereditary ataxias	3219	3213
G119	G11.9 Hereditary ataxia, unspecified	3220	3213
G13	G13 Systemic atrophies primarily affecting central nervous system in diseases classified elsewhere	3227	88
G130	G13.0 Paraneoplastic neuromyopathy and neuropathy	3228	3227
G131	G13.1 Other systemic atrophy primarily affecting central nervous system in neoplastic disease	3229	3227
G132	G13.2 Systemic atrophy primarily affecting central nervous system in myxoedema	3230	3227
G138	G13.8 Systemic atrophy primarily affecting central nervous system in other diseases classified elsewhere	3231	3227
G20	G20 Parkinson's disease	3233	89
G21	G21 Secondary Parkinsonism	3234	89
G210	G21.0 Malignant neuroleptic syndrome	3235	3234
G211	G21.1 Other drug-induced secondary Parkinsonism	3236	3234
G212	G21.2 Secondary Parkinsonism due to other external agents	3237	3234
G213	G21.3 Postencephalitic Parkinsonism	3238	3234
G214	G21.4 Vascular parkinsonism	3239	3234
G218	G21.8 Other secondary Parkinsonism	3240	3234
G219	G21.9 Secondary Parkinsonism, unspecified	3241	3234
G22	G22 Parkinsonism in diseases classified elsewhere	3242	89
G23	G23 Other degenerative diseases of basal ganglia	3243	89
G230	G23.0 Hallervorden-Spatz disease	3244	3243
G231	G23.1 Progressive supranuclear ophthalmoplegia [Steele-Richardson-Olszewski]	3245	3243

ICD10 coding	meaning	Node id	Parent id
G232	G23.2 Striatonigral degeneration	3246	3243
G233	G23.3 Multiple system atrophy, cerebellar type	3247	3243
G238	G23.8 Other specified degenerative diseases of basal ganglia	3248	3243
G239	G23.9 Degenerative disease of basal ganglia, unspecified	3249	3243
G24	G24 Dystonia	3250	89
G240	G24.0 Drug-induced dystonia	3251	3250
G241	G24.1 Idiopathic familial dystonia	3252	3250
G242	G24.2 Idiopathic nonfamilial dystonia	3253	3250
G243	G24.3 Spasmodic torticollis	3254	3250
G244	G24.4 Idiopathic orofacial dystonia	3255	3250
G245	G24.5 Blepharospasm	3256	3250
G248	G24.8 Other dystonia	3257	3250
G249	G24.9 Dystonia, unspecified	3258	3250
G25	G25 Other extrapyramidal and movement disorders	3259	89
G250	G25.0 Essential tremor	3260	3259
G251	G25.1 Drug-induced tremor	3261	3259
G252	G25.2 Other specified forms of tremor	3262	3259
G253	G25.3 Myoclonus	3263	3259
G254	G25.4 Drug-induced chorea	3264	3259
G255	G25.5 Other chorea	3265	3259
G256	G25.6 Drug-induced tics and other tics of organic origin	3266	3259
G258	G25.8 Other specified extrapyramidal and movement disorders	3267	3259
G259	G25.9 Extrapyramidal and movement disorder, unspecified	3268	3259
G26	G26 Extrapyramidal and movement disorders in diseases classified elsewhere	3269	89
G30	G30 Alzheimer's disease	3270	90
G300	G30.0 Alzheimer's disease with early onset	3271	3270
G301	G30.1 Alzheimer's disease with late onset	3272	3270
G308	G30.8 Other Alzheimer's disease	3273	3270
G309	G30.9 Alzheimer's disease, unspecified	3274	3270
G31	G31 Other degenerative diseases of nervous system, not elsewhere classified	3275	90
G310	G31.0 Circumscribed brain atrophy	3276	3275
G311	G31.1 Senile degeneration of brain, not elsewhere classified	3277	3275
G312	G31.2 Degeneration of nervous system due to alcohol	3278	3275
G318	G31.8 Other specified degenerative diseases of nervous system	3279	3275

ICD10 coding	meaning	Node id	Parent id
G319	G31.9 Degenerative disease of nervous system, unspecified	3280	3275
G32	G32 Other degenerative disorders of nervous system in diseases classified elsewhere	3281	90
G320	G32.0 Subacute combined degeneration of spinal cord in diseases classified elsewhere	3282	3281
G328	G32.8 Other specified degenerative disorders of nervous system in diseases classified elsewhere	3283	3281
G35	G35 Multiple sclerosis	3284	91
G36	G36 Other acute disseminated demyelination	3285	91
G360	G36.0 Neuromyelitis optica [Devic]	3286	3285
G361	G36.1 Acute and subacute haemorrhagic leukoencephalitis [Hurst]	3287	3285
G368	G36.8 Other specified acute disseminated demyelination	3288	3285
G369	G36.9 Acute disseminated demyelination, unspecified	3289	3285
G37	G37 Other demyelinating diseases of central nervous system	3290	91
G370	G37.0 Diffuse sclerosis	3291	3290
G371	G37.1 Central demyelination of corpus callosum	3292	3290
G372	G37.2 Central pontine myelinolysis	3293	3290
G373	G37.3 Acute transverse myelitis in demyelinating disease of central nervous system	3294	3290
G374	G37.4 Subacute necrotising myelitis	3295	3290
G375	G37.5 Concentric sclerosis [Balo]	3296	3290
G378	G37.8 Other specified demyelinating diseases of central nervous system	3297	3290
G379	G37.9 Demyelinating disease of central nervous system, unspecified	3298	3290
G40	G40 Epilepsy	3299	92
G400	G40.0 Localisation-related (focal) (partial) idiopathic epilepsy and epileptic syndromes with seizures of localised onset	3300	3299
G401	G40.1 Localisation-related (focal) (partial) symptomatic epilepsy and epileptic syndromes with simple partial seizures	3301	3299
G402	G40.2 Localisation-related (focal) (partial) symptomatic epilepsy and epileptic syndromes with complex partial seizures	3302	3299
G403	G40.3 Generalised idiopathic epilepsy and epileptic syndromes	3303	3299
G404	G40.4 Other generalised epilepsy and epileptic syndromes	3304	3299
G405	G40.5 Special epileptic syndromes	3305	3299
G406	G40.6 Grand mal seizures, unspecified (with or without petit mal)	3306	3299
G407	G40.7 Petit mal, unspecified, without grand mal seizures	3307	3299

ICD10 coding	meaning	Node id	Parent id
G408	G40.8 Other epilepsy	3308	3299
G409	G40.9 Epilepsy, unspecified	3309	3299
G41	G41 Status epilepticus	3310	92
G410	G41.0 Grand mal status epilepticus	3311	3310
G411	G41.1 Petit mal status epilepticus	3312	3310
G412	G41.2 Complex partial status epilepticus	3313	3310
G418	G41.8 Other status epilepticus	3314	3310
G419	G41.9 Status epilepticus, unspecified	3315	3310
G45	G45 Transient cerebral ischaemic attacks and related syndromes	3330	92
G450	G45.0 Vertebro-basilar artery syndrome	3331	3330
G451	G45.1 Carotid artery syndrome (hemispheric)	3332	3330
G452	G45.2 Multiple and bilateral precerebral artery syndromes	3333	3330
G453	G45.3 Amaurosis fugax	3334	3330
G454	G45.4 Transient global amnesia	3335	3330
G458	G45.8 Other transient cerebral ischaemic attacks and related syndromes	3336	3330
G459	G45.9 Transient cerebral ischaemic attack, unspecified	3337	3330
G46	G46 Vascular syndromes of brain in cerebrovascular diseases	3338	92
G460	G46.0 Middle cerebral artery syndrome	3339	3338
G461	G46.1 Anterior cerebral artery syndrome	3340	3338
G462	G46.2 Posterior cerebral artery syndrome	3341	3338
G463	G46.3 Brain stem stroke syndrome	3342	3338
G464	G46.4 Cerebellar stroke syndrome	3343	3338
G465	G46.5 Pure motor lacunar syndrome	3344	3338
G466	G46.6 Pure sensory lacunar syndrome	3345	3338
G467	G46.7 Other lacunar syndromes	3346	3338
G468	G46.8 Other vascular syndromes of brain in cerebrovascular diseases	3347	3338
G80	G80 Infantile cerebral palsy	3484	96
G800	G80.0 Spastic cerebral palsy	3485	3484
G801	G80.1 Spastic diplegia	3486	3484
G802	G80.2 Infantile hemiplegia	3487	3484
G803	G80.3 Dyskinetic cerebral palsy	3488	3484
G804	G80.4 Ataxic cerebral palsy	3489	3484
G808	G80.8 Other infantile cerebral palsy	3490	3484
G809	G80.9 Infantile cerebral palsy, unspecified	3491	3484
G81	G81 Hemiplegia	3492	96
G810	G81.0 Flaccid hemiplegia	3493	3492
G811	G81.1 Spastic hemiplegia	3494	3492
G819	G81.9 Hemiplegia, unspecified	3495	3492

ICD10 coding	meaning	Node id	Parent id
G82	G82 Paraplegia and tetraplegia	3496	96
G820	G82.0 Flaccid paraplegia	3497	3496
G821	G82.1 Spastic paraplegia	3498	3496
G822	G82.2 Paraplegia, unspecified	3499	3496
G823	G82.3 Flaccid tetraplegia	3500	3496
G824	G82.4 Spastic tetraplegia	3501	3496
G825	G82.5 Tetraplegia, unspecified	3502	3496
G83	G83 Other paralytic syndromes	3503	96
G830	G83.0 Diplegia of upper limbs	3504	3503
G831	G83.1 Monoplegia of lower limb	3505	3503
G832	G83.2 Monoplegia of upper limb	3506	3503
G833	G83.3 Monoplegia, unspecified	3507	3503
G834	G83.4 Cauda equina syndrome	3508	3503
G835	G83.5 Locked-in syndrome	3509	3503
G838	G83.8 Other specified paralytic syndromes	3510	3503
G839	G83.9 Paralytic syndrome, unspecified	3511	3503
G91	G91 Hydrocephalus	3520	97
G910	G91.0 Communicating hydrocephalus	3521	3520
G911	G91.1 Obstructive hydrocephalus	3522	3520
G912	G91.2 Normal-pressure hydrocephalus	3523	3520
G913	G91.3 Posttraumatic hydrocephalus, unspecified	3524	3520
G918	G91.8 Other hydrocephalus	3525	3520
G919	G91.9 Hydrocephalus, unspecified	3526	3520
G92	G92 Toxic encephalopathy	3527	97
G93	G93 Other disorders of brain	3528	97
G930	G93.0 Cerebral cysts	3529	3528
G931	G93.1 Anoxic brain damage, not elsewhere classified	3530	3528
G932	G93.2 Benign intracranial hypertension	3531	3528
G933	G93.3 Postviral fatigue syndrome	3532	3528
G934	G93.4 Encephalopathy, unspecified	3533	3528
G935	G93.5 Compression of brain	3534	3528
G936	G93.6 Cerebral oedema	3535	3528
G937	G93.7 Reye's syndrome	3536	3528
G938	G93.8 Other specified disorders of brain	3537	3528
G939	G93.9 Disorder of brain, unspecified	3538	3528
G94	G94 Other disorders of brain in diseases classified elsewhere	3539	97
G940	G94.0 Hydrocephalus in infectious and parasitic diseases classified elsewhere	3540	3539
G941	G94.1 Hydrocephalus in neoplastic disease	3541	3539
G942	G94.2 Hydrocephalus in other diseases classified elsewhere	3542	3539

ICD10 coding	meaning	Node id	Parent id
G948	G94.8 Other specified disorders of brain in diseases classified elsewhere	3543	3539
I60	I60 Subarachnoid haemorrhage	4266	119
I600	I60.0 Subarachnoid haemorrhage from carotid siphon and bifurcation	4267	4266
I601	I60.1 Subarachnoid haemorrhage from middle cerebral artery	4268	4266
I602	I60.2 Subarachnoid haemorrhage from anterior communicating artery	4269	4266
I603	I60.3 Subarachnoid haemorrhage from posterior communicating artery	4270	4266
I604	I60.4 Subarachnoid haemorrhage from basilar artery	4271	4266
I605	I60.5 Subarachnoid haemorrhage from vertebral artery	4272	4266
I606	I60.6 Subarachnoid haemorrhage from other intracranial arteries	4273	4266
I607	I60.7 Subarachnoid haemorrhage from intracranial artery, unspecified	4274	4266
I608	I60.8 Other subarachnoid haemorrhage	4275	4266
I609	I60.9 Subarachnoid haemorrhage, unspecified	4276	4266
I61	I61 Intracerebral haemorrhage	4277	119
I610	I61.0 Intracerebral haemorrhage in hemisphere, subcortical	4278	4277
I611	I61.1 Intracerebral haemorrhage in hemisphere, cortical	4279	4277
I612	I61.2 Intracerebral haemorrhage in hemisphere, unspecified	4280	4277
I613	I61.3 Intracerebral haemorrhage in brain stem	4281	4277
I614	I61.4 Intracerebral haemorrhage in cerebellum	4282	4277
I615	I61.5 Intracerebral haemorrhage, intraventricular	4283	4277
I616	I61.6 Intracerebral haemorrhage, multiple localised	4284	4277
I618	I61.8 Other intracerebral haemorrhage	4285	4277
I619	I61.9 Intracerebral haemorrhage, unspecified	4286	4277
I62	I62 Other nontraumatic intracranial haemorrhage	4287	119
I620	I62.0 Subdural haemorrhage (acute) (nontraumatic)	4288	4287
I621	I62.1 Nontraumatic extradural haemorrhage	4289	4287
I629	I62.9 Intracranial haemorrhage (nontraumatic), unspecified	4290	4287
I63	I63 Cerebral infarction	4291	119
I630	I63.0 Cerebral infarction due to thrombosis of precerebral arteries	4292	4291
I631	I63.1 Cerebral infarction due to embolism of precerebral arteries	4293	4291
I632	I63.2 Cerebral infarction due to unspecified occlusion or stenosis of precerebral arteries	4294	4291

ICD10 coding	meaning	Node id	Parent id
I633	I63.3 Cerebral infarction due to thrombosis of cerebral arteries	4295	4291
I634	I63.4 Cerebral infarction due to embolism of cerebral arteries	4296	4291
I635	I63.5 Cerebral infarction due to unspecified occlusion or stenosis of cerebral arteries	4297	4291
I636	I63.6 Cerebral infarction due to cerebral venous thrombosis, nonpyogenic	4298	4291
I638	I63.8 Other cerebral infarction	4299	4291
I639	I63.9 Cerebral infarction, unspecified	4300	4291
I64	I64 Stroke, not specified as haemorrhage or infarction	4301	119
I66	I66 Occlusion and stenosis of cerebral arteries, not resulting in cerebral infarction	4309	119
I660	I66.0 Occlusion and stenosis of middle cerebral artery	4310	4309
I661	I66.1 Occlusion and stenosis of anterior cerebral artery	4311	4309
I662	I66.2 Occlusion and stenosis of posterior cerebral artery	4312	4309
I663	I66.3 Occlusion and stenosis of cerebellar arteries	4313	4309
I664	I66.4 Occlusion and stenosis of multiple and bilateral cerebral arteries	4314	4309
I668	I66.8 Occlusion and stenosis of other cerebral artery	4315	4309
I669	I66.9 Occlusion and stenosis of unspecified cerebral artery	4316	4309
I67	I67 Other cerebrovascular diseases	4317	119
I670	I67.0 Dissection of cerebral arteries, nonruptured	4318	4317
I671	I67.1 Cerebral aneurysm, nonruptured	4319	4317
I672	I67.2 Cerebral atherosclerosis	4320	4317
I673	I67.3 Progressive vascular leukoencephalopathy	4321	4317
I674	I67.4 Hypertensive encephalopathy	4322	4317
I675	I67.5 Moyamoya disease	4323	4317
I676	I67.6 Nonpyogenic thrombosis of intracranial venous system	4324	4317
I677	I67.7 Cerebral arteritis, not elsewhere classified	4325	4317
I678	I67.8 Other specified cerebrovascular diseases	4326	4317
I679	I67.9 Cerebrovascular disease, unspecified	4327	4317
I68	I68 Cerebrovascular disorders in diseases classified elsewhere	4328	119
I680	I68.0 Cerebral amyloid angiopathy	4329	4328
I681	I68.1 Cerebral arteritis in infectious and parasitic diseases classified elsewhere	4330	4328
I682	I68.2 Cerebral arteritis in other diseases classified elsewhere	4331	4328

ICD10 coding	meaning	Node id	Parent id
I688	I68.8 Other cerebrovascular disorders in diseases classified elsewhere	4332	4328
I69	I69 Sequelae of cerebrovascular disease	4333	119
I690	I69.0 Sequelae of subarachnoid haemorrhage	4334	4333
I691	I69.1 Sequelae of intracerebral haemorrhage	4335	4333
I692	I69.2 Sequelae of other nontraumatic intracranial haemorrhage	4336	4333
I693	I69.3 Sequelae of cerebral infarction	4337	4333
I694	I69.4 Sequelae of stroke, not specified as haemorrhage or infarction	4338	4333
I698	I69.8 Sequelae of other and unspecified cerebrovascular diseases	4339	4333
P10	P10 Intracranial laceration and haemorrhage due to birth injury	11239	186
P100	P10.0 Subdural haemorrhage due to birth injury	11240	11239
P101	P10.1 Cerebral haemorrhage due to birth injury	11241	11239
P102	P10.2 Intraventricular haemorrhage due to birth injury	11242	11239
P103	P10.3 Subarachnoid haemorrhage due to birth injury	11243	11239
P104	P10.4 Tentorial tear due to birth injury	11244	11239
P108	P10.8 Other intracranial lacerations and haemorrhages due to birth injury	11245	11239
P109	P10.9 Unspecified intracranial laceration and haemorrhage due to birth injury	11246	11239
P11	P11 Other birth injuries to central nervous system	11247	186
P110	P11.0 Cerebral oedema due to birth injury	11248	11247
P111	P11.1 Other specified brain damage due to birth injury	11249	11247
P112	P11.2 Unspecified brain damage due to birth injury	11250	11247
P113	P11.3 Birth injury to facial nerve	11251	11247
P114	P11.4 Birth injury to other cranial nerves	11252	11247
P115	P11.5 Birth injury to spine and spinal cord	11253	11247
P119	P11.9 Birth injury to central nervous system, unspecified	11254	11247
Q00	Q00 Anencephaly and similar malformations	11560	194
Q000	Q00.0 Anencephaly	11561	11560
Q001	Q00.1 Craniorachischisis	11562	11560
Q002	Q00.2 Iniencephaly	11563	11560
Q01	Q01 Encephalocele	11564	194
Q010	Q01.0 Frontal encephalocele	11565	11564
Q011	Q01.1 Nasofrontal encephalocele	11566	11564
Q012	Q01.2 Occipital encephalocele	11567	11564
Q018	Q01.8 Encephalocele of other sites	11568	11564
Q019	Q01.9 Encephalocele, unspecified	11569	11564

ICD10 coding	meaning	Node id	Parent id
Q02	Q02 Microcephaly	11570	194
Q03	Q03 Congenital hydrocephalus	11571	194
Q030	Q03.0 Malformations of aqueduct of Sylvius	11572	11571
Q031	Q03.1 Atresia of foramina of Magendie and Luschka	11573	11571
Q038	Q03.8 Other congenital hydrocephalus	11574	11571
Q039	Q03.9 Congenital hydrocephalus, unspecified	11575	11571
Q04	Q04 Other congenital malformations of brain	11576	194
Q040	Q04.0 Congenital malformations of corpus callosum	11577	11576
Q041	Q04.1 Arhinencephaly	11578	11576
Q042	Q04.2 Holoprosencephaly	11579	11576
Q043	Q04.3 Other reduction deformities of brain	11580	11576
Q044	Q04.4 Septo-optic dysplasia	11581	11576
Q045	Q04.5 Megalencephaly	11582	11576
Q046	Q04.6 Congenital cerebral cysts	11583	11576
Q048	Q04.8 Other specified congenital malformations of brain	11584	11576
Q049	Q04.9 Congenital malformation of brain, unspecified	11585	11576
Q05	Q05 Spina bifida	11586	194
Q050	Q05.0 Cervical spina bifida with hydrocephalus	11587	11586
Q051	Q05.1 Thoracic spina bifida with hydrocephalus	11588	11586
Q052	Q05.2 Lumbar spina bifida with hydrocephalus	11589	11586
Q053	Q05.3 Sacral spina bifida with hydrocephalus	11590	11586
Q054	Q05.4 Unspecified spina bifida with hydrocephalus	11591	11586
Q055	Q05.5 Cervical spina bifida without hydrocephalus	11592	11586
Q056	Q05.6 Thoracic spina bifida without hydrocephalus	11593	11586
Q057	Q05.7 Lumbar spina bifida without hydrocephalus	11594	11586
Q058	Q05.8 Sacral spina bifida without hydrocephalus	11595	11586
Q059	Q05.9 Spina bifida, unspecified	11596	11586
Q06	Q06 Other congenital malformations of spinal cord	11597	194
Q060	Q06.0 Amyelia	11598	11597
Q061	Q06.1 Hypoplasia and dysplasia of spinal cord	11599	11597
Q062	Q06.2 Diastematomyelia	11600	11597
Q063	Q06.3 Other congenital cauda equina malformations	11601	11597
Q064	Q06.4 Hydromyelia	11602	11597
Q068	Q06.8 Other specified congenital malformations of spinal cord	11603	11597

ICD10 coding	meaning	Node id	Parent id
Q069	Q06.9 Congenital malformation of spinal cord, unspecified	11604	11597
Q07	Q07 Other congenital malformations of nervous system	11605	194
Q070	Q07.0 Arnold-Chiari syndrome	11606	11605
Q078	Q07.8 Other specified congenital malformations of nervous system	11607	11605
Q079	Q07.9 Congenital malformation of nervous system, unspecified	11608	11605
S061	S06.1 Traumatic cerebral oedema	12755	12751
S0610	S06.10 Traumatic cerebral oedema (without open intracranial wound)	12756	12755
S0611	S06.11 Traumatic cerebral oedema (with open intracranial wound)	12757	12755
S062	S06.2 Diffuse brain injury	12758	12751
S0620	S06.20 Diffuse brain injury (without open intracranial wound)	12759	12758
S0621	S06.21 Diffuse brain injury (with open intracranial wound)	12760	12758
S063	S06.3 Focal brain injury	12761	12751
S0630	S06.30 Focal brain injury (without open intracranial wound)	12762	12761
S0631	S06.31 Focal brain injury (with open intracranial wound)	12763	12761
S064	S06.4 Epidural haemorrhage	12764	12751
S0640	S06.40 Epidural haemorrhage (without open intracranial wound)	12765	12764
S0641	S06.41 Epidural haemorrhage (with open intracranial wound)	12766	12764
S065	S06.5 Traumatic subdural haemorrhage	12767	12751
S0650	S06.50 Traumatic subdural haemorrhage (without open intracranial wound)	12768	12767
S0651	S06.51 Traumatic subdural haemorrhage (with open intracranial wound)	12769	12767
S066	S06.6 Traumatic subarachnoid haemorrhage	12770	12751
S0660	S06.60 Traumatic subarachnoid haemorrhage (without open intracranial wound)	12771	12770
S0661	S06.61 Traumatic subarachnoid haemorrhage (with open intracranial wound)	12772	12770
S067	S06.7 Intracranial injury with prolonged coma	12773	12751
S0670	S06.70 Intracranial injury with prolonged coma (without open intracranial wound)	12774	12773
S0671	S06.71 Intracranial injury with prolonged coma (with open intracranial wound)	12775	12773
S068	S06.8 Other intracranial injuries	12776	12751
S0680	S06.80 Other intracranial injuries (without open intracranial wound)	12777	12776
S0681	S06.81 Other intracranial injuries (with open intracranial wound)	12778	12776

ICD10 coding	meaning	Node id	Parent id
S069	S06.9 Intracranial injury, unspecified	12779	12751
S0690	S06.90 Intracranial injury, unspecified (without open intracranial wound)	12780	12779
S0691	S06.91 Intracranial injury, unspecified (with open intracranial wound)	12781	12779

Table S3. PheWAS results for HICAS_pop. Only the bonferroni significant results are shown.

varName	varType	n	beta	lower	upper	pval	resType	description	Cat1 Title	Cat2 Title	Cat3 Title	Category
1707-1	CAT-SIN	27261/30651	-999	-999	-999	8E-120	MULTINOMIAL-LOGISTIC	Handedness (chirality/laterality)	UK Biobank Assessment Centre	Touchscreen	Early life factors	100033
25744	CONTINUOUS	30655	0.12	0.11	0.13	5.2E-118	LINEAR	Inverted temporal signal-to-noise ratio in artefact-cleaned pre-processed rfMRI	UK Biobank Assessment Centre	Imaging	Brain MRI	111
25042	CONTINUOUS	26737	0.08	0.07	0.09	1.69E-39	LINEAR	Median z-statistic (in group-defined mask) for shapes activation	UK Biobank Assessment Centre	Imaging	Brain MRI	106
25040	CONTINUOUS	26737	0.08	0.07	0.09	2.81E-37	LINEAR	Median BOLD effect (in group-defined mask) for shapes activation	UK Biobank Assessment Centre	Imaging	Brain MRI	106
25762	CONTINUOUS	26737	0.08	0.06	0.09	1.15E-36	LINEAR	90th percentile of z-statistic (in group-defined mask) for shapes activation	UK Biobank Assessment Centre	Imaging	Brain MRI	106
25046	CONTINUOUS	26737	0.08	0.06	0.09	2.4E-36	LINEAR	Median z-statistic (in group-defined mask) for faces activation	UK Biobank Assessment Centre	Imaging	Brain MRI	106

12143	INTEGER	30655	0.06	0.05	0.07	1.17E-34	LINEAR	Weight (pre-imaging)	UK Biobank Assessment Centre	Physical measures	Anthropometry	100010
25044	CONTINUOUS	26737	0.07	0.06	0.08	9.71E-33	LINEAR	Median BOLD effect (in group-defined mask) for faces activation	UK Biobank Assessment Centre	Imaging	Brain MRI	106
25761	CONTINUOUS	26737	0.07	0.06	0.08	1.26E-32	LINEAR	90th percentile of BOLD effect (in group-defined mask) for shapes activation	UK Biobank Assessment Centre	Imaging	Brain MRI	106
25764	CONTINUOUS	26737	0.07	0.06	0.08	1.68E-31	LINEAR	90th percentile of z-statistic (in group-defined mask) for faces activation	UK Biobank Assessment Centre	Imaging	Brain MRI	106
25746	INTEGER	30236	-0.07	-0.08	-0.06	4.86E-31	LINEAR	Number of dMRI outlier slices detected and corrected	UK Biobank Assessment Centre	Imaging	Brain MRI	107
21001	CONTINUOUS	30626	0.06	0.05	0.07	4.39E-29	LINEAR	Body mass index (BMI)	UK Biobank Assessment Centre	Physical measures	Anthropometry	100010

Table S3. PheWAS results for HICAS_pop. Only the bonferroni significant results are shown [Continued]

varName	varType	n	beta	lower	upper	pval	resType	description	Cat1 Title	Cat2 Title	Cat3 Title	Category
21002	CONTINUOUS	30628	0.05	0.04	0.06	1.74E-27	LINEAR	Weight	UK Biobank Assessment Centre	Physical measures	Anthropometry	100010
25763	CONTINUOUS	26737	0.07	0.05	0.08	1.02E-26	LINEAR	90th percentile of BOLD effect (in group-defined mask) for faces activation	UK Biobank Assessment Centre	Imaging	Brain MRI	106
25089	CONTINUOUS	30236	0.06	0.05	0.07	4.3E-26	LINEAR	Mean FA in external capsule on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
48	CONTINUOUS	30642	0.05	0.04	0.06	7.98E-25	LINEAR	Waist circumference	UK Biobank Assessment Centre	Physical measures	Anthropometry	100010
25091	CONTINUOUS	30236	0.05	0.04	0.07	2.55E-23	LINEAR	Mean FA in cingulum cingulate gyrus on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
12188	CAT-SIN	4019/26529(30548)	0.17	0.13	0.20	7.92E-23	LOGISTIC-BINARY	Operator indicated brain MRI measurement completed	UK Biobank Assessment Centre	Imaging	Brain MRI	100

25097	CONTINUOUS	30236	0.05	0.04	0.07	3.02E-22	LINEAR	Mean FA in superior longitudinal fasciculus on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25425	CONTINUOUS	30236	-0.06	-0.07	-0.04	3.51E-22	LINEAR	Mean OD in external capsule on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25096	CONTINUOUS	30236	0.05	0.04	0.06	1.39E-21	LINEAR	Mean FA in superior longitudinal fasciculus on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
49	CONTINUOUS	30642	0.05	0.04	0.06	1.54E-20	LINEAR	Hip circumference	UK Biobank Assessment Centre	Physical measures	Anthropometry	100010
25427	CONTINUOUS	30236	-0.05	-0.06	-0.04	7.7E-20	LINEAR	Mean OD in cingulum cingulate gyrus on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25331	CONTINUOUS	30236	-0.05	-0.06	-0.04	1.95E-19	LINEAR	Mean L3 in cingulum cingulate gyrus on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134

Table S3. PheWAS results for HICAS_pop. Only the bonferroni significant results are shown [Continued]

varName	varType	n	beta	lower	upper	pval	resType	description	Cat1 Title	Cat2 Title	Cat3 Title	Category
25733	CONTINUOUS	30655	-0.05	-0.06	-0.04	2.68E-19	LINEAR	Amount of warping applied to non-linearly align T1 brain image to standard-space	UK Biobank Assessment Centre	Imaging	Brain MRI	110
25056	CONTINUOUS	30236	0.05	0.04	0.06	7.8E-19	LINEAR	Mean FA in middle cerebellar peduncle on FA skeleton	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25088	CONTINUOUS	30236	0.05	0.04	0.06	1.31E-18	LINEAR	Mean FA in external capsule on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25503	CONTINUOUS	30236	0.05	0.04	0.06	8.49E-18	LINEAR	Weighted-mean FA in tract inferior longitudinal fasciculus (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25296	CONTINUOUS	30236	-0.05	-0.06	-0.04	9.6E-18	LINEAR	Mean L3 in middle cerebellar peduncle on FA skeleton	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25433	CONTINUOUS	30236	-0.05	-0.06	-0.04	1.11E-17	LINEAR	Mean OD in superior longitudinal fasciculus on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134

25509	CONTINUOUS	30236	0.05	0.04	0.06	1.49E-17	LINEAR	Weighted-mean FA in tract superior longitudinal fasciculus (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25090	CONTINUOUS	30236	0.05	0.04	0.06	1.98E-17	LINEAR	Mean FA in cingulum cingulate gyrus on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25508	CONTINUOUS	30236	0.05	0.04	0.06	5.93E-17	LINEAR	Weighted-mean FA in tract posterior thalamic radiation (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25283	CONTINUOUS	30236	-0.05	-0.06	-0.03	1.08E-16	LINEAR	Mean L2 in cingulum cingulate gyrus on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25072	CONTINUOUS	30236	0.05	0.04	0.06	1.79E-16	LINEAR	Mean FA in anterior limb of internal capsule on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25638	CONTINUOUS	30236	-0.04	-0.05	-0.03	9.64E-16	LINEAR	Weighted-mean L3 in tract inferior longitudinal fasciculus (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135

Table S3. PheWAS results for HICAS_pop. Only the bonferroni significant results are shown [Continued]

varName	varType	n	beta	lower	upper	pval	resType	description	Cat1 Title	Cat2 Title	Cat3 Title	Category
25332	CONTINUOUS	30236	-0.05	-0.06	-0.03	1.11E-15	LINEAR	Mean L3 in cingulum hippocampus on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25300	CONTINUOUS	30236	-0.04	-0.05	-0.03	1.38E-15	LINEAR	Mean L3 in splenium of corpus callosum on FA skeleton	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25510	CONTINUOUS	30236	0.04	0.03	0.06	1.55E-15	LINEAR	Weighted-mean FA in tract superior longitudinal fasciculus (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25500	CONTINUOUS	30236	0.04	0.03	0.05	2.02E-15	LINEAR	Weighted-mean FA in tract inferior fronto-occipital fasciculus (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25647	CONTINUOUS	30236	-0.04	-0.05	-0.03	2.58E-15	LINEAR	Weighted-mean L3 in tract superior thalamic radiation (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
47	INTEGER	30566	0.03	0.02	0.04	3.98E-15	LINEAR	Hand grip strength (right)	UK Biobank Assessment Centre	Physical measures	Hand grip strength	100019

