The Biology of General Cognitive Ability Driving Methodological Progress

Furtjes, Anna

Awarding institution:
King's College London

DOCTORAL THESIS

_____

The Biology of General Cognitive Ability:

Driving Methodological Progress

_____

ANNA ELISABETH FÜRTJES

Submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

Social, Genetic and Developmental Psychiatry Centre

Institute of Psychology, Psychiatry and Neuroscience

King's College London

March 2023

Statement of authorship

I hereby declare that all the work presented in this thesis is my own, except where explicit acknowledgement is made. I conceived and carried out all the empirical and written work as the first author, in co-operation with collaborators listed at the start of each chapter. The specific contributions of each author are outlined below. The data used for statistical analysis was collected and partially quality controlled by the UK Biobank and the Adolescent Brain Cognitive Development study teams.

The following contributions are acknowledged throughout the thesis:

In Chapter 3: Dr. Javier de la Fuente provided the R function to perform parallel analysis.

In Chapter 4: Scripts used for genetic quality control and GWAS calculation were partially adopted from scripts provided by Dr. Jonathan Coleman and Dr. Ryan Arathimos. Chapter 4 was partially adapted from a pre-print and has undergone peer review. As of the 8th of March 2023, it has been accepted for publication in *Human Brain Mapping*. Listed co-authors gave feedback on the final written manuscripts.

My work adheres to Open Science practices as I pre-registered analysis plans and published analysis code on GitHub. I believe Open Science can help us make science better reproducible, more generalisable and independent of prestige or monetary agendas.

Acknowledgements

Abstract

Cognitive ability (or the *g*-factor) enables individuals to make sense of the world and to navigate it appropriately. Cognitive ability is typically assessed through various cognitive tests of memory, processing speed, abstract reasoning, reading abilities and more. Research has suggested that individuals with good cognitive ability tend to be healthier, have higher income, live longer, and retain more autonomy in old age. Hence, a comprehensive understanding of the drivers of cognitive ability and its age-related decline may lay the foundations for interventions that have the potential to considerably improve quality of life across the general population.

Research has devoted much attention to so-called brain networks of interconnected brain regions that may collectively underpin cognitive ability. Brain morphometry in the central executive brain network, for example, was demonstrated in a well-powered study to play a centrally important role in cognitive ability (Madole et al., 2021). However, different brain morphometry studies have delivered conflicting evidence and it remains unresolved whether structural brain networks reliably correlate with cognitive ability. One reason for the differences in results might be the lack of methodological consensus between studies. Methodological choices that dictate study results include covariate adjustment for brain size – which some studies perform, and others do not – as well as brain atlases used to subdivide the participants brain images into regions. Therefore, there is a need for hypothesis-free exploratory work to help make optimal analytical decisions and to perform meaningful hypothesis tests. This exploratory groundwork should help establish more reliable correlations between brain morphometry and cognitive ability.

This thesis has two major aims: The first aim is to perform exploratory studies to deliver evidence in support of study designs optimal for the investigation of the relationship between cognitive ability and brain morphometry. The second aim is to investigate the multidimensional relationships among structural brain networks, cognitive ability, and age-related processes. This thesis aims to present comprehensive models of cognitive ability and its associated biology by systematically evaluating the impact of certain methodological decisions. Genetic analysis techniques are employed to triangulate phenotypic analyses using innovative and biologically-informed technology. Presented studies analysed genetic data and structural MRI data in two large samples: the UK Biobank ($N \sim 40,000$) and the Adolescent Cognitive Brain Development study ($N \sim 10,000$).

In Chapter 2 I aim to characterise the impact of brain size covariate adjustment on the relationships between cortical brain volumes and cognitive ability. Results indicated that the relationship between regional volumes and cognitive ability is closely entangled with brain size to the extent that their relationship cannot be reliably modelled when made statistically independent of brain size. This study delivered evidence that instead of assessing region-by-region correlations between brain morphometry and cognitive ability, multivariate study designs may help to account holistically for the complex biological underpinnings of cognitive ability. The study provides empirical and theoretical arguments that brain size adjustment induces collider bias in genome-wide analyses.

In Chapter 3 I derive and validate a novel multivariate framework – Genomic Principal Component Analysis (PCA) – that integrates multiple traits as well as genome-wide information. Genomic PCA takes genome-wide association data as input to extract genetic principal components (PCs) underlying multiple phenotypes

(Genomic PCA does *not* capture ancestral PCs of shared genetic make-up between individuals). I use Genomic PCA in Chapter 4 to model genome-wide PCs underlying multiple brain regions that are part of canonical brain networks. The study demonstrates that genome-wide PCs underlying nine canonical brain networks – unadjusted for brain size – are significantly correlated with cognitive ability and brain ageing. However, this study finds no evidence for localised brain network-specific correlates of cognitive ability, as the central executive network is not any more associated with cognitive ability than other brain networks. The results suggest that genetic correlates of brain morphometry relate to cognitive ability through general brain-wide features shared among multiple regions that are not specific to brain networks.

In Chapter 5 I compare brain atlases that are commonly used to subdivide study participants' brain images into regions-of-interest. The brain atlas comparison in this study relies on multivariate prediction models of multiple behavioural traits, including cognitive ability. This study finds that using fine-grained brain atlases maximises the relationship between brain morphometry and cognitive ability. Future studies may be able to model more robust brain trait associations by adopting multivariate models of hundreds of thousands of vertex-wise brain measurements. The final chapter discusses the overarching implications, limitations, and future directions of these findings that should motivate multidisciplinary approaches to more comprehensively account for the complex web of biological factors that give humans advanced cognitive ability.

Contents

List of Tables

List of Figures

Chapter 3:

Chapter 4:

Unpublished addition to Chapter 4:

Chapter 5:

Chapter 1

General Introduction

1.1 General cognitive ability

Life itself is a test of our cognitive abilities. Understanding what drives differences in cognitive abilities between people across the population is at the core of understanding human behaviour. In the scientific literature, cognitive ability is understood as an umbrella term for a human aptitude to make sense of the world. Cognitive ability enables an individual to "reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly and learn from experience" (Gottfredson, 1997 cited in Deary, Penke, et al., 2010). Advanced cognitive ability sets humans apart from animals and requires different kinds of skills. These skills can be broadly grouped into fluid and crystalised cognitive abilities. Fluid abilities describe processing aspects of cognition, for example, processing speed in reaction time tasks (Deary et al., 2011), visual declarative memory, and abstract reasoning. By contrast, crystalised abilities are declarative and procedural knowledge acquired based on past experiences, for example, vocabulary, and reading abilities. Individuals in a population vary in those cognitive abilities which can be quantified with various cognitive ability tests (outlined in Table 1).

Table 1. Popular cognitive tests that may be used to extract a general factor of cognitive ability as described by Spearman (1904).

Cognitive domain	Cognitive test name	Description
Fluid cognitive abilities		
Processing speed	Deary-Liewald reaction time task (Deary et al., 2011)	A stimulus appears on a computer screen. The tested individual reacts to it by pressing a key as promptly as possible. Test performance is judged better, the shorter the measured time between the appearance of the stimulus on the screen and the key press by the individual.
Visual declarative memory	Pairs Matching test (Wechsler, 2010)	Word pairs are presented auditorily. The tested individual is asked to remember the word pairs, and to recall the matched word when prompted with a word from one of the pairs. Test performance is judged better, the fewer incorrect matches are recorded.
Abstract reasoning	Ravens Progressive Matrices (Bilker et al., 2012)	A logically ordered abstract grid pattern is presented on a screen. The tested individual is asked to select the correct multiple-choice pattern that follows the grid logic. Test performance is judged better, the more patterns are correctly solved.
Crystalised cognitive abilities		
Language abilities	Picture Vocabulary test (Dunn & Dunn, 1965)	A word is presented auditorily simultaneously with 4 images on a screen. The tested individual is asked to pick the image that matches the spoken word. Test performance is judged better, the more correct matches are recorded.
Ability to pronounce words and recognise letters	Oral reading recognition task (Akshoomoff et al., 2013)	The tested individual is asked to read aloud words presented on a screen. Test performance is rated and recorded by a trained examiner.

Performances in different cognitive tests positively correlate with each other. For example, people who can memorise knowledge more accurately also tend to process content faster. According to theory, correlated performance across different cognitive abilities is caused by an underlying "talent", a *latent factor of general cognitive ability*, or "intelligence", sometimes simply referred to as the *g*-factor (Fig.1). General cognitive ability enables individuals to perform well across multiple cognitive tasks (Spearman, 1904; Tucker-Drob, 2019). This hierarchical structure (Fig.1) was reliably confirmed by factor analysis, which extracts an error-free latent variable that tends to be more reliable than individual cognitive tests (Deary, Penke, et al., 2010). The general cognitive ability factor typically explains about 40% of the total variance in peoples' performances across different cognitive tests (Deary, Penke, et al., 2010). This thesis is focused on this latent construct underlying cognitive abilities (i.e., the variance common between cognitive test performances) rather than task-specific cognitive ability. Throughout this thesis, I will use the term *cognitive ability* (or *g*-factor in Chapter 4) to refer to this latent factor.
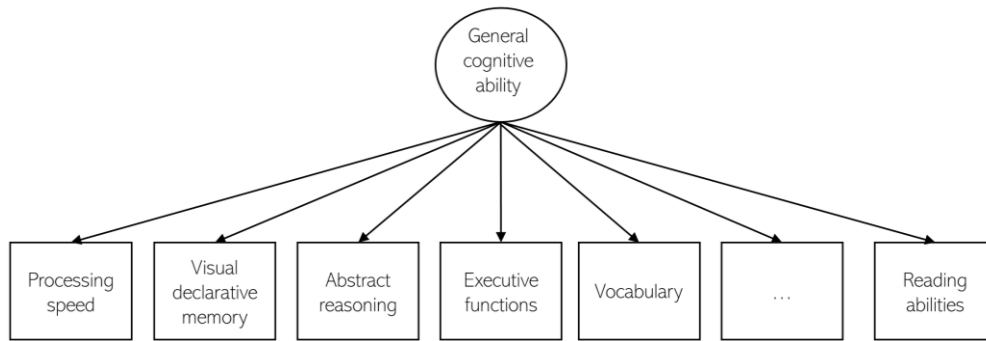
Fig.1. A latent factor of general cognitive ability (_g_-factor) that causes a persons' performance in different cognitive tasks.

The _g_-factor is a latent dimension that cannot be directly observed but can be extracted from variance shared among multiple intercorrelated cognitive tests. The ellipsis in the penultimate box in this Figure indicates that the _g_-factor should be independent of the form, version, wording, or specific cognitive function included cognitive tests are testing. This definition was shown to be highly replicable and to explain ~40% of the variance among cognitive tests (Deary, Penke, et al., 2010).

Cognitive ability evolves across the lifespan. It increases through childhood and early adulthood but declines throughout later adulthood ("cognitive ageing"; Salthouse, 2010). However, individuals of approximately the same age tend to maintain a ranked order among one another in their cognitive ability, and cognitive ability is therefore understood as a time-stable trait. As this ranked order seems to be maintained throughout life, cognitive ability measured in childhood can predict favourable outcomes in later life, such as longevity, health, and old age autonomy (Batty et al., 2007; Deary, Weiss, et al., 2010; Gottfredson & Deary, 2004; Tucker-Drob, 2011). In general, individuals with higher cognitive ability tend to report higher income and better well-being (Davies et al., 2019; Furnham & Cheng, 2017). A 2019 study that was designed sensitive to directions of effects corroborated that higher cognitive ability likely causes these favourable socioeconomic and health outcomes, and not vice versa (Davies et al., 2019)[1].

Due to this link between high cognitive ability and favourable lifelong outcomes, cognitive ability research has the potential to lay the foundations for interventions with important societal implications for public health, well-being, and prosperity. The only intervention – so far discovered – to improve cognitive ability on a population level is years of schooling ("educational attainment"), which was shown to increase an individuals' intelligence score by 1-5 points for every year of education (Ritchie & Tucker-Drob, 2018). This improvement in cognitive ability through longer

---

[1] Davies et al. (2019) employed a Mendelian Randomisation (MR) approach allowing to infer directions of effects between cognitive ability and health outcomes. The MR design simulates a natural experiment by leveraging genetic data in observational studies as instrumental variables. Under certain assumptions, this can control for unmeasured confounding and can deliver evidence for causal relationships between two variables. MR is not used in this thesis, but I briefly mention the MR study here because it underpins cognitive ability as a reliably measured trait that may act as a lifelong factor that precedes and influences later life outcomes. This makes cognitive ability a promising trait to explore human biology, health, and socioeconomic outcomes.

schooling is, however, relatively small in magnitude at a cognitive ability population mean of 100 and a standard deviation of 15 points. Interventions that may help maintain adult cognitive ability or mitigate old age cognitive decline remain active areas of research.

The fact that cognitive ability is time-stable implies that cognitive ability is likely biologically underpinned (e.g., the brain, or genetic factors), as opposed to mainly relying on the environment that tends to be more variable across the lifespan. Hence, potential interventions may be most impactful if they acted on a biological level, rather than an environmental one like longer schooling. This motivates the investigation of reliable biological correlates of healthy cognitive ability and healthy cognitive ageing which are both the focus of this thesis. Such investigations may enable early identification of individuals most in need of support for day-to-day tasks in old age which would help assign caring resources more efficiently. They may also lead to interventions capable of delaying cognitive decline, for example, in the form of drugs, food supplements, or brain cell stimulation using transcranial magnetic stimulation. In our ageing societies, identifying reliable biological correlates of healthy cognitive ageing may also benefit dementia research where learning and memory are core aspects. For example, it may help screen and identify individuals with accelerated cognitive decline who may be at higher risk of developing dementia later in life.

1.2 Brain correlates of cognitive ability

The human brain has 86 billion neurons and about the same number of glial cells (Herculano-Houzel, 2009). Over the past 2-3 million years of human evolution, the expansion of advanced human cognitive ability coincided with a tripling in human brain size (Herculano-Houzel, 2016). Compared with other animals, humans have larger brains, more neurons (Herculano-Houzel, 2009), and increased modularity. Modularity describes the locally interconnected organisation of neurons within regions ("modules") that are more sparsely linked with other regions in humans than in other animals. This is thought to create functional specialisation necessary for efficient brain-wide neuronal communication and complex neuronal processes underpinning language, long-term memory and advanced cognitive ability more generally (Changeux et al., 2020). Therefore, the size of the brain has been a primary target in searching for biological correlates of cognitive ability.

Across the general modern population, a positive association between higher cognitive ability and total brain size recorded by magnetic resonance imaging (MRI)[2] has consistently replicated (Nave et al., 2018; Pietschnig et al., 2015). The brain size cognitive ability correlation, however, only explains about 5% of the variance in cognitive ability (i.e., r ~0.25), illustrating that total brain size as a measure does not convey enough information to reliably account for brain size-based correlates of cognitive ability. Hence, many studies have attempted to find more specific brain

---

[2] MRI is an imaging technique that utilises strong magnetic fields and radio waves to visualise brain anatomy and physiological processes. In this example, it was used to estimate the size of the whole brain across thousands of individuals. More details on how studies estimate brain morphometry and functional activity using MRI is in Table 2.

correlates by considering brain regions, rather than the whole brain. Brain regions seem to be differentially important for cognitive ability (Cox et al., 2019).

Table 2. Brain imaging techniques used across the literature to identify brain-based correlates of general cognitive ability.

Brain imaging technique	Description
Structural MRI (T1-weighted imaging)	Structural MRI makes use of water molecules containing protons in human brain tissue. The main magnetic field in an MRI scanner causes those protons to line up and spin at a certain frequency. A radiofrequency pulse is temporarily turned on, which causes the protons to line up in a different direction; until the pulse is turned off again, and the protons realign with the main magnetic field. During this process, protons emit radio frequencies that can be detected and localised by the scanner. Because the rate at which the protons realign with the main magnet varies between tissue types, the MRI scanner can generate a three-dimensional image of where grey matter, white matter, cerebrospinal fluid, and the skull are located (Lerch et al., 2017). Using imaging processing software such as FreeSurfer (Fischl, 2012), studies use the three-dimensional images to estimate brain morphometry through estimates of cortical thickness, surface area, and grey-matter volume.
Functional MRI	Functional MRI relies on different scanner settings than structural MRI and makes use of the fact that performing a cognitive task (or even resting) induces supply of oxygenated blood to certain brain regions. Oxygenated blood has different magnetic properties than deoxygenated blood, and the scanner records local differences in those magnetic properties (Xue et al., 2010), which are thought to indirectly reflect neuronal activity (Logothetis et al., 2001). Hence, functional MRI is "blood-oxygen-level dependent" (BOLD). Compared with structural MRI, functional MRI is time-sensitive, but it is delayed (2-6 seconds), noisy, can show weak signal, and was reported to have poorer reliability (Elliott et al., 2020).
Diffusion MRI	Diffusion MRI is also performed in MRI scanners whereby moving water molecules in the brain are tracked. The molecules are in motion (they "diffuse"), but they are restricted by physical obstacles, such as myelinated neurons, for example. DTI records the direction of diffusion of the water molecules at a voxel-level. The directional information can be used to infer where larger neuronal tracts are located throughout the brain ("tractography") (Lerch et al., 2017).
Lesion studies	Lesion studies investigate brain lesions in individuals who suffered brain damage, for example, through illness, infection, or trauma. Brain matter losses may be identified through structural MRI. Lesion studies investigate links between brain matter losses and specific losses in cognitive ability to infer whether certain brain regions are necessary for the performance of certain cognitive functions. Lesion studies are typically based on small clinical between-group differences.

Note that this table is meant to add explanations of imaging techniques mentioned in the text. It does not represent an exhaustive list of existing imaging techniques.

To investigate brain correlates of cognitive ability, previous studies have visualised the brains of their study participants using multiple brain imaging techniques. Some of those techniques included for example structural MRI, functional MRI, diffusion MRI, and lesion studies (briefly explained in Table 2). In a review summarising evidence across brain imaging techniques, Jung and Haier (2007) proposed the parieto-frontal integration theory (*P-FIT*) whereby specific brain regions collectively give rise to cognitive ability. Frontal and parietal regions were proposed to play a particularly important role. Briefly summarised, the P-FIT suggests that temporal and occipital brain regions are involved in reasoning and intelligent processing of a cognitive task, in that they enable an individual to perceive sensory information, for example through visual input. Sensory information is communicated to parietal brain regions for abstraction. Subsequently, parietal regions interact with frontal regions to evaluate different solutions to given tasks, and frontal areas are involved in selecting the best solution, which is acted upon. According to Jung and Haier (2007), advanced cognitive abilities depend on efficient between-region communications enabled through strong white matter connections across the brain. Fig.2 visualises the brain lobes to put the P-FIT and the frontal and parietal correlates of cognitive ability into context.
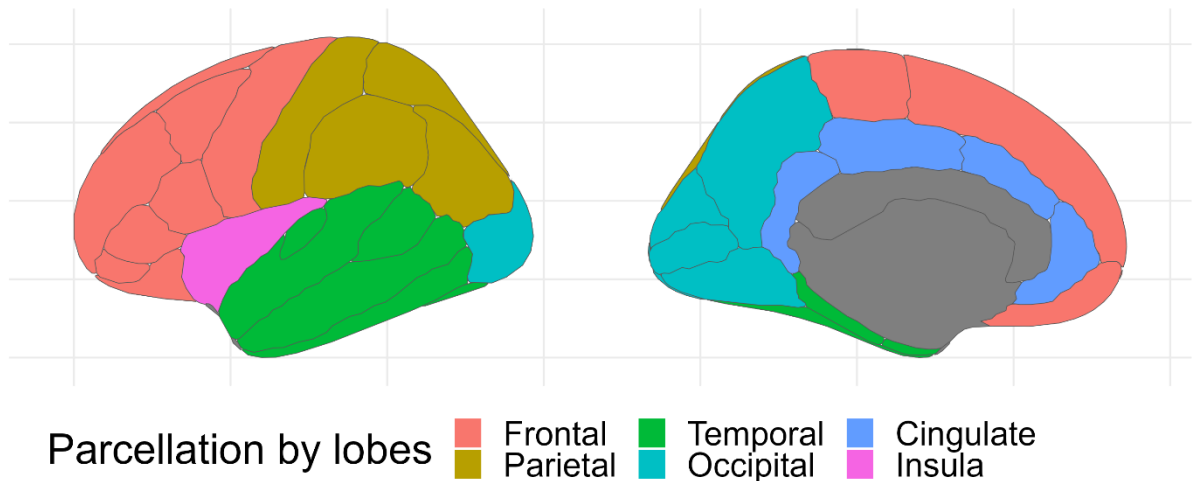
Fig.2. Desikan-Killiany brain parcellation coloured by brain lobes.

The P-FIT has since been criticised due to evidence in its favour being statistically underpowered. A more recent meta-analysis concluded that the highlighted frontal and parietal correlates of cognitive ability do *not* reliably map onto structural measures of grey matter (Basten et al., 2015). Conceptually, however, the P-FIT adds value by postulating that several brain regions are collectively involved in cognitive ability – as opposed to isolated regions underlying cognitive ability. This is in line with a continuously growing body of evidence supporting the idea that structural and functional *brain networks* constrain and give rise to cognitive ability as a complex neuronal process (Bressler & Menon, 2010).

1.2.1 Functional brain networks

Functional brain networks are groups of distributed and functionally interconnected brain regions that are understood to collectively produce brain function more broadly (and could lead to cognitive ability, for example). Typically, studies infer functional connectivity by statistically analysing which brain regions

demonstrate synchronised neuronal activity in functional MRI scans (as described in Table 2). More specifically, these studies identify brain regions that are correlated in their neuronal activity across time, and they assume that simultaneously activated regions communicate and form a functionally interconnected brain network. Functional brain networks can be characterised based on functional MRI data recorded either when participants are resting or conducting specific cognitive tasks (e.g., Power et al., 2011; Yeo et al., 2011).

Functional brain networks that are identified through analysing functional MRI data are typically named after the cognitive task participants conducted while in the MRI scanner. Hence, a functional network is assumed to underpin the cognitive task it is named after (e.g., the default mode network is found when scanned individuals are resting). The functional brain network literature is heterogeneous (Uddin et al., 2019), and the cognitive tasks used to identify brain networks do not typically consider general cognitive ability (as defined in Section 1.1). Instead, brain networks are identified in association with more elementary brain functions, such as sensory information processing and attention.