25490	CONTINUOUS	30236	0.04	0.03	0.05	4.49E-15	LINEAR	Weighted-mean FA in tract anterior thalamic radiation (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25501	CONTINUOUS	30236	0.04	0.03	0.05	4.84E-15	LINEAR	Weighted-mean FA in tract inferior fronto-occipital fasciculus (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25060	CONTINUOUS	30236	0.04	0.03	0.06	7.11E-15	LINEAR	Mean FA in splenium of corpus callosum on FA skeleton	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25674	CONTINUOUS	30236	0.04	0.03	0.05	1.33E-14	LINEAR	Weighted-mean ICVF in tract superior thalamic radiation (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25644	CONTINUOUS	30236	-0.04	-0.05	-0.03	1.95E-14	LINEAR	Weighted-mean L3 in tract superior longitudinal fasciculus (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25616	CONTINUOUS	30236	-0.04	-0.05	-0.03	1.99E-14	LINEAR	Weighted-mean L2 in tract posterior thalamic radiation (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135

Table S3. PheWAS results for HICAS_pop. Only the bonferroni significant results are shown [Continued]

varName	varType	n	beta	lower	upper	pval	resType	description	Cat1 Title	Cat2 Title	Cat3 Title	Category
25332	CONTINUOUS	30236	-0.05	-0.06	-0.03	1.11E-15	LINEAR	Mean L3 in cingulum hippocampus on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25300	CONTINUOUS	30236	-0.04	-0.05	-0.03	1.38E-15	LINEAR	Mean L3 in splenium of corpus callosum on FA skeleton	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25510	CONTINUOUS	30236	0.04	0.03	0.06	1.55E-15	LINEAR	Weighted-mean FA in tract superior longitudinal fasciculus (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25500	CONTINUOUS	30236	0.04	0.03	0.05	2.02E-15	LINEAR	Weighted-mean FA in tract inferior fronto-occipital fasciculus (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25647	CONTINUOUS	30236	-0.04	-0.05	-0.03	2.58E-15	LINEAR	Weighted-mean L3 in tract superior thalamic radiation (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
47	INTEGER	30566	0.03	0.02	0.04	3.98E-15	LINEAR	Hand grip strength (right)	UK Biobank Assessment Centre	Physical measures	Hand grip strength	100019

25490	CONTINUOUS	30236	0.04	0.03	0.05	4.49E-15	LINEAR	Weighted-mean FA in tract anterior thalamic radiation (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25501	CONTINUOUS	30236	0.04	0.03	0.05	4.84E-15	LINEAR	Weighted-mean FA in tract inferior fronto-occipital fasciculus (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25060	CONTINUOUS	30236	0.04	0.03	0.06	7.11E-15	LINEAR	Mean FA in splenium of corpus callosum on FA skeleton	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25674	CONTINUOUS	30236	0.04	0.03	0.05	1.33E-14	LINEAR	Weighted-mean ICVF in tract superior thalamic radiation (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25644	CONTINUOUS	30236	-0.04	-0.05	-0.03	1.95E-14	LINEAR	Weighted-mean L3 in tract superior longitudinal fasciculus (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25616	CONTINUOUS	30236	-0.04	-0.05	-0.03	1.99E-14	LINEAR	Weighted-mean L2 in tract posterior thalamic radiation (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135

Table S3. PheWAS results for HICAS_pop. Only the bonferroni significant results are shown [Continued]

varName	varType	n	beta	lower	upper	pval	resType	description	Cat1 Title	Cat2 Title	Cat3 Title	Category
25019	CONTINUOUS	30655	0.04	0.03	0.05	2.26E-14	LINEAR	Volume of hippocampus (left)	UK Biobank Assessment Centre	Imaging	Imaging	1102
25759	CONTINUOUS	6992/23663 (30655)	-0.10	-0.13	-0.08	3.01E-14	LOGISTIC-BINARY	Scanner table position	UK Biobank Assessment Centre	Imaging	Brain MRI	110
25289	CONTINUOUS	30236	-0.04	-0.05	-0.03	3.16E-14	LINEAR	Mean L2 in superior longitudinal fasciculus on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25731	CONTINUOUS	30655	-0.04	-0.05	-0.03	3.27E-14	LINEAR	Discrepancy between T1 brain image and standard-space brain template (linearly-aligned)	UK Biobank Assessment Centre	Imaging	Brain MRI	110
25611	CONTINUOUS	30236	-0.04	-0.05	-0.03	4.53E-14	LINEAR	Weighted-mean L2 in tract inferior longitudinal fasciculus (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135

25789	CONTINUOUS	30655	0.04	0.03	0.05	5.35E-14	LINEAR	Volume of grey matter in Middle Frontal Gyrus (right)	UK Biobank Assessment Centre	Imaging	Imaging	1101
25498	CONTINUOUS	30236	0.04	0.03	0.05	5.73E-14	LINEAR	Weighted-mean FA in tract forceps major	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25514	CONTINUOUS	30236	0.04	0.03	0.05	6.64E-14	LINEAR	Weighted-mean FA in tract uncinate fasciculus (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25766	CONTINUOUS	26737	0.04	0.03	0.06	7.18E-14	LINEAR	90th percentile of z-statistic (in group-defined mask) for faces-shapes contrast	UK Biobank Assessment Centre	Imaging	Brain MRI	106
25336	CONTINUOUS	30236	-0.04	-0.05	-0.03	1.02E-13	LINEAR	Mean L3 in superior longitudinal fasciculus on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25282	CONTINUOUS	30236	-0.04	-0.05	-0.03	1.05E-13	LINEAR	Mean L2 in cingulum cingulate gyrus on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134

Table S3. PheWAS results for HICAS_pop. Only the bonferroni significant results are shown [Continued]

varName	varType	n	beta	lower	upper	pval	resType	description	Cat1 Title	Cat2 Title	Cat3 Title	Category
25489	CONTINUOUS	30236	0.04	0.03	0.05	1.2E-13	LINEAR	Weighted-mean FA in tract acoustic radiation (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25248	CONTINUOUS	30236	-0.04	-0.05	-0.03	1.77E-13	LINEAR	Mean L2 in middle cerebellar peduncle on FA skeleton	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25288	CONTINUOUS	30236	-0.04	-0.05	-0.03	1.81E-13	LINEAR	Mean L2 in superior longitudinal fasciculus on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25673	CONTINUOUS	30236	0.04	0.03	0.05	2.15E-13	LINEAR	Weighted-mean ICVF in tract superior thalamic radiation (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25311	CONTINUOUS	30236	-0.04	-0.05	-0.03	2.67E-13	LINEAR	Mean L3 in cerebral peduncle on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25252	CONTINUOUS	30236	-0.04	-0.05	-0.03	2.87E-13	LINEAR	Mean L2 in splenium of corpus callosum on FA skeleton	UK Biobank Assessment Centre	Imaging	Brain MRI	134

25359	CONTINUOUS	30236	0.04	0.03	0.05	2.89E-13	LINEAR	Mean ICVF in cerebral peduncle on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25330	CONTINUOUS	30236	-0.04	-0.05	-0.03	3.07E-13	LINEAR	Mean L3 in cingulum cingulate gyrus on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25140	CONTINUOUS	30236	-0.04	-0.05	-0.03	3.54E-13	LINEAR	Mean MD in cingulum hippocampus on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25185	CONTINUOUS	30236	0.04	0.03	0.05	4.1E-13	LINEAR	Mean MO in external capsule on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25408	CONTINUOUS	30236	-0.04	-0.05	-0.03	4.26E-13	LINEAR	Mean OD in anterior limb of internal capsule on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25513	CONTINUOUS	30236	0.04	0.03	0.05	4.35E-13	LINEAR	Weighted-mean FA in tract uncinate fasciculus (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135

Table S3. PheWAS results for HICAS_pop. Only the bonferroni significant results are shown [Continued]

varName	varType	n	beta	lower	upper	pval	resType	description	Cat1 Title	Cat2 Title	Cat3 Title	Category
25085	CONTINUOUS	30236	0.04	0.03	0.05	4.48E-13	LINEAR	Mean FA in posterior thalamic radiation on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25101	CONTINUOUS	30236	0.04	0.03	0.05	5.15E-13	LINEAR	Mean FA in uncinate fasciculus on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25502	CONTINUOUS	30236	0.04	0.03	0.05	7.58E-13	LINEAR	Weighted-mean FA in tract inferior longitudinal fasciculus (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25624	CONTINUOUS	30236	-0.04	-0.05	-0.03	7.77E-13	LINEAR	Weighted-mean L3 in tract acoustic radiation (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25304	CONTINUOUS	30236	-0.04	-0.05	-0.03	8.3E-13	LINEAR	Mean L3 in medial lemniscus on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25087	CONTINUOUS	30236	0.04	0.03	0.05	8.84E-13	LINEAR	Mean FA in sagittal stratum on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134

25637	CONTINUOUS	30236	-0.04	-0.05	-0.03	9.06E-13	LINEAR	Weighted-mean L3 in tract inferior longitudinal fasciculus (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25432	CONTINUOUS	30236	-0.04	-0.05	-0.03	9.28E-13	LINEAR	Mean OD in superior longitudinal fasciculus on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25493	CONTINUOUS	30236	0.04	0.03	0.05	1.04E-12	LINEAR	Weighted-mean FA in tract cingulate gyrus part of cingulum (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25333	CONTINUOUS	30236	-0.04	-0.05	-0.03	1.42E-12	LINEAR	Mean L3 in cingulum hippocampus on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25643	CONTINUOUS	30236	-0.04	-0.05	-0.03	1.43E-12	LINEAR	Weighted-mean L3 in tract posterior thalamic radiation (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25337	CONTINUOUS	30236	-0.04	-0.05	-0.03	1.82E-12	LINEAR	Mean L3 in superior longitudinal fasciculus on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134

Table S3. PheWAS results for HICAS_pop. Only the bonferroni significant results are shown [Continued]

varName	varType	n	beta	lower	upper	pval	resType	description	Cat1 Title	Cat2 Title	Cat3 Title	Category
25609	CONTINUOUS	30236	-0.04	-0.05	-0.03	1.9E-12	LINEAR	Weighted-mean L2 in tract inferior fronto-occipital fasciculus (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25376	CONTINUOUS	30236	0.04	0.03	0.05	1.99E-12	LINEAR	Mean ICVF in external capsule on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25063	CONTINUOUS	30236	0.04	0.03	0.05	2.01E-12	LINEAR	Mean FA in corticospinal tract on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25334	CONTINUOUS	30236	-0.04	-0.05	-0.03	2.09E-12	LINEAR	Mean L3 in fornix cres+stria terminalis on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25651	CONTINUOUS	30236	0.04	0.03	0.05	2.48E-12	LINEAR	Weighted-mean ICVF in tract acoustic radiation (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25297	CONTINUOUS	30236	-0.04	-0.05	-0.03	3.24E-12	LINEAR	Mean L3 in pontine crossing tract on FA skeleton	UK Biobank Assessment Centre	Imaging	Brain MRI	134

25645	CONTINUOUS	30236	-0.04	-0.05	-0.03	3.67E-12	LINEAR	Weighted-mean L3 in tract superior longitudinal fasciculus (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25492	CONTINUOUS	30236	0.04	0.03	0.05	5.78E-12	LINEAR	Weighted-mean FA in tract cingulate gyrus part of cingulum (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25424	CONTINUOUS	30236	-0.04	-0.05	-0.03	8.2E-12	LINEAR	Mean OD in external capsule on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25530	CONTINUOUS	30236	-0.04	-0.05	-0.03	9.44E-12	LINEAR	Weighted-mean MD in tract inferior longitudinal fasciculus (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25380	CONTINUOUS	30236	0.04	0.03	0.05	1.12E-11	LINEAR	Mean ICVF in cingulum hippocampus on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25358	CONTINUOUS	30236	0.04	0.03	0.05	1.21E-11	LINEAR	Mean ICVF in cerebral peduncle on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134

Table S3. PheWAS results for HICAS_pop. Only the bonferroni significant results are shown [Continued]

varName	varType	n	beta	lower	upper	pval	resType	description	Cat1 Title	Cat2 Title	Cat3 Title	Category
25108	CONTINUOUS	30236	-0.04	-0.05	-0.03	1.39E-11	LINEAR	Mean MD in splenium of corpus callosum on FA skeleton	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25646	CONTINUOUS	30236	-0.04	-0.05	-0.03	1.59E-11	LINEAR	Weighted-mean L3 in tract superior thalamic radiation (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25606	CONTINUOUS	30236	-0.04	-0.05	-0.03	1.62E-11	LINEAR	Weighted-mean L2 in tract forceps major	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25650	CONTINUOUS	30236	0.04	0.03	0.05	1.85E-11	LINEAR	Weighted-mean ICVF in tract acoustic radiation (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25094	CONTINUOUS	30236	0.03	0.02	0.04	2.26E-11	LINEAR	Mean FA in fornix cres+stria terminalis on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25633	CONTINUOUS	30236	-0.04	-0.05	-0.03	2.29E-11	LINEAR	Weighted-mean L3 in tract forceps major	UK Biobank Assessment Centre	Imaging	Brain MRI	135

25608	CONTINUOUS	30236	-0.04	-0.05	-0.03	2.78E-11	LINEAR	Weighted-mean L2 in tract inferior fronto-occipital fasciculus (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25635	CONTINUOUS	30236	-0.04	-0.05	-0.02	4.32E-11	LINEAR	Weighted-mean L3 in tract inferior fronto-occipital fasciculus (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25184	CONTINUOUS	30236	0.04	0.02	0.05	4.45E-11	LINEAR	Mean MO in external capsule on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25341	CONTINUOUS	30236	-0.04	-0.05	-0.03	5.05E-11	LINEAR	Mean L3 in uncinate fasciculus on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25389	CONTINUOUS	30236	0.04	0.03	0.05	5.1E-11	LINEAR	Mean ICVF in uncinate fasciculus on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25057	CONTINUOUS	30236	0.04	0.03	0.05	5.76E-11	LINEAR	Mean FA in pontine crossing tract on FA skeleton	UK Biobank Assessment Centre	Imaging	Brain MRI	134

Table S3. PheWAS results for HICAS_pop. Only the bonferroni significant results are shown [Continued]

varName	varType	n	beta	lower	upper	pval	resType	description	Cat1 Title	Cat2 Title	Cat3 Title	Category
25636	CONTINUOUS	30236	-0.04	-0.05	-0.02	6.49E-11	LINEAR	Weighted-mean L3 in tract inferior fronto-occipital fasciculus (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25539	CONTINUOUS	30236	-0.03	-0.04	-0.02	7.32E-11	LINEAR	Weighted-mean MD in tract superior thalamic radiation (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25377	CONTINUOUS	30236	0.03	0.02	0.05	7.63E-11	LINEAR	Mean ICVF in external capsule on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25381	CONTINUOUS	30236	0.04	0.03	0.05	7.78E-11	LINEAR	Mean ICVF in cingulum hippocampus on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25421	CONTINUOUS	30236	-0.04	-0.05	-0.03	8.55E-11	LINEAR	Mean OD in posterior thalamic radiation on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25280	CONTINUOUS	30236	-0.03	-0.04	-0.02	1.12E-10	LINEAR	Mean L2 in external capsule on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134

25100	CONTINUOUS	30236	0.04	0.03	0.05	1.13E-10	LINEAR	Mean FA in uncinate fasciculus on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25050	CONTINUOUS	26737	0.04	0.03	0.05	1.16E-10	LINEAR	Median z-statistic (in group-defined mask) for faces-shapes contrast	UK Biobank Assessment Centre	Imaging	Brain MRI	106
25327	CONTINUOUS	30236	-0.03	-0.05	-0.02	1.28E-10	LINEAR	Mean L3 in sagittal stratum on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25348	CONTINUOUS	30236	0.04	0.02	0.05	1.35E-10	LINEAR	Mean ICVF in splenium of corpus callosum on FA skeleton	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25281	CONTINUOUS	30236	-0.03	-0.04	-0.02	1.39E-10	LINEAR	Mean L2 in external capsule on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25084	CONTINUOUS	30236	0.03	0.02	0.05	1.39E-10	LINEAR	Mean FA in posterior thalamic radiation on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134

Table S3. PheWAS results for HICAS_pop. Only the bonferroni significant results are shown [Continued]

varName	varType	n	beta	lower	upper	pval	resType	description	Cat1 Title	Cat2 Title	Cat3 Title	Category
25658	CONTINUOUS	30236	0.04	0.03	0.05	1.48E-10	LINEAR	Weighted-mean ICVF in tract corticospinal tract (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25516	CONTINUOUS	30236	-0.04	-0.05	-0.03	1.52E-10	LINEAR	Weighted-mean MD in tract acoustic radiation (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25383	CONTINUOUS	30236	0.04	0.03	0.05	1.6E-10	LINEAR	Mean ICVF in fornix cres+stria terminalis on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25355	CONTINUOUS	30236	0.04	0.02	0.05	1.61E-10	LINEAR	Mean ICVF in inferior cerebellar peduncle on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25054	CONTINUOUS	26737	0.04	0.03	0.05	1.71E-10	LINEAR	Median z-statistic (in group-defined amygdala activation mask) for faces-shapes contrast	UK Biobank Assessment Centre	Imaging	Brain MRI	106

25073	CONTINUOUS	30236	0.04	0.02	0.05	1.71E-10	LINEAR	Mean FA in anterior limb of internal capsule on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25284	CONTINUOUS	30236	-0.04	-0.05	-0.03	1.76E-10	LINEAR	Mean L2 in cingulum hippocampus on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25627	CONTINUOUS	30236	-0.04	-0.05	-0.02	1.77E-10	LINEAR	Weighted-mean L3 in tract cingulate gyrus part of cingulum (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25568	CONTINUOUS	30236	0.04	0.03	0.05	2.08E-10	LINEAR	Weighted-mean MO in tract uncinate fasciculus (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25628	CONTINUOUS	30236	-0.04	-0.05	-0.02	2.48E-10	LINEAR	Weighted-mean L3 in tract cingulate gyrus part of cingulum (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25504	CONTINUOUS	30236	0.04	0.02	0.05	2.85E-10	LINEAR	Weighted-mean FA in tract middle cerebellar peduncle	UK Biobank Assessment Centre	Imaging	Brain MRI	135

Table S3. PheWAS results for HICAS_pop. Only the bonferroni significant results are shown [Continued]

varName	varType	n	beta	lower	upper	pval	resType	description	Cat1 Title	Cat2 Title	Cat3 Title	Category
25620	CONTINUOUS	30236	-0.03	-0.05	-0.02	3.18E-10	LINEAR	Weighted-mean L2 in tract superior thalamic radiation (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25877	CONTINUOUS	30655	0.03	0.02	0.04	3.35E-10	LINEAR	Volume of grey matter in Occipital Pole (right)	UK Biobank Assessment Centre	Imaging	Imaging	1101
25660	CONTINUOUS	30236	0.04	0.02	0.05	3.72E-10	LINEAR	Weighted-mean ICVF in tract forceps major	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25426	CONTINUOUS	30236	-0.04	-0.05	-0.02	3.76E-10	LINEAR	Mean OD in cingulum cingulate gyrus on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25379	CONTINUOUS	30236	0.03	0.02	0.04	4.06E-10	LINEAR	Mean ICVF in cingulum cingulate gyrus on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134

25630	CONTINUOUS	30236	-0.04	-0.05	-0.02	4.55E-10	LINEAR	Weighted-mean L3 in tract parahippocampal part of cingulum (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25671	CONTINUOUS	30236	0.03	0.02	0.05	5.25E-10	LINEAR	Weighted-mean ICVF in tract superior longitudinal fasciculus (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25597	CONTINUOUS	30236	-0.04	-0.05	-0.02	5.41E-10	LINEAR	Weighted-mean L2 in tract acoustic radiation (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25491	CONTINUOUS	30236	0.03	0.02	0.05	5.53E-10	LINEAR	Weighted-mean FA in tract anterior thalamic radiation (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25305	CONTINUOUS	30236	-0.04	-0.05	-0.02	5.7E-10	LINEAR	Mean L3 in medial lemniscus on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25314	CONTINUOUS	30236	-0.03	-0.05	-0.02	5.82E-10	LINEAR	Mean L3 in posterior limb of internal capsule on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25114	CONTINUOUS	30236	-0.03	-0.05	-0.02	5.92E-10	LINEAR	Mean MD in inferior cerebellar peduncle on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134

Table S3. PheWAS results for HICAS_pop. Only the bonferroni significant results are shown [Continued]

varName	varType	n	beta	lower	upper	pval	resType	description	Cat1 Title	Cat2 Title	Cat3 Title	Category
25070	CONTINUOUS	30236	0.03	0.02	0.04	7.18E-10	LINEAR	Mean FA in cerebral peduncle on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25495	CONTINUOUS	30236	0.03	0.02	0.05	7.8E-10	LINEAR	Weighted-mean FA in tract parahippocampal part of cingulum (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25535	CONTINUOUS	30236	-0.03	-0.04	-0.02	8.24E-10	LINEAR	Weighted-mean MD in tract posterior thalamic radiation (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25276	CONTINUOUS	30236	-0.03	-0.04	-0.02	8.85E-10	LINEAR	Mean L2 in posterior thalamic radiation on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25618	CONTINUOUS	30236	-0.03	-0.04	-0.02	9.2E-10	LINEAR	Weighted-mean L2 in tract superior longitudinal fasciculus (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135

25020	CONTINUOUS	30655	0.03	0.02	0.04	9.33E-10	LINEAR	Volume of hippocampus (right)	UK Biobank Assessment Centre	Imaging	Imaging	1102
25600	CONTINUOUS	30236	-0.03	-0.04	-0.02	9.52E-10	LINEAR	Weighted-mean L2 in tract cingulate gyrus part of cingulum (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25112	CONTINUOUS	30236	-0.03	-0.05	-0.02	1.31E-09	LINEAR	Mean MD in medial lemniscus on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25610	CONTINUOUS	30236	-0.03	-0.04	-0.02	1.45E-09	LINEAR	Weighted-mean L2 in tract inferior longitudinal fasciculus (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25622	CONTINUOUS	30236	-0.03	-0.04	-0.02	1.49E-09	LINEAR	Weighted-mean L2 in tract uncinate fasciculus (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25667	CONTINUOUS	30236	0.03	0.02	0.04	1.5E-09	LINEAR	Weighted-mean ICVF in tract medial lemniscus (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25512	CONTINUOUS	30236	0.03	0.02	0.05	1.57E-09	LINEAR	Weighted-mean FA in tract superior thalamic radiation (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135

Table S3. PheWAS results for HICAS_pop. Only the bonferroni significant results are shown [Continued]

varName	varType	n	beta	lower	upper	pval	resType	description	Cat1 Title	Cat2 Title	Cat3 Title	Category
25279	CONTINUOUS	30236	-0.03	-0.04	-0.02	1.74E-09	LINEAR	Mean L2 in sagittal stratum on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25886	CONTINUOUS	30655	0.03	0.02	0.04	1.76E-09	LINEAR	Volume of grey matter in Hippocampus (left)	UK Biobank Assessment Centre	Imaging	Imaging	1101
25665	CONTINUOUS	30236	0.03	0.02	0.04	2.61E-09	LINEAR	Weighted-mean ICVF in tract inferior longitudinal fasciculus (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25664	CONTINUOUS	30236	0.03	0.02	0.04	2.77E-09	LINEAR	Weighted-mean ICVF in tract inferior longitudinal fasciculus (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25617	CONTINUOUS	30236	-0.03	-0.04	-0.02	2.96E-09	LINEAR	Weighted-mean L2 in tract superior longitudinal fasciculus (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25652	CONTINUOUS	30236	0.03	0.02	0.04	3.3E-09	LINEAR	Weighted-mean ICVF in tract anterior thalamic radiation (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135

25071	CONTINUOUS	30236	0.03	0.02	0.04	3.42E-09	LINEAR	Mean FA in cerebral peduncle on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
6348	INTEGER	20189	-0.04	-0.05	-0.03	3.52E-09	LINEAR	Duration to complete numeric path (trail #1)	UK Biobank Assessment Centre	Cognitive function	Trail making	505
25525	CONTINUOUS	30236	-0.03	-0.04	-0.02	5.05E-09	LINEAR	Weighted-mean MD in tract forceps major	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25703	CONTINUOUS	30236	-0.03	-0.04	-0.02	5.2E-09	LINEAR	Weighted-mean OD in tract uncinate fasciculus (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25115	CONTINUOUS	30236	-0.03	-0.04	-0.02	5.4E-09	LINEAR	Mean MD in inferior cerebellar peduncle on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25529	CONTINUOUS	30236	-0.03	-0.04	-0.02	5.53E-09	LINEAR	Weighted-mean MD in tract inferior longitudinal fasciculus (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25352	CONTINUOUS	30236	0.03	0.02	0.04	5.91E-09	LINEAR	Mean ICVF in medial lemniscus on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134

Table S3. PheWAS results for HICAS_pop. Only the bonferroni significant results are shown [Continued]

varName	varType	n	beta	lower	upper	pval	resType	description	Cat1 Title	Cat2 Title	Cat3 Title	Category
25353	CONTINUOUS	30236	0.03	0.02	0.04	6.15E-09	LINEAR	Mean ICVF in medial lemniscus on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25329	CONTINUOUS	30236	-0.03	-0.04	-0.02	6.15E-09	LINEAR	Mean L3 in external capsule on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
6772	INTEGER	20228	-0.04	-0.05	-0.02	6.43E-09	LINEAR	Interval between previous point and current one in numeric path (trail #1)	UK Biobank Assessment Centre	Cognitive function	Trail making	505
25601	CONTINUOUS	30236	-0.03	-0.04	-0.02	6.71E-09	LINEAR	Weighted-mean L2 in tract cingulate gyrus part of cingulum (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25069	CONTINUOUS	30236	0.03	0.02	0.04	7.22E-09	LINEAR	Mean FA in superior cerebellar peduncle on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25663	CONTINUOUS	30236	0.03	0.02	0.04	8.1E-09	LINEAR	Weighted-mean ICVF in tract inferior fronto-occipital fasciculus (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135

25104	CONTINUOUS	30236	-0.03	-0.04	-0.02	8.61E-09	LINEAR	Mean MD in middle cerebellar peduncle on FA skeleton	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25326	CONTINUOUS	30236	-0.03	-0.04	-0.02	8.75E-09	LINEAR	Mean L3 in sagittal stratum on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25277	CONTINUOUS	30236	-0.03	-0.04	-0.02	1.19E-08	LINEAR	Mean L2 in posterior thalamic radiation on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25625	CONTINUOUS	30236	-0.03	-0.04	-0.02	1.2E-08	LINEAR	Weighted-mean L3 in tract anterior thalamic radiation (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25659	CONTINUOUS	30236	0.03	0.02	0.04	1.32E-08	LINEAR	Weighted-mean ICVF in tract corticospinal tract (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25320	CONTINUOUS	30236	-0.03	-0.04	-0.02	1.41E-08	LINEAR	Mean L3 in superior corona radiata on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25536	CONTINUOUS	30236	-0.03	-0.04	-0.02	1.43E-08	LINEAR	Weighted-mean MD in tract superior longitudinal fasciculus (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135

Table S3. PheWAS results for HICAS_pop. Only the bonferroni significant results are shown [Continued]

varName	varType	n	beta	lower	upper	pval	resType	description	Cat1 Title	Cat2 Title	Cat3 Title	Category
25193	CONTINUOUS	30236	0.03	0.02	0.04	1.5E-08	LINEAR	Mean MO in superior longitudinal fasciculus on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25603	CONTINUOUS	30236	-0.03	-0.04	-0.02	1.77E-08	LINEAR	Weighted-mean L2 in tract parahippocampal part of cingulum (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25083	CONTINUOUS	30236	0.03	0.02	0.04	1.79E-08	LINEAR	Mean FA in posterior corona radiata on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25681	CONTINUOUS	30236	-0.03	-0.04	-0.02	1.85E-08	LINEAR	Weighted-mean OD in tract cingulate gyrus part of cingulum (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25361	CONTINUOUS	30236	0.03	0.02	0.04	1.87E-08	LINEAR	Mean ICVF in anterior limb of internal capsule on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25654	CONTINUOUS	30236	0.03	0.02	0.04	2.07E-08	LINEAR	Weighted-mean ICVF in tract cingulate gyrus part of cingulum (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135

25378	CONTINUOUS	30236	0.03	0.02	0.04	2.08E-08	LINEAR	Mean ICVF in cingulum cingulate gyrus on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25668	CONTINUOUS	30236	0.03	0.02	0.04	2.29E-08	LINEAR	Weighted-mean ICVF in tract medial lemniscus (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25655	CONTINUOUS	30236	0.03	0.02	0.04	2.39E-08	LINEAR	Weighted-mean ICVF in tract cingulate gyrus part of cingulum (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25309	CONTINUOUS	30236	-0.03	-0.04	-0.02	2.42E-08	LINEAR	Mean L3 in superior cerebellar peduncle on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25528	CONTINUOUS	30236	-0.03	-0.04	-0.02	2.47E-08	LINEAR	Weighted-mean MD in tract inferior fronto-occipital fasciculus (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25256	CONTINUOUS	30236	-0.03	-0.04	-0.02	2.77E-08	LINEAR	Mean L2 in medial lemniscus on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25378	CONTINUOUS	30236	0.03	0.02	0.04	2.08E-08	LINEAR	Mean ICVF in cingulum cingulate gyrus on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134

Table S3. PheWAS results for HICAS_pop. Only the bonferroni significant results are shown [Continued]

varName	varType	n	beta	lower	upper	pval	resType	description	Cat1 Title	Cat2 Title	Cat3 Title	Category
25668	CONTINUOUS	30236	0.03	0.02	0.04	2.29E-08	LINEAR	Weighted-mean ICVF in tract medial lemniscus (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25655	CONTINUOUS	30236	0.03	0.02	0.04	2.39E-08	LINEAR	Weighted-mean ICVF in tract cingulate gyrus part of cingulum (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25309	CONTINUOUS	30236	-0.03	-0.04	-0.02	2.42E-08	LINEAR	Mean L3 in superior cerebellar peduncle on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25528	CONTINUOUS	30236	-0.03	-0.04	-0.02	2.47E-08	LINEAR	Weighted-mean MD in tract inferior fronto-occipital fasciculus (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25256	CONTINUOUS	30236	-0.03	-0.04	-0.02	2.77E-08	LINEAR	Mean L2 in medial lemniscus on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25316	CONTINUOUS	30236	-0.03	-0.04	-0.02	3.03E-08	LINEAR	Mean L3 in retrolenticular part of internal capsule on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134

25672	CONTINUOUS	30236	0.03	0.02	0.04	3.26E-08	LINEAR	Weighted-mean ICVF in tract superior longitudinal fasciculus (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25382	CONTINUOUS	30236	0.03	0.02	0.04	3.27E-08	LINEAR	Mean ICVF in fornix cres+stria terminalis on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25656	CONTINUOUS	30236	0.03	0.02	0.04	3.38E-08	LINEAR	Weighted-mean ICVF in tract parahippocampal part of cingulum (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25113	CONTINUOUS	30236	-0.03	-0.04	-0.02	3.47E-08	LINEAR	Mean MD in medial lemniscus on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25064	CONTINUOUS	30236	0.03	0.02	0.04	3.52E-08	LINEAR	Mean FA in medial lemniscus on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25306	CONTINUOUS	30236	-0.03	-0.04	-0.02	3.52E-08	LINEAR	Mean L3 in inferior cerebellar peduncle on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25632	CONTINUOUS	30236	-0.03	-0.04	-0.02	3.77E-08	LINEAR	Weighted-mean L3 in tract corticospinal tract (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135

Table S3. PheWAS results for HICAS_pop. Only the bonferroni significant results are shown [Continued]

varName	varType	n	beta	lower	upper	pval	resType	description	Cat1 Title	Cat2 Title	Cat3 Title	Category
46	INTEGER	30578	0.02	0.01	0.03	4.16E-08	LINEAR	Hand grip strength (left)	UK Biobank Assessment Centre	Physical measures	Hand grip strength	100019
1990	CAT-SIN	25665/4331 (29996)	-0.09	-0.12	-0.06	4.43E-08	LOGISTIC-BINARY	Tense / 'highly strung'	UK Biobank Assessment Centre	Touchscreen	Psychosocial factors	100060
25837	CONTINUOUS	30655	0.03	0.02	0.04	4.64E-08	LINEAR	Volume of grey matter in Paracingulate Gyrus (right)	UK Biobank Assessment Centre	Imaging	Imaging	1101
25257	CONTINUOUS	30236	-0.03	-0.04	-0.02	5.1E-08	LINEAR	Mean L2 in medial lemniscus on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25649	CONTINUOUS	30236	-0.03	-0.04	-0.02	5.34E-08	LINEAR	Weighted-mean L3 in tract uncinate fasciculus (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25522	CONTINUOUS	30236	-0.03	-0.04	-0.02	5.63E-08	LINEAR	Weighted-mean MD in tract parahippocampal part of cingulum (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135

25629	CONTINUOUS	30236	-0.03	-0.04	-0.02	5.84E-08	LINEAR	Weighted-mean L3 in tract parahippocampal part of cingulum (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25354	CONTINUOUS	30236	0.03	0.02	0.04	6.01E-08	LINEAR	Mean ICVF in inferior cerebellar peduncle on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25537	CONTINUOUS	30236	-0.03	-0.04	-0.02	6.24E-08	LINEAR	Weighted-mean MD in tract superior longitudinal fasciculus (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25423	CONTINUOUS	30236	-0.03	-0.04	-0.02	6.26E-08	LINEAR	Mean OD in sagittal stratum on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25657	CONTINUOUS	30236	0.03	0.02	0.04	6.48E-08	LINEAR	Weighted-mean ICVF in tract parahippocampal part of cingulum (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25328	CONTINUOUS	30236	-0.03	-0.04	-0.02	6.52E-08	LINEAR	Mean L3 in external capsule on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25629	CONTINUOUS	30236	-0.03	-0.04	-0.02	5.84E-08	LINEAR	Weighted-mean L3 in tract parahippocampal part of cingulum (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135

Table S3. PheWAS results for HICAS_pop. Only the bonferroni significant results are shown [Continued]

varName	varType	n	beta	lower	upper	pval	resType	description	Cat1 Title	Cat2 Title	Cat3 Title	Category
25354	CONTINUOUS	30236	0.03	0.02	0.04	6.01E-08	LINEAR	Mean ICVF in inferior cerebellar peduncle on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25537	CONTINUOUS	30236	-0.03	-0.04	-0.02	6.24E-08	LINEAR	Weighted-mean MD in tract superior longitudinal fasciculus (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25423	CONTINUOUS	30236	-0.03	-0.04	-0.02	6.26E-08	LINEAR	Mean OD in sagittal stratum on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25657	CONTINUOUS	30236	0.03	0.02	0.04	6.48E-08	LINEAR	Weighted-mean ICVF in tract parahippocampal part of cingulum (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25328	CONTINUOUS	30236	-0.03	-0.04	-0.02	6.52E-08	LINEAR	Mean L3 in external capsule on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134

25538	CONTINUOUS	30236	-0.03	-0.04	-0.02	7.03E-08	LINEAR	Weighted-mean MD in tract superior thalamic radiation (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25599	CONTINUOUS	30236	-0.03	-0.04	-0.02	7.13E-08	LINEAR	Weighted-mean L2 in tract anterior thalamic radiation (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25255	CONTINUOUS	30236	-0.03	-0.04	-0.02	7.17E-08	LINEAR	Mean L2 in corticospinal tract on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25086	CONTINUOUS	30236	0.03	0.02	0.04	7.44E-08	LINEAR	Mean FA in sagittal stratum on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25099	CONTINUOUS	30236	0.03	0.02	0.04	7.55E-08	LINEAR	Mean FA in superior fronto-occipital fasciculus on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25436	CONTINUOUS	30236	-0.03	-0.04	-0.02	8E-08	LINEAR	Mean OD in uncinate fasciculus on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25388	CONTINUOUS	30236	0.03	0.02	0.04	8.25E-08	LINEAR	Mean ICVF in uncinate fasciculus on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134

Table S3. PheWAS results for HICAS_pop. Only the bonferroni significant results are shown [Continued]

varName	varType	n	beta	lower	upper	pval	resType	description	Cat1 Title	Cat2 Title	Cat3 Title	Category
25124	CONTINUOUS	30236	-0.03	-0.04	-0.02	8.37E-08	LINEAR	Mean MD in retrolenticular part of internal capsule on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25440	CONTINUOUS	30236	-0.03	-0.04	-0.02	8.62E-08	LINEAR	Mean ISOVF in middle cerebellar peduncle on FA skeleton	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25142	CONTINUOUS	30236	-0.03	-0.04	-0.02	8.65E-08	LINEAR	Mean MD in fornix cres+stria terminalis on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25141	CONTINUOUS	30236	-0.03	-0.04	-0.02	8.91E-08	LINEAR	Mean MD in cingulum hippocampus on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25303	CONTINUOUS	30236	-0.03	-0.04	-0.02	9.04E-08	LINEAR	Mean L3 in corticospinal tract on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25374	CONTINUOUS	30236	0.03	0.02	0.04	9.13E-08	LINEAR	Mean ICVF in sagittal stratum on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134

25511	CONTINUOUS	30236	0.03	0.02	0.04	1.03E-07	LINEAR	Weighted-mean FA in tract superior thalamic radiation (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25310	CONTINUOUS	30236	-0.03	-0.04	-0.02	1.04E-07	LINEAR	Mean L3 in cerebral peduncle on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25105	CONTINUOUS	30236	-0.03	-0.04	-0.02	1.14E-07	LINEAR	Mean MD in pontine crossing tract on FA skeleton	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25619	CONTINUOUS	30236	-0.03	-0.04	-0.02	1.16E-07	LINEAR	Weighted-mean L2 in tract superior thalamic radiation (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25662	CONTINUOUS	30236	0.03	0.02	0.04	1.16E-07	LINEAR	Weighted-mean ICVF in tract inferior fronto-occipital fasciculus (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25384	CONTINUOUS	30236	0.03	0.02	0.04	1.2E-07	LINEAR	Mean ICVF in superior longitudinal fasciculus on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25092	CONTINUOUS	30236	0.03	0.02	0.04	1.37E-07	LINEAR	Mean FA in cingulum hippocampus on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134

Table S3. PheWAS results for HICAS_pop. Only the bonferroni significant results are shown [Continued]

varName	varType	n	beta	lower	upper	pval	resType	description	Cat1 Title	Cat2 Title	Cat3 Title	Category
25598	CONTINUOUS	30236	-0.03	-0.04	-0.02	1.42E-07	LINEAR	Weighted-mean L2 in tract anterior thalamic radiation (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25499	CONTINUOUS	30236	0.03	0.02	0.04	1.55E-07	LINEAR	Weighted-mean FA in tract forceps minor	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25286	CONTINUOUS	30236	-0.03	-0.04	-0.02	1.62E-07	LINEAR	Mean L2 in fornix cres+stria terminalis on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25181	CONTINUOUS	30236	0.03	0.02	0.04	1.7E-07	LINEAR	Mean MO in posterior thalamic radiation on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25261	CONTINUOUS	30236	-0.03	-0.04	-0.02	1.92E-07	LINEAR	Mean L2 in superior cerebellar peduncle on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25385	CONTINUOUS	30236	0.03	0.02	0.04	2.08E-07	LINEAR	Mean ICVF in superior longitudinal fasciculus on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134

25527	CONTINUOUS	30236	-0.03	-0.04	-0.02	2.08E-07	LINEAR	Weighted-mean MD in tract inferior fronto-occipital fasciculus (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25889	CONTINUOUS	30655	0.03	0.02	0.04	2.16E-07	LINEAR	Volume of grey matter in Amygdala (right)	UK Biobank Assessment Centre	Imaging	Imaging	1101
25488	CONTINUOUS	30236	0.03	0.02	0.04	2.28E-07	LINEAR	Weighted-mean FA in tract acoustic radiation (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25356	CONTINUOUS	30236	0.03	0.02	0.04	2.29E-07	LINEAR	Mean ICVF in superior cerebellar peduncle on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25264	CONTINUOUS	30236	-0.03	-0.04	-0.02	2.34E-07	LINEAR	Mean L2 in anterior limb of internal capsule on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25484	CONTINUOUS	30236	0.03	0.02	0.04	2.45E-07	LINEAR	Mean ISOVF in uncinate fasciculus on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25278	CONTINUOUS	30236	-0.03	-0.04	-0.02	2.46E-07	LINEAR	Mean L2 in sagittal stratum on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134

Table S3. PheWAS results for HICAS_pop. Only the bonferroni significant results are shown [Continued]

varName	varType	n	beta	lower	upper	pval	resType	description	Cat1 Title	Cat2 Title	Cat3 Title	Category
25814	CONTINUOUS	30655	0.03	0.02	0.04	2.61E-07	LINEAR	Volume of grey matter in Postcentral Gyrus (left)	UK Biobank Assessment Centre	Imaging	Imaging	1101
25888	CONTINUOUS	30655	0.03	0.02	0.04	2.96E-07	LINEAR	Volume of grey matter in Amygdala (left)	UK Biobank Assessment Centre	Imaging	Imaging	1101
25365	CONTINUOUS	30236	0.03	0.02	0.04	3.23E-07	LINEAR	Mean ICVF in retrolenticular part of internal capsule on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25653	CONTINUOUS	30236	0.03	0.02	0.04	3.34E-07	LINEAR	Weighted-mean ICVF in tract anterior thalamic radiation (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25670	CONTINUOUS	30236	0.03	0.02	0.04	3.39E-07	LINEAR	Weighted-mean ICVF in tract posterior thalamic radiation (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25545	CONTINUOUS	30236	0.03	0.02	0.04	3.55E-07	LINEAR	Weighted-mean MO in tract anterior thalamic radiation (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135

25786	CONTINUOUS	30655	0.03	0.02	0.04	3.64E-07	LINEAR	Volume of grey matter in Superior Frontal Gyrus (left)	UK Biobank Assessment Centre	Imaging	Imaging	1101
25144	CONTINUOUS	30236	-0.03	-0.04	-0.02	4.12E-07	LINEAR	Mean MD in superior longitudinal fasciculus on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25682	CONTINUOUS	30236	-0.03	-0.04	-0.02	4.19E-07	LINEAR	Weighted-mean OD in tract cingulate gyrus part of cingulum (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25322	CONTINUOUS	30236	-0.03	-0.04	-0.02	4.43E-07	LINEAR	Mean L3 in posterior corona radiata on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25001	CONTINUOUS	30655	0.02	0.01	0.03	4.5E-07	LINEAR	Volume of peripheral cortical grey matter (normalised for head size)	UK Biobank Assessment Centre	Imaging	Brain MRI	110
25357	CONTINUOUS	30236	0.03	0.02	0.04	4.83E-07	LINEAR	Mean ICVF in superior cerebellar peduncle on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134

Table S3. PheWAS results for HICAS_pop. Only the bonferroni significant results are shown [Continued]

varName	varType	n	beta	lower	upper	pval	resType	description	Cat1 Title	Cat2 Title	Cat3 Title	Category
6383	INTEGER	20386	0.06	0.04	0.09	4.92E-07	ORDERED-LOGISTIC	Number of puzzles attempted	UK Biobank Assessment Centre	Cognitive function	Tower rearranging	503
25065	CONTINUOUS	30236	0.03	0.02	0.04	5.36E-07	LINEAR	Mean FA in medial lemniscus on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25634	CONTINUOUS	30236	-0.03	-0.04	-0.02	5.69E-07	LINEAR	Weighted-mean L3 in tract forceps minor	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25180	CONTINUOUS	30236	0.03	0.02	0.04	6.14E-07	LINEAR	Mean MO in posterior thalamic radiation on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25059	CONTINUOUS	30236	0.03	0.02	0.04	6.2E-07	LINEAR	Mean FA in body of corpus callosum on FA skeleton	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25062	CONTINUOUS	30236	0.03	0.02	0.04	6.23E-07	LINEAR	Mean FA in corticospinal tract on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134

25122	CONTINUOUS	30236	-0.03	-0.04	-0.02	6.46E-07	LINEAR	Mean MD in posterior limb of internal capsule on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25186	CONTINUOUS	30236	0.03	0.02	0.04	6.48E-07	LINEAR	Mean MO in cingulum cingulate gyrus on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25441	CONTINUOUS	30236	-0.03	-0.04	-0.02	6.56E-07	LINEAR	Mean ISOVF in pontine crossing tract on FA skeleton	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25360	CONTINUOUS	30236	0.03	0.02	0.04	7.28E-07	LINEAR	Mean ICVF in anterior limb of internal capsule on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25134	CONTINUOUS	30236	-0.03	-0.04	-0.02	7.45E-07	LINEAR	Mean MD in sagittal stratum on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25890	CONTINUOUS	30655	0.03	0.02	0.04	7.72E-07	LINEAR	Volume of grey matter in Ventral Striatum (left)	UK Biobank Assessment Centre	Imaging	Imaging	1101
25642	CONTINUOUS	30236	-0.03	-0.04	-0.02	7.97E-07	LINEAR	Weighted-mean L3 in tract posterior thalamic radiation (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135

Table S3. PheWAS results for HICAS_pop. Only the bonferroni significant results are shown [Continued]

varName	varType	n	beta	lower	upper	pval	resType	description	Cat1 Title	Cat2 Title	Cat3 Title	Category
25340	CONTINUOUS	30236	-0.03	-0.04	-0.02	8.03E-07	LINEAR	Mean L3 in uncinate fasciculus on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25679	CONTINUOUS	30236	-0.03	-0.04	-0.02	8.64E-07	LINEAR	Weighted-mean OD in tract anterior thalamic radiation (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25363	CONTINUOUS	30236	0.03	0.02	0.04	9.2E-07	LINEAR	Mean ICVF in posterior limb of internal capsule on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25615	CONTINUOUS	30236	-0.03	-0.04	-0.02	9.57E-07	LINEAR	Weighted-mean L2 in tract posterior thalamic radiation (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25364	CONTINUOUS	30236	0.03	0.02	0.04	1.07E-06	LINEAR	Mean ICVF in retrolenticular part of internal capsule on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25307	CONTINUOUS	30236	-0.03	-0.04	-0.02	1.11E-06	LINEAR	Mean L3 in inferior cerebellar peduncle on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134

404	INTEGER	30496	-0.03	-0.04	-0.02	1.2E-06	LINEAR	Duration to first press of snap-button in each round	UK Biobank Assessment Centre	Cognitive function	Reaction time	100032
25251	CONTINUOUS	30236	-0.03	-0.04	-0.02	1.22E-06	LINEAR	Mean L2 in body of corpus callosum on FA skeleton	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25375	CONTINUOUS	30236	0.03	0.02	0.04	1.22E-06	LINEAR	Mean ICVF in sagittal stratum on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25675	CONTINUOUS	30236	0.03	0.02	0.04	1.29E-06	LINEAR	Weighted-mean ICVF in tract uncinate fasciculus (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25494	CONTINUOUS	30236	0.03	0.02	0.04	1.29E-06	LINEAR	Weighted-mean FA in tract parahippocampal part of cingulum (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
1970	CAT-SIN	24064/6016(30080)	-0.07	-0.10	-0.04	1.31E-06	LOGISTIC-BINARY	Nervous feelings	UK Biobank Assessment Centre	Touchscreen	Psychosocial factors	100060
25093	CONTINUOUS	30236	0.03	0.02	0.04	1.32E-06	LINEAR	Mean FA in cingulum hippocampus on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134

Table S3. PheWAS results for HICAS_pop. Only the bonferroni significant results are shown [Continued]

varName	varType	n	beta	lower	upper	pval	resType	description	Cat1 Title	Cat2 Title	Cat3 Title	Category
25139	CONTINUOUS	30236	-0.03	-0.04	-0.02	1.41E-06	LINEAR	Mean MD in cingulum cingulate gyrus on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25698	CONTINUOUS	30236	-0.03	-0.04	-0.02	1.5E-06	LINEAR	Weighted-mean OD in tract superior longitudinal fasciculus (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25555	CONTINUOUS	30236	0.03	0.02	0.04	1.51E-06	LINEAR	Weighted-mean MO in tract inferior fronto-occipital fasciculus (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25187	CONTINUOUS	30236	0.03	0.02	0.04	1.54E-06	LINEAR	Mean MO in cingulum cingulate gyrus on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
23323	INTEGER	20231	0.03	0.02	0.04	1.65E-06	LINEAR	Number of symbol digit matches attempted	UK Biobank Assessment Centre	Cognitive function	Symbol digit substitution	502
25275	CONTINUOUS	30236	-0.03	-0.04	-0.02	1.84E-06	LINEAR	Mean L2 in posterior corona radiata on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134

25249	CONTINUOUS	30236	-0.03	-0.04	-0.02	1.88E-06	LINEAR	Mean L2 in pontine crossing tract on FA skeleton	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25098	CONTINUOUS	30236	0.03	0.02	0.04	2.15E-06	LINEAR	Mean FA in superior fronto-occipital fasciculus on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25002	CONTINUOUS	30655	0.02	0.01	0.03	2.15E-06	LINEAR	Volume of peripheral cortical grey matter	UK Biobank Assessment Centre	Imaging	Brain MRI	110
25145	CONTINUOUS	30236	-0.03	-0.04	-0.02	2.17E-06	LINEAR	Mean MD in superior longitudinal fasciculus on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25420	CONTINUOUS	30236	-0.03	-0.04	-0.02	2.33E-06	LINEAR	Mean OD in posterior thalamic radiation on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25768	CONTINUOUS	26737	0.03	0.02	0.04	2.35E-06	LINEAR	90th percentile of z-statistic (in group-defined amygdala activation mask) for faces-shapes contrast	UK Biobank Assessment Centre	Imaging	Brain MRI	106

Table S3. PheWAS results for HICAS_pop. Only the bonferroni significant results are shown [Continued]

varName	varType	n	beta	lower	upper	pval	resType	description	Cat1 Title	Cat2 Title	Cat3 Title	Category
25299	CONTINUOUS	30236	-0.03	-0.04	-0.01	2.63E-06	LINEAR	Mean L3 in body of corpus callosum on FA skeleton	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25312	CONTINUOUS	30236	-0.02	-0.03	-0.01	3E-06	LINEAR	Mean L3 in anterior limb of internal capsule on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25605	CONTINUOUS	30236	-0.03	-0.04	-0.02	3.06E-06	LINEAR	Weighted-mean L2 in tract corticospinal tract (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25135	CONTINUOUS	30236	-0.03	-0.04	-0.01	3.47E-06	LINEAR	Mean MD in sagittal stratum on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25233	CONTINUOUS	30236	0.03	0.01	0.04	3.92E-06	LINEAR	Mean L1 in external capsule on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25080	CONTINUOUS	30236	0.03	0.01	0.04	4.04E-06	LINEAR	Mean FA in superior corona radiata on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134

25497	CONTINUOUS	30236	0.03	0.01	0.04	4.06E-06	LINEAR	Weighted-mean FA in tract corticospinal tract (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25082	CONTINUOUS	30236	0.03	0.01	0.04	4.6E-06	LINEAR	Mean FA in posterior corona radiata on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25794	CONTINUOUS	30655	0.02	0.01	0.03	4.71E-06	LINEAR	Volume of grey matter in Precentral Gyrus (left)	UK Biobank Assessment Centre	Imaging	Imaging	1101
25626	CONTINUOUS	30236	-0.02	-0.03	-0.01	4.79E-06	LINEAR	Weighted-mean L3 in tract anterior thalamic radiation (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25349	CONTINUOUS	30236	0.03	0.01	0.04	4.81E-06	LINEAR	Mean ICVF in fornix on FA skeleton	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25371	CONTINUOUS	30236	0.03	0.01	0.04	5.18E-06	LINEAR	Mean ICVF in posterior corona radiata on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25211	CONTINUOUS	30236	-0.03	-0.04	-0.01	5.76E-06	LINEAR	Mean L1 in inferior cerebellar peduncle on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134

Table S3. PheWAS results for HICAS_pop. Only the bonferroni significant results are shown [Continued]

varName	varType	n	beta	lower	upper	pval	resType	description	Cat1 Title	Cat2 Title	Cat3 Title	Category
25816	CONTINUOUS	30655	0.03	0.01	0.04	5.86E-06	LINEAR	Volume of grey matter in Superior Parietal Lobule (left)	UK Biobank Assessment Centre	Imaging	Imaging	1101
25570	CONTINUOUS	30236	-0.03	-0.04	-0.01	5.91E-06	LINEAR	Weighted-mean L1 in tract acoustic radiation (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25689	CONTINUOUS	30236	-0.03	-0.04	-0.01	6.09E-06	LINEAR	Weighted-mean OD in tract inferior fronto-occipital fasciculus (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25128	CONTINUOUS	30236	-0.02	-0.03	-0.01	6.55E-06	LINEAR	Mean MD in superior corona radiata on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25323	CONTINUOUS	30236	-0.02	-0.03	-0.01	6.64E-06	LINEAR	Mean L3 in posterior corona radiata on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25607	CONTINUOUS	30236	-0.02	-0.04	-0.01	6.74E-06	LINEAR	Weighted-mean L2 in tract forceps minor	UK Biobank Assessment Centre	Imaging	Brain MRI	135

25435	CONTINUOUS	30236	-0.03	-0.04	-0.01	6.82E-06	LINEAR	Mean OD in superior fronto-occipital fasciculus on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25437	CONTINUOUS	30236	-0.03	-0.04	-0.01	7.14E-06	LINEAR	Mean OD in uncinate fasciculus on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25293	CONTINUOUS	30236	-0.03	-0.04	-0.01	7.16E-06	LINEAR	Mean L2 in uncinate fasciculus on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25192	CONTINUOUS	30236	0.03	0.01	0.04	7.93E-06	LINEAR	Mean MO in superior longitudinal fasciculus on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25521	CONTINUOUS	30236	-0.02	-0.04	-0.01	8.29E-06	LINEAR	Weighted-mean MD in tract parahippocampal part of cingulum (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25676	CONTINUOUS	30236	0.02	0.01	0.04	8.7E-06	LINEAR	Weighted-mean ICVF in tract uncinate fasciculus (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25554	CONTINUOUS	30236	0.03	0.01	0.04	8.72E-06	LINEAR	Weighted-mean MO in tract inferior fronto-occipital fasciculus (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135

Table S3. PheWAS results for HICAS_pop. Only the bonferroni significant results are shown [Continued]

varName	varType	n	beta	lower	upper	pval	resType	description	Cat1 Title	Cat2 Title	Cat3 Title	Category
25562	CONTINUOUS	30236	0.03	0.01	0.04	9.13E-06	LINEAR	Weighted-mean MO in tract posterior thalamic radiation (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25130	CONTINUOUS	30236	-0.02	-0.03	-0.01	9.24E-06	LINEAR	Mean MD in posterior corona radiata on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25546	CONTINUOUS	30236	0.02	0.01	0.03	9.83E-06	LINEAR	Weighted-mean MO in tract cingulate gyrus part of cingulum (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25325	CONTINUOUS	30236	-0.02	-0.03	-0.01	1E-05	LINEAR	Mean L3 in posterior thalamic radiation on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
23324	INTEGER	20231	0.03	0.02	0.04	1.01E-05	LINEAR	Number of symbol digit matches made correctly	UK Biobank Assessment Centre	Cognitive function	Symbol digit substitution	502
25058	CONTINUOUS	30236	0.02	0.01	0.03	1.02E-05	LINEAR	Mean FA in genu of corpus callosum on FA skeleton	UK Biobank Assessment Centre	Imaging	Brain MRI	134

25117	CONTINUOUS	30236	-0.03	-0.04	-0.01	1.05E-05	LINEAR	Mean MD in superior cerebellar peduncle on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25461	CONTINUOUS	30236	0.03	0.01	0.04	1.09E-05	LINEAR	Mean ISOVF in retrolenticular part of internal capsule on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25534	CONTINUOUS	30236	-0.02	-0.03	-0.01	1.11E-05	LINEAR	Weighted-mean MD in tract posterior thalamic radiation (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25669	CONTINUOUS	30236	0.02	0.01	0.03	1.13E-05	LINEAR	Weighted-mean ICVF in tract posterior thalamic radiation (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25507	CONTINUOUS	30236	0.02	0.01	0.03	1.14E-05	LINEAR	Weighted-mean FA in tract posterior thalamic radiation (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25409	CONTINUOUS	30236	-0.03	-0.04	-0.01	1.15E-05	LINEAR	Mean OD in anterior limb of internal capsule on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25461	CONTINUOUS	30236	0.03	0.01	0.04	1.09E-05	LINEAR	Mean ISOVF in retrolenticular part of internal capsule on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134

Table S3. PheWAS results for HICAS_pop. Only the bonferroni significant results are shown [Continued]

varName	varType	n	beta	lower	upper	pval	resType	description	Cat1 Title	Cat2 Title	Cat3 Title	Category
25534	CONTINUOUS	30236	-0.02	-0.03	-0.01	1.11E-05	LINEAR	Weighted-mean MD in tract posterior thalamic radiation (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25669	CONTINUOUS	30236	0.02	0.01	0.03	1.13E-05	LINEAR	Weighted-mean ICVF in tract posterior thalamic radiation (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25507	CONTINUOUS	30236	0.02	0.01	0.03	1.14E-05	LINEAR	Weighted-mean FA in tract posterior thalamic radiation (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25409	CONTINUOUS	30236	-0.03	-0.04	-0.01	1.15E-05	LINEAR	Mean OD in anterior limb of internal capsule on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25732	CONTINUOUS	30655	-0.02	-0.03	-0.01	1.18E-05	LINEAR	Discrepancy between T1 brain image and standard-space brain template (nonlinearly-aligned)	UK Biobank Assessment Centre	Imaging	Brain MRI	110

25324	CONTINUOUS	30236	-0.02	-0.03	-0.01	1.24E-05	LINEAR	Mean L3 in posterior thalamic radiation on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25368	CONTINUOUS	30236	0.02	0.01	0.03	1.31E-05	LINEAR	Mean ICVF in superior corona radiata on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25648	CONTINUOUS	30236	-0.02	-0.03	-0.01	1.39E-05	LINEAR	Weighted-mean L3 in tract uncinate fasciculus (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25861	CONTINUOUS	30655	0.02	0.01	0.03	1.39E-05	LINEAR	Volume of grey matter in Occipital Fusiform Gyrus (right)	UK Biobank Assessment Centre	Imaging	Imaging	1101
25702	CONTINUOUS	30236	-0.02	-0.04	-0.01	1.4E-05	LINEAR	Weighted-mean OD in tract uncinate fasciculus (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
20156	CONTINUOUS	14322	-0.03	-0.05	-0.02	1.48E-05	LINEAR	Duration to complete numeric path (trail #1)	Online follow-up	Cognitive function online	Trail making	121
25191	CONTINUOUS	30236	-0.02	-0.03	-0.01	1.49E-05	LINEAR	Mean MO in fornix cres+stria terminalis on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	134

Table S4. PheWAS results for Rs_DC_asymm. Only Bonferroni significant results are shown.

varName	varType	n	beta	lower	upper	pvalue	resType	description	Cat1 Title	Cat2 Title	Cat3 Title	Category
25739	CONTINUOUS	30655	-0.05	-0.06	-0.04	1.28E-26	LINEAR	Discrepancy between rfMRI brain image and T1 brain image	UK Biobank Assessment Centre	Imaging	Brain MRI	111
25740	CONTINUOUS	26737	-0.05	-0.06	-0.04	8.56E-23	LINEAR	Discrepancy between tfMRI brain image and T1 brain image	UK Biobank Assessment Centre	Imaging	Brain MRI	106
25737	CONTINUOUS	30655	-0.05	-0.06	-0.04	2.75E-19	LINEAR	Discrepancy between dMRI brain image and T1 brain image	UK Biobank Assessment Centre	Imaging	Brain MRI	107
1707-1	CAT-SIN	27261/30651	-999	-999	-999	3.04E-17	MULTINOMIAL-LOGISTIC	Handedness (chirality/laterality)	UK Biobank Assessment Centre	Touch screen	Early life factors	100033
25872	CONTINUOUS	30655	0.04	0.03	0.05	1.46E-14	LINEAR	Volume of grey matter in Planum Temporale (left)	UK Biobank Assessment Centre	Imaging	Imaging	1101
25846	CONTINUOUS	30655	0.04	0.03	0.05	2.27E-13	LINEAR	Volume of grey matter in Frontal Orbital Cortex (left)	UK Biobank Assessment Centre	Imaging	Imaging	1101
25697	CONTINUOUS	30236	-0.04	-0.05	-0.03	1.52E-12	LINEAR	Weighted-mean OD in tract posterior thalamic radiation (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135

25687	CONTINUOUS	30236	-0.04	-0.05	-0.03	1.23E-11	LINEAR	Weighted-mean OD in tract forceps major	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25719	CONTINUOUS	30236	0.03	0.02	0.05	3.44E-10	LINEAR	Weighted-mean ISOVF in tract inferior longitudinal fasciculus (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25696	CONTINUOUS	30236	-0.04	-0.05	-0.02	5.52E-10	LINEAR	Weighted-mean OD in tract posterior thalamic radiation (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25561	CONTINUOUS	30236	0.03	0.02	0.04	4.03E-09	LINEAR	Weighted-mean MO in tract posterior thalamic radiation (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25822	CONTINUOUS	30655	-0.03	-0.04	-0.02	4.51E-09	LINEAR	Volume of grey matter in Angular Gyrus (left)	UK Biobank Assessment Centre	Imaging	Imaging	1101
25468	CONTINUOUS	30236	0.03	0.02	0.04	5.75E-09	LINEAR	Mean ISOVF in posterior thalamic radiation on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25784	CONTINUOUS	30655	0.03	0.02	0.04	5.96E-09	LINEAR	Volume of grey matter in Insular Cortex (left)	UK Biobank Assessment Centre	Imaging	Imaging	1101
25736	CONTINUOUS	30192	-0.03	-0.04	-0.02	8.65E-09	LINEAR	Discrepancy between T2 FLAIR brain image and T1 brain image	UK Biobank Assessment Centre	Imaging	Brain MRI	112

Table S4. PheWAS results for Rs_DC_asymm. Only Bonferroni significant results are shown [continued]

varName	varType	n	beta	lower	upper	pvalue	resType	description	Cat1 Title	Cat2 Title	Cat3 Title	Category
25736	CONTINUOUS	30192	-0.03	-0.04	-0.02	8.65E-09	LINEAR	Discrepancy between T2 FLAIR brain image and T1 brain image	UK Biobank Assessment Centre	Imaging	Brain MRI	111
25582	CONTINUOUS	30236	0.03	0.02	0.04	3.5E-08	LINEAR	Weighted-mean L1 in tract inferior fronto-occipital fasciculus (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	106
25794	CONTINUOUS	30655	0.03	0.02	0.04	4.87E-08	LINEAR	Volume of grey matter in Precentral Gyrus (left)	UK Biobank Assessment Centre	Imaging	Imaging	107
25584	CONTINUOUS	30236	0.03	0.02	0.04	6.95E-08	LINEAR	Weighted-mean L1 in tract inferior longitudinal fasciculus (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	100033
25562	CONTINUOUS	30236	0.03	0.02	0.04	3.55E-07	LINEAR	Weighted-mean MO in tract posterior thalamic radiation (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	1101
25690	CONTINUOUS	30236	-0.03	-0.04	-0.02	4.94E-07	LINEAR	Weighted-mean OD in tract inferior fronto-occipital fasciculus (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	1101

25238	CONTINUOUS	30236	0.03	0.02	0.04	5.33E-07	LINEAR	Mean L1 in fornix cres+stria terminalis on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25717	CONTINUOUS	30236	0.03	0.02	0.04	6.89E-07	LINEAR	Weighted-mean ISOVF in tract inferior fronto-occipital fasciculus (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25866	CONTINUOUS	30655	0.03	0.02	0.04	8.36E-07	LINEAR	Volume of grey matter in Parietal Operculum Cortex (left)	UK Biobank Assessment Centre	Imaging	Imaging	135
25589	CONTINUOUS	30236	0.03	0.02	0.04	8.81E-07	LINEAR	Weighted-mean L1 in tract posterior thalamic radiation (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	135
25236	CONTINUOUS	30236	0.03	0.02	0.04	1.19E-06	LINEAR	Mean L1 in cingulum hippocampus on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	1101
25552	CONTINUOUS	30236	0.03	0.02	0.04	1.27E-06	LINEAR	Weighted-mean MO in tract forceps major	UK Biobank Assessment Centre	Imaging	Brain MRI	134
25239	CONTINUOUS	30236	0.03	0.02	0.04	1.48E-06	LINEAR	Mean L1 in fornix cres+stria terminalis on FA skeleton (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	1101

Table S4. PheWAS results for Rs_DC_asymm. Only Bonferroni significant results are shown [continued]

varName	varType	n	beta	lower	upper	pvalue	resType	description	Cat1 Title	Cat2 Title	Cat3 Title	Category
25726	CONTINUOUS	30236	0.03	0.02	0.04	1.83E-06	LINEAR	Weighted-mean ISOVF in tract superior longitudinal fasciculus (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	111
25228	CONTINUOUS	30236	0.03	0.02	0.04	2.78E-06	LINEAR	Mean L1 in posterior thalamic radiation on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	106
25472	CONTINUOUS	30236	0.03	0.01	0.04	7.37E-06	LINEAR	Mean ISOVF in external capsule on FA skeleton (right)	UK Biobank Assessment Centre	Imaging	Brain MRI	107
25444	CONTINUOUS	30236	0.03	0.01	0.04	7.89E-06	LINEAR	Mean ISOVF in splenium of corpus callosum on FA skeleton	UK Biobank Assessment Centre	Imaging	Brain MRI	100033
25588	CONTINUOUS	30236	0.02	0.01	0.03	1.18E-05	LINEAR	Weighted-mean L1 in tract posterior thalamic radiation (left)	UK Biobank Assessment Centre	Imaging	Brain MRI	1101
25789	CONTINUOUS	30655	0.02	0.01	0.03	1.27E-05	LINEAR	Volume of grey matter in Middle Frontal Gyrus (right)	UK Biobank Assessment Centre	Imaging	Imaging	1101
20075	INTEGER	30375	-0.02	-0.04	-0.01	1.29E-05	LINEAR	Home location at assessment - north co-ordinate (rounded)	UK Biobank Assessment Centre	Recruitment	Reception	100024

25238	20240	INTEGER	15352	0.06	0.03	0.09	1.35E-05	ORDERED-LOGISTIC	Maximum digits remembered correctly	Online follow-up	Cognitive function online	Numeric memory
25717	25768	CONTINUOUS	26737	0.03	0.01	0.04	1.53E-05	LINEAR	90th percentile of z-statistic (in group-defined amygdala activation mask) for faces-shapes contrast	UK Biobank Assessment Centre	Imaging	Brain MRI

Table S5. Top 10 GWAS loci for HICAS_pop.

chr	rsid	pos	a_0	a_1	af	info	beta	se	t	log10p	Z	P
9	rs35898523	136029645	G	T	0.066	1.000	0.084	0.016	5.270	6.862	5.270	1.38E-07
9	rs35902535	136029138	G	A	0.065	1.000	0.085	0.016	5.256	6.829	5.256	1.48E-07
5	rs79338527	113289063	G	A	0.022	0.996	0.134	0.028	4.862	5.931	4.861	1.17E-06
23	rs768114498	116559133	C	CGT	0.015	0.950	0.135	0.028	4.858	5.923	4.858	1.19E-06
18	rs149865802	46958052	G	A	0.016	0.982	0.161	0.033	4.858	5.923	4.858	1.19E-06
23	rs773824644	116559136	G	A	0.015	0.950	0.135	0.028	4.855	5.917	4.855	1.21E-06
3	rs35909324	184011515	C	CA	0.246	0.998	-0.045	0.009	-4.814	5.828	-4.814	1.49E-06
7	rs12533572	14646225	A	T	0.497	0.992	-0.039	0.008	-4.779	5.751	-4.778	1.77E-06
23	rs142008769	116571911	T	C	0.014	0.959	0.138	0.029	4.777	5.747	4.777	1.79E-06
17	rs62068973	27356949	G	C	0.232	0.999	-0.044	0.009	-4.712	5.608	-4.712	2.47E-06

Chr=chromosome;pos=position;a_0=non-effect allele; a_1=effect allele; af=allele frequency; se=standard error; t=t-value, Z=z-score,P =p-value.