Certain brain networks are considered *canonical* because they were repeatedly identified by studies over the past two decades: For example, the *central executive* network shows neuronal activity in functional MRI scans when participants are given cognitive tasks requiring attention and working memory (Menon & Uddin, 2010; Sridharan et al., 2008). The *default mode* network is found when participants are resting, or solving tasks demanding internally directed thought (Buckner & DiNicola, 2019). The *salience* network activates during cognitive tasks that require participants to allocate attention to competing sensory information (Downar et al., 2002), which was shown to integrate brain functions associated with the central executive and

default mode networks described above (Li et al., 2018; Sridharan et al., 2008). Exactly which regions form those functional networks remains debated (Uddin et al., 2019), and criticisms exist about functional brain network reliability (e.g., Singhal et al., 2020).

Building on this functional brain network literature, Barbey (2018) argues that cognitive ability reflects the capacity of someone's brain to activate canonical brain networks (e.g., the central executive network), and the ease with which this persons' brain can transition in its neuronal activity between multiple brain networks ("network flexibility"). Barbey (2018) theorised that this network flexibility supports higher cognitive ability by enabling fast information transfer across functional brain networks. Although my empirical chapters do not analyse functional MRI data, my analytical decisions aim to acknowledge the extensive functional brain network literature by modelling canonical brain networks described above using structural MRI data.

1.2.2 Structural brain networks

Some studies analysing structural, instead of functional MRI data have also adopted a view of the brain that considers brain networks rather than isolated regions. Bullmore et al. (1998) argued that correlations between regional grey matter indicates anatomical connections among brain regions. One possible explanation may be that the developing brain builds interregional connections early and when regions are connected, they share certain growth factors. This concurrent development may make their macroscopic features more similar. Accordingly, researchers may identify *structurally* connected brain networks based on correlation structures among grey matter regions.
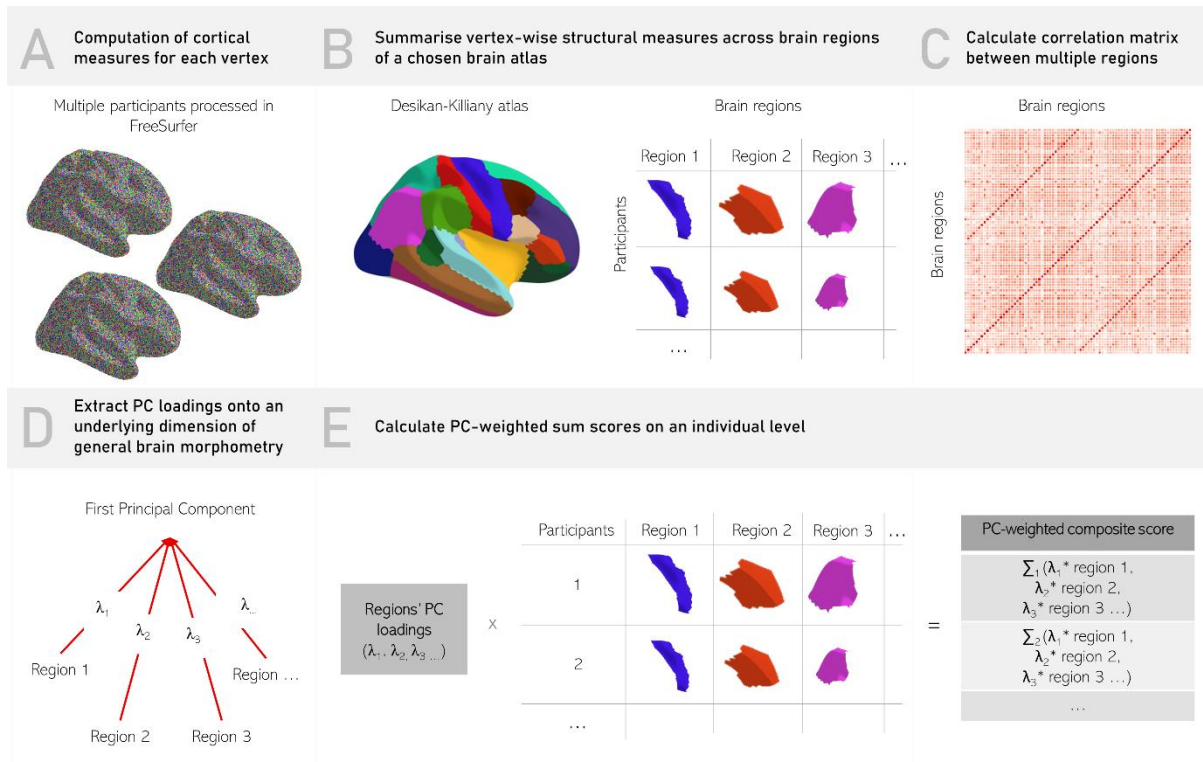
Fig.3. One possible analysis pipeline to extract individual-level structural brain network indices.

(A) Structural MRI data from multiple participants are pre-processed in FreeSurfer to obtain cortical measures for each vertex. This step provides brain volume measures at each vertex (or surface area or cortical thickness). **(B)** The vertex-wise measures are summarised for each brain region of a chosen brain atlas. Here, I depict the Desikan-Killinay atlas (Desikan et al., 2006) that dictates which vertices belong to each of the 68 regions the atlas delineates. **(C)** Structural connectivity is inferred by calculating correlations among structural measures of brain regions. The correlations indicate the strength of connectivity between considered regions. **(D)** Principal Component Analysis (PCA) linearly decomposes the correlation matrix obtained in step C. It provides principal component (PC) loadings of each brain region onto an underlying dimension of general brain morphometry. (E) Finally, the pipeline calculates PC-weighted sum scores whereby a participants' brain network connectivity is approximated by the sum of each regions' brain volume weighted by a regions' PC loading obtained in step D.

15

Typically, a structural brain network analysis pipeline involves several complex steps. Fig.3 illustrates one possible pipeline that is of relevance to this thesis because it was used in a previous study (Madole et al., 2021) on which my empirical work built (Chapter 3 & Chapter 4). It is output from this pipeline that I refer to when discussing structural brain networks. Below, I expand on this pipeline in detail because it allows me to introduce concepts such as vertex-wise MRI data, brain atlases, Principal Component Analysis and how principal components may be interpreted when applied to structural MRI data. These concepts are central to all my empirical chapters.

The pipeline in Fig.3 generates interindividual indices of structural brain network connectivity: first, measures of cortical morphometry (e.g., cortical volumes) are extracted from structural MRI images on a vertex-wise basis using toolkits such as FreeSurfer (Fischl et al., 2002) (Fig.3A). In surface-based analyses (as implemented in FreeSurfer), vertices are thousands of points in a coordinate system (the "cortical mesh") that are projected onto brain scans in order to make brain scans comparable between study participants. Subsequently, vertex-wise cortical measures of an individual are summarised for every region of a chosen brain atlas that dictates which vertices collectively form a brain region (Fig.3B). Brain atlases aim to delineate boundaries of functionally, micro-structurally, macro-structurally, or otherwise distinct brain regions. It is at the researcher's discretion which of the many existing brain atlases to employ. Popularly-used brain atlases include the Desikan-Killiany atlas (68 regions, Desikan et al., 2006), and the Schaefer atlas (500 regions, Schaefer et al., 2017).

Subsequently, structural connectivity across brain networks is inferred by calculating correlations among morphometric measures (e.g., volume) of brain regions. Correlations between brain volumes are assumed to indicate the strength of

connectivity within a structural brain network ([Fig.3](#)C). The pipeline uses Principal Component Analysis (PCA) to reduce dimensionality and extract general dimensions of brain morphometry [3] underlying a structural brain network. PCA linearly decomposes the structural correlations across brain regions and provides principal component (PC) loadings to quantify how much each region contributes to general dimensions of brain morphometry underlying a brain network. PC loadings are interpreted to indicate how representative (or important) each region is to overall brain morphometry in the brain network ([Fig.3](#)D). For example, Madole et al. (2021) find that frontal and parietal brain regions have particularly large PC loadings onto a general dimension of brain morphometry[3] underlying the whole brain which indicates that frontal and parietal regions are particularly representative of (or important for) overall brain morphometry.

Finally, [Fig.3](#)E constructs PC-weighted sum scores to represent how well integrated brain networks are in each sample participant. These PC-weighted scores sum up multiple regional brain measures multiplied by each regions' PC loading, to serve as a composite score of structural brain connectivity (or integrity). Out-of-sample predictions showed that those sum scores are associated with multiple domains of individual-level cognitive ability (Madole et al., 2021). It is an advantage of this pipeline that general dimensions of brain morphometry extracted using PCA are less noisy and more reliable than considering isolated brain regions.

---

[3] The term *general dimensions of brain morphometry* is used in Chapter 4 and Chapter 6 to describe principal components underlying multiple brain volumes extracted using PCA. General dimensions of brain morphometry may be interpreted as indirect indices of brain connectivity or brain integrity within the considered structural brain network.

1.2.3 Combining evidence from studies considering structural and functional MRI data

Structural and functional MRI research are often separate research fields, but this thesis posits that considering evidence from both structural and functional MRI data may allow studies to more accurately approximate brain organisation. After all, it is reasonable to expect that functionally connected brain regions should also be physically connected, or at least be largely constrained by their surrounding anatomy. This idea motivates more comprehensive study designs of structural MRI data to take into account the functional MRI literature that characterised canonical brain networks (as described in Section 1.2.1).

Studies have systematically examined the role of *functionally* defined brain networks in interindividual differences of *structural* grey matter. These studies used approaches detailed in Section 1.2.2 to model structural brain networks (Fig.3). The selection of brain regions included in a structural brain network was decided with reference to functional MRI studies (Hilger et al., 2020; Madole et al., 2021). That is, according to functional MRI studies, the central executive network comprises eight frontal and parietal regions, and based on these functional MRI results, researchers defined the central executive network in structural MRI investigations. This approach has produced conflicting evidence about how these functionally-informed structural brain networks are associated with cognitive ability. Section 1.2.3 below focuses on Madole et al. (2021), a 2021 study with the largest sample size ($N = 8,185$). I will revisit how the findings by Madole et al. conflict with other studies in Section 1.2.5.

Following the pipeline in Fig.3, Madole et al. (2021) modelled PCs underlying the whole brain, as well as PCs underlying nine structural brain networks whose

corresponding regions reflected canonical functional brain networks. Among others, the canonical networks included the default mode, central executive, and salience networks. To obtain even more robust and multidisciplinary-informed brain network characterisation, the authors converged additional evidence from DTI and lesion-based studies (e.g., Bressler & Menon, 2010).

The study by Madole et al. (2021) showed that the central executive network played a disproportionately important role – relative to other brain networks – in multiple domains of cognitive ability, including processing speed, visuospatial ability, and memory (Madole et al., 2021). More specifically, the first PC underlying the whole brain (83 regions) and the first PC underlying the central executive network (8 regions) were equally strongly associated with cognitive ability ($r_{visuospatial\ ability}$ ~0.4, $r_{processing\ speed}$ ~0.25). The central executive network was most strongly associated with cognitive ability when the analysis was adjusted for its small brain network size. This supported the notion that the central executive network – with its eight frontal and parietal regions – is a neuronal correlate of cognitive ability more biologically specific than the whole brain. Additionally, central executive regions had high PC loadings onto a whole brain PC indicating that central executive regions well represent overall brain morphometry and brain organisation.

The Madole et al. (2021) study also considered age as an influential factor for the relationship between brain morphometry and cognitive ability. It demonstrated, using cross-sectional data, that central executive regions were more strongly associated with age than other brain regions. Counter-intuitively, this finding suggests that regions more centrally important to overall brain morphometry (i.e., regions that have large PC loadings onto a PC underlying the whole brain) are also most important for cognitive ability, while they tend to decline in volume most rapidly. A possible

19

explanation may be that central executive regions tend to be used more intensely throughout the lifespan which may put central executive regions under distinctive metabolic strain, accelerating their age-related atrophy compared with other regions. In sum, Madole et al. (2021) demonstrated that the central executive network is, considering its small size, most associated with cognitive ability, and that age-related brain atrophy may contribute to this association. It is the aim of Chapter 4 of my thesis to triangulate these insights using novel technology.

1.2.4. Brain ageing

Madole et al. (2021) demonstrate that comprehensively describing the relationship between brain morphometry and cognitive ability must consider ageing as they simultaneously decline throughout later adulthood (Salthouse, 2010). Older age is associated with larger ventricles (Resnick et al., 2003), and weaker connectivity in functional and structural brain networks (Ferreira & Busatto, 2013; Wu et al., 2013). Different brain regions atrophy at different speeds, whereby frontal and temporal brain regions may decline the most with increasing age (Raz et al., 2010).

Structural MRI research has devoted much attention to the *brain age gap*, a proposed brain biomarker of ageing and health. The brain age gap is derived from the relationship among age and brain morphometry including grey and white matter. Machine learning approaches are typically used to characterise the relationship between age and MRI data which allows out-of-sample predictions of an individuals' brain age based on their brain scans (Cole & Franke, 2017). The brain age gap expresses how much older (or younger) an individual's brain may appear relative to their chronological age. It is calculated as the difference between chronological age and age predicted from MRI scans. In healthy populations, older brain age (i.e., a

larger brain age gap) indicates a person's ageing is accelerated compared with their peers, and older brain age predicts worse cognitive ability (e.g., Cole et al., 2018).

Due to issues with data availability, only few longitudinal studies of brain ageing exist. This is unfortunate as they allow more reliable models of brain change than the cross-sectional investigations discussed above. As opposed to MRI scans at only one point in time, repeated MRI scans enable disentangling brain atrophy (i.e., decline in brain volume occurring from one measured time point to the next), from lifelong brain features that remain stable between measured time points. Using repeated MRI measures ($N = 1,091$), Cox et al. (2021) discovered that age-related brain atrophy across the 8th decade of life was common between brain regions, and that one principal dimension accounted for 66% of brain-wide changes across the Lothian Birth Cohort. This showed that much of the age-related atrophy happens in parallel and is coordinated across brain regions. This principal brain atrophy dimension was correlated with longitudinally measured decline in general cognitive ability ($r = 0.43$), visuospatial ability ($r = 0.41$), processing speed ($r = 0.38$), and memory ($r = 0.37$). This illustrates the fundamental role age plays in the relationship between brain morphometry and cognitive ability, and that it must be considered when exploring their relationship.

1.2.5 Lack of methodological consensus in the investigation of structural brain network correlates of cognitive ability

To establish the central executive network as a reliable neuronal correlate of cognitive ability, it is important to replicate their small to moderate correlation (r ~0.25-0.4) (Madole et al., 2021). However, results are inconsistent across studies which may be due to a lack of methodological consensus. Hilger et al. (2020) ($N = 308$)

showed that whether cognitive ability is significantly correlated with structural brain networks depends on two things: First, the detection of a significant correlation was conditional upon the brain atlas that was used to subdivide brain images into regions (Fig.2B). Hilger et al. (2020) demonstrated that an out-of-sample prediction of cognitive ability was only significant when brain images were subdivided into 400 Schaefer atlas regions (Schaefer et al., 2017), as opposed to hundreds of thousands of brain regions (~500k voxels[4]). Second, a statistically significant correlation between structural brain networks and cognitive ability was conditional upon whether total brain size was included in statistical analyses as a covariate.

Structural MRI studies commonly adjust for total brain size as a covariate. Researchers argue that brain size adjustment enables them to determine whether the relationship between certain brain regions and cognitive ability is independent of total brain size (i.e., whether the correlation persists above and beyond variance accounted for by the whole brain). Studies adjusting for total brain size aim to avoid misattributing a significant association between cognitive ability and a brain region, when instead the whole brain may confound or entirely drive the association. Hilger et al. (2020) found no evidence for an association between structural brain networks and cognitive ability with total brain size adjustment. Hence, Hilger et al. (2020) concluded that cognitive ability is associated with global rather than region-specific or network-specific brain features.

Conclusions by Hilger et al. (2020), that brain network-specific correlates of cognitive ability do not exist, stand in conflict with conclusions by Madole et al. (2021)

---

[4] Hilger et al. used voxel-based morphometry (VBM) to estimate local grey matter volume. In VBM, grey matter is estimated at each of ~500,000 voxels, which are three-dimensional building blocks that represent the smallest image resolution from a brain image. VBM is an alternative cortical modelling approach to surface-based modelling employed in FreeSurfer, for example.

([Section 1.2.3](#)). Madole et al. argue that the central executive network harbours cognitive ability. Hilger et al. (2020) and Madole et al. (2021) methodologically differ by adjusting or not adjusting for total brain size, and in their brain atlas choice: Madole et al. (2021) did not adjust for brain size and used the Desikan-Killiany brain atlas (Desikan et al., 2006). This illustrates the importance of seemingly trivial methodological choices leading to differing conclusions of these studies, where Madole et al. (2021) report localised neuronal correlates of cognitive ability, but Hilger et al. (2020) only report global brain-wide correlates.

The lack of methodological consensus – surrounding covariate control and brain atlas choice – is one of many challenges faced by structural MRI studies. Neuroimaging grapples with inadequate statistical power and large sampling variances due to small sample sizes. A study from 2022 demonstrated that reproducible MRI analyses require thousands of participants, but that the median sample size is only ~25 participants (Marek et al., 2022). Sufficiently large samples are rare because they are expensive to collect (~£10 million for 10,000 participants; www.ukbiobank.ac.uk). Only a few initiatives link cognitive data with MRI scans of thousands of participants including the UK Biobank (UKB; Littlejohns et al., 2020), and the Adolescent Brain Cognitive Development (ABCD) study (Casey et al., 2018). My thesis focuses on large-scale samples with thousands of participants, and I will therefore assume, by and large throughout the thesis that analyses are sufficiently powered.

Artefact correction in structural MRI data is another methodological choice that can alter associations found in structural MRI data (e.g., Botvinik-Nezer et al., 2020). Artefacts impact the quality of MRI images and their suitability for interindividual comparisons, for example, through participants' movement in the scanner, through

differences in brain sizes and shapes between participants, and through low signal to noise ratios. To make MRI scans comparable between participants, technical artefacts are corrected using motion correction, spatial normalisation, and smoothing. An in-depth discussion of these artefact correction techniques is beyond the scope of this thesis. The UK Biobank and ABCD studies provide researchers with structural brain measures already pre-processed which ensures uniform artefact correction. Hence, I assume that technical artefacts have successfully been dealt with in these large-scale samples. Instead, I will focus on issues that arise at a later stage of analysis, namely the statistical modelling of brain morphometry and its association with cognitive ability.

1.2.6 A focus on exploratory structural MRI research

In this thesis, I put emphasis on exploring potential reasons for conflicting evidence about the relationship between brain morphometry and cognitive ability. The fact that studies have produced conflicting conclusions may highlight a major problem: structural MRI studies may not be sufficiently equipped to test very specific hypotheses through confirmatory research. As brain organisation remains incompletely understood, we may be testing premature and unspecific hypotheses. When research fails to reproduce results and hence reach coherent conclusions, the underlying problem may lie in the setup of statistical models. Essential knowledge may be missing to make optimal analytical decisions which impedes meaningful hypothesis testing (Scheel et al., 2021).

Researchers must make many analytical decisions including selecting appropriate covariates and optimal brain atlases, and those decisions directly impact study results. For example, I outlined above (Section 1.2.5) that decisions about

whether to adjust for brain size as a covariate and decisions about which brain atlas to use in a study dictate whether structural brain networks are found to be significantly correlated with cognitive ability (Hilger et al., 2020). These important decisions lack consensus and are rarely explicitly justified in publications giving the impression that some decisions may be pure guesswork. For instance, studies investigating the association between brain morphometry and cognitive ability commonly adjust for brain size as a covariate (Williams et al., 2022), but studies also commonly do *not* include brain size as a covariate (Cox et al., 2019). Studies widely differ in their chosen brain atlas, and they do not tend to justify why a certain brain atlas seemed the most appropriate. Hence, in addition to testing for associations between structural brain networks and cognitive ability, empirical work presented in this thesis aims to address arbitrary decision-making surrounding covariate control and brain atlas choice. Chapter 2, Chapter 3 and Chapter 5 perform exploratory, and hypothesis-free – rather than confirmatory – groundwork to optimise the meaningful inputs, which future studies may feed to the hypothesis-testing machinery. In the future, work presented in this thesis may inform more transparent and testable hypotheses, which may be the key to determining *how* to ask where cognitive ability is anchored in the brain.

1.3 Challenges and opportunities of a genetic level of analysis

Investigating the relationship between cognitive ability and brain morphometry by using genetic analysis designs presents opportunities to perform this exploratory groundwork from a novel biologically informed perspective. However, integrating millions of genetic components for each study participant adds another dimension of complexity to the investigation of the brain and its many regions. It is the aim of the following sections to describe genetic analysis techniques that allow straightforwardly incorporating genetic information in the analysis of structural MRI data. Section 1.3.1 describes genome-wide association studies, and Section 1.3.2 explains the concept behind Linkage Disequilibrium Score Regression (LDSC). These two methods facilitate the calculation of heritability[5] and genetic overlap[6] which have previously been used to quantify shared genetic aetiology between cognitive ability and brain morphometry (Section 1.3.3). Section 1.3.4 outlines how the integration of LDSC into multivariate methods promises opportunities for innovative analysis of cognitive and structural MRI data. These opportunities have shaped the study designs chosen in Chapter 2, Chapter 3 and Chapter 4.

---

[5] Heritability is the extent to which genetic differences between people can account for trait differences between people. For example, cognitive ability is 50% heritable (Deary et al., 2009) which means that 50% of the differences between peoples' cognitive ability systematically cooccur with genetic differences between them.

[6] Genetic overlap is the extent to which two traits share genetic aetiology. How heritability and genetic overlap are calculated is outlined in Section 1.3.2.

1.3.1 Genome-wide association studies

Genome-wide association studies (GWAS) consider large-scale samples of unrelated individuals to identify genetic markers associated with human traits. In GWAS we perform mass-univariate testing across the genome, whereby millions of genetic markers – that are known to vary between people ("single nucleotide polymorphisms"; SNPs) – are tested for their statistical association with a trait; for example, cognitive ability. SNPs are intercorrelated which means that GWAS tend to index clusters of SNPs that may collectively, in a gene for example, influence the manifestation of a trait (Visscher et al., 2017).

It is a well-established characteristic of the human genome that SNPs have a complex correlation structure among them. This complex correlation structure – also referred to as *linkage disequilibrium* (LD) – is the product of historical evolutionary pressures, including mechanisms involving gene mutations, gene recombination, and natural selection (Visscher et al., 2017). LD is mathematically expressed as the squared *LD score* correlation coefficient (r^2) of a SNP with all other SNPs across the genome, which means that a SNP that frequently co-occurs with many other SNPs receives a larger r^2 estimate.