Table S6. Top 10 GWAS loci for Rs_DC_asymm.

chr	rsid	pos	a_0	a_1	af	info	beta	se	t	log10p	Z	P
3	rs116534877	183285382	A	T	0.021	0.984	-0.159	0.030	-5.383	7.132	-5.383	7.38E-08
20	rs78787870	55170750	G	A	0.053	0.977	0.098	0.019	5.160	6.604	5.160	2.49E-07
10	rs471869	97326393	T	C	0.986	0.995	-0.180	0.035	-5.116	6.504	-5.116	3.14E-07
7	rs2103406	7339441	A	G	0.458	0.999	0.040	0.008	5.002	6.244	5.002	5.70E-07
3	rs71318325	183257513	A	G	0.025	0.988	-0.132	0.027	-4.942	6.109	-4.942	7.77E-07
6	rs9345345	94011129	A	G	0.138	0.995	0.058	0.012	4.922	6.065	4.922	8.61E-07
10	rs146809076	133202476	C	T	0.011	0.998	-0.187	0.038	-4.873	5.956	-4.873	1.11E-06
5	rs138360726	33439414	T	TG AG	0.092	0.997	-0.068	0.014	-4.837	5.878	-4.837	1.32E-06
10	rs114652501	133204904	C	T	0.011	0.998	-0.185	0.038	-4.833	5.870	-4.833	1.35E-06
6	6:43131347_A AAAAAG_A	43131347	AAAA AAG	A	0.129	0.973	0.062	0.013	4.796	5.788	4.796	1.63E-06

Table S7. GWAS results for HICAs_pop for those SNPs that had shown associations with brain structural asymmetry in Sha *et al.*.

chr	rsid	pos	a_0	a_1	af	info	beta	se	t	log10p	Z	P
1	rs6658111	47980916	T	G	0.36	1.00	0.02	0.01	2.62	2.06	2.62	**0.01**
2	rs62130503	27215973	C	T	0.05	1.00	-0.01	0.02	-0.63	0.27	-0.63	0.53
2	rs12617392	27336827	C	A	0.44	1.00	0.00	0.01	-0.47	0.19	-0.47	0.64
2	rs368536282	210615598	C	T	0.03	1.00	0.03	0.02	1.36	0.76	1.36	0.17
3	rs2279829	147106319	C	T	0.22	1.00	0.01	0.01	1.03	0.52	1.03	0.30
4	rs9307052	89896891	C	T	0.11	1.00	0.02	0.01	1.44	0.83	1.44	0.15
5	rs869219775	92941083	TTTG	T	0.14	1.00	0.01	0.01	1.02	0.51	1.02	0.31
6	rs7781	30693121	A	G	0.24	1.00	0.01	0.01	1.16	0.61	1.16	0.25
6	rs9385385	126122280	C	T	0.45	1.00	-0.01	0.01	-0.79	0.37	-0.79	0.43
7	rs6947352	33262550	G	A	0.31	1.00	-0.01	0.01	-1.04	0.52	-1.04	0.30
9	rs911934	101637477	T	G	0.70	1.00	0.00	0.01	0.00	0.00	0.00	1.00
10	rs41298373	7622009	G	A	0.10	1.00	0.01	0.01	0.39	0.16	0.39	0.70
12	rs10783306	49516946	G	C	0.33	1.00	-0.01	0.01	-1.34	0.74	-1.34	0.18
14	rs160459	59074136	A	C	0.46	1.00	0.00	0.01	-0.63	0.27	-0.63	**14**

14	rs201816193	59631075	GTTGT	G	0.12	1.00	0.01	0.01	0.70	0.31	0.70	0.49
16	rs72813426	89933007	A	G	0.24	1.00	-0.01	0.01	-1.27	0.69	-1.27	0.20
16	rs111398992	89992613	G	T	0.13	1.00	0.01	0.01	0.58	0.25	0.58	0.56
17	rs55938136	43798360	A	G	0.22	1.00	0.02	0.01	1.59	0.95	1.59	0.11
17	rs35908989	43994021	T	C	0.23	0.99	0.00	0.01	0.32	0.12	0.32	0.75
17	rs35853889	44104612	T	TG	0.19	1.00	-0.02	0.01	-1.85	1.19	-1.85	0.06
17	rs80103986	44224940	A	T	0.20	0.99	0.01	0.01	1.03	0.52	1.03	0.30
17	rs568039055	44632463	G	C	0.20	0.96	-0.02	0.01	-1.41	0.80	-1.41	0.16
19	rs11672092	6501998	C	T	0.22	0.98	0.00	0.01	0.04	0.01	0.04	0.97
20	rs6135555	15796237	C	A	0.39	1.00	0.01	0.01	1.78	1.12	1.78	0.08
21	rs7283026	46853259	T	C	0.27	1.00	-0.01	0.01	-1.21	0.64	-1.21	0.23
22	rs9615351	46760481	C	G	0.25	1.00	0.00	0.01	0.44	0.18	0.44	0.66
X	rs12400461	3336996	A	C	0.58	0.99	-0.01	0.01	-1.24	0.67	-1.24	0.21

Table S8. GWAS results for Rs_DC_asymm for those SNPs that had shown associations with brain structural asymmetry in Sha *et al*.

chr	rsid	pos	a_0	a_1	af	info	beta	se	t	log10p	Z	P
1	rs6658111	47980916	T	G	0.365	0.999	0.015	0.008	1.753	1.099	1.753	0.080
2	rs62130503	27215973	C	T	0.053	1.000	0.013	0.018	0.754	0.346	0.754	0.451
2	rs12617392	27336827	C	A	0.442	0.999	0.007	0.008	0.862	0.410	0.862	0.389
2	rs368536282	210615598	C	T	0.035	0.998	0.008	0.022	0.371	0.148	0.371	0.711
3	rs2279829	147106319	C	T	0.218	0.999	0.010	0.010	1.022	0.513	1.022	0.307
4	rs9307052	89896891	C	T	0.110	1.000	0.018	0.013	1.408	0.799	1.408	0.159
5	rs869219775	92941083	TTTG	T	0.14	1.00	-0.02	0.01	-1.51	0.88	-1.51	0.13
6	rs7781	30693121	A	G	0.240	0.995	0.021	0.010	2.220	1.578	2.220	**0.026**
6	rs9385385	126122280	C	T	0.451	0.997	-0.002	0.008	-0.230	0.087	-0.230	0.818
7	rs6947352	33262550	G	A	0.307	0.999	0.003	0.009	0.384	0.154	0.384	0.701
9	rs911934	101637477	T	G	0.699	0.999	0.000	0.009	0.024	0.009	0.024	0.980
10	rs41298373	7622009	G	A	0.101	1.000	0.009	0.013	0.661	0.294	0.661	0.508
12	rs10783306	49516946	G	C	0.333	1.000	-0.002	0.009	-0.256	0.098	-0.256	0.798

14	rs160459	59074136	A	C	0.464	1.000	-0.011	0.008	-1.365	0.764	-1.365	0.172
14	rs201816193	59631075	GTTGT	G	0.12	1.00	-0.01	0.01	-0.77	0.35	-0.77	0.44
16	rs72813426	89933007	A	G	0.240	0.995	0.007	0.009	0.695	0.313	0.695	0.487
16	rs111398992	89992613	G	T	0.134	0.998	-0.019	0.012	-1.637	0.993	-1.637	0.102
17	rs55938136	43798360	A	G	0.223	1.000	-0.011	0.010	-1.126	0.585	-1.126	0.260
17	rs35908989	43994021	T	C	0.231	0.991	0.018	0.010	1.796	1.139	1.796	0.073
17	rs35853889	44104612	T	TG	0.190	0.999	-0.018	0.010	-1.776	1.120	-1.775	0.076
17	rs80103986	44224940	A	T	0.20	0.99	-0.01	0.01	-1.17	0.62	-1.17	0.24
17	rs568039055	44632463	G	C	0.200	0.958	-0.013	0.011	-1.153	0.604	-1.153	0.249
19	rs11672092	6501998	C	T	0.224	0.982	-0.002	0.010	-0.188	0.070	-0.188	0.851
20	rs6135555	15796237	C	A	0.387	0.998	0.007	0.008	0.812	0.380	0.812	0.417
21	rs7283026	46853259	T	C	0.266	0.999	-0.008	0.009	-0.864	0.412	-0.864	0.387
22	rs9615351	46760481	C	G	0.252	0.998	0.005	0.009	0.590	0.255	0.590	0.555
23	rs12400461	3336996	A	C	0.584	0.993	-0.001	0.007	-0.146	0.054	-0.146	0.884

Table S9. Top 10 genes from MAGMA gene-based association analysis of HICAs_pop.

Entrez gene ID	hgnc_symbol	CHR	START	STOP	NSNPS	NPARAM	N	ZSTAT	P	BONF
83648	FAM167A	8	11228972	11374276	725	40	30660	4.3939	5.57E-06	0.11
10485	C1orf61	1	1.56E+08	1.56E+08	275	46	30660	4.3429	7.03E-06	0.14
640	BLK	8	11301521	11472108	764	36	30660	4.319	7.84E-06	0.15
640	BLK	8	11301521	11472108	764	36	30660	4.319	7.84E-06	0.15
1607	DGKB	7	14134674	14993281	2943	165	30660	4.0302	2.79E-05	0.54
57127	RHBG	1	1.56E+08	1.56E+08	213	37	30660	3.7232	9.84E-05	1.00
79660	PPP1R3B	8	8943764	9059152	414	40	30660	3.7169	0.000101	1.00
23274	CLEC16A	16	10988345	11326046	1063	49	30660	3.6385	0.000137	1.00
359	AQP2	12	50294524	50402664	245	46	30660	3.4654	0.000265	1.00
286046	XKR6	8	10703654	11108875	1491	44	30660	3.4637	0.000266	1.00

Table S10. Top 10 genes from MAGMA gene-based association analysis of Rs_DC_asymm.

Entrez gene ID	hgnc_symbol	CHR	START	STOP	NSNPS	NPARAM	N	ZSTAT	P	BONF
654231	OCM	7	5870429	5975994	564	46	30660	3.7322	0.00009	1.00
1896	EDA	X	68785911	69309322	1001	43	30660	3.7142	0.00010	1.00
27112	FAM155B	X	68675078	68802351	319	39	30660	3.6254	0.00014	1.00
89953	KLC4	6	42977372	43092833	264	22	30660	3.621	0.00015	1.00
5212	VIT	2	36873833	37091937	885	86	30660	3.5805	0.00017	1.00
5719	PSMD13	11	186808	302984	541	37	30660	3.5733	0.00018	1.00
63826	SRR	17	2157244	2278553	396	29	30660	3.5256	0.00021	1.00
55720	TSR1	17	2175972	2290678	362	31	30660	3.4134	0.00032	1.00
9820	CUL7	6	42955355	43071683	287	20	30660	3.402	0.00033	1.00
171389	NLRP6	11	228570	335388	538	58	30660	3.3252	0.00044	1.00

Table S11. Top 10 gene-sets from the MAGMA GO gene-set association analysis of HICAs_pop.

VARIABLE	TYPE	NGENES	BETA	BETA STD	SE	P	FDR
GO CAMERA TYPE EYE MORPHOGENESIS	SET	110	0.28	0.02	0.08	0.0003	0.80
GO EYE MORPHOGENESIS	SET	146	0.23	0.02	0.07	0.0007	0.80
GO POSITIVE REGULATION OF STRIATED MUSCLE CELL DIFFERENTIATION	SET	58	0.34	0.02	0.11	0.0008	0.80
GO POSITIVE REGULATION OF CARBOHYDRATE METABOLIC PROCESS	SET	63	0.35	0.02	0.11	0.0008	0.80
GO POSITIVE REGULATION OF SMALL MOLECULE METABOLIC PROCESS	SET	99	0.26	0.02	0.08	0.0010	0.80
GO CARBOHYDRATE CATION SYMPORTER ACTIVITY	SET	19	0.51	0.02	0.17	0.0013	0.80
GO SEX DIFFERENTIATION	SET	260	0.15	0.02	0.05	0.0015	0.80
GO POSITIVE REGULATION OF CELLULAR CARBOHYDRATE METABOLIC PROCESS	SET	44	0.39	0.02	0.13	0.0017	0.80
GO COENZYME BIOSYNTHETIC PROCESS	SET	143	0.20	0.02	0.07	0.0024	0.80
GO POSITIVE REGULATION OF ALCOHOL BIOSYNTHETIC PROCESS	SET	25	0.42	0.02	0.15	0.0028	0.80

Table S12. Top 10 gene-sets from the MAGMA GO gene-set association analysis of Rs_DC_asymm.

VARIABLE	TYPE	NGENES	BETA	BETA STD	SE	P	FDR
GO PLASMA MEMBRANE RAFT	SET	109	0.30	0.02	0.08	0.0001	0.28
GO POSITIVE REGULATION OF ACTIN CYTOSKELETON REORGANIZATION	SET	18	0.64	0.02	0.18	0.0002	0.53
GO RESPIRATORY BURST	SET	31	0.51	0.02	0.15	0.0003	0.53
GO VITAMIN TRANSMEMBRANE TRANSPORTER ACTIVITY	SET	24	0.66	0.02	0.20	0.0006	0.79
GO FAT SOLUBLE VITAMIN METABOLIC PROCESS	SET	43	0.34	0.02	0.11	0.0014	0.98
GO ADRENAL GLAND DEVELOPMENT	SET	25	0.46	0.02	0.16	0.0017	0.98
GO ARACHIDONIC ACID MONOOXYGENASE ACTIVITY	SET	15	0.82	0.02	0.28	0.0018	0.98
GO INTERLEUKIN 1 RECEPTOR BINDING	SET	15	0.79	0.02	0.27	0.0019	0.98
GO CAVEOLA	SET	81	0.27	0.02	0.09	0.0020	0.98
GO RESPONSE TO TYPE I INTERFERON	SET	94	0.24	0.02	0.09	0.0024	0.98

Table S13. Results of the candidate gene-set association analyses of HICAs_pop for GO microtubule (*top*), visceral laterality (*middle*) and handedness (*bottom*).

VARIABLE	TYPE	NGENES	BETA	BETA STD	SE	P[1]
GO microtubule						
GO_MICROTUBULE_PLUS_END_BINDING	SET	17	0.385	0.011	0.188	0.020
GO_MICROTUBULE_ANCHORING	SET	25	0.221	0.008	0.160	0.083
GO_MICROTUBULE_PLUS_END	SET	20	0.276	0.009	0.208	0.092
GO_NEGATIVE_REGULATION_OF_MICROTUBULE_POLYMERIZATION_OR_DEPOLYMERIZATION	SET	35	0.185	0.008	0.146	0.102
GO_REGULATION_OF_MICROTUBULE_POLYMERIZATION_OR_DEPOLYMERIZATION	SET	76	0.121	0.008	0.096	0.104
GO_MICROTUBULE_ORGANIZING_CENTER_ORGANIZATION	SET	133	0.078	0.006	0.071	0.137
GO_REGULATION_OF_MICROTUBULE_POLYMERIZATION	SET	45	0.132	0.006	0.127	0.149
GO_POSITIVE_REGULATION_OF_MICROTUBULE_POLYMERIZATION_OR_DEPOLYMERIZATION	SET	31	0.160	0.006	0.158	0.155
GO_MICROTUBULE_CYTOSKELETON_ORGANIZATION_INVOLVED_IN_MITOSIS	SET	125	0.065	0.005	0.074	0.192
GO_PROTEIN_TRANSPORT_ALONG_MICROTUBULE	SET	66	0.083	0.005	0.099	0.202
GO_MICROTUBULE_POLYMERIZATION	SET	71	0.074	0.004	0.097	0.224
GO_CYTOPLASMIC_MICROTUBULE	SET	65	0.077	0.004	0.105	0.232
GO_MICROTUBULE_END	SET	28	0.108	0.004	0.165	0.256
GO_NEGATIVE_REGULATION_OF_MICROTUBULE_DEPOLYMERIZATION	SET	24	0.103	0.004	0.169	0.271
GO_REGULATION_OF_MICROTUBULE_BASED_PROCESS	SET	211	0.033	0.003	0.057	0.281
GO_MICROTUBULE_ORGANIZING_CENTER_LOCALIZATION	SET	28	0.082	0.003	0.155	0.298
GO_CYTOPLASMIC_MICROTUBULE_ORGANIZATION	SET	53	0.049	0.003	0.117	0.339

GO_MICROTUBULE_NUCLEATION	SET	26	0.059	0.002	0.153	0.350
GO_REGULATION_OF_MICROTUBULE_DEPOLYMERIZATION	SET	29	0.052	0.002	0.147	0.362
GO_ATTACHMENT_OF_SPINDLE_MICROTUBULES_TO_KINETOCHORE	SET	32	0.042	0.002	0.138	0.380
GO_PROTEIN_LOCALIZATION_TO_MICROTUBULE_ORGANIZING_CENTER	SET	31	0.004	0.000	0.136	0.488
GO_MICROTUBULE_BUNDLE_FORMATION	SET	89	0.003	0.000	0.090	0.489
GO_ATP_DEPENDENT_MICROTUBULE_MOTOR_ACTIVITY_MINUS_END_DIRECTED	SET	18	0.004	0.000	0.188	0.492
GO_MICROTUBULE_POLYMERIZATION_OR_DEPOLYMERIZATION	SET	109	0.001	0.000	0.080	0.495
GO_PROTEIN_LOCALIZATION_TO_MICROTUBULE	SET	16	- 0.015	0.000	0.200	0.530
GO_MICROTUBULE_ORGANIZING_CENTER_PART	SET	179	- 0.014	- 0.001	0.059	0.595
GO_REGULATION_OF_MICROTUBULE_BASED_MOVEMENT	SET	26	- 0.054	- 0.002	0.183	0.615
GO_MICROTUBULE_BASED_MOVEMENT	SET	274	- 0.014	- 0.002	0.049	0.617
GO_ATP_DEPENDENT_MICROTUBULE_MOTOR_ACTIVITY	SET	43	- 0.069	- 0.003	0.132	0.699
GO_ATP_DEPENDENT_MICROTUBULE_MOTOR_ACTIVITY_PLUS_END_DIRECTED	SET	25	- 0.141	- 0.005	0.186	0.775
GO_MICROTUBULE_MOTOR_ACTIVITY	SET	83	- 0.072	- 0.005	0.090	0.789
GO_MICROTUBULE_DEPOLYMERIZATION	SET	44	- 0.101	- 0.005	0.123	0.795
GO_ORGANELLE_TRANSPORT_ALONG_MICROTUBULE	SET	79	- 0.082	- 0.005	0.090	0.819

Table S13. Results of the candidate gene-set association analyses of HICAs_pop for GO microtubule (*top*), visceral laterality (*middle*) and handedness (*bottom*) [continued]

VARIABLE	TYPE	NGENES	BETA	BETA STD	SE	P[1]
GO_SPINDLE_MICROTUBULE	SET	54	-0.13	-0.007	0.113	0.868
GO_MICROTUBULE_ASSOCIATED_COMPLEX	SET	148	-0.09	-0.008	0.070	0.893
GO_MICROTUBULE_BINDING	SET	235	-0.08	-0.009	0.053	0.935
GO_MICROTUBULE	SET	406	-0.07	-0.009	0.041	0.945
Visceral laterality						
MP:0000531	SET	21	0.24	0.008	0.179	0.092
MP:0000508	SET	26	0.13	0.005	0.155	0.208
MP:0001706	SET	72	0.05	0.003	0.093	0.287
MP:0002766	SET	45	0.04	0.002	0.119	0.384
PCD	SET	34	0.01	0.001	0.153	0.464
MP:0011252	SET	22	-0.02	-0.001	0.179	0.542
MP:0010808	SET	25	-0.04	-0.001	0.156	0.591
MP:0000644	SET	37	-0.06	-0.003	0.135	0.664
MP:0003178	SET	32	-0.06	-0.003	0.146	0.670
MP:0004133	SET	28	-0.07	-0.003	0.154	0.675
MP:0011250	SET	15	-0.10	-0.003	0.182	0.701
MP:0000542	SET	35	-0.11	-0.005	0.139	
Handedness						
Partida	SET	36	0.022	0.001	0.121	0.429

[1] P-value are uncorrected and sorted from lowest to highest per category. None of these survived multiple testing correction for the number of tests per candidate category

Table S14. Results of the candidate gene-set association analyses of Rs_DC_asymm for GO microtubule (*top*), visceral laterality (*middle*) and handedness (*bottom*).

VARIABLE	TYPE	NGENES	BETA	BETA STD	SE	P[1]
GO microtubule						
GO_MICROTUBULE_ORGANIZING_CENTER_LOCALIZATION	SET	28	0.261	0.010	0.154	0.045
GO_MICROTUBULE_PLUS_END_BINDING	SET	17	0.308	0.009	0.186	0.049
GO_MICROTUBULE_BINDING	SET	235	0.085	0.009	0.053	0.053
GO_MICROTUBULE_BUNDLE_FORMATION	SET	89	0.118	0.008	0.090	0.095
GO_ATP_DEPENDENT_MICROTUBULE_MOTOR_ACTIVITY_MINUS_END_DIRECTED	SET	18	0.240	0.007	0.186	0.099
GO_MICROTUBULE	SET	406	0.046	0.007	0.040	0.128
GO_NEGATIVE_REGULATION_OF_MICROTUBULE_DEPOLYMERIZATION	SET	24	0.179	0.006	0.168	0.143
GO_MICROTUBULE_MOTOR_ACTIVITY	SET	83	0.085	0.006	0.089	0.170
GO_ATP_DEPENDENT_MICROTUBULE_MOTOR_ACTIVITY	SET	43	0.106	0.005	0.131	0.209
GO_MICROTUBULE_ANCHORING	SET	25	0.112	0.004	0.158	0.240
GO_REGULATION_OF_MICROTUBULE_DEPOLYMERIZATION	SET	29	0.091	0.004	0.146	0.266
GO_NEGATIVE_REGULATION_OF_MICROTUBULE_POLYMERIZATION_OR_DEPOLYMERIZATION	SET	35	0.087	0.004	0.145	0.275
GO_MICROTUBULE_DEPOLYMERIZATION	SET	44	0.071	0.003	0.122	0.281
GO_MICROTUBULE_CYTOSKELETON_ORGANIZATION_INVOLVED_IN_MITOSIS	SET	125	0.024	0.002	0.074	0.375
GO_MICROTUBULE_END	SET	28	0.037	0.001	0.164	0.410
GO_POSITIVE_REGULATION_OF_MICROTUBULE_POLYMERIZATION_OR_DEPOLYMERIZATION	SET	31	0.029	0.001	0.157	0.427
GO_SPINDLE_MICROTUBULE	SET	54	0.017	0.001	0.112	0.440
GO_MICROTUBULE_ASSOCIATED_COMPLEX	SET	148	0.009	0.001	0.070	0.448

Table S14. Results of the candidate gene-set association analyses of Rs_DC_asymm for GO microtubule (*top*), visceral laterality (*middle*) and handedness (*bottom*) [continued]

VARIABLE	TYPE	NGENES	BETA	BETA STD	SE	P[1]
GO_ORGANELLE_TRANSPORT_ALONG_MICROTUBULE	SET	79	0.010	0.001	0.089	0.457
GO_MICROTUBULE_POLYMERIZATION_OR_DEPOLYMERIZATION	SET	109	0.000	0.000	0.079	0.501
GO_MICROTUBULE_ORGANIZING_CENTER_ORGANIZATION	SET	133	-0.005	0.000	0.071	0.525
GO_PROTEIN_LOCALIZATION_TO_MICROTUBULE_ORGANIZING_CENTER	SET	31	-0.011	0.000	0.136	0.533
GO_REGULATION_OF_MICROTUBULE_POLYMERIZATION	SET	45	-0.014	-0.001	0.126	0.543
GO_MICROTUBULE_POLYMERIZATION	SET	71	-0.011	-0.001	0.097	0.544
GO_ATP_DEPENDENT_MICROTUBULE_MOTOR_ACTIVITY_PLUS_END_DIRECTED	SET	25	-0.026	-0.001	0.185	0.555
GO_REGULATION_OF_MICROTUBULE_BASED_PROCESS	SET	211	-0.012	-0.001	0.057	0.582
GO_REGULATION_OF_MICROTUBULE_POLYMERIZATION_OR_DEPOLYMERIZATION	SET	76	-0.021	-0.001	0.095	0.588
GO_PROTEIN_LOCALIZATION_TO_MICROTUBULE	SET	16	-0.052	-0.001	0.198	0.603
GO_CYTOPLASMIC_MICROTUBULE	SET	65	-0.032	-0.002	0.105	0.622
GO_MICROTUBULE_PLUS_END	SET	20	-0.075	-0.002	0.206	0.641
GO_REGULATION_OF_MICROTUBULE_BASED_MOVEMENT	SET	26	-0.068	-0.003	0.182	0.646
GO_MICROTUBULE_BASED_MOVEMENT	SET	274	-0.026	-0.003	0.048	0.702
GO_MICROTUBULE_ORGANIZING_CENTER_PART	SET	179	-0.040	-0.004	0.058	0.756
GO_CYTOPLASMIC_MICROTUBULE_ORGANIZATION	SET	53	-0.081	-0.004	0.116	0.757
GO_ATTACHMENT_OF_SPINDLE_MICROTUBULES_TO_KINETOCHORE	SET	32	-0.105	-0.004	0.137	0.779
GO_MICROTUBULE_NUCLEATION	SET	26	-0.219	-0.008	0.152	0.925
GO_PROTEIN_TRANSPORT_ALONG_MICROTUBULE	SET	66	-0.164	-0.010	0.098	0.953

Visceral latearlity						
MP:0003178	SET	32	0.118	0.005	0.145	0.208
MP:0004133	SET	28	0.116	0.004	0.153	0.224
MP:0000542	SET	35	0.063	0.003	0.138	0.323
MP:0011250	SET	15	-0.012	0.000	0.181	0.525
MP:0000644	SET	37	-0.051	-0.002	0.134	0.649
MP:0001706	SET	72	-0.036	-0.002	0.093	0.653
MP:0000531	SET	21	-0.082	-0.003	0.178	0.677
PCD	SET	34	-0.078	-0.003	0.151	0.696
MP:0011252	SET	22	-0.095	-0.003	0.177	0.704
MP:0002766	SET	45	-0.098	-0.005	0.118	0.797
MP:0010808	SET	25	-0.163	-0.006	0.155	0.854
MP:0000508	SET	26	-0.169	-0.006	0.154	0.864
Handedness						
Partida	SET	36	0.064	0.003	0.120	0.296

[1] P-value are uncorrected and sorted from lowest to highest per category. None of these survived multiple testing correction for the number of tests per candidate category.

Table S15. Top 10 TWAS genes for HICAS_pop based on expression data from GteX Brain Frontal Cortex BA9.

Gene name	z-score	Effect size	p-value	FDR	Var g	Pred perf r2	Pred perf pval	N snps used	N snps in cov	N snps in model
ZNF462	3.937	0.071	0.0001	0.15	0.096	0.109	0.0007	25	26	26
RP11-481A20.4	-3.930	-0.135	0.0001	0.15	0.027	0.068	0.0077	15	15	15
RP11-481A20.11	-3.683	-0.104	0.0002	0.23	0.040	0.156	0.0001	7	7	7
ARSA	-3.659	-0.049	0.0003	0.23	0.124	0.173	0.00002	6	6	6
RP11-351I21.6	-3.557	-0.050	0.0004	0.27	0.160	0.155	0.0001	24	25	25
RP11-481A20.10	-3.394	-0.037	0.0007	0.41	0.276	0.373	1.21E-11	28	28	28
DUS2	-3.217	-0.054	0.0013	0.67	0.109	0.148	0.0001	30	30	30
PCOLCE2	3.111	0.064	0.0019	0.84	0.076	0.131	0.0003	17	17	17
CDK10	3.057	0.070	0.0022	0.87	0.048	0.057	0.0187	8	8	8
NKD2	-3.032	-0.066	0.0024	0.87	0.068	0.175	0.00002	16	16	16

Table S16. Top 10 TWAS genes for HICAS_pop based on expression data from GteX Brain Cortex.

Gene name	z-score	Effect size	p-value	FDR	Var g	Pred perf r2	Pred perf pval	N snps used	N snps in cov	N snps in model
RP11-351I21.7	-3.868	-0.082	0.0001	0.27	0.074	0.051	0.01913	18	18	18
TDH	3.708	0.097	0.0002	0.27	0.046	0.121	0.00018	9	9	9
ALG1L11P	-3.685	-0.049	0.0002	0.27	0.182	0.056	0.01317	51	52	52
FAM90A2P	-3.655	-0.109	0.0003	0.27	0.036	0.120	0.00017	13	15	15
PGR	-3.601	-0.121	0.0003	0.27	0.030	0.050	0.01852	16	16	16
ARSA	-3.471	-0.068	0.0005	0.37	0.074	0.074	0.00286	13	13	13
ADAT2	3.280	0.105	0.0010	0.62	0.029	0.038	0.03830	8	8	8
TMEM150C	-3.212	-0.074	0.0013	0.62	0.056	0.116	0.00025	5	5	5
RP11-351I21.6	-3.165	-0.039	0.0016	0.62	0.217	0.195	1.05E-06	41	43	43
ZNF497	3.161	0.037	0.0016	0.62	0.245	0.168	3.59E-06	45	45	45

Table S17. Top 10 TWAS genes for HICAS_pop based on expression data from GteX Whole Blood.

Gene name	z-score	Effect size	p-value	FDR	Var g	Pred perf r2	Pred perf pval	N snps used	N snps in cov	N snps in model
INSC	3.772	0.066	0.0002	0.650	0.101	0.164	7.70E-14	37	37	37
HIST1H4H	-3.668	-0.097	0.0002	0.650	0.042	0.051	0.0001	27	27	27
FAM167A	-3.555	-0.054	0.0004	0.650	0.134	0.228	1.73E-19	14	14	14
AHI1	-3.367	-0.052	0.0008	0.650	0.131	0.201	3.01E-17	18	19	19
SLC45A4	3.320	0.054	0.0009	0.650	0.120	0.235	2.26E-20	9	9	9
ENDOG	3.302	0.130	0.0010	0.650	0.020	0.060	1.05E-05	12	13	13
PDLIM4	3.287	0.108	0.0010	0.650	0.027	0.024	0.0059	11	11	11
PNRC1	3.265	0.077	0.0011	0.650	0.060	0.055	3.23E-05	37	37	37
RPL14	3.228	0.054	0.0012	0.650	0.122	0.163	5.92E-14	33	34	34
EVI2A	-3.213	-0.082	0.0013	0.650	0.045	0.087	1.08E-07	24	26	26

Table S18. Top 10 TWAS genes for Rs_DC asymmetry based on expression data from GteX Brain Frontal Cortex BA9.

Gene name	z-score	Effect size	p-value	FDR	Var g	Pred perf r2	Pred perf pval	N snps used	N snps in cov	N snps in model
PDCL3	-3.378	-0.084	0.0007	0.990	0.062	0.076	0.006	27	27	27
ERAP1	3.262	0.034	0.0011	0.990	0.283	0.415	0.000	30	43	43
FAM180B	3.157	0.058	0.0016	0.990	0.080	0.140	0.000	11	12	12
STAU1	3.133	0.101	0.0017	0.990	0.030	0.110	0.001	31	31	31
IFITM4P	3.093	0.049	0.0020	0.990	0.120	0.119	0.000	37	37	37
PBLD	-3.074	-0.101	0.0021	0.990	0.033	0.055	0.022	16	16	16
AC108868.6	3.007	0.046	0.0026	0.990	0.122	0.051	0.027	52	53	53
GNPDA2	-2.998	-0.064	0.0027	0.990	0.074	0.158	0.000	17	17	17
NID1	2.991	0.101	0.0028	0.990	0.026	0.059	0.019	12	12	12
BLCAP	-2.977	-0.030	0.0029	0.990	0.360	0.101	0.002	65	69	69

Table S19. Top 15 TWAS genes for Rs_DC asymmetry based on expression data from GteX Brain Cortex.

Gene name	z-score	Effect size	p-value	FDR	Var g	Pred perf r2	Pred perf pval	N snps used	N snps in cov	N snps in model
ZNF784	3.544	0.093	0.0004	0.992	0.046	0.044	0.0248	25	25	25
RP11-347I19.8	3.286	0.054	0.0010	0.992	0.135	0.167	0.0000	20	20	20
TNNI3K	3.140	0.044	0.0017	0.992	0.158	0.194	0.0000	97	97	97
RP11-567C2.1	-3.130	-0.091	0.0017	0.992	0.037	0.048	0.0184	22	22	22
ACAD9	-3.076	-1.158	0.0021	0.992	0.000	0.095	0.0000	2	16	16
ENTPD3	-3.034	-0.062	0.0024	0.992	0.092	0.088	0.0008	23	24	24
FLCN	3.029	0.047	0.0024	0.992	0.123	0.171	0.0000	22	22	22
GPALPP1	-3.005	-0.070	0.0027	0.992	0.074	0.094	0.0012	22	22	22
DUSP11	2.992	0.059	0.0028	0.992	0.093	0.082	0.0020	21	32	32
STOML1	-2.979	-0.091	0.0029	0.992	0.032	0.117	0.0002	5	5	5

Supplementary Table 20. Top 10 TWAS genes for Rs_DC asymmetry based on expression data from GteX Whole Blood.

Gene name	z-score	Effect size	p-value	FDR	Var g	Pred perf r2	Pred perf pval	N snps used	N snps in cov	N snps in model
CUL7	4.276	0.071	0.00002	0.119	0.098	0.171	1.65E-14	24	24	24
UGT2B17	3.540	0.105	0.0004	0.865	0.036	0.082	1.95E-07	21	21	21
VN1R1	3.468	0.207	0.0005	0.865	0.009	0.026	4.43E-03	1	1	1
SDCBP	3.397	0.187	0.0007	0.865	0.013	0.035	0.001	5	5	5
HLA-S	3.342	0.033	0.0008	0.865	0.316	0.368	1.81E-34	59	60	60
MYBPC3	-3.337	-0.054	0.0008	0.865	0.134	0.164	5.98E-14	38	38	38
MYO16	3.250	0.074	0.0012	0.865	0.063	0.144	1.79E-12	8	8	8
SLC39A13	3.244	0.161	0.0012	0.865	0.011	0.029	0.002	9	9	9
EFCAB14	3.184	0.066	0.0015	0.865	0.069	0.134	2.72E-11	31	31	31
SLC18B1	-3.135	-0.035	0.0017	0.865	0.273	0.332	1.96E-31	36	38	38

CHAPTER

General discussion 6

Left-right asymmetry is a key organizing feature of the human brain, with relevance to higher order cognitive functioning and the evolution of our species. Yet little is known about the genetic basis of brain asymmetry, or its structural and functional changes in people with cognitive disorders. Two of the studies in this thesis aimed to identify genes involved in human brain asymmetry, one of which also explored potential genetic links between brain and body asymmetries. The two other studies of this thesis aimed to measure the extent of altered brain anatomical asymmetry in attention-deficit/hyperactivity disorder (ADHD) and autism spectrum disorder (ASD). All of the studies were based on the largest collections of individuals ever analyzed in relation to these questions, and involved a wide range of different statistical and bioinformatic analyses.

In this final chapter, common themes arising from the studies will be discussed with respect to previous literature and implications for future research.

6.1 Biological insights through small effect sizes

Chapter 2 described how various aspects of brain structural asymmetry are subtly altered in ASD, particularly regional cortical thickness asymmetries. Chapter 3 described a similar study of ADHD, in which alterations of brain asymmetry were similarly subtle. Indeed, when looking at the range of the Cohen's *d* effects reported in chapters 2-3, it is clear that there is largely overlapping variation between cases and controls. These small effects are likely to be accurately measured, given the large statistical power of these

studies (Munafo & Flint, 2010). Chapter 4 reported that more than half of people with non-PCD *situs inversus* (SI), including most of those with left-handedness, did not have obvious monogenic mutations that could explain their condition. Finally, chapter 5 showed that only around 3% of variation in the hemispheric asymmetry of functional connectivity was influenced by common genetic variation.

Thus it is clear that, in each of the studies in this thesis, the majority of variance in brain or behavioural asymmetry remained unexplained by the effects tested. For example, considered in isolation, altered asymmetry is unlikely to be a useful biomarker for ASD or ADHD. Nonetheless, the observed effects, despite their small sizes, can be informative about the biology of the healthy and disordered brain. Subtle alterations of structural brain asymmetry in neurodevelopmental disorders indicates that an aspect of brain organization that is important for many functions is disrupted in these conditions. Magnetic resonance imaging (MRI)-based measures of macro-structural asymmetry (0.5 mm - 200 mm) are relatively coarse biological readouts, but slightly altered asymmetry at the macro-level may point to finer-scale structural and functional disruption in these disorders, for example in the hemispheric asymmetry of specific neuronal types, gene expression levels, or circuit-level connectivity. It is largely unknown how such deeper levels of organization relate to variation in macro-structure, for example in terms of cortical thickness asymmetries. Therefore it remains possible that subtle alterations of asymmetrical macro-structure indicate substantial rearrangements of asymmetrical micro-structural and functional organization (100 μm – 1 μm), and further studies in this direction are therefore motivated by the findings of chapters 2 and 3.

Limited genetic effects on the variability of brain and behavioural asymmetries are broadly consistent with developmental randomness being the primary driver of trait variance. Previous studies have also pointed to generally low heritabilities for various structural, functional and behavioural asymmetry measures, as well as limited effects of the environmental factors studied so far, even when considered in combination (see Introduction). The development of population-level brain asymmetry is most likely regulated by

a genetically encoded program, which can bias towards asymmetry in a certain direction, on average (Francks, 2015; Kong, Postema, et al., 2020). Yet it appears that deviations from the default, asymmetrical course of development occur largely without a contribution from genetic or environmental variation, which leaves random events in early embryonic development as a likely source of trait variance (Kong, Postema, et al., 2020). Note that such random effects are measured as 'non-shared environment' in twin studies. It may only require small variations of early developmental processes, such as lateralized distributions of embryonic signaling molecules, rates of cell growth or cell division, or cell-cell communication (Lens, Van Dongen, Kark, & Matthysen, 2002; Palmer, 1994), to affect the adult outcome.

Early 'developmental noise' can become reinforced through downstream gene expression cascades to affect the adult phenotype (Palmer, 1994). Primary Ciliary Dyskinesia (PCD) gives an example of this, whereby half of people with this condition also have *situs inversus* (Leigh et al., 2009), and similar randomization is seen in *iv-/iv-* mutant mice (Layton, 1976) (see Introduction). These observations indicate that, once ciliary function is disrupted due to a PCD mutation, it becomes a matter of chance whether the visceral organs develop in the typical left-right pattern. The eventual direction of asymmetry is likely determined by minor, random fluctuations in the initial concentrations of signaling molecules. This idea is further supported by the existence of monozygotic twin pairs discordant for *situs inversus* (Segal, 2017). Additionally, a recent study in zebrafish (Moreno-Ayala, Olivares-Chauvet, Schäfer, & Junker, 2021) reported an association between laterality defects and the number of embryonic dorsal forerunner cells, which are precursors of Kupffer's vesicle, i.e. the developmental origin of left-right visceral and brain asymmetry in the fish. Around 50% of the embryos with less than 20 dorsal forerunner cells exhibited reversed heart laterality, again showing that profound phenotypic rearrangements on the left-right axis can arise from early developmental variation. Another example is that the timing of splitting of an embryo during the formation of monozygotic twins can induce differences of laterality between twins, even though the genetic makeup is the same (Scott, 2002).

The biological relevance of small effects has already been widely acknowledged in the field of human genetics, where complex traits are known to be influenced by multiple common variants with small effects (Manolio et al., 2009), but which can combine for substantial overall polygenic effects, and/or implicate specific biological pathways. However, researchers in the field of neuroimaging have only recently started to appreciate that many effects, such as case-control disorder associations, or effects of individual genetic loci, are likely to be much smaller than first anticipated. There were at least 300 neuroimaging studies published between 2017 and 2018 that had sample sizes below 25 (Szucs & Ioannidis, 2020), suggesting that the importance of adequate statistical power in relation to small effects is not yet widely acknowledged in the field. Indeed, the effect sizes reported by many such studies are likely to be exaggerated, if not false positives (Szucs & Ioannidis, 2020). More recent, large-scale studies of brain structure and function by consortia such as ENIGMA have shown that individual effects are rarely large enough to have been detected in the previous generation of smaller studies, whether with regard to brain asymmetry or otherwise (Button et al., 2013; Thompson et al., 2020).

6.2 Issues in the analysis of big data

To date, various efforts have been realized to collect big data in the field of human neuro-genomics, including large-scale international efforts such as the Cohorts for Heart and Ageing Research in Genomic Epidemiology (CHARGE) consortium (Psaty et al., 2009), Psychiatric Genomics Consortium (PGC) (Sullivan et al., 2018) and ENIGMA, and large population-based datasets such as the Chinese Imaging Genetics (CHIMGEN) cohort (Xu et al., 2020) and UK Biobank. Two of the studies in this thesis were based on ENIGMA consortium data, and one was based on UK Biobank data. Even the study of SIT, which included only 15 affected individuals, involved big data in the sense of whole-genome-level sequence coverage, comprising millions of genetic variants.

According to Oxford's learning dictionary (www.oxfordlearnersdictionaries.com), big data is defined as 'sets of

information that are too large or too complex to handle, analyse or use with standard methods'. Three main features, denoted as the '3V's', are generally used to describe big data, namely *volume* (for the amount of data; measured in gigabytes, terabytes, petabytes, exabytes, zettabytes, or even yottabytes), *variety* (for the diversity of data and data structures), and *velocity* (for the speed with which data are generated) (Kinkorova & Topolcan, 2020).

As regards the data analyzed in this thesis, the three Vs apply most directly to the UK Biobank imaging genetics dataset, which included high-dimensional, multimodal brain image data, genome-wide data for millions of genetic variants, and thousands of additional variables used for phenome-wide association scanning, for each of tens of thousands of participants. In total the dataset comprised almost 100 terabytes, and data collection has been on-going for several years. While the ENIGMA data were not voluminous in the sense that the data took up a lot of disk space, they were available for thousands of individuals and were collected over roughly two decades. Even the whole genome sequence data from 30 individuals in Chapter 4 can be considered big with respect to the first V, 'various', in the sense that millions of genetic variants were measured.

6.2.1 Quality versus quantity

Four features that are central to data quality are accuracy (a low rate of entry errors), validity (based on reliable measurement), completeness (a low rate of missing values), and availability (accessibility to potential users) (Kilkenny & Robinson, 2018).

An ideal large-scale data collection would maintain the same level of quality and depth of measurement as is possible through focused efforts in smaller datasets, but sometimes the collection of large-scale data necessarily comes at the expense of some aspects of quality. In principle, a large sample size can at least partly compensate for a high level of measurement error when testing for effects. An example of this was a GWAS study of smoking quantity in roughly 40,000 individuals (Liu et al., 2010), where the phenotype was imperfect, i.e. self-report of cigarettes smoked per day, by recollection over past decades of life. Nonetheless, the sample size enabled statistically

unambiguous detection of a small-effect locus that implicated a cluster of genes encoding nicotinic-cholinergic receptor subunits.

On the other hand, if data are too noisy, results will be unreliable regardless of the sample size. Careful quality control procedures therefore remain important, insofar as they are feasible to carry out in large-scale data. In terms of genetic data, various thresholds were defined in this thesis to exclude poor quality variants from analysis, based on technical parameters related to the method of genotyping, or else statistical parameters such as imputation confidence or deviation from Hardy-Weinberg equilibrium. For neuroimaging data, the ENIGMA quality control protocol first involved the identification of outliers, as these were most likely to arise from measurement error, followed by visual re-checking and possible removal of specific data points. This approach was feasible for contributing sites, whereas detailed checking of every image may not have been. Outlier removal was also employed with respect to numerous variables of potential relevance to functional connectivity in the UK biobank data (Chapter 5), whereas again, visual inspection of more than 30,000 resting state images throughout the different stages of the processing pipeline would not have been feasible within the constraint of researcher time available. Another strategy, not pursued in this thesis, could be to visually check a practically feasible proportion of images, in order to estimate error frequencies without regard to outlier data points. However, if the bulk of the dataset remains unchecked, then it is unclear how to apply error frequency rates derived from a minority subset.

Another important quality issue for neuroimaging data – which is particularly crucial for laterality studies – is the possibility for left-right flipping of data held within NIFTI files (i.e., standard format files containing MRI scans). Left-right flipping can sometimes occur during the conversion from DICOM (raw data) to NIFTI, when done using non-standard data processing scripts. Fortunately, the large majority of groups rely on standard, validated software for this conversion (Kong et al., 2018). When both DICOM and NIFTI scans are available, the files can be compared to detect differences. However, when the DICOM is not available, comparison with

NIFTI is not possible, and there is no way to be certain about flipping, unless a fiducial marker was used (a small liquid-filled object placed on one side of the head during scanning). However, reversed population-level lateralities of some specific structural measures, such as lateral ventricle volumes, may be useful indicators of flipping at the level of entire individual datasets (Dick Schijven, personal communication), as most individuals show the same direction of laterality for such measures.

6.2.2 Investment in big data

Analyzing big data takes time, resources, and advanced computational and statistical knowledge. Computers and servers need to have sufficient processing speed and storage memory in order to support practicable analysis of big data, although for multi-site consortium studies the computational load can be partly divided across different sites. Processing speed can be increased by building methods that are highly efficient and scalable (e.g. BGENIE (Bycroft et al., 2018), a program for GWAS analyses that was designed and optimized to handle biobank-scale data).

Storage memory can be extended by moving data to the cloud (Kinkorova & Topolcan, 2020), although this was not possible for the present thesis due to restrictions defined within the relevant collaboration agreements or data transfer agreements, based on concerns about data security and safety. The security and safety of data is pivotal to ensure health privacy, which can be violated, for example, when data fall into the wrong hands, or when the frequency or purpose of data access is unfitting (Price & Cohen, 2019). Violation of health data privacy can potentially impact a person's life negatively (e.g. insurance premium increases, employment discrimination and/or emotional distress). While the importance of data security and safety also holds for 'small data', privacy violations of big data can impact a much larger number of individuals, and therefore potentially have a larger severity of effects (Price & Cohen, 2019).

Another general issue for big data is the questionable value of continued increases in sample sizes, with respect to possibly diminishing returns. With respect to any given biological effect, it would be a waste of time, money and

resources to continue collecting data to measure that effect beyond the point at which it has already been reliably detected, measured and replicated to a high level of confidence (Nayak, 2010). It may also be unethical to recruit more individuals than needed to measure a given effect, because it requires people to unduly give up their time to participate. However, in practice, Biobanks or consortia are designed to facilitate many biological investigations from the same overall set of individuals, and to identify novel dependencies between different types of variables, such that these concerns are of limited relevance to e.g. the UK Biobank or ENIGMA consortium.

Questions may also be raised about the utility of discovering ever smaller effects. Big data permit the identification of small effects that are statistically significant, but which may sometimes be insignificant in terms of biological insight or practical impact (Smith & Nichols, 2018) (see next section for more explanation). This latter point has already been mentioned above with respect to the findings in this thesis (see section 1 'Biological insights through small effect sizes'). In general, neuroimaging genetics studies are still at a point where larger sample sizes will likely have further value. Consider, for example, that the SNP-based heritability of functional connectivity asymmetry is only 3% (Chapter 5), and that roughly 30,000 individuals did not permit the reliable identification of individual SNPs that contribute to this heritability. The value of such discoveries will not be in terms of their effect sizes, but rather in case they implicate specific biological processes, which could be central in establishing the typical pattern of brain asymmetry in the majority of people.

6.2.3 Testing for effects in big data

In this thesis I have relied on hypothesis testing using p-values (often with FDR to correct for multiple testing). P-values can provide a convenient and objective way of interpreting data (Halsey, 2019), when used together with appropriate multiple testing correction, and in the absence of selective reporting. However, it should be acknowledged that p-values have limitations and can be easily misinterpreted (Halsey, 2019). For example, use of the p-value is often based on an arbitrary threshold of 0.05, and can be misused through P-hacking (i.e., repeatedly testing in different ways until significant

effects are found in the data) and selective reporting (i.e., only reporting significant effects) (Jafari & Ansari-Pour, 2019). Importantly, p-values are only informative with respect to the null hypothesis. Not being able to reject the null hypothesis due to a non-significant effect often leaves researchers with an 'open verdict', and when the power is high in large datasets, p-values will often be significant regardless of minimal effect sizes (Halsey, 2019).
An alternative approach to using the p-value, that can be suitable in large datasets, is to focus instead on effect sizes and their 95% confidence intervals. This approach provides information about the strengths of effects and how accurately they have been measured. Thus, for example, I have reported all effects in the large-scale analyses of ADHD and ASD in chapters 2 and 3, regardless of their significance in p-value-based testing.

Another alternative is to calculate the Bayes facto, i.e. a ratio that describes the relative likelihoods that the null and alternative hypotheses are true. However, the Bayes factor often incorporates a degree of subjectivity, as it requires the specification of the prior probability for an effect, which is based on existing information that is often imperfect. I have mostly not used Bayesian statistics in the present thesis due to this limitation. For example, prior literature on disorder case-control differences in brain asymmetry have been based on much smaller datasets than those used here, as well as mostly different brain atlases, so that there was little solid basis on which to specify prior probabilities. Likewise, no prior probabilities could be reliably specified for candidate genes that might affect variance in different aspects of brain asymmetry. However, in chapter 2, I used the Bayesian information criterion (BIC), as well as the Akaike information criterion (AIC), to compare between linear mixed models with different confound variables when assessing the effect of ASD diagnosis on brain asymmetry. These penalized-likelihood criteria can be used as an alternative to methods that rely on p-values. That is to say, rather than selecting the best model via a stepwise procedure that involves p-values, the AIC and/or BIC can be calculated to provide an estimate of model fit. The lower the value, the better the model represents the effect in the data (Halsey, 2019).

Big data collections usually provide many measures, and accordingly a substantial multiple testing burden can arise. This was an issue in chapters 2 and 3 where multiple regional brain asymmetries were measured per individual, and in chapters 4 and 5 with respect to whole genome coverage. Multivariate statistical approaches such as principal component analyses, factor analysis or canonical correlation analyses can sometimes be preferred, and have indeed been used successfully in a recent GWAS study of brain asymmetry (Sha et al., 2020) . In this thesis I have mostly relied on univariate testing with multiple testing correction. Different brain regional asymmetry measures are largely uncorrelated with each other (Chapters 2 and 3), so that the potential benefits of multivariate analysis were not clear, since these methods exploit trait covariance to increase power relative to mass univariate testing. In the genetic analyses (chapters 4 and 5), following mass univariate testing across the genome, I applied various techniques to aggregate the evidence for association at the level of whole genes or biological pathways.

6.2.4 Correcting for confounds in big data

Confounding is an issue for neuroimaging, where many potential confound effects exists (e.g. head motion, scanning options and parameters, position within scanner etc.) (Smith & Nichols, 2018), as well as for genetic analysis, where factors such as population stratification (when different populations systematically have different allele frequencies) or technical aspects of genotyping must be considered (Turner et al., 2011). Generally, multiple regression analyses can correct adequately for such confounding effects. However, in very large datasets, any small, residual confounding effects can bias associations, or create spurious associations, due to high statistical power (Westfall & Yarkoni, 2016). In addition, when conducting multiple regression analyses, one should be careful not to condition on a common effect of predictor and outcome variables that are otherwise independent, because this can induce an artifactual association (i.e., collider bias) (Cole et al., 2010; Day, Loh, Scott, Ong, & Perry, 2016). Furthermore, selection bias can confound study results, as those individuals that participate in a study can differ from the general population with respect to certain factors. For example, individuals from the UK Biobank were reported to have higher education and to smoke less compared to the general population in the UK

(Munafò, Tilling, Taylor, Evans, & Davey Smith, 2018), such that any association between smoking and education may be biased in this dataset (Smith & Nichols, 2018).

In this thesis I have used multiple regression analyses while being mindful of potential colliders. For example, in chapters 2 and 3, fixed effects of age and sex were considered as covariates when testing for mean differences in brain asymmetry between affected individuals and controls, but handedness was not used as a covariate, as differences in handedness may be a common effect of disorder and altered brain asymmetry (**Fig. 6.1**). Likewise, handedness was not considered as a covariate in GWAS for functional connectivity asymmetry (chapter 5), as handedness is partly heritable and may be a common effect of both genetic variation and brain functional variation.

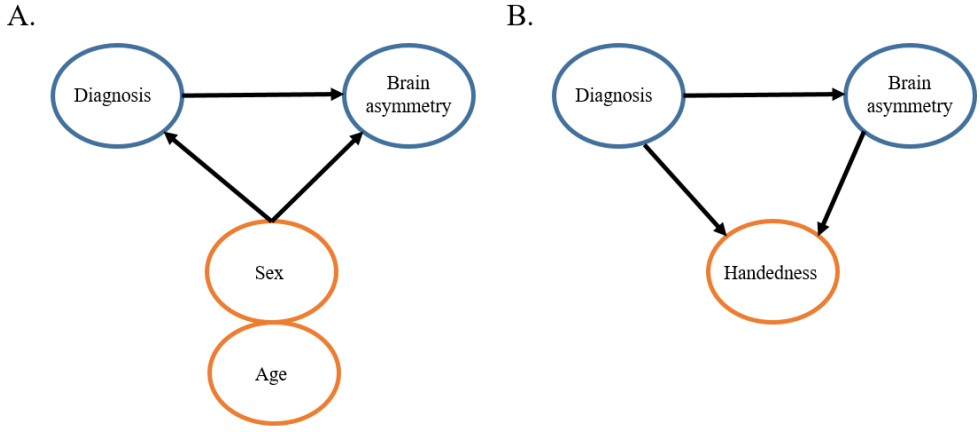


Fig. 6.1. Schematic illustration of possible collider bias. (A) The effect of diagnosis on brain asymmetry can be confounded by e.g. sex or age, which can influence both diagnosis and brain asymmetry. For example, proportionally more males or more children may receive a diagnosis, or brain asymmetry can be different, on average, between males and females, or between children and adults. However, as neither diagnosis nor brain asymmetry are likely to causally influence age and sex, then age and sex can be treated as covariates in testing associations between diagnosis and brain asymmetry. (**B**) Handedness can, in principle, be causally influenced by both diagnosis and brain asymmetry, such that handedness may act as a collider when testing associations between diagnosis and brain asymmetry.

Another general way of dealing with confounding is by stratifying the study sample into groups of matched subjects, rather than, or in addition to, regressing the relevant confounds out. This approach was applied in chapter 5, where principal components that define major axes of population-genetic ancestry were used to define a cluster of 'white British' individuals who had a relatively high degree of genetic homogeneity (and comprised the bulk of the UK Biobank dataset), while excluding any individuals that fell outside of this cluster. Having made this selection, the PCs were then used as continuous covariates in genetic association testing. Indeed, many GWAS studies have focused on people with European ancestry, simply because these individuals form the largest single ancestry group in many of the wealthy countries at the forefront of GWAS efforts. However, from a scientific as well as an ethical point of view, it will be essential to also include sufficient people from other ancestries, not only because it helps to better understand genetic architecture in all populations, but also because results from more inclusive studies will be more broadly applicable (Peterson et al., 2019). Additionally, studying admixed populations can potentially increase sample sizes even further, so long as methods are employed to account for admixture in data analysis.

A particular issue for consortium studies, in which data are collected from multiple separate sites, is data heterogeneity or site effects. In chapters 2 and 3 of this thesis I have used linear mixed effects analysis with *dataset* as a random intercept variable in the models, to correct for potential site-effects in terms of structural MRI data. Another recently applied approach within the ENIGMA consortium is ComBat harmonization, which uses a Bayesian framework to remove site-effects (Fortin et al., 2018; Radua et al., 2020), and can result in increased numbers of findings compared to random-effects meta-analyses or mixed-effects mega-analyses, although performance with respect to the latter was similar (Radua et al., 2020). Alternatively, deep learning algorithms, such as DeepHarmony (Dewey et al., 2019), can be used to produce images with consistent contrast across sites.

An ideal way to reduce site-effects is by means of harmonization of protocols. This was the approach taken by the UK Biobank, who took care to standardize their data collection protocol across a small number of centres in

the UK from the outset of their study (Sudlow et al., 2015). In contrast, such a degree of harmonization was not possible for the ENIGMA consortium, as most datasets were originally collected as individual, separate studies. At least the MRI processing and quality control protocol was harmonized across participating ENIGMA sites, even if the scanning was not. Harmonization across many sites is challenging, as it requires intensive collaboration between many research groups. In addition, with continual progress in technology it is inevitable that scanners and software evolve over time, making harmonization especially difficult for very large studies, which are typically long-term endeavours (Smith & Nichols, 2018).

6.3 Future directions

Over the last decade, the different disciplines of neuroscience, genetics and epidemiology have joined forces to create a field that is sometimes referred to as population neuroscience (Paus, 2010). This field aspires to examine all influences that shape the human brain throughout the lifespan (Paus, 2010). The current thesis is a contribution to this ambitious effort, with regard to brain and behavioural asymmetry in health and disease. Here I will discuss various future research directions for studies of brain asymmetry that have arisen from this work, as well as more general possibilities and principles for the field going forward.

6.3.1 Methods for measuring brain asymmetry

The studies of this thesis were based on either structural, functional (resting state) or behavioural measures of asymmetry. Multi-modal imaging studies, i.e. based on simultaneous analysis of imaging data across different modalities, may aid in obtaining a more complete view of the brain's complexities (Cercignani & Bouyagoub, 2018). For example, aspects of functional network asymmetry may relate to white matter structural connectivity as measured through DTI, such that joint or multivariate analyses of both types of measure may yield greater statistical power to detect case-control disorder effects, or genetic effects. Indeed, advancements in multi-modal imaging techniques and analysis will likely be needed to unravel

structure-function relationships, which are still largely unclear (Batista-Garcia-Ramo & Fernandez-Verdecia, 2018).

In chapter 5, the genetics of functional network asymmetry was analyzed using data from the resting state, with a specific focus on measures that have previously shown correlations with task-defined functional language dominance in another dataset (Joliot, Tzourio-Mazoyer, & Mazoyer, 2016; Labache et al., 2020). However, the UK Biobank does not include language fMRI data, and so it remains an open question whether these resting-state laterality indexes show genetic correlations with task-defined language laterality indexes, or whether the phenotypic correlations are not heritable. This question will need to be addressed in other datasets with genetic information as well as resting-state and language task fMRI. The Human Connectome Project (HCP) dataset may be suitable as it contains twins, although with a total sample size of less than 1500 it is not large enough for GWAS or SNP-based heritability analysis. There is therefore a need for larger scale collections with language task fMRI data, to support population neuroscience of language laterality.

In addition, most functional studies of brain laterality have focused on single tasks within a given cognitive domain (e.g. language) (Hausmann, 2019; Vingerhoets, 2019). Consequently, relationships between different functions (e.g. language production versus listening) and domains (e.g. language versus learning and memory) remain under-investigated (Hausmann, 2019; Vingerhoets, 2019). Again, multivariate studies of effects on functional asymmetry are well motivated. There is also increasing recognition that multi-modal, multivariate neuroimaging data can benefit from data-driven approaches to interrogate them for effects or group differences, e.g. by means of machine learning based on many features simultaneously (Greenstein, Malley, Weisinger, Clasen, & Gogtay, 2012; Yassin et al., 2020), or graph-based network connectivity analyses (Fornito, Zalesky, & Breakspear, 2013; Labache et al., 2020) .

Also there is likely much to be gained from exploring different parcellation schemes. Structural and functional brain atlases vary, for instance, in terms

of reference samples used in their construction (i.e., based on different ages and ethnic groups), numbers of regions defined, and the criteria for defining them. For example, future studies may benefit from more fine definitions of brain regions (e.g. sub-regions of the thalamus), or atlas-free approaches to the cerebral cortex such as hemispheric co-registration at the level of individual vertices (Maingault, Tzourio-Mazoyer, Mazoyer, & Crivello, 2016; Roe et al., 2020).