LD structures across the population have been extensively studied. LD differs substantially between broad ancestry groups (Shi et al., 2020), and a standard record of LD across European ancestry is freely available online (URL). European populations remain, to this date, the most genetically researched ancestry group (Mills & Rahal, 2019). To obtain sound GWAS discoveries, SNPs included in a GWAS must be representative of the population-wide (ancestry-specific) LD landscape surrounding them. Specifically, SNPs included in a GWAS should have correlations with other SNPs

as expected based on the LD correlation structure at their genomic location because it is possible that the GWAS may not have included the SNP driving causal genetic signal. When SNPs are representative of their LD landscape, the GWAS will still pick up the causal genetic signal (of the SNP that was not included) through its correlated SNPs included in the GWAS (Wray, 2005). That many genetic analyses require LD representative SNPs is also the reason why many genetic studies only consider samples of homogeneous ancestry (e.g., White European only).

Typically, GWAS reveal hundreds of SNPs significantly associated with traits, such as cognitive ability and brain morphometry. Hence, cognitive ability and brain morphometry are described as *polygenic* and *complex*. One major endeavour in genetics research is to account for LD, at the same time as disentangling which independent biological pathways may causally lead from genes to the manifestation of a trait (e.g., de Leeuw et al., 2015; van der Meer et al., 2020). This however is not the focus of this thesis.

Instead, I will focus on GWAS *summary statistics*, which are the primary output of a GWAS. GWAS summary statistics are a hypothesis-free summary of genome-wide associations between millions of genetic markers and a trait of interest. Researchers publicly share GWAS summary statistics, which can be safely done because individual-level information is untraceable in the summary-level GWAS summary statistics, overcoming data privacy concerns. GWAS summary statistics are routinely used in follow-up analyses, for example, to infer heritability of one trait, as well as genetic overlap between two traits.

1.3.2 Linkage Disequilibrium Score Regression

Multiple analysis techniques may be used to quantify heritability and genetic overlap, but this thesis focuses on Linkage Disequilibrium Score Regression (LDSC). LDSC utilises GWAS summary statistics to calculate heritability and genetic overlap estimates. Heritability is the proportion of trait variance that can be explained by genetic factors. Genetic overlap estimates the extent to which two traits share genetic aetiology. It marked an influential discovery in the field of statistical genetics when Bulik-Sullivan, Loh, et al. (2015) found that the polygenic signal indexed by GWAS systematically relates to LD structures. They discovered that a trait was heritable and yielded non-zero genome-wide signal in a GWAS, if SNPs with high LD across the genome (i.e., SNPs with a large r^2 LD score estimate) also tended to be the SNPs most associated with the trait. This systematic relationship between LD and GWAS associations reflects that a SNP with high LD has greater chances – than SNPs with low LD – to pick up on signal driven by genetic mechanisms that are causal (i.e., driving the signal of association) on the trait. This phenomenon is due to the fact that SNPs with low LD are less likely to be correlated with other markers in the genome, including the ones driving the association between genetic markers and the trait of interest.

On this basis, the overall extent to which a trait is associated with genetic markers (i.e., heritability) can be simply expressed as a linear regression of LD scores against genetic trait correlates from GWAS summary statistics. This routinely used method is named *univariate* LDSC. It separates signal that systematically maps onto LD structure (regression slope), from signal that does *not* systematically relate to LD structure (regression intercept). The former is often referred to as *SNP-heritability*.

The regression intercept features the error term and well-approximates confounding factors such as population stratification (Bulik-Sullivan, Loh, et al., 2015). LDSC assumes that many SNPs collectively affect complex traits in an additive fashion, and it does not account for SNP-SNP interactions or dominance effects, for example.

Quantifying genetic correlations between two traits using *bivariate* LDSC follows this same principle, but considers the SNP-wise product between GWAS correlates of two traits: Specifically, a linear regression is calculated between LD scores and the product between genetic trait correlations that were taken from two different trait-specific sets of GWAS summary statistics (Bulik-Sullivan, Finucane, et al., 2015). The slope from this bivariate LDSC provides an estimate of the genetic overlap between the two traits and is expressed through a correlation coefficient (r_g) between -1 and 1. In addition to population stratification, the bivariate LDSC intercept also accounts for sample overlap. This means that the two sets of GWAS summary statistics can have been derived from the same sample or two different ones, but the LDSC slope estimate of the genetic trait correlation will not inflate if GWAS samples overlapped.

Heritability and genetic overlap estimates discussed below were almost exclusively derived from White European population because LD structures differ between broad ancestry groups. This means LDSC must consider homogeneous ancestry groups to produce sensible results. Genetic analyses were limited to White Europeans due to data availability at the time of writing. It would have been preferable to include other ethnic groups, but unfortunately data featuring genetic, structural MRI, and cognitive data were not readily available for those populations to ensure sufficient statistical GWAS power. Genetic findings discussed below are therefore limited in their generalisability to more diverse populations.

1.3.3 Genetic correlates of cognitive ability and brain morphometry

Different genetic methods, including but not limited to LDSC and GWAS, have shown that all human traits are heritable (Turkheimer, 2000) which extends to cognitive ability, regional, as well as total brain volume (Table 3). The robust phenotypic relationship between total brain volume and cognitive ability (discussed in Section 1.2) was established to be due to genetic factors (r_g = 0.24), mainly implicating genes involved in regulating cell growth (Jansen et al., 2020). The literature quantifying a genetic relationship between cognitive ability and brain regions, rather than the whole brain, is as inconsistent as the phenotypic literature discussed above. To my best knowledge, Biton et al. (2020) is the only study that considered regional volumes *unadjusted* for total brain size. Biton et al. (2020) reported positive genetic correlations between cognitive ability and seven brain regions at approximately half the magnitudes (range r_g = 0.07-0.13) as Jansen et al. (2020) obtained for total brain volume (r_g = 0.24).

Table 3. Twin and SNP-heritability estimates for intelligence and brain volumes.

Trait	Twin heritability [a]	SNP-heritability [b]	Number of associated GWAS loci
Cognitive ability	50% [1]	25% [2-4]	187 [5]
Total brain volume	83% [6]	66% [7]	5 [8]
Regional brain volumes	~30-80% [†6]	~55-65% [*7,9]	365 (total number for multiple regions) [*8]
Example regions below			
Volume in the inferior parietal gyrus	12% [†10]	21% [*7]	0 [*8]
Volume in the superior parietal gyrus	58% [†10]	49% [*7]	8 [*8]
Volume in the inferior parietal gyrus	12% [†10]	21% [*7]	11 [*8]

[a] Twin heritability indicates the proportion of interindividual trait variance (in cognitive ability, for example) that can be attributed to genetic factors. Twin heritability is quantified based on the comparison of monozygotic and dizygotic twins, who share 100% and ~50% of their genetic makeup, respectively. The twin method gives broad-sense heritability estimates which includes additive, as well as non-additive sources of genetic effects, including genetic dominance and SNP-SNP interactions. [b] SNP-heritability indicates the proportion of interindividual trait variance in the narrow-sense, which refers to additive effects of independent genetic alleles. Narrow-sense heritability can be inferred via LDSC or other methods relying on unrelated individuals, such as Genome-wide Complex Trait Analysis (GCTA) (Yang et al., 2011). Broad sense heritability should always be larger than narrow sense heritability, but the table shows that this is not the case for the volume in the inferior parietal gyrus. I suggest this may be due to statistical power as the broad-sense estimate[10] (N = 486) is from a study 9 years older than the narrow-sense heritability[7] (N ~9,000). Estimates may also not be directly comparable as some have been adjusted for brain size as a covariate and others have not.

[†] Calculated excluding brain size as a covariate. [*] Calculated including brain size as a covariate. [1] Deary et al. (2009); [2] Davies et al. (2015); [3] Plomin et al. (2013), [4] Savage et al. (2018); [5] Hill et al. (2019); [6] Blokland et al. (2012); [7] Zhao, Ibrahim, et al. (2019); [8] Zhao, Luo, et al. (2019); [9] Biton et al. (2020); [10] Winkler et al. (2010)

These positive genetic correlations between cognitive ability and brain regions reported by Biton et al. (2020) stand in conflict with findings from studies adjusting for brain size, as they obtained estimates of genetic correlations between cognitive ability and brain regions near zero, or even negative estimates (e.g., smallest rg = -0.12 between intelligence and caudate; de Vlaming et al., 2021) (see also Grasby et al., 2020; Zhao, Luo, et al., 2019). This demonstrates that adjusting for brain size as a covariate alters genetic correlations as much as phenotypic correlations, producing a largely inconclusive body of evidence. To the best of my knowledge, previous genetic studies have also not yet quantified the genetic overlap between cognitive ability and structural brain networks, as opposed to considering isolated brain regions. Modelling multivariate brain networks at the level of their underlying genetic architecture and assessing the genetic overlap between brain networks and cognitive ability may allow for more robust analyses, than considering isolated regions only. Such multivariate models may also help gain a more biologically informed understanding of their relationships with one another.

1.3.4 Opportunities emerging from adopting a genetically informed level of analysis

Considerable advantages emerge from employing inter-disciplinary, genetically informed analysis techniques to the relationship between cognitive ability and brain morphometry. Firstly, LDSC allows one to straightforwardly integrate genome-wide data into structural MRI analyses. The bivariate LDSC framework reduces genome-wide information to a single estimate of genetic overlap between two traits. As such, LDSC dictates a standardised procedure to reducing dimensionality across millions of genome-wide markers, which is easily applied to GWAS summary statistics of cognitive ability and structural MRI phenotypes.

Secondly, the LDSC framework provides practical advantages that facilitate accessible integration of genome-wide information. LDSC genetic analyses are not restricted by the rarity of large data sets that simultaneously assessed cognitive ability, genetic and structural MRI data, because genetic overlap can be inferred from two sets of GWAS summary statistics calculated in separate samples. LDSC analyses rely on freely available GWAS summary data and do not require sensitive individual-level genetic data often blocked by bureaucracy or paywalls. Thus, researchers can always integrate the newest, and most powerful GWAS summary statistics in their own studies.

Thirdly, genetic studies provide opportunities to replicate and validate known phenotypic correlations using innovative technology. This assumption considerably influenced Chapter 2 and Chapter 4 where I triangulate phenotypic correlations between cognitive ability and brain morphometry using genetically informed methods. Based on a review of mostly animal studies, James Cheverud (Cheverud, 1988)

suggested that phenotypic correlations broadly mirror genetic correlations, and that they can be treated as proxies of one another. Cheverud's Conjecture remains to be widely accepted to apply to different traits and species including humans. Using LDSC, a 2018 study delivered compelling evidence that Cheverud's Conjecture indeed applies to human morphological traits. Specifically, phenotypic and genetic correlations were comparable in magnitude, but on average, genetic correlations tended to be slightly larger than phenotypic ones (Sodini et al., 2018). If Cheverud's Conjecture holds for associations between cognitive ability and brain morphometry, genetic analyses should allow validating previous phenotypic findings from a biologically informed perspective.

Finally, it is an advantage that interpreting results from genetic analyses may be biologically simpler than results from phenotypic analyses. Compared to phenotypic analyses, I suggest that genetic analyses such as LDSC may be more objective as they bring statistical inferences closer to the underlying biology. Theoretically, estimates of genetic overlap should be free of environmental confounding, at least of environments that are not heritable (van Rheenen et al., 2019). Hence, these advantages may allow genetic analyses using LDSC to describe the biological relationship more objectively between cognitive ability and brain morphometry.

To validate phenotypic findings using genetic data, sophisticated methods are required that facilitate complex models of multiple traits on the level of their underlying genetic architecture. Building on genetic correlation matrices derived using bivariate LDSC, GenomicSEM is a genetic multivariate approach that enables testing exploratory as well as confirmatory hypotheses of complex trait interdependencies (Grotzinger et al., 2019). Using GenomicSEM, de la Fuente et al.

(2021) demonstrated that genetic correlations among multiple cognitive abilities mapped well onto one single underlying factor of cognitive ability (*g*-factor). This genetic *g*-factor was very similar to the phenotypic *g*-factor previously described by Spearman (1904) (Fig.1). In line with Cheverud's Conjecture, these genetic analyses helped triangulate that variance across multiple cognitive abilities can be explained best by one latent factor that causes individuals' cognitive test performances. In the same way that de la Fuente et al. (2021) validated the *g*-factor structure using genetic methods, genetic analyses should also help validate phenotypic relationships between cognitive ability and brain morphometry, as well as generate novel biologically informed insights.

1.4 Aims and structure of the thesis

To date, structural MRI studies have delivered conflicting evidence about how brain morphometry relates to cognitive ability (the *g*-factor) which may be due to a lack of methodological consensus. By systematically evaluating methodological decisions, studies in this thesis aim to help reduce some of the arbitrary decisions made across structural MRI studies. Investigations focus on the inclusion of brain size as a covariate in statistical models, as well as brain atlas choices that determine how participants' brain images are subdivided into regions. Informed by such exploratory groundwork, studies aim to characterise the relationships more robustly among structural brain networks, cognitive ability, and age-related processes.

Chapter 2 performs an exploration of how adjusting for brain size as a covariate impacts the statistical relationship between brain regions and cognitive ability. It delivers evidence weighing against brain size adjustment which informs analytical decisions in the following chapters.

Inspired by multivariate genetics techniques such as GenomicSEM (Grotzinger et al., 2019), Chapter 3 derives and validates *Genomic Principal Component Analysis* (Genomic PCA) to add to the existing toolset of multivariate analysis techniques that model associations across multiple traits based on their genetic overlap.

Using Genomic PCA, Chapter 4 mirrors the phenotypic approach used by Madole et al. (2021) to triangulate on a genetic level of analysis whether the central executive brain network plays an important and replicable role in cognitive ability and healthy ageing.

Chapter 5 explores multiple brain atlases that may be used to subdivide participants' brain images into regions. The study compares commonly used atlases to

deliver evidence which atlas may be most appropriate to maximise brain-based predictions of cognitive ability and other behavioural traits.

A final discussion chapter (Chapter 6) considers the overall significance and limitations of the presented work. It also discusses future directions and the benefits of multi-disciplinary studies integrating genetic and MRI data. I elaborate that interdisciplinary studies may be necessary to drive a more comprehensive understanding of the brain as an interconnected system, that depends upon many complex genetic, environmental, and developmental factors.

Chapter 2

Whether to adjust for total brain size: An investigation of the links between regional brain morphometry and cognitive ability

1. Introduction

Brain size is the most robust neuronal correlate of cognitive ability (r = 0.25) (Lee et al., 2019; Pietschnig et al., 2015), but no reliable associations between regional brain morphometry and cognitive ability were found. Studies on healthy individuals investigating associations between regional brain volumes and cognitive ability produced inconsistent results (e.g., Jung & Haier, 2007). Due to these heterogeneous results, a 2020 study concluded that cognitive ability may have no correlates of localised brain morphometry beyond the association between cognitive ability and total brain size (Hilger et al., 2020). Likely reasons for these heterogeneous results are that the underlying studies are statistically underpowered with sample sizes of a median of 25 individuals (Marek et al., 2022). Furthermore, these studies use methodology inconsistently.

One methodological inconsistency is covariate control. Some studies of regional brain correlates of cognitive ability adjust for brain size while others do not (Fig.1). This analytical decision can drastically alter findings such that it can invert the direction of the association (i.e., from positive to negative; Peelle et al., 2012). To improve transparency, it has been suggested that brain imaging studies should report both brain size-adjusted and unadjusted results (O'Brien et al., 2011). Yet simply reporting both is no satisfactory solution because it is unclear how to interpret an association that is positive when unadjusted for brain size but becomes negative when adjusted for brain size.

Fig.1. Model from which correlations of interest are derived in this chapter.

(A) Directed acyclic graph displaying the relationship of interest in the context of the confounder scenario implied by studies adjusting for a covariate. The correlations between 66 regional brain volumes and one of five different cognitive ability definitions is assessed before and after adjusting for brain size as a covariate. The arrows connecting total brain size with regional brain size and cognitive ability are directed outwards to reflect a confounder scenario, which is implied by studies that adjust for total brain size as a covariate. (B) Venn Diagram displaying overlapping variances between cognitive ability, total brain size and regional brain size. This Venn diagram demonstrates how variances may overlap in the relationship of interest. Variance uniquely shared between cognitive ability and regional brain size is considerably smaller after (than before) adjusting for total brain size, because variance associated with total brain size (white circle) is discarded as nuisance when TBV is adjusted for as a covariate.

Researchers rarely explicitly outline their reasons for adjusting for brain size. Not adjusting associations for brain size enables more intuitive interpretation: an unadjusted brain volume should be proportional to the amount of cellular material in mm^3 in this specific region. A measure of unadjusted brain volume is, therefore, directly translatable, for example, to judge whether someone's brain size is healthy compared with population-based brain growth charts (Bethlehem et al., 2022). There seems to be no consensus whether grey matter volume reflects specific neurobiological attributes, but most researchers postulate that volume is a proxy of neuron number per mm^3 (e.g., Genç et al., 2018). Hence, researchers interpret that a positive correlation between brain volume and cognitive ability indicates that greater neural mass leads to higher cognitive ability (Herculano-Houzel, 2017). However, if we considered unadjusted regional brain correlations with cognitive ability in isolation, their magnitude is likely partially driven by the correlated variance between regional and total brain volumes (Fig.1B). Brain size-unadjusted correlations fail to isolate whether those correlations are driven by regional or total brain morphometry.

Following this line of reasoning, studies tend to adjust for brain size as they wish to identify purely regional brain correlates (independent of total brain size). In doing so, they treat inter-individual variance accounted for by total brain size as nuisance variance [7] (e.g., Hilger et al., 2020). Taking this approach, researchers avoid misattributing statistically significant associations between a brain region and

---

[7] I use the term *nuisance variance* in this chapter to refer to the variance indexed by a covariate. I assume that researchers chose to adjust for a covariate because they decided that the covariate confounds the relationship of interest and captures variance that is of no interest to their research question. Hence, researchers adjust for covariates because they understand variance associated with the covariate to be an inconvenience to their statistical model (i.e., nuisance variance). I assume that researchers adjust for covariates in order to discard of the nuisance variance which they believe makes their results less biased and more reliable.

cognitive ability to the region when the association may be confounded by total brain size (Pietschnig et al., 2015) or body size (taller people tend to have larger brains and better cognitive ability; Jerison, 1973). Many studies of healthy brains that adjust for brain size may intend to imitate lesion studies, which can identify whether individual brain regions may be necessary for specific cognitive abilities (e.g., Gläscher et al., 2009). Lesion studies differ from large-scale healthy studies in that they typically examine small groups of patients with localised brain damage.

Compared to an unadjusted regional brain volume, it is more challenging to illustrate the biological interpretation of a brain size-adjusted volume. It has been suggested that it may reflect brain regional deviations in neuron number beyond the expected neuron number with reference to someone's whole brain size (Hilger et al., 2020). Practically, regional and total brain size cannot, however, exist in isolation because brain size is the sum of all regional brain volumes. Based on this definition, regional and total brain size will always correlate substantially. If we accept this fact, adjusting for total brain size introduces multicollinearity leading to statistical issues. The multicollinearity may be so extreme such that directions of associations may get inverted (i.e., flipped correlation signs). Suggestions how to interpret an inverted correlation between regional volume and cognitive ability are ambiguous. For example, Genon et al. (2022) suggest that *"researchers should be aware that the interpretation of results is conditioned by such adjustments"*. How brain size adjustment alters correlations between brain volumes and cognitive ability is unclear and no clear *best practice* on brain size adjustment exists in studies of interindividual differences.

This chapter characterises the impact of brain size adjustment on correlation estimates by exploring associations between cortical brain volumes and cognitive

ability. I calculated correlations between 66 regional volumes and cognitive ability before and after adjusting for total brain volume (TBV) in two large samples (N_{total} ~ 50,000). For this I used five different definitions of cognitive ability. Furthermore, I applied three different analysis techniques to measure associations: phenotypic correlations, genetic correlations, and polygenic scores. The two latter genetic analysis techniques triangulate phenotypic correlations with genetic information. According to Cheverud's Conjecture (refer to Section 1.3.4), genetic correlations should be of similar magnitude as phenotypic correlations (Cheverud, 1988; Sodini et al., 2018). Additionally, I simulated data to benchmark real-world correlational patterns against correlational patterns from simulated data that should reflect the hypothetical scenario that true correlations between regional volumes and cognitive ability only exist if unadjusted for TBV.

Statistical analyses in this chapter are broadly split into two analyses: Analysis 1 and Analysis 2. Analysis 1 describes how brain size adjustment impacts the correlations between 66 brain volumes and cognitive ability. Analysis 1 focuses on the average impact of brain size adjustment across the 66 regions (cross-region approach). Analysis 2 focuses on each brain region individually by testing meta-analytically whether any region is reliably associated with cognitive ability across analysis techniques. I assume in Analysis 2 that a regional volume is truly associated with cognitive ability if I find consistent results across different analyses (multiple samples, cognitive ability definitions, and analysis techniques).

Since the existing literature on the relationship between brain morphometry and cognitive ability is heterogeneous, the working hypothesis of this mass-univariate approach is that there should be no reliable correlations between regional volumes and cognitive ability after TBV adjustment. To illustrate this, Fig.1 displays the relationship

of interest: if my hypothesis was true that there was no true correlation between regional brain volume and cognitive ability after TBV adjustment, the white circle in the Venn diagram (Fig.1B) would overlap with the entire area shared between cognitive ability and regional brain size. In this case, correlations between cognitive ability and regional brain size would either be zero or not consistently replicate across analyses.

2. Methods

2.1. UK Biobank (UKB) cohort

The UK Biobank is an observational study collecting health-related information of half a million community-dwelling participants of middle and old age (40-69 years) across the United Kingdom (Sudlow et al., 2015). The Research Ethics Committee ethically approved the UKB study and participants signed informed consent. Data access was obtained through application 40933. Magnetic resonance imaging (MRI) data were collected using identical MRI scanners and software in Manchester, Newcastle, and Reading. Sixty-six regional Desikan-Killiany (Desikan et al., 2006) volume variables were generated by the UKB team and available for download (Alfaro-Almagro et al., 2018; Smith, Alfaro-Almagro, et al., 2020). The frontal pole region is not provided in the UKB due to high missingness (URL). Total brain volume (TBV) measures were obtained from field ID 25010-2.0. The sample had a mean age of 64 years (range 45.2-82.3) and included 21,217 females (52%; $N = 39,947$).