In fact, a particularly relevant feature of brain atlases for studies of brain asymmetry is whether efforts have been made to achieve cross-hemispheric correspondence. The neuroimaging data from chapters 2 and 3 in this thesis were parcellated using the Desikan-Killiany atlas (Desikan et al., 2006), a structurally asymmetric atlas that subdivides the cerebral cortex into 34 labelled regions per hemisphere. This atlas is asymmetrical because the source population from which it was constructed had average population-level asymmetries, as expected from any healthy population. Nonetheless, efforts were made to define corresponding regions across the two hemispheres, while adapting to the anatomy of each specific hemisphere. The use of an artificially symmetrical atlas (e.g. averaged over the two hemispheres) can safeguard against the influence of asymmetry from the reference sample used in its construction, but the spatial alignment of asymmetric regions to a symmetric atlas can also lead to a reduction of measured asymmetry in the target dataset (Watkins et al., 2001). The use of a 'real-world' asymmetrical atlas has the advantage that regional identification is likely to be more accurate for structures that are asymmetrical both in the atlas and, on average, in the target dataset. For chapter 5, the AICHA atlas (Joliot et al., 2015) was specifically constructed with respect to functionally homotopic regions between the two hemispheres.

As noted further above, macro-structural measures of e.g. cortical thickness are potentially influenced by many different types of micro-structural properties, such as cellular sizes or densities, and across the entire depth of the cortex or more limited to layer-specific variations. The degree of myelination can also affect cortical thickness measures in structural MRI (Natu et al., 2019). Regional interhemispheric differences in such properties

may also relate to functional laterality. In this case, future studies of human brain laterality may benefit from non-invasive imaging techniques such as neurite orientation dispersion and density imaging (Ocklenburg et al., 2018; Schmitz et al., 2019), which is sensitive to cortical grey matter microstructure and laminar fMRI (Lawrence, Formisano, Muckli, & de Lange, 2019), which can resolve differences in the blood oxygen level dependent (BOLD) signal at different depths across the layers of the cerebral cortex. Such efforts will ideally be complemented by post mortem histological studies of laterality, in which different aspects of microstructure, such as cytoarchitecture, can be contrasted between corresponding left and right brain regions (Chance, 2014).

6.3.2 Approaches to asymmetry as a phenotype

After deriving matched left and right hemispheric measures per individual, there are various ways to analyze the data with respect to left-right asymmetry, such as (1) conducting a paired t-test to compare left- and right hemispheric measures at the group level (e.g. see (Yamashita et al., 2011)), (2) performing a repeated measures analysis, making use of a group by hemisphere interaction (e.g. see (Gómez-Gastiasoro et al., 2019)) or (3) calculating the asymmetry index (AI = (Left-Right)/(Left+Right)), examples of which can be found in chapters 2 and 3 of this thesis. The AI is a widely used metric in brain asymmetry studies (Kurth, Gaser, & Luders, 2015) (although its exact formulation can vary, such as multiplying by 2 (Kong et al., 2018; Leroy et al., 2015) or 100 (Goldberg et al., 2013), or using R-L as the numerator). The AI does not scale with the bilateral size of a pair of regions, due to adjustment through its denominator. Calculating only the left-right difference (i.e., without including the denominator of the AI) can induce bias due to bilateral scaling effects (Kurth et al., 2015), although can be used prior to other normalization steps (Carrion-Castillo et al., 2020).

Another variation is the calculation of the absolute (i.e., unsigned) AI (e.g. (Douglas et al., 2018)), which reflects the degree of lateralization irrespective of its direction. For example, after the left-right axis is established in early development, many downstream genetic effects are likely to build on that asymmetry regardless of its direction, such as those that pattern the internal organs in *situs inversus* (i.e. in the same way that they would in *situs solitus*,

just in the mirror form). Any analogous effects in the brain may therefore be of more relevance to the degree rather than the direction of asymmetry. In general, a higher degree of 'fluctuating' asymmetry (i.e. in either direction) in vertebrates is often taken as a measure of developmental instability or perturbation (Varón-González, Pallares, Debat, & Navarro, 2019). In principle, such instability may also be partly heritable or related to disease-causing environmental risk factors, i.e. again there is the possibility of effects on the degree rather than direction of asymmetry. The absolute degree of asymmetry has generally not been considered in this thesis because it is a distinct biological question, and it was desirable to avoid further multiple testing. Also, for many brain measures, the absolute AI is a highly non-normal measure with a floor effect at zero, so that its use should require sophisticated statistical model design and/or selection, with caution for violation of assumptions, or else complex transformation (Šimkovic & Träuble, 2019). Nonetheless, future studies of brain asymmetry in health and disease should include greater consideration of the absolute AI.

Of note, testing separately for effects in each hemisphere, as was done in previous ENIGMA case-control studies of ASD (van Rooij et al., 2018) and ADHD (Hoogman et al., 2017; Hoogman et al., 2019), does not inform about asymmetry *per se*, even if an effect is significant in one hemisphere and not the other. Such patterns can reflect incomplete statistical power in relation to small but uniform bilateral effect sizes, and do not necessarily indicate changes in brain asymmetry. Based on such an approach, there would need to be an additional, post hoc test of whether the effect sizes (e.g. of diagnoses) are significantly different between hemispheres, but even then, the statistical power provided by the paired nature of individual–level left and right data are not utilized. Thus the AI or within-subject repeated measure approaches are preferable.

6.3.3 Future studies of brain asymmetry in disorders

In addition to methodological advances in brain imaging and analysis (above), future studies of brain asymmetry in disorders may benefit from more individual-level approaches, rather than contrasting group means as was done in chapters 2 and 3. For example, ASD is highly heterogeneous disorder

with multiple distinct genetic contributions (An & Claudianos, 2016), which may not be particularly well suited to a 'one size fits all' approach (Lombardo, Lai, & Baron-Cohen, 2019) based on overall case-control group-level comparisons. One promising approach for dealing with clinical heterogeneity is to model individual deviations from the normative pattern of brain change across age (Floris et al., 2020), where the normative model is constructed from large population datasets (Marquand, Rezek, Buitelaar, & Beckmann, 2016). In a recent study using this approach (Floris et al., 2020), autism was associated with distinct, individualized patterns of both extreme right- and leftward deviations, and affected individuals with language delay showed more pronounced rightward deviations than those without language delay. Thus altered brain asymmetry may prove useful as a stratifying biomarker for clinical subtypes of ASD. That study was based on two datasets of hundreds of individuals, but the approach will benefit from being extended to larger-scale data. However, care must be taken when applying the normative modelling approach to data that are aggregated from multiple sites, as these can be biased by site-effects (Kia et al., 2020).

In general, big data will likely be key to examining heterogeneity in ASD, as it may allow reliable etiological subgrouping of individuals, which accordingly might enable personalized medicine (Lombardo et al., 2019). Personalized medicine aims to predict individual susceptibility to a disorder, as well as to provide an accurate diagnosis, consequently promoting an optimal treatment response (Ozomaro, Wahlestedt, & Nemeroff, 2013). The concept is that prediction and diagnosis are ideally based on unique features of the individual, including their genetic and epigenetic makeup, clinical symptoms, environmental exposures, and presence of biomarkers (Ozomaro et al., 2013). Currently, pharmacological treatments can be prescribed for many psychological and or neurodevelopmental disorders, but these act on mechanisms that are not necessarily always relevant in each individual (Huang et al., 2019).

In chapters 2 and 3 of this thesis, I examined brain anatomical data from people with ASD and ADHD separately. This was primarily for practical reasons, as each study involved a different set of collaborations with partners

in different ENIGMA working groups. However, the approach is also well motivated scientifically, because before the publication of DSM-5 in 2013, a diagnosis of ADHD was excluded whenever an ASD diagnosis was present (Hoogman et al., 2020). Many of the datasets were collected before 2013, under DSM-IV diagnostic criteria.

Nonetheless, it is now generally accepted that ASD and ADHD can co-occur (Hoogman et al., 2020). Furthermore, ASD and ADHD have shown a SNP-based genetic correlation of 0.37 (Lee et al., 2019), and twin-based genetic correlation in the range of 0.5-0.75 (Rommelse, Franke, Geurts, Hartman, & Buitelaar, 2010). Accordingly, a recent cross-disorder study from the ENIGMA consortium ASD, ADHD and obsessive compulsive disorder (OCD) Working Groups examined bilateral differences in brain anatomical measures between individuals with these three disorders and healthy controls (Boedhoe et al., 2020). Future joint studies of brain asymmetry in these disorders may also provide increased power to detect effects that are shared between them. It may at first seem counter to the goal of personalized medicine to further combine already heterogeneous disorders into joint analyses. However, if such an approach supports the identification of etiological subtypes that cut across traditional diagnostic boundaries, for example with respect to extreme deviations from normative patterns of brain asymmetry, then personalized medicine may actually be aided. Genetics research is also likely to contribute to this effort (see below).

6.3.4 Future genetic studies of brain asymmetry

As discussed in the Introduction of this thesis (chapter 1), the identification of individual genetic effects on complex traits requires large sample sizes. Small samples are only possible when traits have potentially monogenic backgrounds, at least in subsets of those affected (as for *situs inversus* in Chapter 4). Chapter 5 of this thesis, and other recent studies (Carrion-Castillo et al., 2020; Sha et al., 2020), have shown that measures of brain asymmetry have generally low but significant SNP-based heritabilities, such that larger datasets than those used previously are likely to be successful in identifying more genetic effects and biological pathways. Chapter 5 also included analysis of rare genetic variation with respect to population variance in one

aspect of brain asymmetry (in this case hemispheric differences in functional connectivity). Such analyses have only now become possible through the release of large-scale exome sequencing and MRI data by the UK Biobank. Additional studies of other imaging modalities with respect to exome sequencing data should now be pursued. Ideally, more efforts like the UK Biobank, or e.g. the Million Veteran Program (Gaziano et al., 2016) (a large-scale population study in the USA), will also be initiated in the future, and they should be located at different sites around the world in order to reflect global population heterogeneity.

Genetic analyses may also augment our understanding of altered brain asymmetry in disorders. For example, polygenic risk scoring based on disorder case-control GWAS can be applied in population neuroscience data such as from the UK Biobank, to understand how polygenic predisposition to disorder impacts brain asymmetry. In addition, genetic correlation analysis of brain asymmetries and disorders may indirectly suggest that some causal genetic effects on disorders are partly mediated via altered brain asymmetry, a possibility that can then be investigated through Mendelian Randomization approaches (Bowden & Holmes, 2019). Pooling of data from different biobanks will increase the possibility to study the genetics of rare etiological subtypes, as well as to pinpoint further small-effect associations between causal genetic variants and common traits (Davey Smith et al., 2005).

A complementary approach to studying the genetics of brain asymmetries is post-mortem gene expression analysis of the human brain (Francks, 2015; Karlebach & Francks, 2015). Such studies have so far only been carried out in very small sample sizes, due to limited availability of suitable tissue. However, the approach has already suggested hemispheric differences of gene expression between the left and right superior temporal cortex in adults (Karlebach & Francks, 2015), which may relate to functional laterality of this region, as well as asymmetrical expression of transcriptomic profiles of the developing central nervous system in human embryos (de Kovel, Lisgo, Fisher, & Francks, 2018; Ocklenburg et al., 2017). This approach can also be combined with histological studies of potentially lateralized cytoarchitecture

(Chance, 2014; Hutsler, 2003), and immunohistochemistry to study protein distributions.

Useful existing sources of gene expression data in the human brain are the Allen human brain atlas (for adult data: http://human.brain-map.org/) and BrainSpan atlas (data from across the lifespan: www.brainspan.org). For example, Allen brain atlas data indicated that *DTYMK*, a gene implicated in *planum temporale* asymmetry by GWAS, is also asymmetrically expressed in that region, showing higher expression on the left side compared to the right (Carrion-Castillo et al., 2020). Other genes implicated in brain anatomical asymmetry by multivariate GWAS in adults were also found to be particularly highly expressed in the embryonic and fetal human brain in the Brainspan data (Sha et al., 2020). Such analyses would be greatly aided in future studies by gene expression data collections from larger numbers of individuals, as currently the Allen human brain atlas is based on just six post-mortem brains (with only two having data from both hemispheres), while the Brainspan data comprise only handfuls of brains at any given developmental stage. Further studies of gene expression in the embryonic brain may be especially valuable for future studies of brain asymmetry, given that left-right differences likely arise early in development (see Introduction).

Microtubule-related genes have now been implicated in affecting brain asymmetry by large-scale GWAS of handedness (Cuellar-Partida et al., 2020; de Kovel & Francks, 2019; Wiberg et al., 2019) and brain structural asymmetry (Sha et al., 2020). As mentioned in the Introduction, this may indicate that cellular chirality is involved in creating the left-right axis of the brain (Sha et al., 2020). The term ‘chirality’ comes from the Greek word ‘chier’, meaning hand. A pair of human hands is an example of chirality, as the left and right hand are mirror-images of each other that cannot be superimposed, just like chirality in molecular structures (Obara & Egan, 2013). It is known, for example, that left-right asymmetry of the hindgut in fruit flies (*Drosophila melanogaster*) relies on an organ-intrinsic mechanism that involves cellular chirality due to intracellular actin organization, in terms of unidirectional tilting of radial fibers, and anti-clockwise swirling of transverse fibers (Tee et al., 2015). As early as the 1990s, Brown and Wolpert (Brown & Wolpert, 1990) proposed that initial symmetry breaking in the

embryo - in the absence of ciliary induced nodal flow - can be initiated by molecular chirality. The possibility of a microtubule-mediated mechanism of brain left-right axis formation should therefore be further investigated using *in vitro* and animal models (see below).

Of note, functional clustering of genes identified by GWAS does not necessarily mean that every implicated pathway is crucially involved in a phenotype of interest, especially when the phenotype is genetically complex (Boyle, Li, & Pritchard, 2017). The 'omnigenic' model postulates that associations in GWAS mostly reflect variants of tiny effect in genes that are only peripherally involved in the phenotype, and that a smaller number of directly relevant core genes for that phenotype may not even carry much genetic variation associated with its variance (Boyle et al., 2017). Yet microtubule-related genes seem a promising set to further examine in relation to brain asymmetry, given that associations with such genes have been found in different studies, and because a role for microtubules in left-right patterning makes sense theoretically, with empirical support in relation to other organs in other species.

6.3.5 Model systems

It is now widely appreciated that brain asymmetry is not a uniquely human feature, but rather is found throughout the animal kingdom (see chapter 1, Introduction). Studies of the genetics and development of brain asymmetry in animals may therefore help to gain important insights into the asymmetry of the human brain, to the extent that mechanisms are conserved (Güntürkün, Ströckens, & Ocklenburg, 2020). For example, using mouse models, specific genes can be knocked out entirely or in a restricted manner, to study subsequent alterations in molecular, cellular and developmental processes, as well as behaviour (Ocklenburg, Berretz, Packheiser, & Friedrich, 2020). Various molecular asymmetries have been reported in the mouse hippocampus, with potential consequences for learning, memory and lateralized behaviour (Jordan, 2020). Knockout mice for specific microtubule component genes, or genes encoding microtubule-associated proteins, could therefore be studied with respect to these asymmetries and behaviours.

Such studies can be aided by gene editing techniques such as CRISPR (which stands for clustered regularly interspaced short palindromic repeats), which is now efficient, practical and affordable (Memi, Ntokou, & Papangeli, 2018; Ormond et al., 2019). Of note, adapted CRISPR protocols now also permit sequence-specific chromatin manipulation without altering the genome, including transcriptional activation or repression of genes (Memi et al., 2018). Such applications may be useful for future studies of how left-right epigenetic differences affect embryonic neurodevelopment, as some epigenetic asymmetries have been implicated in the developing human spinal cord (Ocklenburg et al., 2017).

A growing concern for animal welfare and sense of animal rights has increased focus on ethical issues around their use in research (Andersen & Winter, 2019). One promising alternative approach to animal models is organoids (i.e., *in vitro* 3D cell culturing systems) (Giandomenico & Lancaster, 2017). Mammalian organoids have been cultured for different organs, including the brain (Rossi, Manfrin, & Lutolf, 2018). They can be grown from pluripotent stem cells (PSCs), that is to say, embryonic stem cells or induced PSCs, which have the capacity to grow into any cell type of the embryo (Rossi et al., 2018). Human cerebral organoids can be grown from human PSCs, but currently they lack microvasculature, and therefore can only grow up to stages comparable with the second prenatal trimester, as otherwise the innermost cells start to die due to limited oxygen and nutrient supply (Qian, Song, & Ming, 2019). It may therefore be possible to use cerebral organoids in studies of embryonic mechanisms of left-right brain axis formation. For example, it has not yet been studied whether cellular chirality can be observed or induced in the context of cerebral organoids, for example of neuro-epithelial cells. In the Drosophila hindgut, transient cellular chirality in epithelial cells is thought to drive chiral looping of the gut during its early formation as a tube (Hatori et al., 2014).

Human PSCs can also be differentiated into neuron-like and other types of cells outside of the context of organoids, which may also support in vitro analysis of cellular chirality, and how it is influenced by changes in the levels of microtubule-related genes.

6.3.6 Reproducibility

The studies in this thesis had several aspects aimed at enhancing the reproducibility and reliability of findings. These included the use of the largest ever sample sizes for the questions that were addressed, together with multiple testing correction, sensitivity analyses, the reporting of all effect sizes whether significant or not, posting of preprints, and open access publication. Limitations on the sharing of some primary data were in force, mostly due to the use of legacy datasets in which the informed consent by participants did not include agreement for open data sharing.

In general, reproducibility can be enhanced through open science research practices such as sharing of data, code and research materials, double-blind peer review, preprints, open access publishing, and pre-registration (Allen & Mehler, 2019). The latter entails *a priori* reporting of the study hypothesis and analysis pipelines before results are known, and can prevent P-hacking and selective reporting (i.e., searching for significant effects and selecting accordingly) or hypothesizing after the results are known (HARKing). However, the present thesis did not make use of pre-registration. This was because pre-registration is time-consuming up front, which would have been a practical problem for carrying out the consortium projects of chapters 2 and 3 (each step in the study proposal and publication process required obtaining approval from hundreds of co-authors and could take many weeks). Pre-registration is also not well suited to exploratory studies such as described in chapters 4 and 5, where optimal strategies using recently-developed analysis methods and data types could only be developed through working on the data themselves. Instead, P-hacking and HARKing were avoided in this thesis through a combination of careful design and commitment to reporting all results, whether they were significant or not.

6.4 Concluding remarks

The studies in this thesis have advanced our understanding of the extent to which brain asymmetry is altered in neurodevelopmental disorders, and the extent to which genes influence variation in brain asymmetry. These advances were made concurrently with some other notable, related findings during the

last four years, including from large-scale imaging studies of disorder-asymmetry associations (de Kovel et al., 2019; Kong, Boedhoe, et al., 2020), genome-wide screens for effects on brain or behavioural asymmetries (Carrion-Castillo et al., 2020; Cuellar-Partida et al., 2020; Sha et al., 2020), and post-mortem analysis of left-right molecular differences in human developing brain tissue (de Kovel et al., 2017; de Kovel et al., 2018; Ocklenburg et al., 2017). Building on these advances, the field is set for an exciting period of further research on brain asymmetry in health and disease.

6.5 References

Allen, C., & Mehler, D. M. A. (2019). Open science challenges, benefits and tips in early career and beyond. *PLoS Biol, 17*(5), e3000246. doi:10.1371/journal.pbio.3000246

An, J. Y., & Claudianos, C. (2016). Genetic heterogeneity in autism: From single gene to a pathway perspective. *Neurosci Biobehav Rev, 68*, 442-453. doi:10.1016/j.neubiorev.2016.06.013

Andersen, M. L., & Winter, L. M. F. (2019). Animal models in biological and biomedical research - experimental and ethical concerns. *An Acad Bras Cienc, 91*(suppl 1), e20170238. doi:10.1590/0001-3765201720170238

Batista-Garcia-Ramo, K., & Fernandez-Verdecia, C. I. (2018). What We Know About the Brain Structure-Function Relationship. *Behav Sci (Basel), 8*(4). doi:10.3390/bs8040039

Boedhoe, P. S. W., van Rooij, D., Hoogman, M., Twisk, J. W. R., Schmaal, L., Abe, Y., . . . van den Heuvel, O. A. (2020). Subcortical Brain Volume, Regional Cortical Thickness, and Cortical Surface Area Across Disorders: Findings From the ENIGMA ADHD, ASD, and OCD Working Groups. *Am J Psychiatry, 177*(9), 834-843. doi:10.1176/appi.ajp.2020.19030331

Bowden, J., & Holmes, M. V. (2019). Meta-analysis and Mendelian randomization: A review. *Res Synth Methods, 10*(4), 486-496. doi:10.1002/jrsm.1346

Boyle, E. A., Li, Y. I., & Pritchard, J. K. (2017). An Expanded View of Complex Traits: From Polygenic to Omnigenic. *Cell, 169*(7), 1177-1186. doi:10.1016/j.cell.2017.05.038

Brown, N. A., & Wolpert, L. (1990). The development of handedness in left/right asymmetry. *Development, 109*(1), 1-9.

Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., & Munafo, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nat Rev Neurosci, 14*(5), 365-376. doi:10.1038/nrn3475

Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L. T., Sharp, K., . . . Marchini, J. (2018). The UK Biobank resource with deep phenotyping and genomic data. *Nature, 562*(7726), 203-209. doi:10.1038/s41586-018-0579-z

Carrion-Castillo, A., Pepe, A., Kong, X. Z., Fisher, S. E., Mazoyer, B., Tzourio-Mazoyer, N., . . . Francks, C. (2020). Genetic effects on planum temporale asymmetry and their limited relevance to neurodevelopmental disorders, intelligence or educational attainment. *Cortex, 124*, 137-153. doi:10.1016/j.cortex.2019.11.006

Cercignani, M., & Bouyagoub, S. (2018). Brain microstructure by multi-modal MRI: Is the whole greater than the sum of its parts? *Neuroimage, 182*, 117-127. doi:10.1016/j.neuroimage.2017.10.052

Chance, S. A. (2014). The cortical microstructural basis of lateralized cognition: a review. *Front Psychol, 5*, 820. doi:10.3389/fpsyg.2014.00820

Cole, S. R., Platt, R. W., Schisterman, E. F., Chu, H., Westreich, D., Richardson, D., & Poole, C. (2010). Illustrating bias due to conditioning on a collider. *Int J Epidemiol, 39*(2), 417-420. doi:10.1093/ije/dyp334

Cuellar-Partida, G., Tung, J. Y., Eriksson, N., Albrecht, E., Aliev, F., Andreassen, O. A., . . . Medland, S. E. (2020). Genome-wide association study identifies 48 common genetic variants associated with handedness. *Nat Hum Behav*. doi:10.1038/s41562-020-00956-y

Davey Smith, G., Ebrahim, S., Lewis, S., Hansell, A. L., Palmer, L. J., & Burton, P. R. (2005). Genetic epidemiology and public health: hope, hype, and future prospects. *Lancet, 366*(9495), 1484-1498. doi:10.1016/s0140-6736(05)67601-5

Day, F. R., Loh, P. R., Scott, R. A., Ong, K. K., & Perry, J. R. (2016). A Robust Example of Collider Bias in a Genetic Association Study. *Am J Hum Genet, 98*(2), 392-393. doi:10.1016/j.ajhg.2015.12.019

de Kovel, C. G. F., Aftanas, L., Aleman, A., Alexander-Bloch, A. F., Baune, B. T., Brack, I., . . . Francks, C. (2019). No Alterations of Brain Structural Asymmetry in Major Depressive Disorder: An ENIGMA Consortium Analysis. *Am J Psychiatry, 176*(12), 1039-1049. doi:10.1176/appi.ajp.2019.18101144

de Kovel, C. G. F., & Francks, C. (2019). The molecular genetics of hand preference revisited. *Sci Rep, 9*(1), 5986. doi:10.1038/s41598-019-42515-0

de Kovel, C. G. F., Lisgo, S., Karlebach, G., Ju, J., Cheng, G., Fisher, S. E., & Francks, C. (2017). Left-Right Asymmetry of Maturation Rates in Human Embryonic Neural Development. *Biol Psychiatry, 82*(3), 204-212. doi:10.1016/j.biopsych.2017.01.016

de Kovel, C. G. F., Lisgo, S. N., Fisher, S. E., & Francks, C. (2018). Subtle left-right asymmetry of gene expression profiles in embryonic and foetal human brains. *Sci Rep, 8*(1), 12606. doi:10.1038/s41598-018-29496-2

Desikan, R. S., Segonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., . . . Killiany, R. J. (2006). An automated labeling system for subdividing

the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage, 31*(3), 968-980. doi:10.1016/j.neuroimage.2006.01.021

Dewey, B. E., Zhao, C., Reinhold, J. C., Carass, A., Fitzgerald, K. C., Sotirchos, E. S., . . . Prince, J. L. (2019). DeepHarmony: A deep learning approach to contrast harmonization across scanner changes. *Magn Reson Imaging, 64*, 160-170. doi:10.1016/j.mri.2019.05.041

Douglas, P. K., Gutman, B., Anderson, A., Larios, C., Lawrence, K. E., Narr, K., . . . Bookheimer, S. Y. (2018). Hemispheric brain asymmetry differences in youths with attention-deficit/hyperactivity disorder. *Neuroimage Clin, 18*, 744-752. doi:10.1016/j.nicl.2018.02.020

Floris, D. L., Wolfers, T., Zabihi, M., Holz, N. E., Zwiers, M. P., Charman, T., . . . Beckmann, C. F. (2020). Atypical Brain Asymmetry in Autism-A Candidate for Clinically Meaningful Stratification. *Biol Psychiatry Cogn Neurosci Neuroimaging*. doi:10.1016/j.bpsc.2020.08.008

Fornito, A., Zalesky, A., & Breakspear, M. (2013). Graph analysis of the human connectome: promise, progress, and pitfalls. *Neuroimage, 80*, 426-444. doi:10.1016/j.neuroimage.2013.04.087

Fortin, J. P., Cullen, N., Sheline, Y. I., Taylor, W. D., Aselcioglu, I., Cook, P. A., . . . Shinohara, R. T. (2018). Harmonization of cortical thickness measurements across scanners and sites. *Neuroimage, 167*, 104-120. doi:10.1016/j.neuroimage.2017.11.024

Francks, C. (2015). Exploring human brain lateralization with molecular genetics and genomics. *Ann N Y Acad Sci, 1359*, 1-13. doi:10.1111/nyas.12770

Gaziano, J. M., Concato, J., Brophy, M., Fiore, L., Pyarajan, S., Breeling, J., . . . O'Leary, T. J. (2016). Million Veteran Program: A mega-biobank to study genetic influences on health and disease. *J Clin Epidemiol, 70*, 214-223. doi:10.1016/j.jclinepi.2015.09.016

Giandomenico, S. L., & Lancaster, M. A. (2017). Probing human brain evolution and development in organoids. *Curr Opin Cell Biol, 44*, 36-43. doi:10.1016/j.ceb.2017.01.001

Goldberg, E., Roediger, D., Kucukboyaci, N. E., Carlson, C., Devinsky, O., Kuzniecky, R., . . . Thesen, T. (2013). Hemispheric asymmetries of cortical volume in the human brain. *Cortex, 49*(1), 200-210. doi:10.1016/j.cortex.2011.11.002

Gómez-Gastiasoro, A., Zubiaurre-Elorza, L., Peña, J., Ibarretxe-Bilbao, N., Rilo, O., Schretlen, D. J., & Ojeda, N. (2019). Altered frontal white matter asymmetry and its implications for cognition in schizophrenia: A tractography study. *Neuroimage Clin, 22*, 101781. doi:10.1016/j.nicl.2019.101781

Greenstein, D., Malley, J. D., Weisinger, B., Clasen, L., & Gogtay, N. (2012). Using multivariate machine learning methods and structural MRI to classify childhood onset schizophrenia and healthy controls. *Front Psychiatry, 3*, 53. doi:10.3389/fpsyt.2012.00053

Güntürkün, O., Ströckens, F., & Ocklenburg, S. (2020). Brain Lateralization: A Comparative Perspective. *Physiol Rev, 100*(3), 1019-1063. doi:10.1152/physrev.00006.2019

Halsey, L. G. (2019). The reign of the p-value is over: what alternative analyses could we employ to fill the power vacuum? *Biol Lett, 15*(5), 20190174. doi:10.1098/rsbl.2019.0174

Hatori, R., Ando, T., Sasamura, T., Nakazawa, N., Nakamura, M., Taniguchi, K., . . . Matsuno, K. (2014). Left-right asymmetry is formed in individual cells by intrinsic cell chirality. *Mech Dev, 133*, 146-162. doi:10.1016/j.mod.2014.04.002

Hausmann, M. (2019). Variations of hemispheric functional segregation in the laterality spectrum: Comment on "Phenotypes in hemispheric functional segregation? Perspectives and challenges" by Guy Vingerhoets. *Phys Life Rev, 30*, 27-29. doi:10.1016/j.plrev.2019.08.006

Hoogman, M., Bralten, J., Hibar, D. P., Mennes, M., Zwiers, M. P., Schweren, L. S., . . . Franke, B. (2017). Subcortical brain volume differences in participants with attention deficit hyperactivity disorder in children and adults: a cross-sectional mega-analysis. *Lancet Psychiatry, 4*(4), 310-319. doi:10.1016/s2215-0366(17)30049-4

Hoogman, M., Muetzel, R., Guimaraes, J. P., Shumskaya, E., Mennes, M., Zwiers, M. P., . . . Franke, B. (2019). Brain Imaging of the Cortex in ADHD: A Coordinated Analysis of Large-Scale Clinical and Population-Based Samples. *Am J Psychiatry*, appiajp201918091033. doi:10.1176/appi.ajp.2019.18091033

Hoogman, M., van Rooij, D., Klein, M., Boedhoe, P., Ilioska, I., Li, T., . . . Franke, B. (2020). Consortium neuroscience of attention deficit/hyperactivity disorder and autism spectrum disorder: The ENIGMA adventure. *Hum Brain Mapp*. doi:10.1002/hbm.25029

Huang, Y., Furuno, M., Arakawa, T., Takizawa, S., de Hoon, M., Suzuki, H., & Arner, E. (2019). A framework for identification of on- and off-target transcriptional responses to drug treatment. *Sci Rep, 9*(1), 17603. doi:10.1038/s41598-019-54180-4

Hutsler, J. J. (2003). The specialized structure of human language cortex: pyramidal cell size asymmetries within auditory and language-associated regions of the temporal lobes. *Brain Lang, 86*(2), 226-242. doi:10.1016/s0093-934x(02)00531-x

Jafari, M., & Ansari-Pour, N. (2019). Why, When and How to Adjust Your P Values? *Cell J, 20*(4), 604-607. doi:10.22074/cellj.2019.5992

Joliot, M., Jobard, G., Naveau, M., Delcroix, N., Petit, L., Zago, L., . . . Tzourio-Mazoyer, N. (2015). AICHA: An atlas of intrinsic connectivity of homotopic areas. *J Neurosci Methods, 254*, 46-59. doi:10.1016/j.jneumeth.2015.07.013

Joliot, M., Tzourio-Mazoyer, N., & Mazoyer, B. (2016). Intra-hemispheric intrinsic connectivity asymmetry and its relationships with handedness and

language Lateralization. *Neuropsychologia, 93*(Pt B), 437-447. doi:10.1016/j.neuropsychologia.2016.03.013

Jordan, J. T. (2020). The rodent hippocampus as a bilateral structure: A review of hemispheric lateralization. *Hippocampus, 30*(3), 278-292. doi:10.1002/hipo.23188

Karlebach, G., & Francks, C. (2015). Lateralization of gene expression in human language cortex. *Cortex, 67*, 30-36. doi:10.1016/j.cortex.2015.03.003

Kia, S. M., Huijsdens, H., Dinga, R., Wolfers, T., Mennes, M., Andreassen, O. E., . . . Marquand, A. F. (2020). Hierarchical Bayesian Regression for Multi-site Normative Modeling of Neuroimaging Data. In: Martel A.L. et al. (eds) Medical Image Computing and Computer Assisted Intervention – MICCAI 2020. MICCAI 2020. Lecture Notes in Computer Science, vol 12267. Springer, Cham. doi:https://doi.org/10.1007/978-3-030-59728-3_68

Kilkenny, M. F., & Robinson, K. M. (2018). Data quality: "Garbage in - garbage out". *Health Inf Manag, 47*(3), 103-105. doi:10.1177/1833358318774357

Kinkorova, J., & Topolcan, O. (2020). Biobanks in the era of big data: objectives, challenges, perspectives, and innovations for predictive, preventive, and personalised medicine. *Epma j, 11*(3), 333-341. doi:10.1007/s13167-020-00213-2

Kong, X. Z., Boedhoe, P. S. W., Abe, Y., Alonso, P., Ameis, S. H., Arnold, P. D., . . . Francks, C. (2020). Mapping Cortical and Subcortical Asymmetry in Obsessive-Compulsive Disorder: Findings From the ENIGMA Consortium. *Biol Psychiatry, 87*(12), 1022-1034. doi:10.1016/j.biopsych.2019.04.022

Kong, X. Z., Mathias, S. R., Guadalupe, T., Glahn, D. C., Franke, B., Crivello, F., . . . Francks, C. (2018). Mapping cortical brain asymmetry in 17,141 healthy individuals worldwide via the ENIGMA Consortium. *Proc Natl Acad Sci U S A*. doi:10.1073/pnas.1718418115

Kong, X. Z., Postema, M. C., Guadalupe, T., de Kovel, C., Boedhoe, P. S. W., Hoogman, M., . . . Francks, C. (2020). Mapping brain asymmetry in health and disease through the ENIGMA consortium. *Hum Brain Mapp*. doi:10.1002/hbm.25033

Kurth, F., Gaser, C., & Luders, E. (2015). A 12-step user guide for analyzing voxel-wise gray matter asymmetries in statistical parametric mapping (SPM). *Nat Protoc, 10*(2), 293-304. doi:10.1038/nprot.2015.014

Labache, L., Mazoyer, B., Joliot, M., Crivello, F., Hesling, I., & Tzourio-Mazoyer, N. (2020). Typical and atypical language brain organization based on intrinsic connectivity and multitask functional asymmetries. *eLife, 9*. doi:10.7554/eLife.58722

Lawrence, S. J. D., Formisano, E., Muckli, L., & de Lange, F. P. (2019). Laminar fMRI: Applications for cognitive neuroscience. *NeuroImage, 197*, 785-791. doi:https://doi.org/10.1016/j.neuroimage.2017.07.004

Layton, W. M., Jr. (1976). Random determination of a developmental process: reversal of normal visceral asymmetry in the mouse. *J Hered, 67*(6), 336-338.