Individual-level performances in seven cognitive ability tests were provided in the UKB data: Verbal Numerical Reasoning (field ID 20016, 20191), Trail Making – B (field ID 6350, 20157), Matrix Pattern Completion (field ID 6373), Tower Rearranging (field ID 21004), Symbol Digit Substitution (field ID 23324 & 20159), Pairs Matching (field ID 399), and Reaction Time (field ID 200032). The first measurement was considered for each participant. Cognitive tests in the UKB are short and therefore non-standard, but they still have good concurrent validity ($r = .83$) and test-retest reliability (mean Pearson $r = .55$) compared with more comprehensive standard cognitive tests (Fawns-Ritchie & Deary, 2020). Based on the cognitive tests available in UKB, I performed a one-factor confirmatory factor analysis in the lavaan package

in R (Rosseel, 2012) to extract a *g*-factor of cognitive ability. The *g*-factor should explain about 40% of the total variance across cognitive tests (Deary, Penke, et al., 2010). Model fit of the *g*-factor was assessed using standard fit indices: Comparative Fit Index (CFI), Root Mean Square Error of Approximation (RMSEA) (Hu & Bentler, 1998).

2.2. Adolescent Brain Cognitive Development (ABCD) cohort

The ABCD cohort is an observational and longitudinal study tracking developmental trajectories of ~12,000 nine-to-eleven-year-olds. Participants were recruited at 21 sites across the United States of America (Garavan et al., 2018). Individuals were excluded based on practical considerations (e.g., inability to communicate in English), or based on specific developmental, neurological, or psychiatric problems (e.g., schizophrenia, autism). The study was ethically approved by each site's Institutional Review Board. Written informed consent and assent was obtained from the participants and their parents. Analyses presented here used the ABCD Annual Data Release 3.0 (https://abcdstudy.org/scientists/data-sharing/). MRI methods including harmonisation protocols across sites are described in Casey et al. (2018).

FreeSurfer processed cortical volume variables were available for download (file ID: abcd_smrip10201.txt). To match the MRI measures in the UKB, I downloaded 66 Desikan-Killiany (Desikan et al., 2006) volumes and TBV (field ID: smri_vol_cdk_total). I did not include the frontal pole region in ABCD because the frontal pole was not available in UKB. More information on image processing is available online (https://nda.nih.gov/study.html?id=1299).

Fluid and crystallised cognitive ability variables uncorrected for age (file ID: abcd_tbss01) were downloaded as standardised sum scores of test performances across six cognitive abilities measured in the National Institute of Health (NIH) Toolbox Cognition Battery (Akshoomoff et al., 2013). Fluid ability scores were based on vocabulary and oral reading recognition tasks, and crystallised ability scores were based on processing speed, working memory, sequence memory, Flanker, and dimensional change card tasks. Instead of calculating a *g*-factor, I used available fluid and crystalised ability composite scores to increase comparability to previous literature. The composite fluid and crystalised cognitive scores are commonly used (a recent example: Zhao et al., 2022). Siblings were removed at random so that analyses only included unrelated individuals. My analysis included unrelated individuals with available cognitive scores ($N_{crystalised}$ = 8,252; N_{fluid} = 8,147). Mean sample age was 10 years (range 8.9-11.0 years) and it contained 4,273 females (47%).

ABCD genotype data. DNA was extracted from saliva and blood samples and subsequently genotyped on the Affymetrix NIDA SmokeScreen Array (Baurley et al., 2016). The genotype data underwent standard quality control (Lam et al., 2019), and I downloaded the genotype data in PLINK binary format (.fam, .bed, .bim) including 516,598 genetic variants. Ambiguous and mismatched SNPs were cleaned according to the lassosum tutorial (URL). To account for population stratification, I calculated twenty ancestral Principal Components (PCs) based on a set of pruned SNPs using a window of 1500 variants and a shift of 150 variants between windows with an r^2 cut-off of 0.2 (as recommended by Abraham et al., 2017). The first three ancestral PCs are plotted in Fig.2 showing that ancestral PCs captured self-reported race reasonably well. Analyses using this data did not use imputed genetic data.
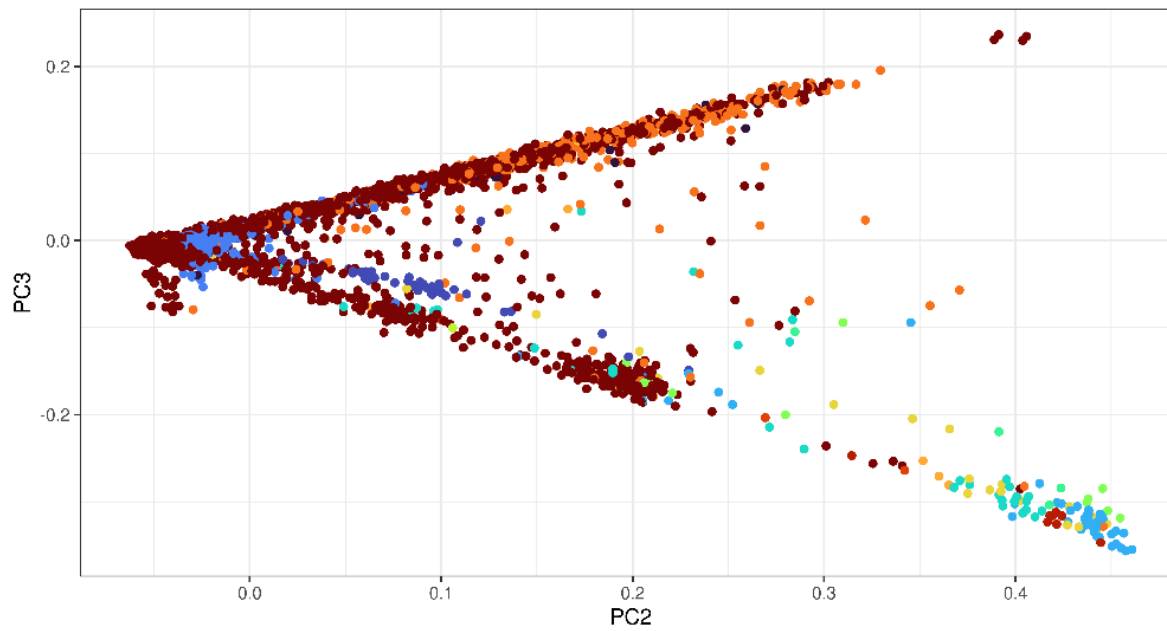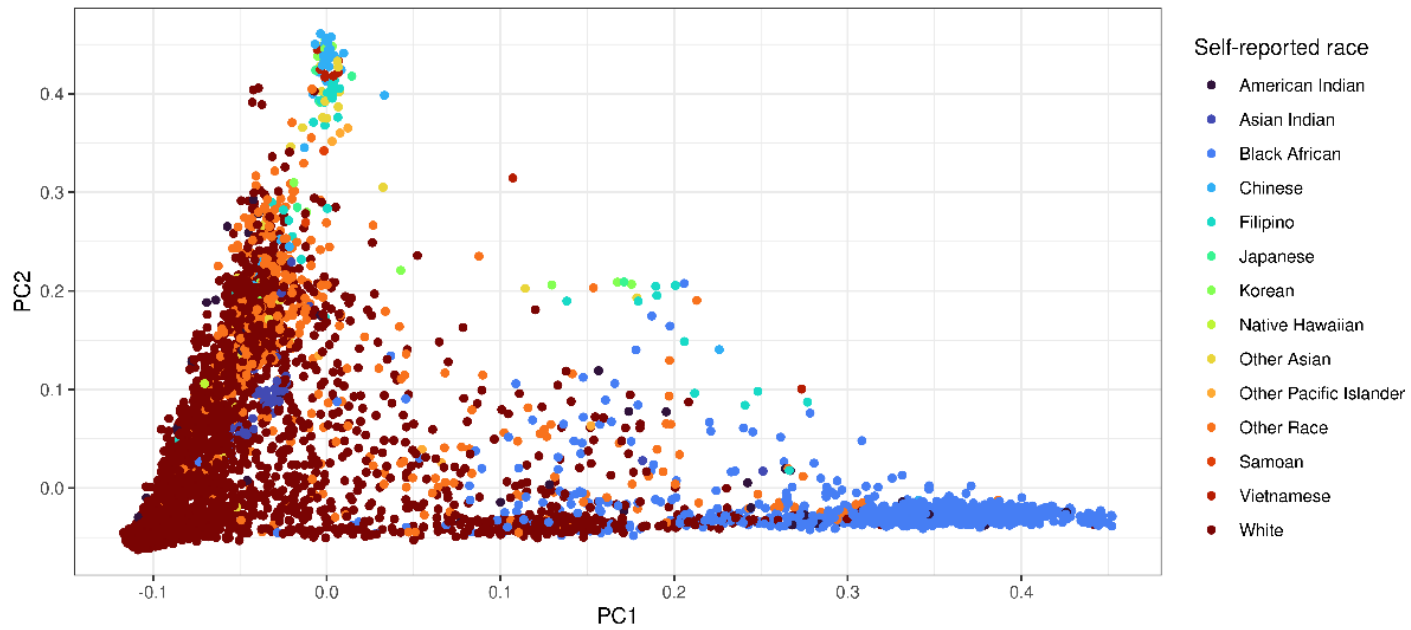
These plots demonstrate that self-reported race (coloured) is broadly clustered along the ancestral PC gradients in the x- and y-axes. For example, self-reported White ancestry tends to score low on PC1, while self-reported Black African ancestry scores high on PC1.

2.3. Simulated data

To benchmark real-world correlations from the real-world observations in UKB and ABCD, I simulated two data sets in which any correlation between regional volumes and cognitive ability was purely driven by TBV. Thus, the simulated data should reflect the hypothetical scenario that true correlations between regional volumes and cognitive ability only exist if unadjusted for TBV. The simulations (Simulation 1 and Simulation 2) both contained 66 random simulated brain volumes that were correlated among each other (N = 10,000). A 67th variable was created as the sum of all 66 simulated regional volumes to represent a simulated measure of TBV. To represent cognitive ability, a 68th random variable was created which was correlated with simulated TBV at a magnitude of 0.25 (as empirically shown in Nave et al., 2018).

Simulation 1 and Simulation 2 differ in their correlation structure among the 66 simulated brain volumes. In *Simulation 1,* the 66 simulated volumes are uniformly correlated at r = 0.4 (Fig.3). In *Simulation 2,* the 66 simulated brain volumes are correlated according to a correlation matrix empirically extracted from UKB data on 66 regional brain volumes (*mean correlation* = 0.4, *SD* = 0.15, range = 0.25 to 0.75; Fig.3). Any association between volumes and cognitive ability in this simulated data must by definition be driven by one region's correlation with TBV. Due to the nature of these simulated data, any significant TBV adjusted correlation reflects a false positive association.

Simulation 1

Correlations between ROIs (all r=0.4)

IQ ~ ROI (simulated)

no TBV adjustment
TBV adjusted

IQ ~ ROI (simulated)

Effect size

Simulation 2

Correlations between ROIs (mean r=0.4; SD = 0.15)

IQ ~ ROI (simulated)

no TBV adjustment
TBV adjusted

IQ ~ ROI (simulated)

Effect size
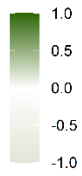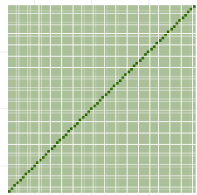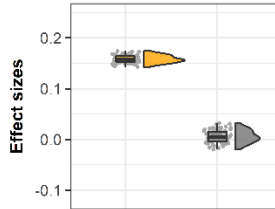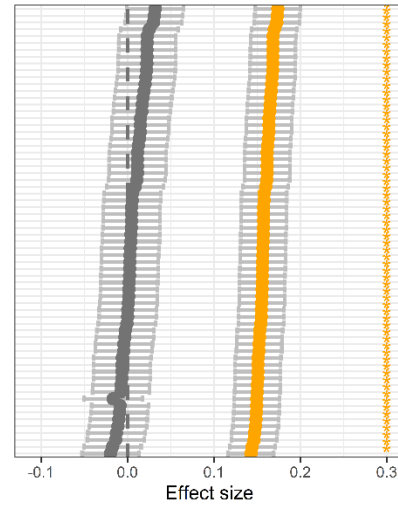
Simulation 2 with increasing N

IQ ~ ROI (N = 20000)

Effect size

IQ ~ ROI (N = 30000)

Effect size

IQ ~ ROI (N = 40000)

Effect size

IQ ~ ROI (N = 50000)

Effect size
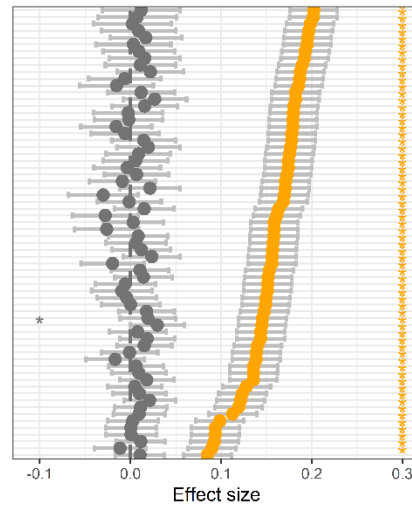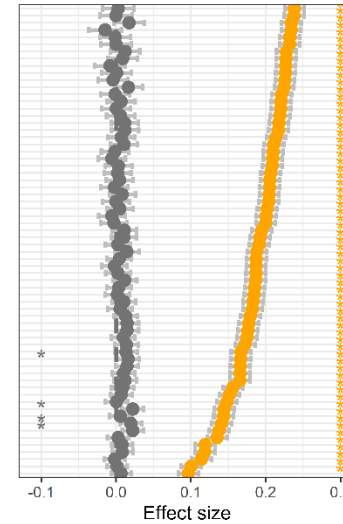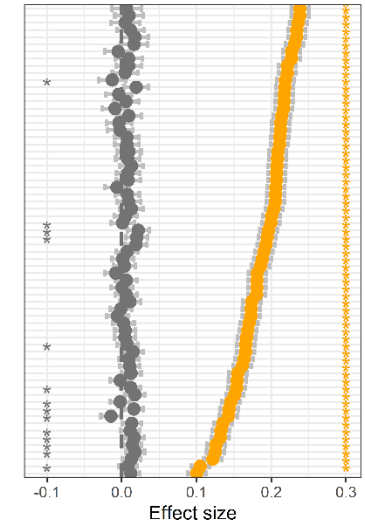
Left panel: Simulation 1 and Simulation 2 both contain 66 regional volumes (ROIs) that are either correlated among each other at a uniform magnitude of $r = 0.4$ (*Simulation 1*), or according to an empirical correlation matrix obtained from UKB MRI data (*Simulation 2*). The regional volumes are summed up to represent TBV, and I generated an additional variable to represent cognitive ability which is correlated with simulated TBV at $r = 0.25$ (as in Pietschnig et al., 2015). In this data, I calculated correlations between cognitive ability and each of the 66 simulated regional volumes. The figure shows effect sizes unadjusted for TBV in orange, and effect sizes adjusted for TBV are in grey. Correlations marked with an asterisk (*) are statistically significant after correcting for multiple testing (99% confidence interval around the correlation). *Right panel:* Simulation 2 in incrementally larger sample sizes ($N = 20{,}000$; $N = 30{,}000$; $N = 40{,}000$, $N = 50{,}000$). Due to the nature of these simulated data, any significant TBV adjusted correlation between a volume and cognitive ability are false positive associations. The simulations with increasing N show that TBV adjusted correlations are more likely positive at larger sample sizes N.

2.4. Genome-wide association data

I used genetically informed analyses to triangulate correlations between regional volumes and cognitive ability. The calculation of genetic correlations required two sets of 66 GWAS summary statistics of regional brain volume: one set adjusted for a proxy of brain size (i.e., volumetric scaling factor) (Smith, Douaud, et al., 2020) and the other that I generated myself as part of Chapter 4 was not adjusted for brain size (Fürtjes et al., 2021). Both sets of GWAS were based on the Desikan-Killiany brain volumes and approximately the same subsample of UKB participants (N~33.000 by Smith, Douaud, et al. (2020) and N~36.000 in Chapter 4). Both sets of GWAS were adjusted for age and sex, because age and sex should not be correlated with the genome but may spuriously induce genetic trait correlations.

I downloaded the GWAS summary statistics adjusted for a proxy of brain size from the *big40* repository (Smith, Douaud, et al., 2020). In addition to brain size, age, and sex, this GWAS set was also adjusted for another ~200 covariates (as outlined in Alfaro-Almagro et al., 2021). Excluding brain size as a covariate, I adjusted the GWAS of regional brain volumes only for imaging centre, X, Y, and Z scanner coordinates, age, and sex (see Appendix).

To calculate genetic correlations between regional volumes and cognitive ability (described below in Section 2.5.2), I downloaded intelligence quotient (IQ) GWAS summary statistics (N ~ 100,000; Davies et al., 2018), and education GWAS summary statistics (N ~ 1.1Mio; Lee, Wedow, et al., 2018). Educational attainment is frequently used as a proxy phenotype for cognitive ability as it can be self-reported allowing the collection of very large samples.

2.5. Obtaining correlations between cognitive ability and regional brain volume from multiple analyses

I obtained correlations between cognitive ability and 66 regional brain volumes from multiple analyses including phenotypic analyses, genetic analyses, and polygenic score analyses (Fig.4). Variables were standardised (mean = 0; *SD* = 1) prior to analyses so effect sizes were interpretable in standard units. Correlations adjusted for TBV were obtained from linear regression models in which cognitive ability was the outcome, regional brain volume the predictor, and TBV the covariate (cognitive ability ~ regional volume + TBV). For the partial regression coefficient of regional volume to be interpretable as a correlation, *p*-values of a regions' regression effect on cognitive ability were converted to correlations via the *t*-statistic [8]. This ensured that correlation magnitudes were comparable across phenotypic correlations, genetic correlations, and polygenic score correlations with and without TBV adjustment.

2.5.1. Obtaining phenotypic correlations

Phenotypic correlations between cognitive ability and 66 regional volumes were calculated using linear regressions. Regressions were performed without adjusting (cognitive ability ~ regional volume) and with adjusting for TBV (cognitive ability ~ regional volume + TBV), which resulted in *2 x 5 x 66 = 660* correlations. Correlations

---

[8] I did this using the *p2cor*() function from the lassosum package (Mak et al., 2017). Taking the regression coefficient *p*-value and degrees of freedom as input, the function works out the *t*-statistic using the *t* distribution *qt*() function in R. The *t*-statistic is then transformed into a correlation using the following formula:

$$r = \frac{t}{\sqrt{(df - 2) + t^2}}$$

were calculated based on my two simulated data sets (Simulation 1 & Simulation 2, $N_{simulated}$ = 10,000), the UKB sample (g-factor, N = 39,947), and the ABCD sample (i.e., crystalised & fluid cognitive ability, $N_{crystalised}$ = 8,252; N_{fluid} = 8,147). No additional covariates were included. Regional brain volumes differ by age as some regions atrophy faster than others (Raz et al., 2010) but including age in these models would prevent this analysis from isolating the impact of TBV adjustment on the correlations of interest. To test if age may alter the correlations of interest, I calculated the correlations in two samples of different age: adults in the UKB (mean age = 64.26 years) and adolescents in the ABCD study (mean age = 9.9 years). I did not include sex as a covariate because the association between regional volumes and cognitive ability did not vary by sex in previous analysis of the UKB (Cox et al., 2019).

2.5.2. Obtaining genetic correlation

Using bivariate LDSC (Bulik-Sullivan, Finucane, et al., 2015), I calculated genetic correlations among cognitive ability, educational attainment and 66 regional brain volumes with and without adjustment for brain size. I used the GWAS of cognitive ability (Davies et al., 2018), educational attainment (Lee, Wedow, et al., 2018), 66 brain volumes adjusted for brain size (Smith, Douaud, et al., 2020), and 66 brain volumes unadjusted for brain size (Fürtjes et al., 2021) (Section 2.4). I obtained *2 x 2 x 66 = 264* genetic correlations.

2.5.3. Obtaining polygenic score correlations

I calculated polygenic scores for 66 regional volumes in unrelated ABCD participants to generate measures of the genetic propensity to 66 regional volumes. Polygenic scores are the sum of genome-wide alleles that predispose an individual to a larger (or smaller) regional brain volume. Based on GWAS adjusted (Smith, Douaud,

et al., 2020) or unadjusted for brain size (Fürtjes et al., 2021), I calculated polygenic scores using lassosum (Mak et al., 2017) which is a frequentist penalty-based shrinkage method. Lassosum is computationally efficient and does not require clumping [9] which is associated with better predictive abilities (Pain et al., 2021). To avoid overfitting, split-validation was performed as implemented in the *splitvaldidate* R function by Mak et al. (2018).

Finally, I calculated linear regressions in ABCD with crystalised and fluid cognitive ability as outcomes and regional volume polygenic scores as predictor variables. The models included age, sex, and 20 ancestral PCs as covariates (cognitive ability ~ polygenic score + age + sex + 20 PCs). These analyses resulted in *2 x 2 x 66 = 264* polygenic score correlations.

2.6. Analysis 1: Exploring correlations between regional volumes and cognitive ability

To describe the distribution of correlations between the 66 regional volumes and cognitive ability, I calculated mean correlations, and 95% confidence intervals from multiple analyses (i.e., analyses outlined in Section 2.5). I compared correlations based on real-world data in ABCD and UKB with correlations from the simulated data to benchmark against the simulated data in which correlations among regional volumes and cognitive ability were entirely driven by brain size (i.e., no regional

---

[9] Clumping means that SNPs are filtered to obtain a selection of LD independent SNPs with the strongest signals in a GWAS. Traditionally, polygenic scores would be calculated as a weighted sum based on those top filtered SNPs. Using lassosum, the SNP feature space is not reduced by clumping but through shrinkage of the least associated SNPs.

variance unique of TBV). I used Wilcoxon signed-rank analysis to quantify whether mean ranks of the real-world correlation distributions differed from the simulated correlations (results reported in Section 3.1.1). I used the non-parametric Wilcoxon method because real-world correlations were not normally distributed (as shown in Fig.4).

Analysis 1 includes a sub-analysis measuring the extent to which correlations between regional volumes and cognitive ability are driven by a region's correlation with TBV. I calculated the linear associations between two vectors: the first vector contained correlations between regional volumes and cognitive ability (obtained from different analyses outlined in Section 2.5) and the second vector contained correlations between regional volumes and TBV. The strength of the association between the two vectors will indicate the extent to which one can predict a regions' correlation with cognitive ability purely by knowing the regions' correlation with TBV (results in Section 3.1.2).

I performed an additional simulation to explore the impact of TBV adjustment on the variance inflation factor (VIF) [10] which tends to be used to decide whether covariate adjustment is problematic. I simulated data keeping constant a hypothetical

---

[10] Exactly how large the VIF must be to cause concern is debated but as a rule of thumb studies tend to accept a VIF <5 (Gareth et al., 2013) as unproblematic. Some studies set this threshold at VIF <2.5 (Johnston et al., 2018) and others at VIF <10 (Vittinghoff et al., 2006). The following formula calculates the VIF when no other covariate is included in the statistical model:

$$VIF = \frac{1}{1 - (r_{predictor\ \&\ covariate})^2}$$

For example, a VIF of 1.33 is equivalent to a regions' correlation with TBV of $r = 0.5$.

correlation between a brain region variable (predictor) and cognitive ability variable (outcome). This constant correlation was at $r = 0.10$ in Fig.6A, $r = 0.15$ in Fig.6B, $r = 0.20$ in Fig.6C. Additionally, I simulated the region's correlation with TBV (covariate) at 19 different values between $r = 0.00$ and 0.95 that each corresponded to a specific VIF [9]. Using these simulated data, I compared the regression coefficients adjusted and unadjusted for TBV and showed how the coefficients change with reference to a regions' correlation with TBV and corresponding VIF. Results of this additional simulation are reported in Section 3.1.3.

2.7. Analysis 2: Region-by-region meta-analyses of correlations between regional volumes and cognitive ability across analyses

It was the aim of Analysis 2 to test using 66 meta-analyses whether any of the 66 regional volumes yielded consistent non-zero correlations with cognitive ability. Each meta-analysis considered five correlations between a specific region and cognitive ability that were obtained from five different real-world analyses: a phenotypic correlation in UKB (g-factor), two phenotypic correlations in ABCD (fluid & crystalised cognitive ability), a regions' genetic correlation with IQ and a regions' genetic correlation with education.

I used the meta package in R (Schwarzer et al., 2015) to perform the meta-analyses and to obtain a random effects estimator for each of the 66 regions. The *metacor* function provides random effects meta-analytic results. I report random-effects instead of fixed-effect correlations because it is reasonable to expect some between-study variance in correlation magnitudes obtained from different definitions of cognitive ability and analysis techniques. As standard errors in genetic correlations

are not directly proportional to the sample size (as is the case in phenotypic correlations), the appropriate weighting of genetic correlations in the meta-analyses was ensured by iteratively finding a representative sample size that is required by the *metacor* function to produce confidence intervals equivalent to those obtained from LDSC [11]. Random-effects meta-analytic correlations and multiple testing corrected 99% confidence intervals are displayed in Fig.8A.

The *metacor* function also tests using the homogeneity statistic Q whether different analyses produced correlations that likely share a common population effect size. Q is chi-squared distributed at *k-1* degrees of freedom (*k* = number of correlations obtained from different analysis techniques that were included in the meta-analysis), and indicates, if significant, that the variance between considered correlations is greater than would have been expected, had all studies shared a population effect size. I^2 indicates the proportion of variability of measures included in the meta-analysis that can be explained by differences between the different analyses.

---

[11] The *metacor* function in R requires the user to input correlations and corresponding sample sizes to calculate the correlations' confidence intervals. Simply inputting GWAS sample sizes however would not have produced accurate LDSC confidence intervals for the genetic correlations. I coded a somewhat inefficient function to iteratively work out the representative sample size that would produce confidence intervals in the *metacor* function equivalent to the confidence intervals obtained in LDSC. Specifically, I would feed the genetic correlation and a randomly generated sample size to the *metacor* function, extract the confidence interval calculated by *metacor*, and compare it with the corresponding confidence interval obtained with LDSC. For example, if the confidence interval given by the *metacor* function was too small, my function would repeat this process by inputting a smaller sample size (and vice versa) until the input sample size would produce the accurate LDSC confidence interval. Therefore, the sample size given in Fig.8 does not represent the real GWAS sample size based on which LDSC was calculated but it is a representative sample size that ensured appropriate weighting of the genetic correlations in the meta-analyses.

3. Results

3.1. Analysis 1: Describing correlations between 66 regional volumes and cognitive ability

As part of Analysis 1 I explored the average distribution of correlations between 66 regional volumes and cognitive ability – before and after brain size adjustment (cross-region approach). The distributions are displayed for each analysis adjusted and unadjusted for TBV (Table 1, Fig.4). Using confirmatory factor analysis, cognitive ability in the UKB sample was quantified by a factor of general cognitive ability (*g*-factor) which accounted for 34% of the variance across cognitive tests. The *g*-factor model demonstrated good model fit (CFI = 0.97, RMSEA = 0.05) (Hu & Bentler, 1998).

A Phenotypic assoc. g-factor (UKB)

B Phenotypic assoc. cryst. IQ (ABCD)

C Phenotypic assoc. fluid IQ (ABCD)

D Genetic corr. education

E Genetic corr. IQ

F Polygenic pred. cryst. IQ

G Polygenic pred. fluid IQ

no TBV adjustment
TBV adjusted
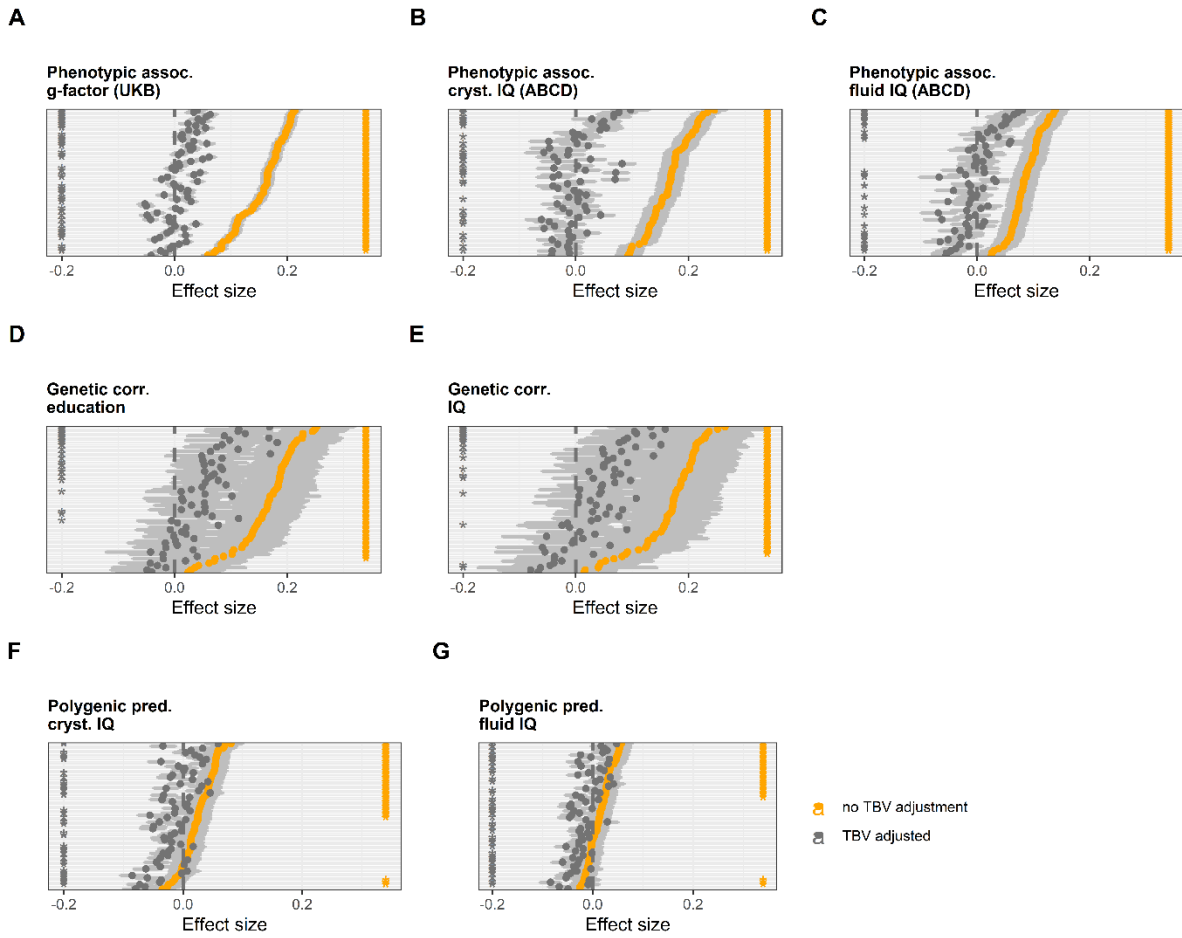
Effect size distributions
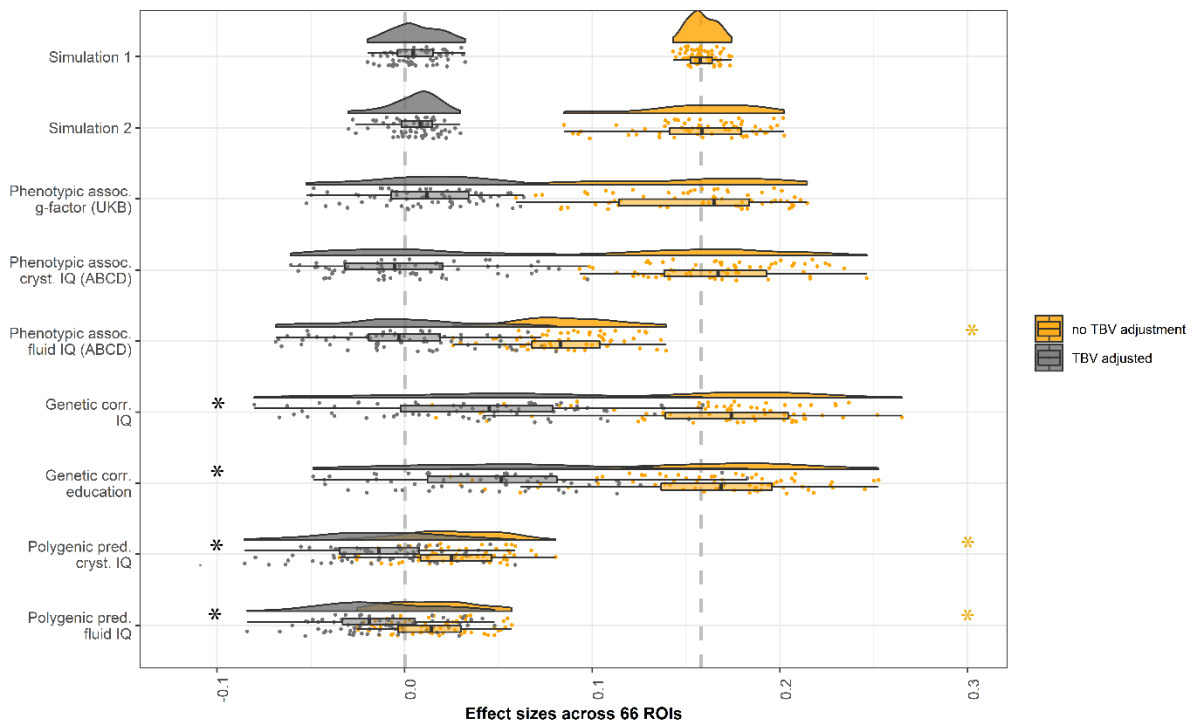
Simulation 1
Simulation 2
Phenotypic assoc. g-factor (UKB)
Phenotypic assoc. cryst. IQ (ABCD)
Phenotypic assoc. fluid IQ (ABCD)
Genetic corr. IQ
Genetic corr. education
Polygenic pred. cryst. IQ
Polygenic pred. fluid IQ

Effect sizes across 66 ROIs

no TBV adjustment
TBV adjusted

62

Upper panel: (A) Phenotypic correlations between 66 regional volumes and the *g*-factor in the UKB (N = 39,947). (B) Phenotypic correlations between 66 regional volumes and crystallised IQ in ABCD (N = 8,252). (C) Phenotypic correlations between 66 regional volumes and fluid IQ in ABCD (N = 8,147). (D) Genetic correlations between 66 regional volumes and IQ calculated using linkage disequilibrium score regression (LDSC) and GWAS summary statistics as input. (E) Genetic correlations between 66 regional volumes and educational attainment. (F) Correlations between 66 regional volume polygenic scores and crystalised IQ in ABCD calculated using lassosum (N = 8,252). (G) Correlations between 66 regional polygenic scores and fluid IQ in ABCD (N = 8,147). Correlations in orange are unadjusted for TBV, in grey adjusted for TBV. Correlations are sorted from largest to smallest unadjusted correlation.

Bottom panel: Distributions of correlations between 66 cortical regions and cognitive ability displayed by analysis type (i.e., different samples, phenotypic and genetic analyses, and multiple cognitive ability definitions). Distributions marked with an asterisk (*) were significantly different in their mean ranks from Simulation 2 (Wilcoxon analysis). Note that the correlations between the 66 volumes and *fluid* cognitive ability as defined in ABCD were smaller than those in the other analyses displayed in the bottom panel. That the fluid cognitive ability definition in ABCD produces smaller correlations with brain volume was previously shown in (Zhao et al., 2022).

Table 1. Describing the distributions of correlations between 66 regional volumes and cognitive ability adjusted and unadjusted for TBV.

Analysis	TBV adjusted			No TBV adjustment		
	Mean correlation [95%]	Minimum - Maximum	*p*	Mean correlation [95%]	Minimum - Maximum	*p*
Simulation 1	0.01 [0.01 to 0.01]	-0.02 - 0.03	0.886	0.16 [0.16 to 0.16]	0.14 - 0.17	0.550
Simulation 2	0.01 [0.01 to 0.01]	-0.03 - 0.03	1	0.16 [0.15 to 0.17]	0.08 - 0.2	1
Phenotypic assoc. g-factor (UKB)	0.01 [0.01 to 0.01]	-0.04 - 0.05	0.177	0.15 [0.14 to 0.16]	0.06 - 0.21	0.976
Phenotypic assoc. cryst. IQ (ABCD)	0.00 [-0.01 to 0.01]	-0.05 - 0.08	0.038	0.17 [0.16 to 0.18]	0.09 - 0.25	0.329
Phenotypic assoc. fluid IQ (ABCD)	0.00 [-0.01 to 0.01]	-0.05 - 0.06	0.062	0.09 [0.08 to 0.1]	0.03 - 0.14	2.26×10^{-19} *
Genetic corr. IQ	0.04 [0.03 to 0.05]	-0.08 - 0.16	2.15×10^{-5} *	0.16 [0.15 to 0.17]	0.01 - 0.26	0.035
Genetic corr. education	0.05 [0.04 to 0.06]	-0.05 - 0.18	1.45×10^{-7} *	0.16 [0.15 to 0.17]	0.02 - 0.25	0.189
Polygenic pred. cryst. IQ	-0.01 [-0.02 to 0.00]	-0.09 - 0.06	0.000 11 *	0.03 [0.02 to 0.04]	-0.03 - 0.08	3.76×10^{-23} *
Polygenic pred. fluid IQ	-0.01 [-0.02 to 0.00]	-0.08 - 0.05	1.09×10^{-5} *	0.01 [0.01 to 0.01]	-0.03 - 0.06	3.76×10^{-23} *

Mean correlations, 95% confidence intervals, minimum and maximum values were calculated for the distribution of correlations between regional volumes and cognitive ability differently defined and measured in different samples. Summary statistics are shown for each analysis separately. *p* indicates the *p*-value for a Wilcoxon signed-rank test comparing ranked means of simulated correlations (Simulation 2) with correlations between regional volumes and cognitive ability obtained from each real-world correlations. Values marked with an asterisk were statistically significant after correcting for multiple testing.

3.1.1. Adjusting for TBV substantially reduces correlations between cognitive ability and 66 regional volumes

Overall, adjusting for TBV substantially reduced correlation magnitudes (Fig.4B): Without TBV adjustment, correlations between regional volumes and cognitive ability were distributed around r ~0.15. With TBV adjustment, these correlations were closely distributed around zero (Table 1). Distributions from some analyses (e.g., genetic correlations unadjusted for brain size) significantly differed from this correlational pattern whereby TBV adjustment did *not* reduce average correlations to r ~0. Those divergent distributions are marked with asterisks in Fig.4B and are further outlined below (remainder of Section 3.1.1).

Genetic correlations. Distributions of genetic correlations (r_g) obtained from GWAS *adjusted* for a proxy of brain size were distributed around a mean value significantly larger than zero (IQ: r_g = 0.04 [0.03-0.05], education: r_g = 0.05 [0.04-0.06]). This was considerably larger than the TBV-adjusted correlations in the other analyses (Fig.4B). By contrast, genetic correlations *unadjusted* for brain size were similar to those from other analyses (both IQ and education r_g = 0.16 [0.15-0.17]).

Polygenic score associations. Correlations between regional volume polygenic scores and cognitive ability were overall smaller than the correlations in other analyses (unadjusted and adjusted for proxy of brain size; Table 1). This is not further discussed below because it is expected that polygenic score associations tend to be smaller than phenotypic and genetic correlations (Dudbridge, 2013). The distributions of correlations between regional volume polygenic scores and cognitive ability still support the overall trend that TBV adjustment reduces volume cognitive ability correlations to an average of zero.

Phenotypic correlations in ABCD (fluid cognitive ability). Without TBV adjustment, cognitive ability brain volume correlations obtained with the fluid cognitive ability definition in ABCD was smaller than expected from other analyses (r = 0.09 [0.08-0.10], Table 1). This is not further discussed below because the distribution obtained from the fluid cognitive ability definition still supports the overall trend that TBV adjustment reduces correlation magnitudes to an average of zero (r = 0 [-0.01-0.01]). That the fluid cognitive ability definition produces smaller brain volume correlations than the crystalised cognitive ability definition in the same sample was previously shown in Zhao et al. (2022).

3.1.2. Correlations between regional volumes and cognitive ability are larger the stronger a region is correlated with TBV
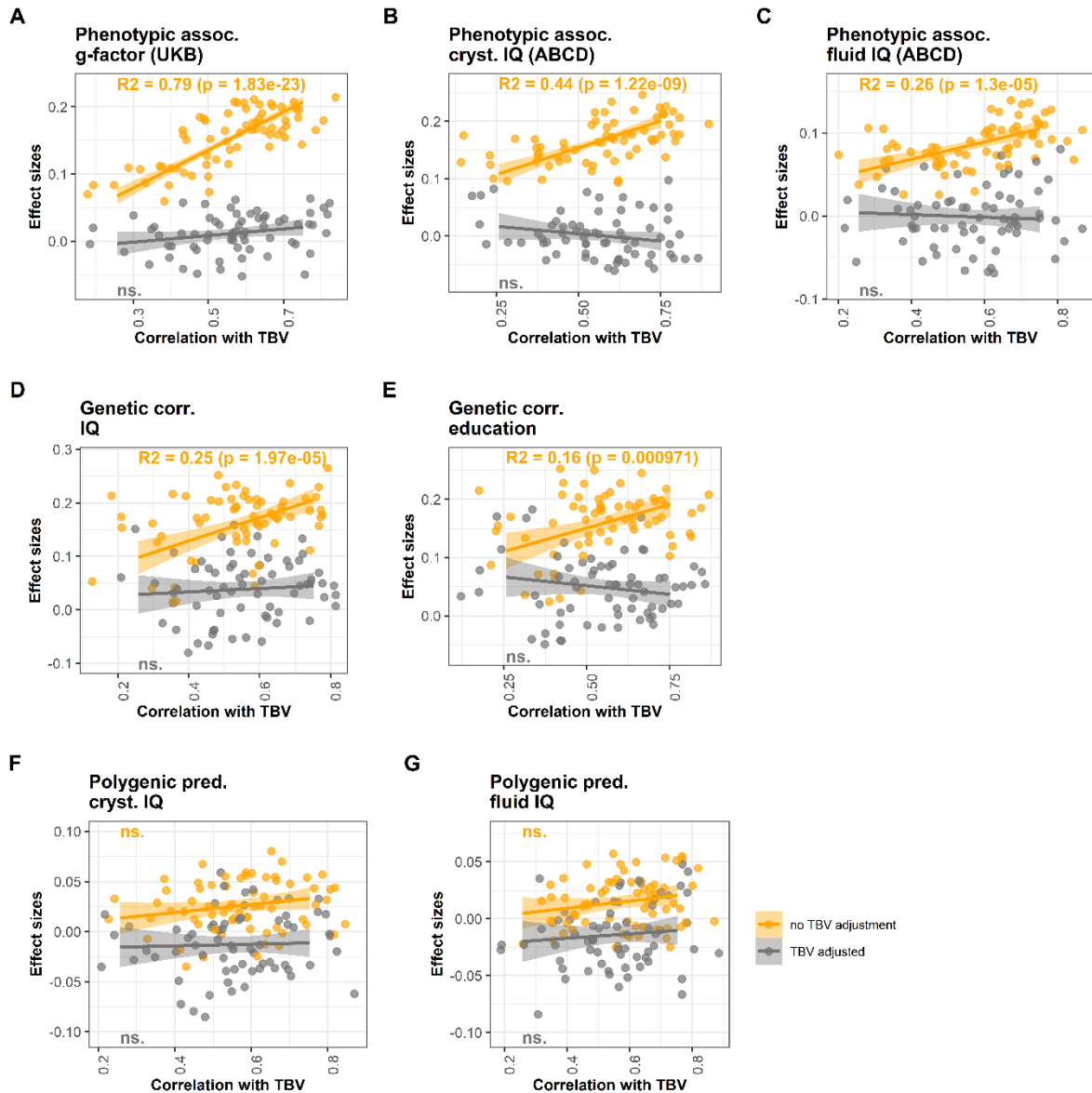
Fig.5. Relationship between a vector of correlations between 66 regional volumes and cognitive ability (y-axis) and a vector of correlations between the same 66 regional volumes and TBV (x-axis).