Lee, P. H., Anttila, V., Won, H., Feng, Y.-C. A., Rosenthal, J., Zhu, Z., . . . Smoller, J. W. (2019). Genomic Relationships, Novel Loci, and Pleiotropic Mechanisms across Eight Psychiatric Disorders. *Cell, 179*(7), 1469-1482.e1411. doi:10.1016/j.cell.2019.11.020

Leigh, M. W., Pittman, J. E., Carson, J. L., Ferkol, T. W., Dell, S. D., Davis, S. D., . . . Zariwala, M. A. (2009). Clinical and genetic aspects of primary ciliary dyskinesia/Kartagener syndrome. *Genet Med, 11*(7), 473-487. doi:10.1097/GIM.0b013e3181a53562

Lens, L., Van Dongen, S., Kark, S., & Matthysen, E. (2002). Fluctuating asymmetry as an indicator of fitness: can we bridge the gap between studies? *Biol Rev Camb Philos Soc, 77*(1), 27-38. doi:10.1017/s1464793101005796

Leroy, F., Cai, Q., Bogart, S. L., Dubois, J., Coulon, O., Monzalvo, K., . . . Dehaene-Lambertz, G. (2015). New human-specific brain landmark: the depth asymmetry of superior temporal sulcus. *Proc Natl Acad Sci U S A, 112*(4), 1208-1213. doi:10.1073/pnas.1412389112

Liu, J. Z., Tozzi, F., Waterworth, D. M., Pillai, S. G., Muglia, P., Middleton, L., . . . Marchini, J. (2010). Meta-analysis and imputation refines the association of 15q25 with smoking quantity. *Nat Genet, 42*(5), 436-440. doi:10.1038/ng.572

Lombardo, M. V., Lai, M. C., & Baron-Cohen, S. (2019). Big data approaches to decomposing heterogeneity across the autism spectrum. *Mol Psychiatry, 24*(10), 1435-1450. doi:10.1038/s41380-018-0321-0

Maingault, S., Tzourio-Mazoyer, N., Mazoyer, B., & Crivello, F. (2016). Regional correlations between cortical thickness and surface area asymmetries: A surface-based morphometry study of 250 adults. *Neuropsychologia, 93*(Pt B), 350-364. doi:10.1016/j.neuropsychologia.2016.03.025

Manolio, T. A., Collins, F. S., Cox, N. J., Goldstein, D. B., Hindorff, L. A., Hunter, D. J., . . . Visscher, P. M. (2009). Finding the missing heritability of complex diseases. *Nature, 461*(7265), 747-753. doi:10.1038/nature08494

Marquand, A. F., Rezek, I., Buitelaar, J., & Beckmann, C. F. (2016). Understanding Heterogeneity in Clinical Cohorts Using Normative Models: Beyond Case-Control Studies. *Biol Psychiatry, 80*(7), 552-561. doi:10.1016/j.biopsych.2015.12.023

Memi, F., Ntokou, A., & Papangeli, I. (2018). CRISPR/Cas9 gene-editing: Research technologies, clinical applications and ethical considerations. *Semin Perinatol, 42*(8), 487-500. doi:10.1053/j.semperi.2018.09.003

Moreno-Ayala, R., Olivares-Chauvet, P., Schäfer, R., & Junker, J. P. (2021). Variability of an Early Developmental Cell Population Underlies Stochastic Laterality Defects. *Cell Rep, 34*(2), 108606. doi:10.1016/j.celrep.2020.108606

Munafo, M. R., & Flint, J. (2010). How reliable are scientific studies? *Br J Psychiatry, 197*(4), 257-258. doi:10.1192/bjp.bp.109.069849

Munafò, M. R., Tilling, K., Taylor, A. E., Evans, D. M., & Davey Smith, G. (2018). Collider scope: when selection bias can substantially influence observed associations. *Int J Epidemiol, 47*(1), 226-235. doi:10.1093/ije/dyx206

Natu, V. S., Gomez, J., Barnett, M., Jeska, B., Kirilina, E., Jaeger, C., . . . Grill-Spector, K. (2019). Apparent thinning of human visual cortex during childhood is associated with myelination. *Proceedings of the National Academy of Sciences, 116*(41), 20750-20759. doi:10.1073/pnas.1904931116

Nayak, B. K. (2010). Understanding the relevance of sample size calculation. *Indian J Ophthalmol, 58*(6), 469-470. doi:10.4103/0301-4738.71673

Obara, S., & Egan, T. D. (2013). Chapter 2 - Pharmacokinetic and Pharmacodynamic Principles for Intravenous Anesthetics, Editor(s): Hugh C. Hemmings, Talmage D. Egan, Pharmacology and Physiology for Anesthesia, W.B. Saunders. Pages 20-42. doi:10.1016/B978-1-4377-1679-5.00002-8

Ocklenburg, S., Berretz, G., Packheiser, J., & Friedrich, P. (2020). Laterality 2020: entering the next decade. *Laterality*, 1-33. doi:10.1080/1357650x.2020.1804396

Ocklenburg, S., Friedrich, P., Fraenz, C., Schluter, C., Beste, C., Gunturkun, O., & Genc, E. (2018). Neurite architecture of the planum temporale predicts neurophysiological processing of auditory speech. *Sci Adv, 4*(7), eaar6830. doi:10.1126/sciadv.aar6830

Ocklenburg, S., Schmitz, J., Moinfar, Z., Moser, D., Klose, R., Lor, S., . . . Gunturkun, O. (2017). Epigenetic regulation of lateralized fetal spinal gene expression underlies hemispheric asymmetries. *eLife, 6*. doi:10.7554/eLife.22784

Ormond, K. E., Bombard, Y., Bonham, V. L., Hoffman-Andrews, L., Howard, H., Isasi, R., . . . Allyse, M. (2019). The clinical application of gene editing: ethical and social issues. *Per Med, 16*(4), 337-350. doi:10.2217/pme-2018-0155

Ozomaro, U., Wahlestedt, C., & Nemeroff, C. B. (2013). Personalized medicine in psychiatry: problems and promises. *BMC Med, 11*, 132. doi:10.1186/1741-7015-11-132

Palmer, A. R. (1994). Fluctuating asymmetry analyses: a primer. In T. A. Markow (Ed.), *Developmental Instability: Its Origins and Evolutionary Implications: Proceedings of the International Conference on Developmental Instability: Its Origins and Evolutionary Implications, Tempe, Arizona, 14–15 June 1993* (pp. 335-364). Dordrecht: Springer Netherlands.

Paus, T. (2010). Population neuroscience: why and how. *Hum Brain Mapp, 31*(6), 891-903. doi:10.1002/hbm.21069

Peterson, R. E., Kuchenbaecker, K., Walters, R. K., Chen, C. Y., Popejoy, A. B., Periyasamy, S., . . . Duncan, L. E. (2019). Genome-wide Association Studies in Ancestrally Diverse Populations: Opportunities, Methods, Pitfalls, and Recommendations. *Cell, 179*(3), 589-603. doi:10.1016/j.cell.2019.08.051

Price, W. N., 2nd, & Cohen, I. G. (2019). Privacy in the age of medical big data. *Nat Med, 25*(1), 37-43. doi:10.1038/s41591-018-0272-7

Psaty, B. M., O'Donnell, C. J., Gudnason, V., Lunetta, K. L., Folsom, A. R., Rotter, J. I., . . . Boerwinkle, E. (2009). Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium: Design of prospective meta-analyses of genome-wide association studies from 5 cohorts. *Circ Cardiovasc Genet, 2*(1), 73-80. doi:10.1161/circgenetics.108.829747

Qian, X., Song, H., & Ming, G. L. (2019). Brain organoids: advances, applications and challenges. *Development, 146*(8). doi:10.1242/dev.166074

Radua, J., Vieta, E., Shinohara, R., Kochunov, P., Quidé, Y., Green, M. J., . . . van Erp, T. (2020). Increased power by harmonizing structural MRI site differences with the ComBat batch adjustment method in ENIGMA. *Neuroimage, 218*, 116956. doi:10.1016/j.neuroimage.2020.116956

Roe, J. M., Vidal-Piñeiro, D., Sørensen, Ø., Brandmaier, A. M., Düzel, S., Gonzalez, H. A., . . . Westerhausen, R. (2020). Asymmetric thinning of the cerebral cortex across the adult lifespan is accelerated in Alzheimer's Disease. *bioRxiv*, 2020.2006.2018.158980. doi:10.1101/2020.06.18.158980

Rommelse, N. N., Franke, B., Geurts, H. M., Hartman, C. A., & Buitelaar, J. K. (2010). Shared heritability of attention-deficit/hyperactivity disorder and autism spectrum disorder. *Eur Child Adolesc Psychiatry, 19*(3), 281-295. doi:10.1007/s00787-010-0092-x

Rossi, G., Manfrin, A., & Lutolf, M. P. (2018). Progress and potential in organoid research. *Nat Rev Genet, 19*(11), 671-687. doi:10.1038/s41576-018-0051-9

Schmitz, J., Fraenz, C., Schluter, C., Friedrich, P., Jung, R. E., Gunturkun, O., . . . Ocklenburg, S. (2019). Hemispheric asymmetries in cortical gray matter microstructure identified by neurite orientation dispersion and density imaging. *NeuroImage*. doi:10.1016/j.neuroimage.2019.01.079

Scott, L. (2002). The origin of monozygotic twinning. *Reprod Biomed Online, 5*(3), 276-284. doi:10.1016/s1472-6483(10)61833-0

Segal, N. L. (2017). Situs Inversus Totalis in Twins: A Brief Review and a Life History / Twin Research: Twin Studies of Trisomy 21; Monozygotic Twin Concordance for Bilateral Coronoid Hyperplasia; Prenatal Hormonal Effects in Mixed-Sex Non-Human Primate Litters; Insurance Mandates and Twinning After In Vitro Fertilization / News Reports: First Report of Identical Twin Puppies; Twins Sisters Turn 100; Remembering an Identical Twin Production Designer; New York City Marathon Quadruplets. *Twin Res Hum Genet, 20*(1), 90-95. doi:10.1017/thg.2016.101

Sha, Z., Schijven, D., Carrion-Castillo, A., Joliot, M., Mazoyer, B., Fisher, S. E., . . . Francks, C. (2020). The genetic architecture of structural left-right asymmetry of the human brain. *bioRxiv*, 2020.2006.2030.179721. doi:10.1101/2020.06.30.179721

Šimkovic, M., & Träuble, B. (2019). Robustness of statistical methods when measure is affected by ceiling and/or floor effect. *PLoS One, 14*(8), e0220889. doi:10.1371/journal.pone.0220889

Smith, S. M., & Nichols, T. E. (2018). Statistical Challenges in "Big Data" Human Neuroimaging. *Neuron, 97*(2), 263-268. doi:10.1016/j.neuron.2017.12.018

Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., . . . Collins, R. (2015). UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med, 12*(3), e1001779. doi:10.1371/journal.pmed.1001779

Sullivan, P. F., Agrawal, A., Bulik, C. M., Andreassen, O. A., Børglum, A. D., Breen, G., . . . O'Donovan, M. C. (2018). Psychiatric Genomics: An Update and an Agenda. *Am J Psychiatry, 175*(1), 15-27. doi:10.1176/appi.ajp.2017.17030283

Szucs, D., & Ioannidis, J. P. (2020). Sample size evolution in neuroimaging research: An evaluation of highly-cited studies (1990-2012) and of latest practices (2017-2018) in high-impact journals. *Neuroimage, 221*, 117164. doi:10.1016/j.neuroimage.2020.117164

Tee, Y. H., Shemesh, T., Thiagarajan, V., Hariadi, R. F., Anderson, K. L., Page, C., . . . Bershadsky, A. D. (2015). Cellular chirality arising from the self-organization of the actin cytoskeleton. *Nat Cell Biol, 17*(4), 445-457. doi:10.1038/ncb3137

Thompson, P. M., Jahanshad, N., Ching, C. R. K., Salminen, L. E., Thomopoulos, S. I., Bright, J., . . . Zelman, V. (2020). ENIGMA and global neuroscience: A decade of large-scale studies of the brain in health and disease across more than 40 countries. *Transl Psychiatry, 10*(1), 100. doi:10.1038/s41398-020-0705-1

Turner, S., Armstrong, L. L., Bradford, Y., Carlson, C. S., Crawford, D. C., Crenshaw, A. T., . . . Ritchie, M. D. (2011). Quality control procedures for genome-wide association studies. *Curr Protoc Hum Genet, Chapter 1*, Unit1.19. doi:10.1002/0471142905.hg0119s68

van Rooij, D., Anagnostou, E., Arango, C., Auzias, G., Behrmann, M., Busatto, G. F., . . . Buitelaar, J. K. (2018). Cortical and Subcortical Brain Morphometry Differences Between Patients With Autism Spectrum Disorder and Healthy Individuals Across the Lifespan: Results From the ENIGMA ASD Working Group. *Am J Psychiatry, 175*(4), 359-369. doi:10.1176/appi.ajp.2017.17010100

Varón-González, C., Pallares, L. F., Debat, V., & Navarro, N. (2019). Mouse Skull Mean Shape and Shape Robustness Rely on Different Genetic Architectures and Different Loci. *Front Genet, 10*, 64. doi:10.3389/fgene.2019.00064

Vingerhoets, G. (2019). Phenotypes in hemispheric functional segregation? Perspectives and challenges. *Phys Life Rev, 30*, 1-18. doi:10.1016/j.plrev.2019.06.002

Watkins, K. E., Paus, T., Lerch, J. P., Zijdenbos, A., Collins, D. L., Neelin, P., . . . Evans, A. C. (2001). Structural asymmetries in the human brain: a voxel-based statistical analysis of 142 MRI scans. *Cereb Cortex, 11*(9), 868-877.

Westfall, J., & Yarkoni, T. (2016). Statistically Controlling for Confounding Constructs Is Harder than You Think. *PLoS One, 11*(3), e0152719. doi:10.1371/journal.pone.0152719

Wiberg, A., Ng, M., Al Omran, Y., Alfaro-Almagro, F., McCarthy, P., Marchini, J., . . . Furniss, D. (2019). Handedness, language areas and neuropsychiatric diseases: insights from brain imaging and genetics. *Brain*. doi:10.1093/brain/awz257

Xu, Q., Guo, L., Cheng, J., Wang, M., Geng, Z., Zhu, W., . . . Yu, C. (2020). CHIMGEN: a Chinese imaging genetics cohort to enhance cross-ethnic and cross-geographic brain research. *Mol Psychiatry, 25*(3), 517-529. doi:10.1038/s41380-019-0627-6

Yamashita, K., Yoshiura, T., Hiwatashi, A., Noguchi, T., Togao, O., Takayama, Y., . . . Honda, H. (2011). Volumetric asymmetry and differential aging effect of the human caudate nucleus in normal individuals: a prospective MR imaging study. *J Neuroimaging, 21*(1), 34-37. doi:10.1111/j.1552-6569.2009.00403.x

Yassin, W., Nakatani, H., Zhu, Y., Kojima, M., Owada, K., Kuwabara, H., . . . Koike, S. (2020). Machine-learning classification using neuroimaging data in schizophrenia, autism, ultra-high risk and first-episode psychosis. *Transl Psychiatry, 10*(1), 278. doi:10.1038/s41398-020-00965-5

Research data management

The research in this thesis has been carried out under the institute research data management policy of the Max Planck Institute for Psycholinguistics (https://www.mpi.nl/page/research-data-management).

Chapter 1:
No data were produced.

Chapter 2:
This study made use of 54 separate data sets collected around the world, under a variety of different consent procedures and regulatory bodies, during the past 25 years. Requests to access the data sets will be considered in relation to the relevant consents, rules and regulations, and can be made via the ENIGMA consortium's ASD working group:
http://enigma.ini.usc.edu/ongoing/enigma-asd-working-group/

Chapter 3:
This study made use of 39 separate data sets collected around the world, under a variety of different consent procedures and regulatory bodies, during the past 25 years. Requests to access the data sets will be considered in relation to the relevant consents, rules and regulations, and can be made via the ENIGMA consortium's ADHD working group:
http://enigma.ini.usc.edu/ongoing/enigma-adhd-working-group/

Chapter 4:
The individual-level genomic data are archived permanently at the following restricted-access location: https://archive.mpi.nl/, under archive/language and genetics/situs inversus/
Requests to access these data will be considered in relation to the consents, relevant rules and regulations, and can be made via study leader Clyde Francks: clyde.francks@mpi.nl

Chapter 5:
The data used for this chapter were downloaded from the UK Biobank as part of research application 16066, with Clyde Francks as principal applicant. See

https://www.ukbiobank.ac.uk/data-management for the data management policies of the UK Biobank.

Chapter 6:
No data were produced.

English Summary

Left-right asymmetries are observed throughout nature, and humans are no exception. While the human body appears roughly symmetrical from the outside, the inside shows left-right differences, most strikingly of the visceral organs (e.g. heart left, liver right). Additionally, the human brain is asymmetric in both structure and function. For example, a left-right skewing of the two hemispheres can be observed in the brain's structure, and a greater activity of the left hemisphere during language tasks is found when studying the brain's function. Functional asymmetries in the brain can also manifest themselves in behaviour. A well-known example hereof is that roughly 90% of people are right handed.

Genes form the instructions for development, but little is known about the genetic basis of human brain asymmetry. Identifying genes involved in brain asymmetry will increase our understanding of the fundamental principles of left-right patterning of this organ. In addition, individuals with neurodevelopmental disorders such as attention-deficit/hyperactivity disorder (ADHD), or autism spectrum disorder (ASD), have lower rates of right-handedness compared to the general population, and there have been reports of altered brain asymmetry in these individuals. Altered brain asymmetry might therefore be a risk factor for some neurodevelopmental disorders, and identifying genes involved in brain asymmetry may yield new insights into the biological mechanisms of these disorders.

Prior genetic studies have shown that *Nodal*, *Lefty* and *Pitx2* are key genes in establishing visceral organ asymmetry in vertebrates. Asymmetric expression of these genes is influenced by a leftward fluid flow within a pitted structure on the surface of very early embryos, called the embryonic node. The asymmetric flow is generated by motile cilia which always rotate in a clockwise direction due to their intrinsic chirality (molecular handedness). This is because all life on earth uses amino acids with only one of two possible mirror forms, to construct proteins. As a result, cells can also exhibit chirality of shape or structure, which can drive the creation of organ asymmetries in

invertebrates, including fruit flies and snails. The extent to which such mechanisms are involved in human brain asymmetries remains unknown.

Human brain asymmetry is highly variable between individuals. All of the studies in this thesis made use of this variation, either to characterize associations of brain asymmetries with neurodevelopmental disorders, or to gain new insights into the biological underpinnings of left-right asymmetries.

In the first two studies of this thesis (chapters 2 and 3) I examined differences of brain structural asymmetry between healthy individuals and people with ASD (chapter 2) or ADHD (chapter 3). These studies were conducted via the ENIGMA consortium, a global collaborative network of researchers, which enables large-scale analyses by combining many datasets from research groups around the world. The aggregation of data allowed for mixed-effects mega-analysis, a statistically powerful approach to detecting and measuring effects. These were the largest ever studies of altered brain asymmetry in either of these disorders. In ASD, significant case-control group differences in asymmetry were observed in various brain regions, particularly the grey matter thickness of the medial orbitofrontal cortex, rostral anterior cingulate cortex, and superior frontal cortex. There was an overall tendency towards reduced asymmetry in ASD compared to controls, although the effects were subtle and region-specific. For ADHD, case-control average differences in asymmetry were non-significant, apart from an association between ADHD and globus pallidus asymmetry, which was found to be more symmetric in adults with ADHD. Overall, these results demonstrate that alterations of structural brain asymmetries are very subtle in ASD and even more so in ADHD. Therefore, brain anatomical asymmetry is not likely to be a useful biomarker for either disorder. However, the findings point to an aspect of brain development that may be altered in ASD particularly, and the effects have been mapped and quantified across the brain. Further studies of brain asymmetry in neurodevelopmental disorders, at the microstructural level, are motivated by the findings.

In an attempt to link variation in brain asymmetry to asymmetry of the visceral organs, I examined whole genome sequencing data from fifteen people with the extremely rare condition of *situs inversus* (SI) (chapter 4),

which features a complete mirror-reversal of visceral organ asymmetry. This condition occurs sometimes together with Primary Ciliary Dyskinesia (PCD), a recessive genetic disorder that impairs motility of the cilia. Five out of nine SI individuals without PCD were naturally left-handed, which is more than expected based on a 10% rate of left-handedness at the population-level. This study therefore had the potential to identify genes with shared effects on both visceral and brain asymmetries, or in other words, to genetically link body and brain asymmetries. I identified clear monogenic causes for the SI individuals with PCD, in genes already known to cause this disorder (i.e., *DNAH5*, *DNAH11*, *DNAAF1*, *CCDC114* and *LRRC6*). However, the left-handed SI individuals without PCD remained largely unsolved by the genome sequencing analysis, such that a clear monogenic link between visceral organ asymmetry and left-handedness was not found. Non-genetic factors, such as environmental influences or random variation during development, are therefore likely to play a key role in SI without PCD.

In chapter 5, I studied both rare and common genetic variation in the population, in relation to hemispheric asymmetry of functional brain networks in healthy adults from the UK Biobank. This analysis was based on functional magnetic resonance imaging (fMRI) data acquired during the resting state, i.e. when people were not required to perform any particular cognitive task during fMRI scanning. Even in the resting state, activity continues within networks of interconnected regions that reflect the brain's functional organization. I focused on two network metrics, one of which measured functional asymmetry at the whole hemisphere level, and the other based specifically on regions that are important for language. Both were previously shown to correlate with hemispheric language dominance assessed during task-based functional imaging. In my study I also found that both phenotypes were associated with handedness. Both metrics were weakly heritable, but despite having a large sample size of more than 30,000 individuals, no single common variants were significantly associated with either metric at a level which survived correction for multiple testing of millions of genetic variants distributed over the whole genome. I also analyzed whole exome sequences from approx. 16,000 individuals, to test for an association with rare genetic variants, but here again no significant associations were found.

In sum, in this thesis I have characterized the extent of altered structural brain asymmetry in two neurodevelopmental disorders, and advanced our knowledge and understanding of the genetic contributions to brain asymmetry, which is an important aspect of healthy brain organization.

Nederlandse samenvatting

Links-rechts asymmetrie is overal in de natuur aanwezig en de mens is geen uitzondering. Alhoewel het menselijk lichaam vrijwel symmetrisch lijkt vanaf de buitenkant, toont de binnenkant wel degelijk links-rechts verschillen, het meest duidelijk te herkennen aan de positie van de viscerale organen (bv. hart links, lever rechts). Daarnaast is het menselijk brein asymmetrisch in zowel structuur als functie. Zo zijn de hersenen lichtelijk gedraaid naar links of rechts, en tonen ze een grotere mate van activiteit in de linker hersenhelft tijdens het uitvoeren van taal-gerelateerde taken. Asymmetrie met betrekking tot de hersenactiviteit wordt ook wel functionele asymmetrie genoemd en deze kan tot uiting komen in het gedrag. Een bekend voorbeeld hiervan is dat 90% van de mensen rechtshandig is.

Genen vormen de instructies voor ontwikkeling, maar er is nog weinig bekend over de genetische basis van de menselijke hersenasymmetrie. Het identificeren van genen die invloed hebben op hersenasymmetrie zal bijdragen aan onze kennis van de fundamentele principes voor links-rechts vorming in dit orgaan. Bovendien hebben individuen met een neurologische ontwikkelingsstoornis als attenion-deficit/hyperactivity disorder (ADHD) of autism spectrum disorder (ASD) een verlaagde mate van rechtshandigheid in vergelijking met de algemene populatie, en hebben verscheidene verslagen benoemd dat de hersenasymmetrie anders is in deze individuen. Een verandering van hersenasymmetrie kan daarom worden gezien als een risicofactor voor sommige neurologische ontwikkelingstoornissen, en het identificeren van genen die betrekking hebben op hersenasymmetrie kunnen dan ook tot nieuwe inzichten leiden met betrekking tot de biologische mechanismen onderliggend aan deze stoornissen.

Voorafgaande genetische studies hebben aangetoond dat *Nodal*, *Lefty* en *Pitx2* belangrijke genen zijn voor het tot stand brengen van de viscerale orgaan asymmetrie in gewervelden. Asymmetrische expressie van deze genen wordt beinvloed door een naar links stromende vloeistofstroom in de embryonale knoop, een put-achtige structuur die te vinden is op het oppervlak van het vroege embryo. Deze asymmetrische stroming wordt gegenereerd door bewegende trilharen (cilia), welke altijd met de klok mee roteren

vanwege hun intrinsieke chiraliteit (hetgeen te vergelijken is met links- of rechtshandigheid van moleculen). Dit is omdat al het leven op aarde bij het bouwen van eiwitten enkel en alleen gebruik maakt van één van de twee mogelijke spiegelvormen van aminozuren. Het resultaat hiervan is dat cellen ook chiraliteit kunnen vertonen, namelijk in vorm of structuur, hetgeen de creatie van orgaan asymmetrie kan aansturen in gewervelden, waaronder fruitvliegjes en slakken. De mate waarin zulke mechanismen invloed hebben op de menselijke hersenasymmetrie blijft vooralsnog onbekend.

Menselijke hersenasymmetrie is zeer variabel tussen personen. Alle onderzoeken in dit proefschrift hebben gebruik gemaakt van deze variatie, ofwel om associaties vast te stellen tussen hersenasymmetrie en neurologische ontwikkelingsstoornissen, ofwel om nieuwe inzichten te krijgen in de biologische theorie onderliggend aan links-rechts asymmetrie.

In de eerste twee onderzoeken in dit proefschrift (hoofdstuk 2 en 3) onderzocht ik de verschillen in structuur van hersenasymmetry tussen gezonde personen en personen met ASD (hoofdstuk 2) of ADHD (hoofdstuk 3). Deze onderzoeken waren uitgevoerd via het ENIGMA consortium, een wereldwijd samenwerkingsverband tussen onderzoekers, wat het mogelijk maakte om groot-schalig onderzoek uit te voeren door de vele datasets van onderzoeksgroepen over de hele wereld samen te voegen. Het samenvoegen van data gaf de gelegenheid tot het uitvoeren van mixed-effects mega-analyse, een statistisch krachtige aanpak voor het detecteren en meten van effecten. Dit waren de grootste onderzoeken ooit gedaan naar veranderingen van hersenasymmetrie in elk van deze aandoeningen.

In ASD werden significante case-control groepverschillen in asymmetrie gevonden in verscheidene hersengebieden, met name in de dikte van de grijze stof van de medial orbitofrontale cortex, de rostral anterior cingulate cortex, en de superior frontale cortex. Er was een algemene tendens te zien van verminderde asymmetrie in patiënten met ASD in vergelijking met controlepersonen, alhoewel de effecten erg klein en regio-specifiek waren. Voor ADHD waren de gemiddele case-control verschillen in asymmetrie non-significant, buiten een associatie tussen ADHD en globus pallidus asymmetrie, welke meer symmetrisch was in volwassenen met ADHD. Over het algemeen laten deze resultaten zien dat veranderingen in asymmetrie van

de hersenstructuur zeer subtiel zijn in ASD en zelfs nog subtieler in ADHD. Daardoor is structurele hersenasymmetrie waarschijnlijk geen goede biomarker voor elk van deze aandoeningen. Echter wijzen de resultaten naar een aspect van de hersenontwikkeling dat specifiek veranderd zou kunnen zijn in ASD, en de effecten zijn in kaart gebracht en gekwantificeerd over de gehele hersenen. Verder onderzoek naar hersenasymmetrie in neurologische aandoeningen, op microstructureel niveau, wordt aangemoedigd door deze resultaten.

In een poging om variatie in hersenasymmetrie te linken aan variatie in asymmetrie van de viscerale organen, heb ik onderzoek gedaan naar DNA sequenties van het hele genoom (whole genome sequences) van 15 mensen met de extreem zeldzame aandoening *situs inversus* (SI) (hoofdstuk 4), gekenmerkt door een algehele spiegeling van de viscerale orgaan asymmetrie. Deze aandoening komt soms samen voor met Primary Ciliary Dyskinesia (PCD), een recessieve genetische ziekte die een immobiliteit van de bewegende cilia veroorzaakt. Vijf van de negen SI individuen zonder PCD waren van nature links-handig. Dit is meer dan verwacht gebaseerd op het feit dat 10% van de algemene populatie linkshandig is. Dit onderzoek had dan ook de mogelijkheid om genen te identificeren die een gezamelijk effect hebben op de asymmetrie van zowel de hersenen als de viscerale organen, ofwel om lichaams- en hersenasymmetrie genetisch aan elkaar te linken. Ik heb duidelijke monogenetische oorzaken gevonden voor de SI individuen met PCD, in genen waarvan al bekend was dat deze PCD veroozaken (namelijk *DNAH5*, *DNAH11*, *DNAAF1*, *CCDC114* en *LRRC6*). De linkshandige individuen met SI zonder PCD bleven echter grotendeels onopgelost door het analyseren van hun genoomsequenties, zodat een duidelijke monogenetische link tussen viscerale orgaan asymmetrie en linkshandigheid uitbleef. Niet-genetische factoren, zoals invloeden vanuit de omgeving of willekeurige variatie tijdens de ontwikkeling, spelen daardoor waarschijnlijk een belangrijke rol in SI zonder PCD.

In hoofdstuk 5 heb ik onderzoek gedaan naar zeldzame alsmede veelvoorkomende genetische varianten in de populatie, in relatie tot hersenhelft asymmetrie van de functionele netwerken in gezonde volwassenen van de UK Biobank. Dit onderzoek was gebasseerd op

functionele MRI (fMRI) data die was verkregen tijdens de rusttoestand (dat is wanneer mensen geen taak hoeven uit te voeren tijdens de fMRI scan). Zelfs in de rusttoestand zijn de hersenen actief, en deze activiteit wordt gekoppeld aan netwerken van regio's die met elkaar verbonden zijn en zodoende de functionele organisatie van de hersenen reflecteren. Ik heb specifiek gekeken naar twee netwerk maten, waarvan de een een reflectie was van de functionele asymmetrie op het niveau van de gehele hersenhelft, terwijl de ander was gebasseerd op gebieden die specifiek betrokken zijn bij taal. Beide maten toonden eerder al een correlatie met hersenhelft dominantie voor taal, gemeten tijdens taak-gebaseerde fMRI. In mijn onderzoek vond ik dat beide fenotypes geassocieerd waren met hand voorkeur, waarbij rechtshandigen over het algemeen meer symmetrie vertoonden in de intrinsieke functionele connectiviteit dan linkshandigen. Beide maten waren lichtelijk erfelijk, maar ondanks de grote steekproefomvang van meer dan 30,000 individuen werden er geen significante associaties gevonden tussen veelvoorkomende genetische varianten en een van beide maten, waarbij een meervoudige testcorrectie was toegepast voor het testen van miljoenen varianten verspreid over het hele genoom. Ik heb ook gehele exoom sequenties geanalyseerd van ongeveer 16,000 individuen, om zo de associatie met zeldzame genetische varianten te testen, maar ook hier werden geen significante associaties gevonden.