Each plotted dot represents a brain region. Effect sizes of correlations between regional volumes and cognitive ability in the y-axis in each panel are taken from different analyses: (A) *g*-factor cognitive ability definition in UKB sample, (B) crystalised IQ cognitive ability definition in ABCD, (C) fluid IQ definition in ABCD, (D) genetic correlations using IQ definition, (E) genetic correlations using education as proxy phenotype for cognitive ability, (F) polygenic score associations with crystalized IQ cognitive ability definition, (E) polygenic score associations with fluid IQ cognitive ability definition. Correlations between regional volumes and TBV in the x-axis were calculated in the UKB sample. The x-axis correlations are the same in panels A-F.

67

This sub-analysis of Analysis 1 quantified correlations between two vectors: **(1)** 66 TBV-unadjusted correlations between cognitive ability and 66 regional volumes, and (y-axes in Fig.5A-G) **(2)** 66 correlations between the same 66 regional volumes and TBV (x-axes in Fig.5A-G). The regions' correlations with TBV were calculated based on UKB data and ranged from 0.26 to 0.75 with a median of 0.58. Substantial associations between the two vectors showed that a correlation between a regional volume and cognitive ability was larger when the region was more strongly associated with TBV. Depending on the analysis type the correlations between cognitive ability and brain volumes were calculated in, the vector associations in this analysis ranged in their R^2 from 16% to 79% (Fig.5). For example, by knowing a region's correlation with TBV one may predict with up to 79% accuracy how correlated this region will be with cognitive ability (Fig.5A). These substantial vector associations indicate that TBV drove the correlations between regional volumes and cognitive ability. TBV adjustment seems to affect a region's correlation with cognitive ability differentially relative to the region's correlation with TBV [12]. This was not the case for polygenic score correlations as they were not associated with a regions' correlation with TBV (Fig.5F-G).

---

[12] I have heard proponents of TBV adjustment argue that they think adjusting for TBV should impact different regions of the brain uniformly in how related they are to a trait. They assumed that even after TBV adjustment we would maintain a ranked order between brain volumes in how associated they are with a trait of interest. However, the analysis displayed in Fig.5 shows that each region is impacted by TBV adjustment differently relative to a regions' correlation with TBV. This correlation with TBV widely ranged between 0.26 – 0.75 for different regions.

3.1.3. A variance inflation factor <2.5 may not be a sufficient criterion to decide whether TBV adjustment is problematic

This additional simulation explored the impact of TBV adjustment on the variance inflation factor (VIF, explained in Section 2.6). VIFs below 2.5 tend to be considered unproblematic (Johnston et al., 2018). Fig.6 illustrates that a correlation between cognitive ability (outcome) and the brain region (predictor) is reduced through TBV (covariate) adjustment as a function of a regions' correlation with TBV. The stronger the correlation with TBV, the more likely the correlation of interest is flipped into a negative estimate. This demonstrates that negative correlations can result as an artefact of covariate over-adjustment whereby covariate adjustment removes all the systematic variance between predictor and outcome.

A

Simulated effect size held constant (r = 0.10)
while increasing the correlation between region and TBV
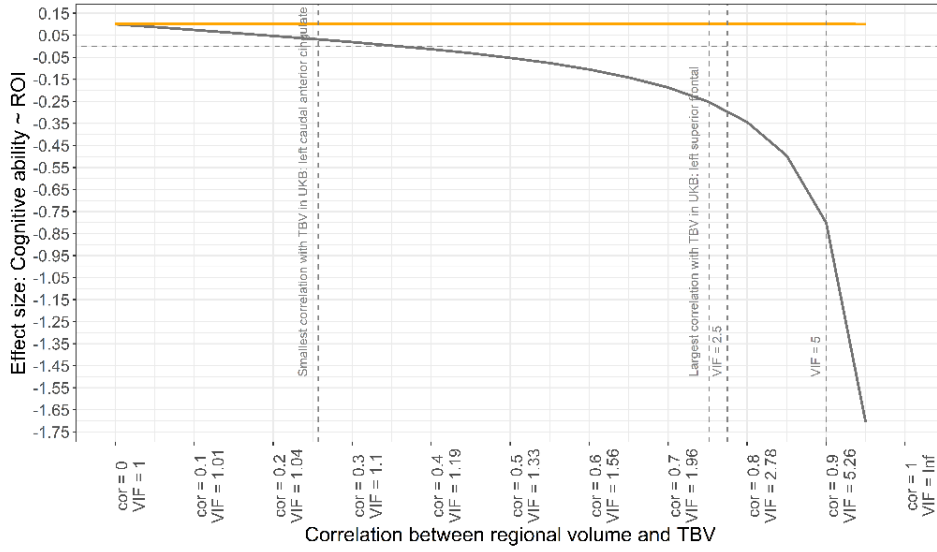
B

Simulated effect size held constant (r = 0.15)
while increasing the correlation between region and TBV
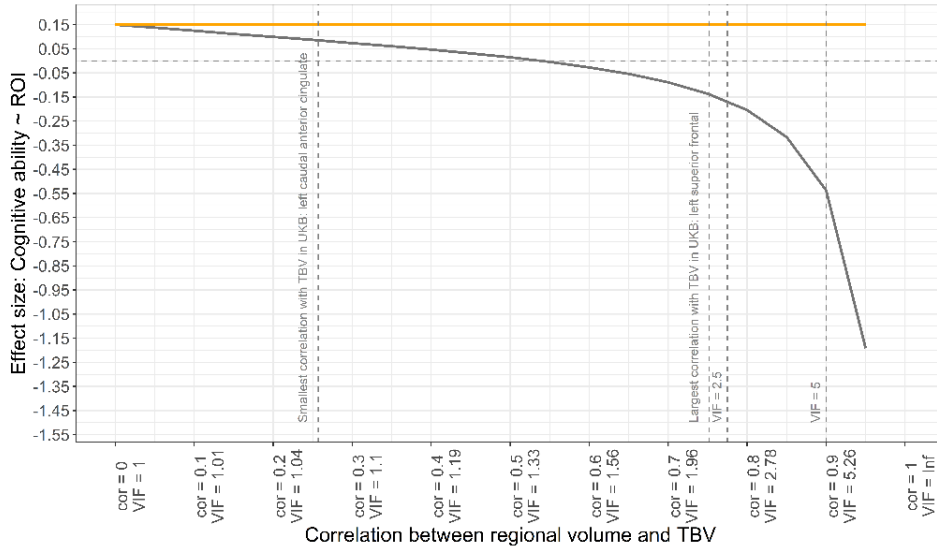
C

Simulated effect size held constant (r = 0.20)
while increasing the correlation between region and TBV
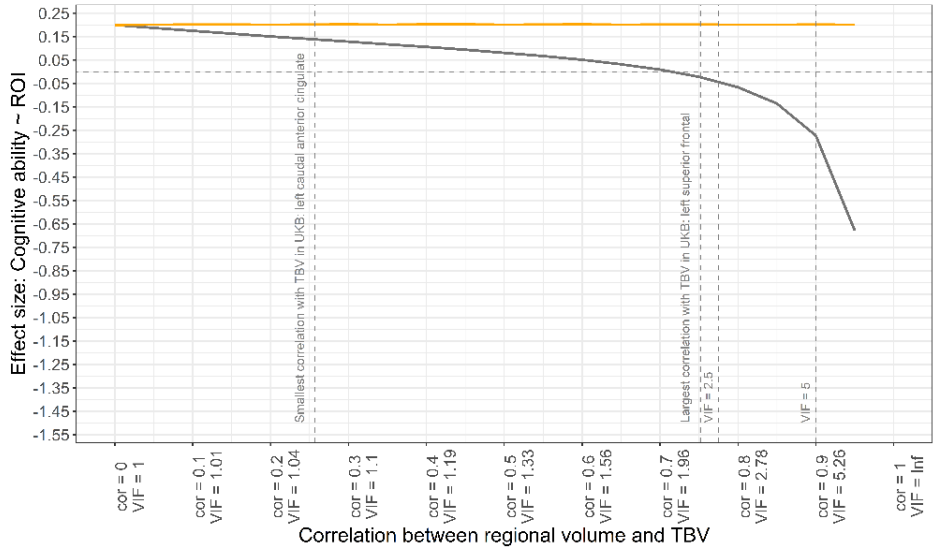
The correlation between regional volume (predictor) and cognitive ability (outcome) is held constant. The correlation between the region (predictor) and TBV (the covariate) is continuously increased from left to the right on the x axis to illustrate the impact of TBV adjustment. The larger a regions' correlation with TBV the more likely it is that the adjusted correlation between a region and cognitive ability is negative. **Panel A** shows the unadjusted correlation between regional volume and cognitive ability constant at 0.10, **Panel B** at 0.15, and **Panel C** at 0.20.

Fig.6A-C shows that VIF <2.5 may not be a sufficient criterion to judge whether TBV is an appropriate covariate: Even at a VIF <2.5, TBV adjustment can invert the direction of a correlation. Whether TBV adjustment inverted the correlation depended on how great the correlation was prior to TBV adjustment (represented by the orange line in Fig.6A-C). When the correlation between the TBV unadjusted region and cognitive ability was small (r = 0.10; Fig.6A), TBV adjustment flipped their correlation at VIFs as small as 1.14 (equivalent to $r_{TBV\ \&\ region}$ = 0.35). When the correlation between the TBV unadjusted region and cognitive ability was larger prior to TBV adjustment (r = -.20; Fig.6C), TBV adjustment flipped their correlation only at larger VIFs (e.g., VIF >1.95, $r_{TBV\ \&\ region}$ = 0.70). This shows that whether a covariate can be safely adjusted for depends on interdependencies between all the variables in the model, not just the relationship between predictor and covariate as is considered by the VIF. Thus, the VIF alone is not a sufficient criterion to decide whether a covariate can be safely adjusted for.

Crucially, the Fig.6 simulation demonstrated that negative correlations between TBV-adjusted volumes and cognitive ability are an artefact of TBV over-adjustment. TBV over-adjustment occurs in cases where the unadjusted correlation between a region and cognitive ability is so negligibly small relative to the variance accounted for by TBV, that no systematic variance remained after discarding of TBV associated variance through covariate adjustment. Hence, TBV adjustment may not be advisable because this brute force approach has the potential to discard most, if not all the variance shared between a regional volume and cognitive ability.

3.2. Analysis 2: Region-by-region meta-analyses of correlations between volumes and cognitive ability

In Analysis 2, I conducted 66 region-specific meta-analyses to identify regional volumes that had reliable correlations with cognitive ability across multiple analysis types (multiple samples, different cognitive ability definitions, and analysis techniques). Without TBV adjustment, the 66 meta-analysed correlations between each regional volume and cognitive ability were all significantly different from zero (mean unadjusted r = 0.14, range = 0.06 to 0.20; Fig.8*B*). These significant correlations were likely largely driven by TBV.

By contrast, only 14 (out of 66) regional volumes *adjusted* for TBV yielded significantly non-zero meta-analysed correlations with cognitive ability (Table 2). As negative correlations likely result from TBV over-adjustment (as demonstrated in Section 3.1.3), Table 2 only includes the eleven regions that yielded positive meta-analytic correlations with cognitive ability after TBV adjustment. The following four regions replicated across hemispheres and were weakly associated with cognitive ability: *parahippocampal* (meta-analysed $r_{\text{left hemisphere}}$ = 0.05 [0.03 to 0.08], right hemisphere r = 0.04 [0.01 to 0.08]), *precentral* (left r = 0.05 [0.04 to 0.06], right r = 0.04 [0.03 to 0.05]), *middle temporal* (left r = 0.03 [0.02 to 0.04], right r = 0.04 [0.01 to 0.06]), and *lateral orbitofrontal gyrus* (left r = 0.02 [0.01 to 0.03], right r = 0.01 [0.00 to 0.02]). Forest plots for the significant meta-analysed effects are displayed in Fig.7.

The fourteen significantly positively associated regions showed low heterogeneity (range Q across the 14 regions = 1.19 to 19.71, *mean* = 5.68) compared with all other regions (range Q across all regions adjusted for TBV = 0.41 to 114.61,

mean = 22.69; Fig.8D). It supports the reliability of these results that meta-analyses of these fourteen regions had low heterogeneity. The fourteen regions were moderately to strongly correlated with TBV ($r_{with\ TBV}$ = 0.30 to 0.72).

Left precentral

Study	Total	Correlation	COR	95%-CI	Weight (common)	Weight (random)
g-factor (UKB)	39947		0.05	[0.04; 0.06]	69.1%	69.1%
rg IQ	605		0.09	[0.01; 0.16]	1.1%	1.1%
rg education	870		0.09	[0.02; 0.15]	1.5%	1.5%
cryst. IQ (ABCD)	8252		0.05	[0.03; 0.07]	14.3%	14.3%
fluid IQ (ABCD)	8147		0.05	[0.03; 0.07]	14.1%	14.1%
Common effect model	57821		0.05	[0.04; 0.06]	100.0%	--
Random effects model			0.05	[0.04; 0.06]	--	100.0%

-0.15 -0.1 -0.05 0 0.05 0.1 0.15

Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0$, $p = 0.69$

Left parahippocampal

Study	Total	Correlation	COR	95%-CI	Weight (common)	Weight (random)
rg education	1445		0.04	[-0.01; 0.09]	2.5%	10.0%
g-factor (UKB)	39947		0.04	[0.03; 0.05]	67.9%	33.1%
cryst. IQ (ABCD)	8252		0.08	[0.06; 0.10]	14.2%	24.9%
fluid IQ (ABCD)	8147		0.06	[0.03; 0.08]	13.9%	24.8%
rg IQ	930		0.03	[-0.03; 0.10]	1.6%	7.1%
Common effect model	58721		0.05	[0.04; 0.05]	100.0%	--
Random effects model			0.05	[0.03; 0.07]	--	100.0%

-0.05 0 0.05

Heterogeneity: $I^2 = 69\%$, $\tau^2 = 0.0003$, $p = 0.01$

Left middle temporal

Study	Total	Correlation	COR	95%-CI	Weight (common)	Weight (random)
cryst. IQ (ABCD)	8252		0.04	[0.01; 0.06]	14.3%	14.7%
rg IQ	685		-0.01	[-0.09; 0.06]	1.2%	1.2%
fluid IQ (ABCD)	8147		0.04	[0.01; 0.06]	14.1%	14.5%
g-factor (UKB)	39947		0.03	[0.02; 0.04]	68.9%	68.0%
rg education	900		0.05	[-0.02; 0.11]	1.6%	1.6%
Common effect model	57931		0.03	[0.02; 0.04]	100.0%	--
Random effects model			0.03	[0.02; 0.04]	--	100.0%

-0.1 -0.05 0 0.05 0.1

Heterogeneity: $I^2 = 0\%$, $\tau^2 < 0.0001$, $p = 0.72$

Left lateral orbitofrontal

Study	Total	Correlation	COR	95%-CI	Weight (common)	Weight (random)
g-factor (UKB)	39947		0.02	[0.01; 0.03]	68.5%	68.5%
rg education	1215		0.01	[-0.05; 0.06]	2.1%	2.1%
fluid IQ (ABCD)	8147		0.03	[0.01; 0.05]	14.0%	14.0%
cryst. IQ (ABCD)	8252		0.02	[-0.00; 0.04]	14.1%	14.1%
rg IQ	775		0.04	[-0.03; 0.12]	1.3%	1.3%
Common effect model	58336		0.02	[0.01; 0.03]	100.0%	--
Random effects model			0.02	[0.01; 0.03]	--	100.0%

-0.1 -0.05 0 0.05 0.1

Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0$, $p = 0.71$

Fig.7. Forest plots for four regions in the left hemisphere that were significantly correlated with cognitive ability after TBV adjustment in their region-specific meta-analysis.

These four regions had positive correlations and replicated across hemispheres. Note that the sample sizes indicated for genetic correlations (rg) are not the sample sizes from the GWAS summary statistics. To ensure appropriate weighting of rg in the meta-analysis, I computed a representative sample size for the genetic correlations that I used as input to the *metacor* function and would produce the same confidence intervals as obtained in LDSC (see Methods).

75

Regional brain volume	*Meta-analysed effect size [99% CI]*	*Q (p-value)*	*I^2*	*Regions' correlation with TBV*
Left parahippocampal	0.05 [0.03 to 0.08]	12.79 (0.0124)	69%	0.30
Left precentral	0.05 [0.04 to 0.06]	2.26 (0.687)	0%	0.65
Right middle temporal	0.04 [0.01 to 0.06]	10.98 (0.0268)	64%	0.71
Right parahippocampal	0.04 [0.01 to 0.08]	19.71 (0.000571)	80%	0.32
Right precentral	0.04 [0.03 to 0.05]	3.01 (0.556)	0%	0.64
Left fusiform	0.03 [0.01 to 0.04]	4.69 (0.32)	15%	0.65
Left middle temporal	0.03 [0.02 to 0.04]	2.07 (0.723)	0%	0.67
Left caudal anterior cingulate	0.02 [0.01 to 0.03]	1.19 (0.88)	0%	0.26
Left caudal middle frontal	0.02 [0.00 to 0.03]	3.83 (0.429)	0%	0.58
Left lateral orbitofrontal	0.02 [0.01 to 0.03]	2.13 (0.712)	0%	0.72
Right lateral orbitofrontal	0.01 [0 to 0.02]	2.00 (0.737)	0%	0.70
Right cuneus	-0.02 [-0.04 to -0.01]	4.59 (0.332)	13%	0.47
Right isthmus cingulate	-0.03 [-0.05 to -0.02]	1.23 (0.873)	0%	0.55
Left isthmus cingulate	-0.04 [-0.05 to -0.03]	9.11 (0.0585)	56%	0.59

The homogeneity statistic *Q* indicates – if statistically non-significant (*p*>0.05) – that different analyses produced correlations that likely share a common population effect size. I^2 indicates the proportion of variability of measures included in the meta-analysis that can be explained by differences between different analyses.

A

Total brain volume (TBV)

Regional brain volume → Cognitive ability

B

Meta-analysed associations

Left_precentral
Right_middle_temporal
Right_precentral
Left_superior_temporal
Left_inferior_temporal
Left_middle_temporal
Left_insula
Left_pars_orbitalis
Left_superior_temporal
Left_superior_frontal
Left_lateral_orbitofrontal
Right_lateral_orbitofrontal
Left_rostral_middle_frontal
Right_fusiform
Right_inferior_temporal
Right_superior_frontal
Right_insula
Left_postcentral
Right_inferior_parietal
Right_precuneus
Right_rostral_middle_frontal
Left_supramarginal
Right_postcentral
Left_precuneus
Right_medial_orbitofrontal
Right_pars_orbitalis
Left_medial_orbitofrontal
Left_lateral_occipital
Left_inferior_parietal
Left_caudal_middle_frontal
Left_paracentral
Right_rostral_anterior_cingulate
Left_superior_parietal
Right_superior_parietal
Left_rostral_anterior_cingulate
Left_bankssts
Right_superior_parietal
Right_posterior_cingulate
Right_caudal_middle_frontal
Right_bankssts
Left_posterior_cingulate
Left_pars_opercularis
Left_entorhinal
Right_supramarginal
Right_paracentral
Right_pars_triangularis
Right_pars_opercularis
Left_pars_triangularis
Right_caudal_anterior_cingulate
Right_parahippocampal
Left_transverse_temporal
Right_transverse_temporal
Left_isthmus_cingulate
Left_parahippocampal
Right_entorhinal
Left_frontal_pole
Left_lingual
Right_isthmus_cingulate
Right_lingual
Left_caudal_anterior_cingulate
Right_cuneus
Left_cuneus
Right_frontal_pole
Left_pericalcarine
Right_pericalcarine

Correlation (99% CI)

● no TBV adjustment
● TBV adjusted

C

no TBV adjustment | TBV adjusted

Meta-analysed effect size
0.20
0.15
0.10
0.05
0.00

no TBV adjustment | TBV adjusted

Heterogeneity (Q)
150
100
50

D

no TBV adjustment

Most heterogeneous
Left pars_triangularis: Q = 187.5 (p-value = 1.82e-39); I2 = 98%

Least heterogeneous
Left pericalcarine: Q = 11.8 (p-value = 0.0192); I2 = 66%

Effect sizes for 66 cortical ROIs

Phenotypic assoc. g-factor (UKB)
Phenotypic assoc. cryst. IQ (ABCD)
Phenotypic assoc. fluid IQ (ABCD)
Genetic corr. IQ
Genetic corr. education
Polygenic pred. cryst. IQ
Polygenic pred. fluid IQ

TBV adjusted

Most heterogeneous
Left pars_triangularis: Q = 114.6 (p-value = 7.54e-24); I2 = 97%

Least heterogeneous
Left supramarginal: Q = 0.4 (p-value = 0.985); I2 = 0%

Effect sizes for 66 cortical ROIs

Phenotypic assoc. g-factor (UKB)
Phenotypic assoc. cryst. IQ (ABCD)
Phenotypic assoc. fluid IQ (ABCD)
Genetic corr. IQ
Genetic corr. education
Polygenic pred. cryst. IQ
Polygenic pred. fluid IQ

Fig.8. Results from 66 meta-analyses testing whether 66 regional volumes were significantly associated with cognitive ability across multiple analyses (before and after TBV adjustment).

(**A**) Association of interest between regional volume and cognitive ability that I calculated adjusted and unadjusted for brain size as a covariate. (**B**) Sixty-six meta-analysed region-by-region correlations with cognitive ability (99% confidence intervals for multiple testing correction). These correlations were obtained from 66 region-specific meta-analyses. Correlations between regional volumes and cognitive ability unadjusted for TBV are in orange, and correlations adjusted for TBV are in grey. The significant correlations after TBV adjustment are in Table 2. (**C**) Meta-analytic correlations and heterogeneity (Q) results plotted onto regions' respective cortical location. Meta-analysed effect sizes are the same as in panel B. Heterogeneity is represented by the Q statistic that ranged between 11.8 and 187.5 for TBV-unadjusted correlations, and between 0.4 and 114.6 for TBV-adjusted correlations. (**D**) Correlation obtained from different analyses that were the basis for the meta-analyses (apart from the polygenic score predictions which are plotted here but were not included in the meta-analysis due to their small magnitudes). Highlighted and named regions are those whose correlations with cognitive ability differed most and least (according to Q statistic) between the different analyses.

4. Discussion

This is the first study to systematically examine the impact of brain size adjustment on the statistical relationship between regional grey matter volume and cognitive ability in two large samples (UKB & ABCD, N_{total} ~50,000). Analysis 1 demonstrated that brain size adjustment substantially reduces correlations between regional volumes and cognitive ability from a mean of $r = 0.15$ to a mean of $r = 0.00$ (Section 3.1.1). This means that a regions' correlation with cognitive ability largely reflects variance accounted for by brain size as opposed to variance unique to each region (Section 3.1.2). Moreover, brain size adjustment can even result in over-adjustment whereby it eliminates most or all variance shared between a regional volume and cognitive ability. Section 3.1.3 demonstrated that brain size adjustment can easily invert their relationship to a negative estimate at VIFs <2.5, which would falsely support the opposite interpretation – to the one made without brain size adjustment – that larger brain volume is harmful for cognitive ability. This illustrates that brain size is central to the relationship between regional volumes and cognitive ability and that discarding variance accounted for by brain size prevents the delivery of robust and straightforwardly interpretable findings. Future studies should carefully consider whether brain size truly indexes nuisance variance in their specific research question and whether its inclusion as a covariate may result in over-adjustment.

Another study also found that a measure of brain size captures regional brain information (Reardon et al., 2018) further supporting that adjusting for brain size will discard valuable regional information. By approximating brain size with total surface area instead of TBV, Reardon et al. (2018) demonstrated that larger brains more likely

preferentially expand in higher-order areas [13] assumed to be involved in higher cognitive ability. This suggests that larger brains are not just larger (i.e., more cellular material), but inherently differ in their organisation in that higher-order regions are more pronounced and selectively enhance cognitive ability.

In Analysis 2, I showed that four regional volumes were consistently correlated with cognitive ability across multiple analyses and the hemispheres. Contrary to my hypothesis, this finding suggests that the *parahippocampal, precentral*, *middle temporal*, and *lateral orbitofrontal* gyrus contribute significantly non-zero amounts of variance relevant to cognitive ability independent of brain size. This region-specific variance is robust because it replicated across multiple definitions of cognitive ability, different large-scale samples, as well as phenotypic and genetic analysis approaches (further regional discussion in Section 4.1).

However, these brain size adjusted correlations were small (r = 0.01-0.05, i.e., R^2 = 0.01-0.25%) which reduces statistical power and puts into question the validity of uni-regional brain models of cognitive ability. More work is needed to determine whether these small population-level effect sizes are practically relevant and to what extent they can be generalised (Anvari et al., 2022). To achieve larger and more reliable correlations, findings from Analysis 1 & 2 motivate avoiding brain size adjustment and to explicitly model brain size instead of discarding it as a covariate. For example, it would capture more biologically plausible brain structure underlying cognitive ability to use multivariate techniques that account for interdependencies between multiple

---

[13] By higher-order areas I mean association areas that integrate sensory information as opposed to areas that are involved in more elementary processing of sensory information (e.g., primary visual cortex). Higher-order areas are assumed to be involved in complex brain functions (e.g., Pandya & Yeterian, 1985).

brain volumes unadjusted for brain size by modelling brain networks (as done in [Chapter 4](#)).

4.1. Meta-analyses reveal reliable associations between cognitive ability and four brain regions after brain size adjustment

Four regional volumes contributed unique variance relevant to cognitive ability independent of brain size: the *precentral*, *parahippocampal*, *lateral orbitofrontal*, and *middle temporal* gyrus. Previous studies have implicated these four regions in different brain functions possibly related to cognitive ability and examples of such studies are presented in [Table 3](#). However, they must be interpreted with caution because the heterogeneous literature would have supported prominent roles in cognitive ability for any of the 66 brain regions.

Table 3. Examples of brain studies delivering evidence for neural links between cognitive ability and the four brain regions identified in the region-specific meta-analyses to be associated with cognitive ability.

Brain regions (gyri)	References	Previously reported brain function that could link this region to cognitive ability
Precentral	Purves et al. (2001)	The precentral gyrus marks the primary motor cortex involved in voluntary movement. Randomised exercise interventions improved cognitive ability in over 50-year-olds (Northey et al., 2018). According to this, the precentral gyrus may be correlated with cognitive ability on a population level because people that exercise have larger primary motor cortices which improves cognitive ability.
Parahippocampal	Hayes et al. (2007)	In a functional MRI study, increased activation in the parahippocampal gyrus was associated with memory encoding and retrieval, which may link the parahippocampal gyrus to higher cognitive ability.
Lateral orbitofrontal	Deng et al. (2017)	Functional MRI activation in the lateral orbitofrontal gyrus was associated with inhibitory executive control in 14-year-olds. Executive control is often assessed by cognitive tests and would positively load onto the cognitive ability *g*-factor.
Middle temporal	Acheson and Hagoort (2013)	A transcranial magnetic stimulation study found some evidence that the middle temporal gyrus modulates language and reading abilities; two cognitive abilities that positively load onto cognitive ability.

Out of the 66 regions, the pericalcarine gyrus had the smallest and least significant correlation with cognitive ability (Fig.8B). Nevertheless, the literature would have also supported an important role of the pericalcarine gyrus in cognitive ability. By including brain size as a covariate, a study reported a negative association between cortical thickness in the pericalcarine gyrus and a memory change score – the difference between two cognitive assessments 12 months apart (Jiang et al., 2016). To make sense of this negative association, it was interpreted as evidence that the pericalcarine gyrus *indirectly* influences an individuals' memory performance (Jiang et al., 2016). However, my explorations suggest that this memory change score likely shared limited variance with the pericalcarine gyrus which resulted in a negative association because it was over-adjusted by including brain size as a covariate. This would indicate that brain size adjustment has directly contributed to the apparent inconclusiveness of the literature. Future studies should transparently report interdependencies between all variables included in statistical models to aid in building scientific consensus.

4.2. Adjusting for brain size implies causal chains that are challenging to defend

The consequences of brain size adjustment may be illustrated best when considering a statistical designs' implied causal chains which correlational studies should be able to theoretically defend (Wysocki et al., 2022). Fig.9 shows the causal chains implied by phenotypic and genetic designs considered in this chapter. In phenotypic analyses (Fig.9A-C), it is equally likely for brain size to be a confounder, a mediator, or a collider. Studies adjusting for brain size as a covariate imply a confounder scenario that suggests that brain size causes regional volume to be larger

as well as cognitive ability to be greater (Fig.9A). In support of the confounder scenario, strong evidence exists that larger brain size causes greater cognitive ability (Lee et al., 2019). Evidence from Reardon et al. (2018) may support that brain size causes regional volumes as they report some higher-order brain regions may be more pronounced in large brain. However, the effect direction of brain size on regional volume may not always be an appropriate assumption because larger brain size does not strictly mean that a particular region is larger. The opposite direction would most likely apply to all brain regions because brain size is the sum of all individual regions (dotted lines Fig.9A).
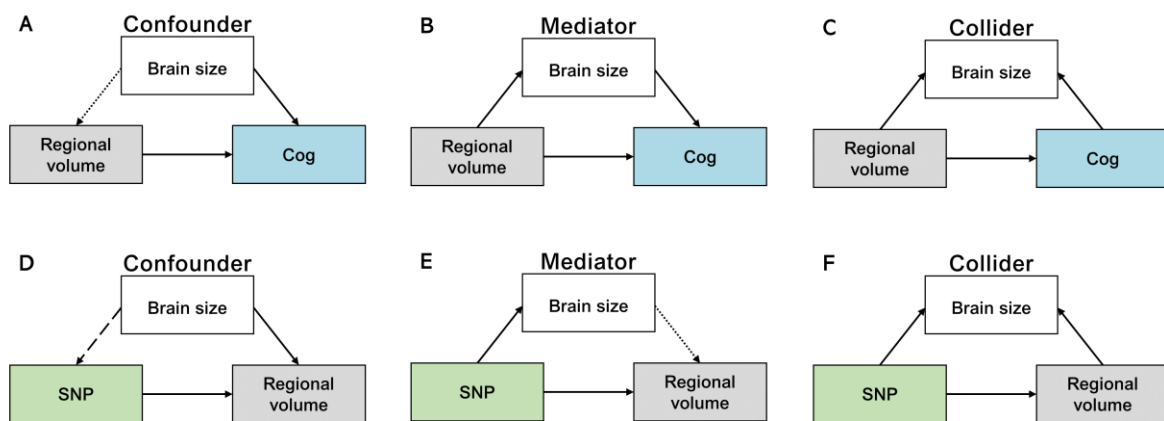
Fig.9. Directed acyclic graph underlying phenotypic (A-C) and genome-wide (D-F) association studies.

Solid lines indicate reasonable directions of effects backed by empirical evidence. Dotted lines indicate less likely directions of effects: a region is more likely to cause brain size to be smaller or larger than vice versa because brain size is the sum of all regions. The dashed line indicates an impossible direction of effects because brain size cannot affect inhered genetic markers which are conceived at birth.

If we were to accept the effect of brain size on regional volume as most likely, we assume that brain size is a *mediator* whereby larger regional volume causes larger total brain size (Fig.9B). Brain size could also be a *collider* whereby higher cognitive ability causes the total brain size to grow (Fig.9C). This would be in line with meta-

analytic reports in healthy older adults of increases in and attenuated losses of brain size through cognitive ability training (Nguyen et al., 2019). In summary, no theoretical argument can defend brain size as a clear confounder in phenotypic analyses (i.e., excluding mediator and collider as reasonable options). This questions the validity of adjusting for brain size and illustrates – should uni-regional brain models be necessary – that mediator and collider scenarios must be modelled and compared.

In GWAS-implied designs (Fig.9D-F), brain size can never be a confounder because we can assume a directional effect of genes on brain size (dashed line Fig.9D). Brain size may be more likely a collider (Fig.9F) than a mediator (Fig.9E) because, as discussed above, larger regional brain volume will always result in larger total brain size because it is the sum of all regions. This logical argument shows that existing GWAS of regional brain morphometry adjusting for brain size (or a proxy of brain size such as a volumetric scaling factor; as in Smith, Douaud et al., 2020) as a confound variable potentially suffer from collider bias (Smith, Douaud, et al., 2020; Zhao, Luo, et al., 2019). Future studies are needed to disentangle the extent and severity of different brain size proxy measures to bias GWAS results (e.g., TBV, intracranial volume, head circumference, or volumetric scaling of T1 head image as used in Smith, Douaud et al., 2020).

4.3. Adjusting brain regional GWAS for brain size (or proxies thereof) may inflate genetic correlations with cognitive ability through collider bias

The following empirical evidence from this chapter may support that GWAS adjusted for a proxy of brain size are likely subject to collider bias. Genetic correlations between cognitive ability measures (IQ and education) and regional volumes derived from GWAS adjusted for a proxy of brain size were distributed around a mean significantly larger than zero (education: r_g = 0.04 [0.03-0.05], IQ: r_g = 0.05 [0.04-0.06]; Fig.4B). These genetic correlations were overall larger than corresponding phenotypic correlations after adjustment. Only *adjusted* genetic correlations (and not *unadjusted* genetic correlations) were larger than their phenotypic counterparts. Therefore, the comparatively larger genetic correlations adjusted for a proxy of brain size cannot be explained by a general tendency for genetic correlations to be slightly larger than phenotypic correlations (as had been previously suggested in Sodini et al., 2018).

As logically argued above (Fig.9F), those comparatively larger genetic correlations adjusted for a proxy of brain size could have been induced by genome-wide collider bias. Aschard et al. (2015) showed that GWAS likely capture false genetic trait correlates when adjusting for a heritable covariate; and brain size and its proxies are highly heritable (broad sense heritability ~80%; Blokland et al., 2012). Collider bias in GWAS can occur when the GWAS trait (the regional volume) and the covariate (brain size) are correlated due to genetic effects: I argue regional volume and brain size must be correlated due to genetic effects as both phenotypes reside in the same organ and regional volume directly contributes to total brain size. Consequently, there

is no certainty whether a GWAS adjusted for brain size indicates true SNP correlates or whether SNP correlates have been falsely induced by conditioning on a collider. Different SNPs may be affected by collider bias differentially. Considering brain morphometry and cognitive ability are highly polygenic (Jansen et al., 2020), even false positive genetic correlates of the brain may (at least partly) map onto genetic correlates of cognitive ability, which could have falsely inflated genetic correlations between IQ, education and brain volumes adjusted for brain size (Fig.4B). An unpublished addition to Chapter 4 delivers more empirical evidence that may support that GWAS adjusted for a proxy of brain size (Smith, Douaud, et al., 2020) are subject to collider bias.

4.5. Limitations

This study has limitations in its brain size adjustment method because brain size was simply modelled as a linear covariate. This was intended to enable straightforward comparisons between analyses. Future studies are needed to examine whether using alternative adjustment techniques (e.g., non-linear covariate adjustment) would alter results presented here. Future studies should also investigate how results would change if models were adjusted for intracranial volume (ICV) instead of TBV. TBV is subject to ageing and pathology, while ICV should remain stable across the lifespan, and is therefore likely to alter the relationship between regional volume and cognitive ability to a lesser degree (especially in older samples but not in younger samples). The exact roles played by ageing and brain development also remain to be discerned.

For simplicity this study used within-sample regression models, but the presented correlations are likely over-fitted and future out-of-sample predictions are

needed to determine *true* effect sizes. Future studies should also explore the transferability of these findings to surface area and cortical thickness measures, as well as non-European populations. Results likely depend on the parcellation scheme used to subdivide the cortex into regions ("brain atlases") which I specifically address in [Chapter 5](#).

4.6. Conclusion

This study examined the impact of brain size covariate adjustment on the relationship between regional grey matter and cognitive ability. Evidence presented here weighs against brain size covariate adjustment because it disposed of most of the variance shared between regional volumes and cognitive ability. Variance associated with brain size is *not* nuisance variance because it accounted for regional information that is lost through brain size adjustment. For most brain regions, this loss of information resulted in small effect sizes that could not be reliably modelled when made independent of brain size. Compared to those brain size adjusted uni-regional brain models, multivariate analyses *unadjusted* for brain size have better potential to uncover larger and more robust brain morphometric correlates of cognitive ability. Without disposing of brain size variance and the regional information that it carries, future multivariate studies promise more holistic, better interpretable, and biologically plausible accounts of brain organisation underlying cognitive ability (e.g., [Chapter 4](#)).

Finally, there is a lack of theoretical arguments to defend causal chains implied by adjusting for brain size as a covariate in phenotypic as well as genetic association studies. Specifically in genetic designs, brain size confounding is logically impossible and there are strong reasons to suggest that brain size adjustment induces collider

bias. Future genome-wide analyses should calculate genetic correlates of regional brain morphometry without adjusting for brain size, which will make GWAS summary data more broadly usable in multivariate follow-up analyses.

Chapter 3

Genomic Principal Component Analysis

1. Background

The study of *genetic overlap* between complex traits promises a quantitative representation of the extent to which two traits have either the same underlying causal loci, or have associated genes with shared pleiotropic action (van Rheenen et al., 2019). GWAS summary statistics are routinely leveraged across the literature to detect genetic overlap between any two traits which has helped uncover that most complex traits share genetic bases. For example, higher cognitive ability, as measured using cognitive tests, is co-inherited with large brain volume to a moderate degree (r_g = 0.24) (Jansen et al., 2020).

Multivariate techniques already exist to investigate genetic correlations across multiple traits more systematically, and they were widely adopted given the large public availability of GWAS summary statistics. Methods such as GenomicSEM (Grotzinger et al., 2019) and Genomic Independent Component Analysis (*Genomic ICA*) (Soheili-Nezhad et al., 2021) integrate multiple traits to model their genetic correlations, test hypotheses at the level of their underlying genetic architecture, and meaningfully reduce dimensionality on a SNP-wise level. This promises more interpretable, more robust, biologically plausible insights into genetic trait variance than inferences from univariate GWAS. However, GenomicSEM and Genomic ICA were both designed to serve specialised purposes and have specific limitations.

GenomicSEM (Grotzinger et al., 2019) provides a flexible framework to apply user-defined structural equation models (SEMs) to genetic correlations across multiple traits. No access to individual-level data is needed, as GenomicSEM takes advantage of genetic correlation matrices derived from bivariate LDSC of multiple GWAS phenotypes. Model fit indices allow users to test how well a proposed model

accounts for observed genetic correlations, and how this model compares to competing models. For example, excellent model fit was found for a genetic general factor of intelligence that accounted for 58% of the genetic variance across seven cognitive traits, but the genetic data did not support a hierarchical structure including more specific cognitive domains (e.g., crystalised and fluid ability), as is often found in phenotypic cognitive testing data (de la Fuente et al., 2021).

GenomicSEM is often used to integrate SNPs into SEMs (for example, into a one-factor model) to represent broad genetic propensity underpinning the proposed factor structure on a SNP-wise level. Though the quality of the results depends on model fit. Consequently, GenomicSEM is not feasible when users attempt to fit too complicated factor structures which could arise when considering too large numbers of GWAS phenotypes or complex loading structures. If model fit is inadequate, GenomicSEM forces SNPs to be associated with an inaccurate factor, resulting in unreliable SNP-wise correlates.

Genomic ICA (Soheili-Nezhad et al., 2021) is a data-driven technique that reduces the dimensional space of multiple GWAS phenotypes and their associated genetic signal into more interpretable, statistically independent sources by which genomic variants collectively influence complex traits. Genomic ICA applies probabilistic ICA to trait-by-SNP matrices. Independent components of genomic variance isolated through Genomic ICA were validated by showing that they relate to distinct, previously-known molecular mechanisms through which the genome acts on the brain (Soheili-Nezhad et al., 2021). With a focus on biological specificity of genetic signal underlying GWAS phenotypes, Genomic ICA can integrate thousands of GWAS phenotypes, though its unsupervised nature means that derived components are less clearly labelled (than factors in GenomicSEM, for example), and that many

components are needed to comprehensively account for genetic variance across the considered GWAS phenotypes (e.g., 200 components to explain 80% of the genetic variance across 1,448 phenotypes in Soheili-Nezhad et al. (2021)).

Here, I present *Genomic PCA* to integrate multiple traits and genome-wide information and to overcome some of the limitations of GenomicSEM and Genomic ICA. Genomic PCA extracts genetic principal components (PCs) underlying multiple GWAS phenotypes to reduce trait dimensionality to one set of genome-wide SNP effects. It decomposes genetic correlation matrices derived from LDSC and has fewer statistical assumptions than GenomicSEM: Genomic PCA permits modelling large numbers of GWAS phenotypes by extracting the first dimension of maximal variation without assuming that there is only one dimension, which is what a one-factor model in GenomicSEM would require. This focus on maximal variation also means it typically accounts for much of the genetic variance with few components, rather than isolating many biologically specific components of genetic variation like Genomic ICA. Genomic PCA enables targeted hypothesis testing of associations between genetic PCs underlying user-selected groups of traits and other traits in follow-up analyses. It is the purpose of this chapter to derive and validate Genomic PCA by demonstrating that a genetic PC extracted with Genomic PCA produces equivalent results to calculating a GWAS on a phenotypic PC.

2. Methods

2.1. Input GWAS summary data

I downloaded freely available GWAS summary statistics from Linnér et al. (2019) on risky behaviours (https://www.thessgac.org/data). Linnér et al. (2019) performed GWAS on four self-reported behaviours: automobile speeding ($n = 404{,}291$), alcohol consumption (i.e., average alcoholic drinks per week, "drinking"; $n = 414{,}343$), lifetime number of sexual partners ($n = 370{,}711$), and smoking (i.e., ever been smoker; $n = 518{,}633$). In addition to the four traits, they extracted the first phenotypic PC (PC1) underlying these four traits in a PCA of phenotypic data and performed a GWAS on PC1 to obtain SNP-wise correlates of a general dimension underlying all four behaviours ($n = 315{,}894$). Speeding, drinking, number of sexual partners, and the PC1 GWAS were performed in the UK Biobank sample (Bycroft et al., 2018). The smoking GWAS was meta-analysed across UK Biobank and the MTAG consortium (Furberg et al., 2010). I validated Genomic PCA by modelling a genetic PC1 underlying the four risky behaviours from GWAS summary statistics, which I compared to the PC1 GWAS extracted from phenotypic data in Linnér et al. (2019).

2.2. Genomic PCA

Below, I describe the step-by-step procedure required to perform Genomic PCA using publicly available tools in *R* (Fig.1). Analysis code can be found at https://annafurtjes.github.io/genomicPCA/.

A. Input data:
Publicly available GWAS summary statistics

B. Genetic correlation matrix:
inferred through LDSC in GenomicSEM (Grotzinger et al. 2019)

C. Principal Component Analysis:
extract principal component (PC) loadings on the first PC of genetic correlation matrix with eigen() R function

First Principal Component

0.7   0.8   0.5   0.6

Trait 1   Trait 2   Trait 3   Trait 4

D. Summary statistics representing underlying genetic PC:
Genome-wide SNP effects on the underlying PC calculated as average of all SNP effects contributed by multiple volumes, weighted by respective PC loadings

$$Z_k \frac{\sum_{i=1}^{x}(w_{ik} * Z_{ik})}{\sqrt{\sum_{i=1}^{x}(w_{ik} * V_i) + \sum_{i=1}^{x}\sum_{j=1}^{x}(\sqrt{w_{ik} * w_{jk}} * C_{i,j})(j \neq i)}}$$

$w_{ik} = \sqrt{N}$*PC loadings for trait i and SNP k, $Z_{i,k}$= test statistic of SNP k on trait I. V_i=univariate LDSC intercept for trait I, $C_{i,j}$= bivariate LDSC intercept between trait i and j. x = number of traits PCs are extracted from

Fig.1. Genomic PCA pipeline.

(A) Genomic PCA takes genome-wide association data from multiple, user-selected traits as input. (B) A genetic correlation matrix is calculated quantifying the extent to which the input GWAS traits genetically overlap, for which it uses LDSC as implemented in GenomicSEM. This also provides LDSC intercepts which are needed in step 4 to adjust for unknown degrees of sample overlap. (C) Through eigen decomposition (*eigen* function in R), Genomic PCA decomposes the genetic correlation matrix into eigenvectors and eigenvalues, from which PC loadings are calculated. (D) Finally, genome-wide SNP effects are obtained through the displayed formula to obtain one set of GWAS summary statistics of a PC underlying the input GWAS phenotypes.

2.2.1 Obtain Genetic Correlation Matrix

Genomic PCA requires the calculation of a genetic correlation matrix which is inferred through LDSC as implemented in GenomicSEM (Grotzinger et al., 2019). First, munging (using the *munge* function) filters GWAS SNPs according to the HapMap 3 reference excluding the MHC region that has complicated linkage disequilibrium structures (URL). Second, LDSC is performed (using the *ldsc* function) to obtain a symmetrical genetic correlation matrix of x dimensions, where x is the number of considered traits. Genetic correlations quantify the extent to which GWAS traits share genome-wide polygenic signal.