Samengevat heb ik in dit proefschrift de mate van veranderingen in structurele hersenasymmetrie in twee neurologische stoornissen geanalyseerd, en heb ik onze kennis en ons begrip van de genetische bijdragen aan hersenasymmetrie vergroot, welke een belangrijk aspect vormen van de organisatie van gezonde hersenen.

Biography

Merel Charlotte Postema was born in Soest, the Netherlands, on October 30, 1991. She went to high school at het Baarnsch Lyceum in Baarn, where she graduated in 2010. Merel continued her education at VU University Amsterdam, where she obtained a Bachelor's degree in Health and Life Sciences in 2014, followed by a research Master's degree in Neurosciences in 2016. As part of her Master's internship she moved to Sheffield, UK, for 6 months, to study the reliability of the default mode network in healthy elderly and elderly with mild cognitive impairment and Alzheimer's disease. In 2017, Merel started working as a PhD student at the Max Planck Institute for Psycholinguistics in Nijmegen, the Netherlands, where she worked within the Language and Genetics department, led by Prof. Simon E. Fisher, as a member of Dr. Clyde Francks' Imaging Genomics research group. Here, Merel studied left-right asymmetry of the human brain and its associations with neurodevelopmental disorders and genetic factors.

Publication list

Kong X**., Postema M.C.,** Carrion-Castillo A., Pepe A., Crivello F., Joliot M., Mazoyer B., Fisher S.E., Francks C., Handedness and other variables associated with human brain asymmetrical skew, *Cereb Cortex*. 2021 Apr 9. doi: 10.1093/cercor/bhab075. Epub ahead of print. PMID: 33836062.

Postema M.C., Hoogman M., Ambrosino S., Asherson P., Banaschewski T., Bandeira C. E., Baranov A., Bau C.H.D. ,Baumeister S., Baur-Streubel R. , Bellgrove M. A., Biederman J., Bralten J., Brandeis D., Brem S., Buitelaar J. K., Busatto G.F., Castellanos F.X., Cercignani M., Chaim-Avancini T.M., Chantiluke K.C., Christakou A., Coghill D., Conzelmann A., Cubillo A. I., Cupertino R. B. ,de Zeeuw P., Doyle A. E., Durston S., Earl E. A. ,Epstein J. N., Ethofer T., Fair D. A., Fallgatter A.J., Faraone S.V., Frodl T., Gabel M. C., Gogberashvili T., Grevet E.H., Haavik J., Harrison N.A., Hartman C. A., Heslenfeld D.J., Hoekstra P.J., Hohmann S., Hövik M.F., Jernigan T. L., Kardatzki B., Karkashadze G., Kelly C., Kohls G., Konrad K., Kuntsi J., Lazaro L., Lera-Miguel S., Lesch K., Louza M. R., Lundervold A. J., Malpas C. B, Mattos P. , McCarthy H. , Namazova-Baranova L., Nicolau R., Nigg J. T, Novotny S. E., Oberwelland Weiss E., O'Gorman Tuura R. L., Oosterlaan J. , Oranje B.,Paloyelis Y. , Pauli P. ,Picon F.A., Plessen K. J., Ramos-Quiroga J. , Reif A. ,Reneman L., Rosa P.G.P., Rubia K., Schrantee A., Schweren L.J.S., Seitz J., Shaw P., Silk T. J., Skokauskas N., Soliva Vila J., Stevens M.C., Sudre G.,Tamm L.,Tovar-Moll F.,van Erp T.G.M., Vance A., Vilarroya O.,Vives-Gilabert Y. ,von Polier G.G., Walitza S., Yoncheva Y. N., Zanetti M. V., Ziegler G. C., Glahn D.C., Jahanshad N., Medland S.E., Thompson P.M., Fisher S.E., Franke B., Francks C., Analysis of structural brain asymmetries in Attention-Deficit/Hyperactivity Disorder in 39 datasets, *J Child Psychol Psychiatry*. 2021 Mar 22. doi: 10.1111/jcpp.13396. Epub ahead of print. PMID: 33748971.

Hoogman M, van Rooij D, Klein M, Boedhoe P, Ilioska I, Li T, Patel Y, **Postema MC**, Zhang-James Y, Anagnostou E, Arango C, Auzias G, Banaschewski T, Bau CHD, Behrmann M, Bellgrove MA, Brandeis D, Brem S, Busatto GF, Calderoni S, Calvo R, Castellanos FX, Coghill D, Conzelmann

A, Daly E, Deruelle C, Dinstein I, Durston S, Ecker C, Ehrlich S, Epstein JN, Fair DA, Fitzgerald J, Freitag CM, Frodl T, Gallagher L, Grevet EH, Haavik J, Hoekstra PJ, Janssen J, Karkashadze G, King JA, Konrad K, Kuntsi J, Lazaro L, Lerch JP, Lesch KP, Louza MR, Luna B, Mattos P, McGrath J, Muratori F, Murphy C, Nigg JT, Oberwelland-Weiss E, O'Gorman Tuura RL, O'Hearn K, Oosterlaan J, Parellada M, Pauli P, Plessen KJ, Ramos-Quiroga JA, Reif A, Reneman L, Retico A, Rosa PGP, Rubia K, Shaw P, Silk TJ, Tamm L, Vilarroya O, Walitza S, Jahanshad N, Faraone SV, Francks C, van den Heuvel OA, Paus T, Thompson PM, Buitelaar JK, Franke B. Consortium neuroscience of attention deficit/hyperactivity disorder and autism spectrum disorder: The ENIGMA adventure. *Hum Brain Mapp*. 2020 May 18. doi: 10.1002/hbm.25029. Epub ahead of print. PMID: 32420680.

Kong XZ, **Postema MC**, Guadalupe T, de Kovel C, Boedhoe PSW, Hoogman M, Mathias SR, van Rooij D, Schijven D, Glahn DC, Medland SE, Jahanshad N, Thomopoulos SI, Turner JA, Buitelaar J, van Erp TGM, Franke B, Fisher SE, van den Heuvel OA, Schmaal L, Thompson PM, Francks C. Mapping brain asymmetry in health and disease through the ENIGMA consortium. *Hum Brain Mapp*. 2020 May 18. doi: 10.1002/hbm.25033. Epub ahead of print. PMID: 32420672.

Thompson PM, Jahanshad N, Ching CRK, Salminen LE, Thomopoulos SI, Bright J, Baune BT, Bertolín S, Bralten J, Bruin WB, Bülow R, Chen J, Chye Y, Dannlowski U, de Kovel CGF, Donohoe G, Eyler LT, Faraone SV, Favre P, Filippi CA, Frodl T, Garijo D, Gil Y, Grabe HJ, Grasby KL, Hajek T, Han LKM, Hatton SN, Hilbert K, Ho TC, Holleran L, Homuth G, Hosten N, Houenou J, Ivanov I, Jia T, Kelly S, Klein M, Kwon JS, Laansma MA, Leerssen J, Lueken U, Nunes A, Neill JO, Opel N, Piras F, Piras F, **Postema MC**, Pozzi E, Shatokhina N, Soriano-Mas C, Spalletta G, Sun D, Teumer A, Tilot AK, Tozzi L, van der Merwe C, Van Someren EJW, van Wingen GA, Völzke H, Walton E, Wang L, Winkler AM, Wittfeld K, Wright MJ, Yun JY, Zhang G, Zhang-James Y, Adhikari BM, Agartz I, Aghajani M, Aleman A, Althoff RR, Altmann A, Andreassen OA, Baron DA, Bartnik-Olson BL, Marie Bas-Hoogendam J, Baskin-Sommers AR, Bearden CE, Berner LA, Boedhoe PSW, Brouwer RM, Buitelaar JK, Caeyenberghs K, Cecil CAM, Cohen RA, Cole JH, Conrod PJ, De Brito SA, de Zwarte SMC, Dennis EL,

Desrivieres S, Dima D, Ehrlich S, Esopenko C, Fairchild G, Fisher SE, Fouche JP, Francks C, Frangou S, Franke B, Garavan HP, Glahn DC, Groenewold NA, Gurholt TP, Gutman BA, Hahn T, Harding IH, Hernaus D, Hibar DP, Hillary FG, Hoogman M, Hulshoff Pol HE, Jalbrzikowski M, Karkashadze GA, Klapwijk ET, Knickmeyer RC, Kochunov P, Koerte IK, Kong XZ, Liew SL, Lin AP, Logue MW, Luders E, Macciardi F, Mackey S, Mayer AR, McDonald CR, McMahon AB, Medland SE, Modinos G, Morey RA, Mueller SC, Mukherjee P, Namazova-Baranova L, Nir TM, Olsen A, Paschou P, Pine DS, Pizzagalli F, Rentería ME, Rohrer JD, Sämann PG, Schmaal L, Schumann G, Shiroishi MS, Sisodiya SM, Smit DJA, Sønderby IE, Stein DJ, Stein JL, Tahmasian M, Tate DF, Turner JA, van den Heuvel OA, van der Wee NJA, van der Werf YD, van Erp TGM, van Haren NEM, van Rooij D, van Velzen LS, Veer IM, Veltman DJ, Villalon-Reina JE, Walter H, Whelan CD, Wilde EA, Zarei M, Zelman V; ENIGMA Consortium. ENIGMA and global neuroscience: A decade of large-scale studies of the brain in health and disease across more than 40 countries. *Transl Psychiatry.* 2020 Mar 20;10(1):100. doi: 10.1038/s41398-020-0705-1. PMID: 32198361; PMCID: PMC7083923.

Postema MC, Carrion-Castillo A, Fisher SE, Vingerhoets G, Francks C. The genetics of situs inversus without primary ciliary dyskinesia. *Sci Rep.* 2020 Feb 28;10(1):3677. doi: 10.1038/s41598-020-60589-z. PMID: 32111882; PMCID: PMC7048929.

Postema MC, van Rooij D, Anagnostou E, Arango C, Auzias G, Behrmann M, Filho GB, Calderoni S, Calvo R, Daly E, Deruelle C, Di Martino A, Dinstein I, Duran FLS, Durston S, Ecker C, Ehrlich S, Fair D, Fedor J, Feng X, Fitzgerald J, Floris DL, Freitag CM, Gallagher L, Glahn DC, Gori I, Haar S, Hoekstra L, Jahanshad N, Jalbrzikowski M, Janssen J, King JA, Kong XZ, Lazaro L, Lerch JP, Luna B, Martinho MM, McGrath J, Medland SE, Muratori F, Murphy CM, Murphy DGM, O'Hearn K, Oranje B, Parellada M, Puig O, Retico A, Rosa P, Rubia K, Shook D, Taylor MJ, Tosetti M, Wallace GL, Zhou F, Thompson PM, Fisher SE, Buitelaar JK, Francks C. Altered structural brain asymmetry in autism spectrum disorder in a study of 54 datasets. *Nat Commun.* 2019 Oct 31;10(1):4958. doi: 10.1038/s41467-019-13005-8. PMID: 31673008; PMCID: PMC6823355.

Postema MC, De Marco M, Colato E, Venneri A. A study of within-subject reliability of the brain's default-mode network. *MAGMA*. 2019 Jun;32(3):391-405. doi: 10.1007/s10334-018-00732-0. Epub 2019 Feb 7. PMID: 30730023; PMCID: PMC6525123.

Acknowledgements

Clyde, my first word of thanks goes, of course, to you. You were a great co-supervisor, giving me enough space to work independently and enough guidance when I needed to. Apart from our weekly team/project meetings, I could contact you with any question/issue, and you would reply very quickly. With your keen eye you could spot any abnormalities in the data/figures that I hadn't seen myself, ensuring good quality of the results, and I thank you for that. Also thank you for editing my draft manuscripts (and the current thesis) and making sure that important aspects were included in a logical order. Finally, I am grateful for the opportunities you gave me to expand my (international) network, which I really appreciated.

Simon, thank you for giving me the opportunity to work as a PhD student in your department, and for providing input and feedback on my manuscripts/thesis/department talks. I very much valued the opportunities to participate in international conferences and courses. Also, thank you for the yearly department outings and the New Year's dinner parties at your house, which were great.

My thanks also goes to the members of the Doctoral Thesis Committee, namely prof. Guillén Fernandez, prof. Chris McManus and Dr. Emma Sprooten, for their time and willingness to review this thesis.

Also many thanks to my collaborators. For the ENIGMA projects I had the pleasure to work with many researchers around the world. My thanks goes especially to Paul Thompson, the director of the ENIGMA consortium, as well as to Daan van Rooij and Jan Buitelaar, the leaders of the ENIGMA-ASD working group, and to Martine Hoogman and Barbara Franke, the leaders of the ENIGMA-ADHD working group. For the *situs inversus* project, I would like to thank Guy Vingerhoets for recruiting the participants and obtaining their DNA samples, and for providing feedback on the manuscript. Finally, I would like to thank the Bordeaux group, including Marc Joliot, Antonietta Pepe, Fabrice Crivello, Nathalie Tzourio-Mazoyer, and Bernard Mazoyer, for their collaboration on the functional connectivity asymmetry

project. Also many thanks for the warm welcome I received in Bordeaux when meeting there.

Some of my collaborators were close colleagues/friends, namely Amaia Carrion-Castillo, Dick Schijven and Xiangzhen Kong. Amaia, you have been my support and help when I just started my PhD. In the beginning I had no experience with coding or script writing whatsoever, but you were there to teach me. Even later, when I had learnt how to code, you continued being there for me when I encountered an unknown scripting error or when I didn't understand something. Thank you so much for your time, patience, and help throughout the first years of my PhD. Also many thanks to Dick for his generous help with the UK Biobank exome data. Dick, thank you for sharing and explaining your clearly written scripts, and for your clear answers on any of my genetics questions. It was great to work with you during the last years of my PhD. Finally my thanks goes to Xiangzhen, who was my colleague for almost the entire four years of my PhD trajectory. Xiangzhen, thank you, among others, for your inputs on the ENIGMA projects, for computing the functional connectivity matrices from Chapter 5, and of course for the nice collaboration!

Other colleagues/friends I would like to thank are Barbara, Beate, Ellen, Else, Fenja, Marjolein, Martina, Zhiqiang and all other members of the Language & Genetics department. Thank you all for contributing to a great PhD experience! It is unfortunate we were not able to see each other much last year due to COVID-19. I have missed our lunches together, our occasional drinks at the cultuur cafe, dropping by your offices for a chat, and/or the monthly board game nights at Else's place. However, apart from the difficulties we all experienced, I can look back on a great time with a lot of good memories to cherish. Thank you all for that! A special thanks to Martina for being there to guide me before, during and after my PhD trajectory (dankjewel Martina!), as well as to my paranymphs, Ellen and Dick, for guiding me towards the PhD ceremony and making the last miles less heavy.

Also, to those who have left the MPI during my PhD, including Amanda, Carolien, Chin Yang, Lara, Margot, Rick and Tulio, thank you for the nice collaboration.

Also many thanks to the TG, including Tobias, for their support on technical issues and ensuring that the servers are running smoothly, and to all the members of team Operations for ensuring all operations are functioning well.

Thank you to the members of the committees/teams I was part of during my PhD, including the volleyball team (although briefly), the IMPRS conference 2018 organizing committee, the PhD representatives, and the emergency response team, for enriching my PhD experience.

Kevin Lam, thank you for your interest in my PhD progression and personal life, and for being the 'person-to-go-to' for almost any issue.

Verder wil ik mijn vriendinnen, familie en schoonfamilie bedanken voor hun steun en interesse in mijn PhD onderzoeken, maar vooral ook voor alle gezelligheid naast mijn PhD! Voldoende ontspanning is belangrijk, en jullie hebben daar zeker aan bijgedragen.

Tot slot dank aan mijn partner, Tim. Dankjewel voor al je steun en dat je me altijd hebt aangemoedigd om voldoende pauzes en vakantiedagen te nemen. Ik ben blij en dankbaar dat ik jou al 10 jaar in mijn leven heb. We hebben al vele mijlpalen met elkaar gedeeld en ik hoop dat er nog vele zullen volgen!

MPI series in Psycholinguistics

1. The electrophysiology of speaking: Investigations on the time course of semantic, syntactic, and phonological processing. *Miranda van Turennout*
2. The role of the syllable in speech production: Evidence from lexical statistics, metalinguistics, masked priming, and electromagnetic midsagittal articulography. *Niels O. Schiller*
3. Lexical access in the production of ellipsis and pronouns. *Bernadette M. Schmitt*
4. The open-/closed-class distinction in spoken-word recognition. *Alette Haveman*
5. The acquisition of phonetic categories in young infants: A self-organising artificial neural network approach. *Kay Behnke*
6. Gesture and speech production. *Jan-Peter de Ruiter*
7. Comparative intonational phonology: English and German. *Esther Grabe*
8. Finiteness in adult and child German. *Ingeborg Lasser*
9. Language input for word discovery. *Joost van de Weijer*
10. Inherent complement verbs revisited: Towards an understanding of argument structure in Ewe. *James Essegbey*
11. Producing past and plural inflections. *Dirk Janssen*
12. Valence and transitivity in Saliba: An Oceanic language of Papua New Guinea. *Anna Margetts*
13. From speech to words. *Arie van der Lugt*
14. Simple and complex verbs in Jaminjung: A study of event categorisation in an Australian language. *Eva Schultze-Berndt*
15. Interpreting indefinites: An experimental study of children's language comprehension. *Irene Krämer*
16. Language-specific listening: The case of phonetic sequences. *Andrea Weber*
17. Moving eyes and naming objects. *Femke van der Meulen*
18. Analogy in morphology: The selection of linking elements in Dutch compounds. *Andrea Krott*
19. Morphology in speech comprehension. *Kerstin Mauth*

20. Morphological families in the mental lexicon. *Nivja H. de Jong*
21. Fixed expressions and the production of idioms. *Simone A. Sprenger*
22. The grammatical coding of postural semantics in Goemai (a West Chadic language of Nigeria). *Birgit Hellwig*
23. Paradigmatic structures in morphological processing: Computational and cross-linguistic experimental studies. *Fermín Moscoso del Prado Martín*
24. Contextual influences on spoken-word processing: An electrophysiological approach. *Daniëlle van den Brink*
25. Perceptual relevance of prevoicing in Dutch. *Petra M. van Alphen*
26. Syllables in speech production: Effects of syllable preparation and syllable frequency. *Joana Cholin*
27. Producing complex spoken numerals for time and space. *Marjolein Meeuwissen*
28. Morphology in auditory lexical processing: Sensitivity to fine phonetic detail and insensitivity to suffix reduction. *Rachèl J. J. K. Kemps*
29. At the same time...: The expression of simultaneity in learner varieties. *Barbara Schmiedtová*
30. A grammar of Jalonke argument structure. *Friederike Lüpke*
31. Agrammatic comprehension: An electrophysiological approach. *Marlies Wassenaar*
32. The structure and use of shape-based noun classes in Miraña (North West Amazon). *Frank Seifart*
33. Prosodically-conditioned detail in the recognition of spoken words. *Anne Pier Salverda*
34. Phonetic and lexical processing in a second language. *Mirjam Broersma*
35. Retrieving semantic and syntactic word properties. *Oliver Müller*
36. Lexically-guided perceptual learning in speech processing. *Frank Eisner*
37. Sensitivity to detailed acoustic information in word recognition. *Keren B. Shatzman*
38. The relationship between spoken word production and comprehension. *Rebecca Özdemir*

39. Disfluency: Interrupting speech and gesture. *Mandana Seyfeddinipur*
40. The acquisition of phonological structure: Distinguishing contrastive from non-contrastive variation. *Christiane Dietrich*
41. Cognitive cladistics and the relativity of spatial cognition. *Daniel B.M. Haun*
42. The acquisition of auditory categories. *Martijn Goudbeek*
43. Affix reduction in spoken Dutch. *Mark Pluymaekers*
44. Continuous-speech segmentation at the beginning of language acquisition: Electrophysiological evidence. *Valesca Kooijman*
45. Space and iconicity in German Sign Language (DGS). *Pamela Perniss*
46. On the production of morphologically complex words with special attention to effects of frequency. *Heidrun Bien*
47. Crosslinguistic influence in first and second languages: Convergence in speech and gesture. *Amanda Brown*
48. The acquisition of verb compounding in Mandarin Chinese. *Jidong Chen*
49. Phoneme inventories and patterns of speech sound perception. *Anita Wagner*
50. Lexical processing of morphologically complex words: An information-theoretical perspective. *Victor Kuperman*
51. A grammar of Savosavo, a Papuan language of the Solomon Islands. *Claudia Wegener*
52. Prosodic structure in speech production and perception. *Claudia Kuzla*
53. The acquisition of finiteness by Turkish learners of German and Turkish learners of French: Investigating knowledge of forms and functions in production and comprehension. *Sarah Schimke*
54. Studies on intonation and information structure in child and adult German. *Laura de Ruiter*
55. Processing the fine temporal structure of spoken words. *Eva Reinisch*
56. Semantics and (ir)regular inflection in morphological processing. *Wieke Tabak*

57. Processing strongly reduced forms in casual speech. *Susanne Brouwer*
58. Ambiguous pronoun resolution in L1 and L2 German and Dutch. *Miriam Ellert*
59. Lexical interactions in non-native speech comprehension: Evidence from electro-encephalography, eye-tracking, and functional magnetic resonance imaging. *Ian FitzPatrick*
60. Processing casual speech in native and non-native language. *Annelie Tuinman*
61. *Split intransitivity in Rotokas, a Papuan language of Bougainville. Stuart Robinson*
62. Evidentiality and intersubjectivity in Yurakaré: An interactional account. *Sonja Gipper*
63. The influence of information structure on language comprehension: A neurocognitive perspective. *Lin Wang*
64. The meaning and use of ideophones in Siwu. *Mark Dingemanse*
65. The role of acoustic detail and context in the comprehension of reduced pronunciation variants. *Marco van de Ven*
66. Speech reduction in spontaneous French and Spanish. *Francisco Torreira*
67. The relevance of early word recognition: Insights from the infant brain. *Caroline Junge*
68. Adjusting to different speakers: Extrinsic normalization in vowel perception. *Matthias J. Sjerps*
69. Structuring language. Contributions to the neurocognition of syntax. *Katrien R. Segaert*
70. Infants' appreciation of others' mental states in prelinguistic communication: A second person approach to mindreading. *Birgit Knudsen*
71. Gaze behavior in face-to-face interaction. *Federico Rossano*
72. Sign-spatiality in Kata Kolok: how a village sign language of Bali inscribes its signing space. *Conny de Vos*
73. Who is talking? Behavioural and neural evidence for norm-based coding in voice identity learning. *Attila Andics*
74. Lexical processing of foreign-accented speech: Rapid and flexible adaptation. *Marijt Witteman*

75. The use of deictic versus representational gestures in infancy. *Daniel Puccini*

76. Territories of knowledge in Japanese conversation. *Kaoru Hayano*

77. Family and neighbourhood relations in the mental lexicon: A cross-language perspective. *Kimberley Mulder*

78. Contributions of executive control to individual differences in word production. *Zeshu Shao*

79. Hearing speech and seeing speech: Perceptual adjustments in auditory-visual processing. *Patrick van der Zande*

80. High pitches and thick voices: The role of language in space-pitch associations. *Sarah Dolscheid*

81. Seeing what's next: Processing and anticipating language referring to objects. *Joost Rommers*

82. Mental representation and processing of reduced words in casual speech. *Iris Hanique*

83. The many ways listeners adapt to reductions in casual speech. *Katja Poellmann*

84. Contrasting opposite polarity in Germanic and Romance languages: Verum Focus and affirmative particles in native speakers and advanced L2 learners. *Giuseppina Turco*

85. Morphological processing in younger and older people: Evidence for flexible dual-route access. *Jana Reifegerste*

86. Semantic and syntactic constraints on the production of subject-verb agreement. *Alma Veenstra*

87. The acquisition of morphophonological alternations across languages. *Helen Buckler*

88. The evolutionary dynamics of motion event encoding. *Annemarie Verkerk*

89. Rediscovering a forgotten language. *Jiyoun Choi*

90. The road to native listening: Language-general perception, language-specific input. *Sho Tsuji*

91. Infants'understanding of communication as participants and observers. *Gudmundur Bjarki Thorgrímsson*

92. Information structure in Avatime. *Saskia van Putten*

93. Switch reference in Whitesands. *Jeremy Hammond*

94. Machine learning for gesture recognition from videos. *Binyam Gebrekidan Gebre*
95. Acquisition of spatial language by signing and speaking children: a comparison of Turkish sign language (TID) and Turkish. *Beyza Sümer*
96. An ear for pitch: on the effects of experience and aptitude in processing pitch in language and music. *Salomi Savvatia Asaridou*
97. lncrementality and Flexibility in Sentence Production. *Maartje van de Velde*
98. Social learning dynamics in chimpanzees: Reflections on (nonhuman) animal culture. *Edwin van Leeuwen*
99. The request system in Italian interaction. *Giovanni Rossi*
100. Timing turns in conversation: A temporal preparation account. *Lilla Magyari*
101. Assessing birth language memory in young adoptees. *Wencui Zhou*
102. A social and neurobiological approach to pointing in speech and gesture. *David Peeters*
103. Investigating the genetic basis of reading and language skills. *Alessandro Gialluisi*
104. Conversation Electrified: The Electrophysiology of Spoken Speech Act Recognition. *Rósa Signý Gisladottir*
105. Modelling Multimodal Language Processing. *Alastair Smith*
106. Predicting language in different contexts: The nature and limits of mechanisms in anticipatory language processing. *Florian Hintz*
107. Situational variation in non-native communication. *Huib Kouwenhoven*
108. Sustained attention in language production. *Suzanne Jongman*
109. Acoustic reduction in spoken-word processing: Distributional, syntactic, morphosyntatic, and orthographic effects. *Malte Viebahn*
110. Nativeness, dominance, and the flexibility of listening to spoken language. *Laurence Bruggeman*
111. Semantic specificity of perception verbs in Maniq. *Ewelina Wnuk*
112. On the identification of FOXP2 gene enhancers and their role in brain development. *Martin Becker*

113. Events in language and thought: The case of serial verb constructions in Avatime. *Rebecca Defina*

114. Deciphering common and rare genetic effects on reading ability. *Amaia Carrión Castillo*

115. Music and language comprehension in the brain. *Richard Kunert*

116. Comprehending Comprehension: Insights from neuronal oscillations on the neuronal basis of language. *Nietzsche H.L. Lam*

117. The biology of variation in anatomical brain asymmetries. *Tulio Guadalupe*

118. Language processing in a conversation context. *Lotte Schoot*

119. Achieving mutual understanding in Argentine Sign Language. *Elizabeth Manrique*

120. Talking Sense: the behavioural and neural correlates of sound symbolism. *Gwilym Lockwood*

121. Getting under your skin: The role of perspective and simulation of experience in narrative comprehension. *Franziska Hartung*

122. Sensorimotor experience in speech perception. *Will Schuerman*

123. Explorations of beta-band neural oscillations during language comprehension: Sentence processing and beyond. *Ashley Lewis*

124. Influences on the magnitude of syntactic priming. *Evelien Heyselaar*

125. Lapse organization in interaction. *Elliott Hoey*

126. The processing of reduced word pronunciation variants by natives and foreign language learners: Evidence from French casual speech. *Sophie Brand*

127. The neighbors will tell you what to expect: Effects of aging and predictability on language processing. *Cornelia Moers*

128. The role of voice and word order in incremental sentence processing. *Sebastian Sauppe*

129. Learning from the (un)expected: Age and individual differences in statistical learning and perceptual learning in speech. *Thordis Neger*

130. Mental representations of Dutch regular morphologically complex neologisms. *Laura de Vaan*

131. Speech production, perception, and input of simultaneous bilingual preschoolers: Evidence from voice onset time. *Antje Stoehr*

132. A holistic approach to understanding pre-history. *Vishnupriya Kolipakam*
133. Characterization of transcription factors in monogenic disorders of speech and language. *Sara Busquets Estruch*
134. Indirect request comprehension in different contexts. *Johanne Tromp*
135. Envisioning Language - An Exploration of Perceptual Processes in Language Comprehension. *Markus Ostarek*
136. Listening for the WHAT and the HOW: Older adults' processing of semantic and affective information in speech. *Juliane Kirsch*
137. Let the agents do the talking: on the influence of vocal tract anatomy on speech during ontogeny and glossogeny. *Rick Janssen*
138. Age and hearing loss effects on speech processing. *Xaver Koch*
139. Vocabulary knowledge and learning: Individual differences in adult native speakers. *Nina Mainz*
140. The face in face-to-face communication: Signals of understanding and non-understanding. *Paul Hömke*
141. Person reference and interaction in Umpila/Kuuku Ya'u narrative. *Clair Hill*
142. Beyond the language given: The neurobiological infrastructure for pragmatic inferencing. *Jana Bašnáková*
143. From Kawapanan to Shawi: Topics in language variation and change. *Luis Miguel Rojas-Berscia*
144. On the oscillatory dynamics underlying speech-gesture integration in clear and adverse listening conditions. *Linda Drijvers*
145. Understanding temporal overlap between production and comprehension. *Amie Fairs*
146. The role of exemplars in speech comprehension. *Annika Nijveld*
147. A network of interacting proteins disrupted in language-related disorders. *Elliot Sollis*
148. Fast speech can sound slow: Effects of contextual speech rate on word recognition. *Merel Maslowski*
149. Reason-giving in everyday activities. *Julija Baranova*
150. Speech planning in dialogue - Psycholinguistic studies of the timing of turn taking. *Mathias Barthel*

151. The role of neural feedback in language unification: How awareness affects combinatorial processing. *Valeria Mongelli*
152. Exploring social biases in language processing. *Sara Iacozza*
153. Vocal learning in the pale spear-nosed bat, *Phyllostomus discolor*. *Ella Lattenkamp*
154. The effect of language contact on speech and gesture: The case of Turkish-Dutch bilinguals in the Netherlands. *Elif Zeynep Azar*
155. Language and society: How social pressures shape grammatical structure. *Limor Raviv*
156. The moment in between: Planning speech while listening. *Svetlana-Lito Gerakaki*
157. How speaking fast is like running: Modelling control of speaking rate. *Joe Rodd*
158. The power of context: How linguistic contextual information shapes brain dynamics during sentence processing. *René Terporten*
159. Neurobiological models of sentence processing. *Marvin Uhlmann*
160. Individual differences in syntactic knowledge and processing: The role of literacy experience. *Saoradh Favier*
161. Memory for speaking and listening. *Eirini Zormpa*
162. Masculine generic pronouns: Investigating the processing of an unintended gender cue. *Theresa Redl*
163. Properties, structures and operations: Studies on language processing in the brain using computational linguistics and naturalistic stimuli. *Alessandro Lopopolo*
164. Investigating spoken language comprehension as perceptual inference. *Greta Kaufeld*
165. What was that Spanish word again? Investigations into the cognitive mechanisms underlying foreign language attrition. *Anne Mickan*
166. A tale of two modalities: How modality shapes language production and visual attention. *Francie Manhardt*
167. Why do we change how we speak? Multivariate genetic analyses of language and related traits across development and disorder. *Ellen Verhoef*
168. Variation in form and meaning across the Japonic language family with a focus on the Ryukyuan languages. *John Huisman*

169. Bilingual sentence production and code-switching: Neural network simulations. *Chara Tsoukala*
170. Effects of aging and cognitive abilities on multimodal language production and comprehension in context. *Louise Schubotz*
171. Speaking while listening: Language processing in speech shadowing and translation. *Jeroen van Paridon*
172. Left-right asymmetry of the human brain: Associations with neurodevelopmental disorders and genetic factors. *Merel Postema*