2.2.2 Eigen decomposition

Using the *eigen* function in R, Genomic PCA linearly decomposes the genetic correlation matrix into eigenvectors (its principal components) which are sorted by their corresponding eigenvalues (explained variance). Dimensionality reduction is achieved by choosing a principal component and discarding the other components explaining less variance. Here I model PC1 but users can also choose to model another component, e.g., PC2. PC loadings for each contributing trait are calculated as displayed in *Equation 1*:

$$Loadings = Eigenvectors \; x \; \sqrt{Eigenvalues}$$

2.2.3 SNP effects of the underlying genetic PC1

To obtain genome-wide SNP effects of an underlying genetic PC1, Genomic PCA calculates SNP-wise effects according to the equation displayed in Fig.1D. Briefly explained, each SNP effect is the average of all SNP effects from contributing GWAS

traits weighted by a traits PC1 loading as calculated in *Equation 1*. Users may weight SNP effects by loadings onto the second PC, should they wish to obtain SNP effects onto PC2 (or any other principal component). Genome-wide PCs can be extracted from multiple GWAS phenotypes of unknown degrees of overlap as SNP effects are adjusted for univariate and bivariate LDSC intercepts.

Computationally, Genomic PCA relies on the multivariate GWAMA R package (Baselmans et al., 2019). It takes genome-wide association data as input and performs genome-wide meta-analysis across multiple traits while adjusting for sample overlap using LDSC intercepts. To perform Genomic PCA (instead of genome-wide meta-analysis), I modified the *multivariate_GWAMA* function in order to weight contributing SNP effects by their PC1 loading rather than their heritability (https://github.com/baselmans/multivariate_GWAMA/). As a result, the modified R function produces one set of GWAS summary statistics of the underlying genetic PC. It is available on GitHub (https://annafurtjes.github.io/genomicPCA/).

2.2.4. Parallel analysis

To test whether genetic PCs explained more variance than expected by chance, my collaborator Dr. Javier de la Fuente (University of Texas at Austin, US) developed a version of parallel analysis. Parallel analysis aims to determine how many components explain more variance than would have been expected by chance, by comparing variance explained in the actual data with variance explained in random null data. Null distributions of eigenvalues are generated by simulating null correlation matrices sampled from a diagonal population correlation matrix, where the multivariate sampling distribution is specified to take the form of the sampling distribution of the standardised empirical genetic correlation matrix (the V_{STD} matrix,

as estimated using GenomicSEM). This sampling correlation matrix serves as an index of the precision of the elements in the empirical genetic correlation matrix (i.e., heritabilities and genetic overlap across traits) and the sampling dependencies among these when generating the random null models. I specified 1,000 replications to simulate the null correlation matrices and used a 95% threshold for distinguishing true eigenvalues from noise. As the parallel analysis R function was contributed by a collaborator, and I simply applied and described it here.

2.3. Validation of Genomic PCA

To validate that Genomic PCA produces valid genetic PCs, I calculated genetic correlations as a measure of similarity between a genetic PC1 underlying four risky behaviours obtained from Genomic PCA ("genetically extracted") and a GWAS of a phenotypic PC1 of the same four risky behaviours ("phenotypically extracted") calculated by Linnér et al. (2019). A high correlation would indicate that both approaches capture the same polygenic signal while Genomic PCA does not require access to individual-level data. I also report genetic correlations with both phenotypically and genetically extracted PC1 GWAS and the individual risky behaviours.

3. Results

Genetic correlation matrices and LDSC intercepts for the four risky behaviours used as input to Genomic PCA are displayed in Fig.2A-B. Using Genomic PCA, I extracted PC1 underlying four risky behaviours GWAS, which explained 50% of the genetic variance contained across GWAS phenotypes (Fig.2C). Parallel analysis demonstrated in Fig.3 that only PC1 (and not any other components) explained more variance than expected by chance given the empirical sampling dependencies among the four risky behaviour GWAS phenotypes.

The genetically extracted PC1 (using Genomic PCA) was genetically correlated at r_g = 0.99 with the phenotypically extracted PC1 GWAS from Linnér et al. (2019). This suggests that the phenotypically and genetically extracted PC1s capture equivalent polygenic signal, that systematically co-varies with LD (LDSC slope). They were also genetically correlated with their underlying risky behaviours at comparable magnitudes (Fig.4A). For example, the phenotypically extracted PC1 had about the same genetic correlation with self-reported number of sexual partners (r_g = 0.80) as the genetically extracted PC1 (r_g = 0.79). Though, phenotypically and genetically extracted PC1 GWAS somewhat differed in the confounding signal they captured (LDSC intercept = 0.89; Fig.4B). Presented LDSC intercepts capture population stratification, environmental confounding, and sample overlap.
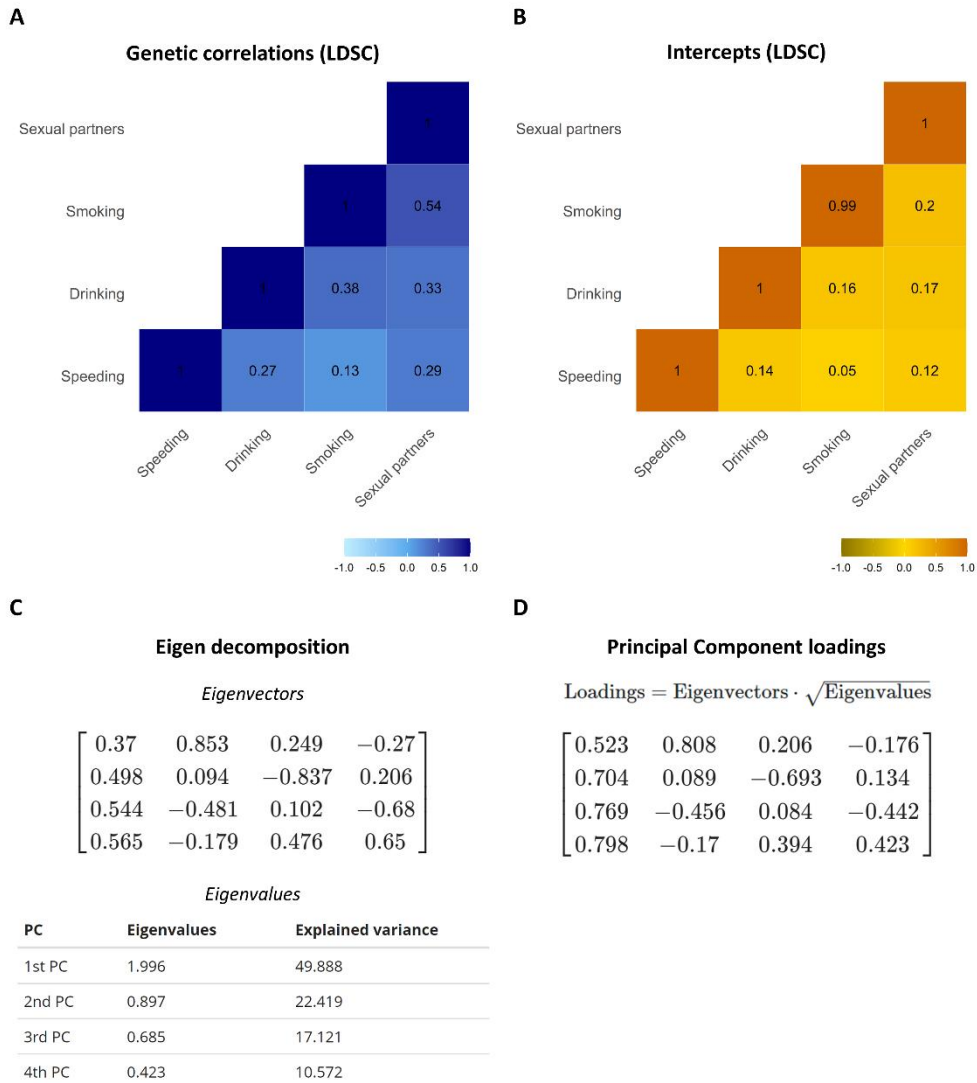
Fig.2. Descriptive statistics.

A. Genetic correlations obtained through cross-trait LDSC between the four risky behaviour GWAS phenotypes by (Linnér et al., 2019). **B.** Corresponding LDSC intercepts. **C.** The genetic correlation matrix is decomposed into eigenvectors and eigenvalues. **D.** PC loadings are extracted from eigenvectors and eigenvalues. Each displayed column contains PC loadings for four traits onto an underlying PC, that is, column 1 from the left is PC1, column 2 is PC2 etc.

100
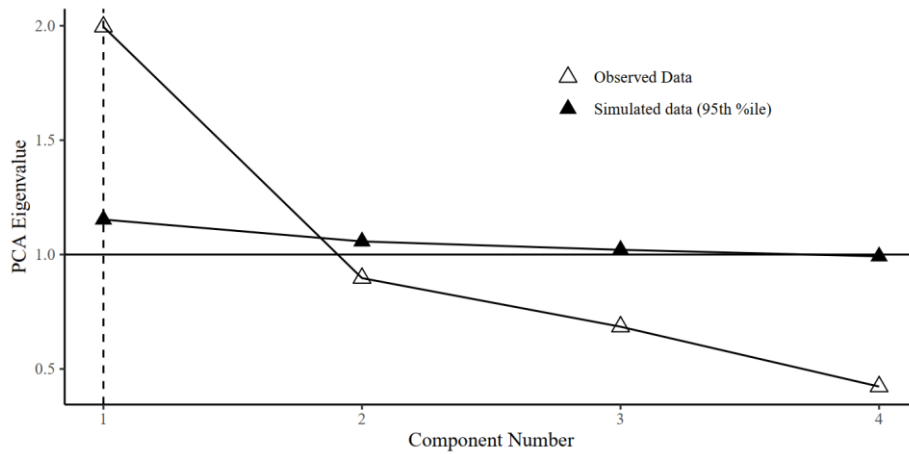
Fig.3. Parallel analysis.

Parallel analysis compares simulated null data against observed data to quantify whether Genomic PCA extracted PCs that explained more variance than expected by chance. With the observed data based on the risky behaviours GWAS (Linnér et al., 2019), PC1 explained more variance than expected by chance, but this was not the case for the other components.
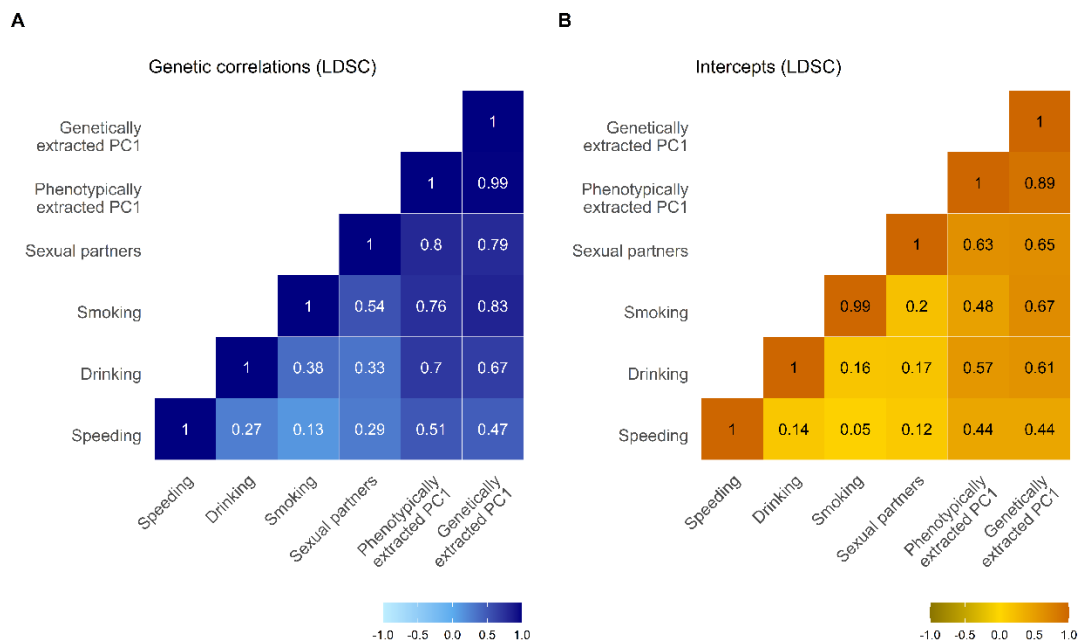
Fig.4. Cross-trait LDSC results.

Genetic correlations and intercepts between genetically and phenotypically extracted PC1 as well as four risky behaviour GWAS phenotypes (Linnér et al., 2019).

4. Discussion

Genomic PCA integrates multiple traits and their genome-wide associations by producing one set of GWAS summary statistics of a PC underlying the input GWAS phenotypes. To validate that Genomic PCA indeed extracts valid genetic PCs, I demonstrated that a genetic PC1 extracted with Genomic PCA captures equivalent polygenic signal (r_g = 0.99) to phenotypically extracting PC1 from individual-level data.

Given the widespread availability of GWAS summary statistics, Genomic PCA is a convenient dimensionality reduction tool because it takes genome-wide association data (as opposed to individual-level data) as input to retain maximal genetic variance shared between multiple traits. I demonstrate above that the genomic PC1 extracted with Genomic PCA explained 50% of the total genetic variance underlying four risky behaviours. Users of Genomic PCA may obtain a robust dimension of genetic signal shared between multiple traits, and perform follow-up analyses to model associations with other traits on the level of their underlying genetic architecture. For example, they may theorise that a group of brain volumes form a brain network, model genetic variance shared between those volumes using Genomic PCA, and test whether their theorised brain network is genetically correlated with other traits such as cognitive ability. Future investigations are needed to test whether individual-level genetic propensity scores ("*polygenic scores*") (Lewis & Vassos, 2022) derived from genetic PCs improve prediction accuracy of phenotypic trait variance compared with scores derived from univariate GWAS.

Genomic PCA should be used as an alternative to other techniques, such as GenomicSEM or Genomic ICA, should researchers wish to extract maximal genome-

wide variation underlying large numbers of GWAS phenotypes. It must be considered that Genomic PCA output is limited by the quality and power of the GWAS summary statistics used as input. If input GWAS contain low precision and limited systematic variance, genetic correlations can become unstable (Lee, McGue, et al., 2018). Furthermore, because *eigen* decomposition is based on a genetic correlation matrix obtained from LDSC, the same limitations that apply to LDSC are relevant to Genomic PCA. For example, LDSC estimates can be downwardly biased with larger standard errors as compared with genome-wide complex trait analysis, for example (Lee, McGue, et al., 2018). Genomic PCA assumes that genetic correlations obtained from LDSC are genome-wide representative.

The next chapter of this thesis uses Genomic PCA to model genetic dimensions of brain morphometry underlying canonical brain networks. Chapter 4 describes canonical brain networks via the variance explained by PC1s (R^2) underlying multiple brain volumes and their corresponding PC1 loadings. As we model genetic PCs purely based on GWAS input – independent of any prior phenotypic information on trait relationships – we can compare independently modelled genetic and phenotypic correlation structures to systematically investigate whether phenotypes are organised similarly to their underlying genetic architecture. According to Cheverud's Conjecture, it is expected for phenotypic and genetic correlations to show strong correspondence (Cheverud, 1988; Sodini et al., 2018). Such a comparison of independent phenotypic and genetic models would not have been possible using PCA applied to GWAS summary statistics as implemented in the 'genome-wise inferred statstics' method (GWIS; Nieuwboer et al., 2016; Shin et al., 2020) because GWIS requires prior information on phenotypic trait relationships to extract genetic PCs.

It would validate previously described phenotypic brain network structure, (Madole et al., 2021) if phenotypic and genetic PCA results were equivalent. It would support that genetics present a promising tool to triangulate phenotypic research findings, since it brings inference closer to underlying biological information. This will help shed light on whether genetics operate along the same dimensions as are evident phenotypically, and it may help better understand whether co-occurrence of two traits more likely originates from genetic or environmental influences. For example, if phenotypic and genetic correlations were of opposite signs (e.g., phenotypic correlation positive, genetic correlation negative), it would imply that there are environmental (or developmental) processes at play that act in opposite directions to the genetically-driven influences, and are in turn likely to exert substantial impacts on the phenotypic manifestation of a trait.

Beyond describing brain-wide correlation structures, the next chapter also assesses shared genetic aetiology between cognitive ability, aging, and multiple brain networks. To the best of my knowledge, this is the first study to systematically investigate theoretically-informed brain networks in interindividual differences of structural grey matter and their significance for cognitive aging on a genetically-informed level of analysis.

Chapter 4

General dimensions of human brain morphometry inferred from genome-wide association data

This project was pre-registered on the Open Science Framework (https://osf.io/7n4qj), it was previously published as a pre-print on bioRxiv (doi: 10.1101/2021.10.22.465437) and has undergone one round of revisions with *Human Brain Mapping*. The version presented below was adopted from the bioRxiv version and is the revised version resubmitted to *Human Brain Mapping* on the 1st of February 2023. As of the 8th of March 2023, we have received confirmation of publication. Supplementary Materials, including supplementary methods, can be found in the Appendix.

General dimensions of human brain morphometry inferred from genome-wide association data

Anna E. Fürtjes, MSc[1*], Ryan Arathimos, PhD[1,2], Jonathan R. I. Coleman, PhD[1,2], James H. Cole, PhD[3,4,5], Simon R. Cox, PhD[6,7], Ian J. Deary, PhD [6,7], Javier de la Fuente, PhD [8,9], James W. Madole, PhD [8], Elliot M. Tucker-Drob, PhD [8,9], Stuart J. Ritchie, PhD [1]

The names of the co-authors are listed in alphabetical order and grouped by affiliation.

[1] Social, Genetic and Developmental Psychiatry (SGDP) Centre, Institute of Psychiatry, Psychology & Neuroscience, King's College London, SE5 8AF, UK

[2] National Institutes for Health Research Maudsley Biomedical Research Centre, South London and Maudsley NHS Trust, London, SE5 8AF, UK

[3] Department of Neuroimaging, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, SE5 8AF, UK

[4] Centre for Medical Image Computing, Department of Computer Science, University College London, London, WC1V 6LJ, UK

[5] Dementia Research Centre, Institute of Neurology, University College London, London, WC1N 3BG, UK

[6] Department of Psychology, The University of Edinburgh, Edinburgh, EH8 9JZ, UK

[7] Lothian Birth Cohorts, University of Edinburgh, Edinburgh, EH8 9JZ, UK

[8] Department of Psychology, University of Texas at Austin, Austin, TX 78712-1043,

USA

[9] Population Research Center and Center on Aging and Population Sciences,

University of Texas at Austin, Austin, TX 78712-1043, USA

Abstract

Background: Understanding the neurodegenerative mechanisms underlying cognitive decline in the general population may facilitate early detection of adverse health outcomes in late life. This study investigates genetic links between brain morphometry, ageing, and cognitive ability.

Methods: We develop Genomic Principal Components Analysis *(Genomic PCA)* to model general dimensions of brain-wide morphometry at the level of their underlying genetic architecture. Genomic PCA is applied to genome-wide association data for 83 brain-wide volumes (36,778 UK Biobank participants), and we extract genomic principal components to capture global dimensions of genetic covariance across brain regions (unlike ancestral principal components that index genetic similarity between participants). Using linkage disequilibrium score regression, we estimate genetic overlap between those general brain dimensions and cognitive ageing.

Results: The first genetic principal components underlying morphometric organisation of 83 brain-wide regions accounted for substantial genetic variance (R^2 = 40%) with the pattern of component loadings corresponding closely to obtained from phenotypic analyses. Genetically more central regions to overall brain structure--specifically frontal and parietal volumes thought to be part of the central executive network--tended to be somewhat more susceptible towards age (r = -0.27). We demonstrate moderate genetic overlap between the first principal component underlying each of several structural brain networks and general cognitive ability (r_g = 0.17-0.21), which was not specific to a particular subset of the canonical networks examined.

Conclusion: We provide a multivariate framework integrating covariance across multiple brain regions and the genome, revealing moderate shared genetic aetiology between brain-wide morphometry and cognitive aging.

Key words: Complex traits genetics, statistical modelling, structural brain networks, structural neuroimaging, genetics, cognitive ability, brain age, statistical modelling

1. Introduction

Progressive ageing-related neurodegenerative processes in human brain are well-documented across the micro- and macro-scales within otherwise healthy adults, and are linked to ageing-related declines in multiple domains of cognitive function (Cox et al., 2016; Fjell & Walhovd, 2010; Madole et al., 2021). Understanding the biological processes underlying these links is paramount for identifying mechanisms of cognitive ageing that can ultimately be targeted by intervention. The human brain is a complex network of partially functionally and anatomically overlapping and interconnected regions (Bressler & Menon, 2010; Power et al., 2011; Sporns, 2011; Yeo et al., 2011), whose components age unevenly over time (Raz et al., 2010), and may be differentially relevant to adult cognitive ageing (Cox et al., 2019; Fjell & Walhovd, 2010; Madole et al., 2021).

Whereas considerable attention has been devoted separately to the genetic architecture of human brain morphometry (Anderson et al., 2021; Meer et al., 2021; Zhao, Luo, et al., 2019) and the genetic architecture of adult cognitive ability (de la Fuente et al., 2021), relatively less work has explicitly linked investigations of the genetic architecture of human brain morphometry to the putative organisation of brain networks (although see (Arnatkevičiūtė et al., 2021) for a recent exception). In addition, there have been few investigations of how genetic links between components of human brain networks relate to ageing and cognition.

To model the underlying genetic architecture of brain organisation, we developed Genomic Principal Component Analysis (*Genomic PCA*), a multivariate approach in which we integrate multiple regional brain volumes and the genome to model general dimensions of brain structure. Using genome-wide association study

(GWAS) summary statistics as input, Genomic PCA extracts genetic principal components (PCs) underlying multiple GWAS phenotypes (unlike the ancestry-based PCs commonly used in genomic research that index genetic similarity between participants). Subsequently, genetic PCs underlying the whole brain, as well as nine groups of regional brain volumes that reflect canonical brain networks (Fig.1) are tested for associations with cognitive ability and aging. This genetically-informed approach parallels a previous study modelling *phenotypic* PCs underlying the same canonical brain networks, which showed that frontal and parietal brain volumes – part of the central executive network – were more important to overall brain structure (i.e., higher loadings onto a PC underlying the whole brain), and tended to have stronger cross-sectional associations with age than other regions of the brain (N = 8,185) (Madole et al., 2021).
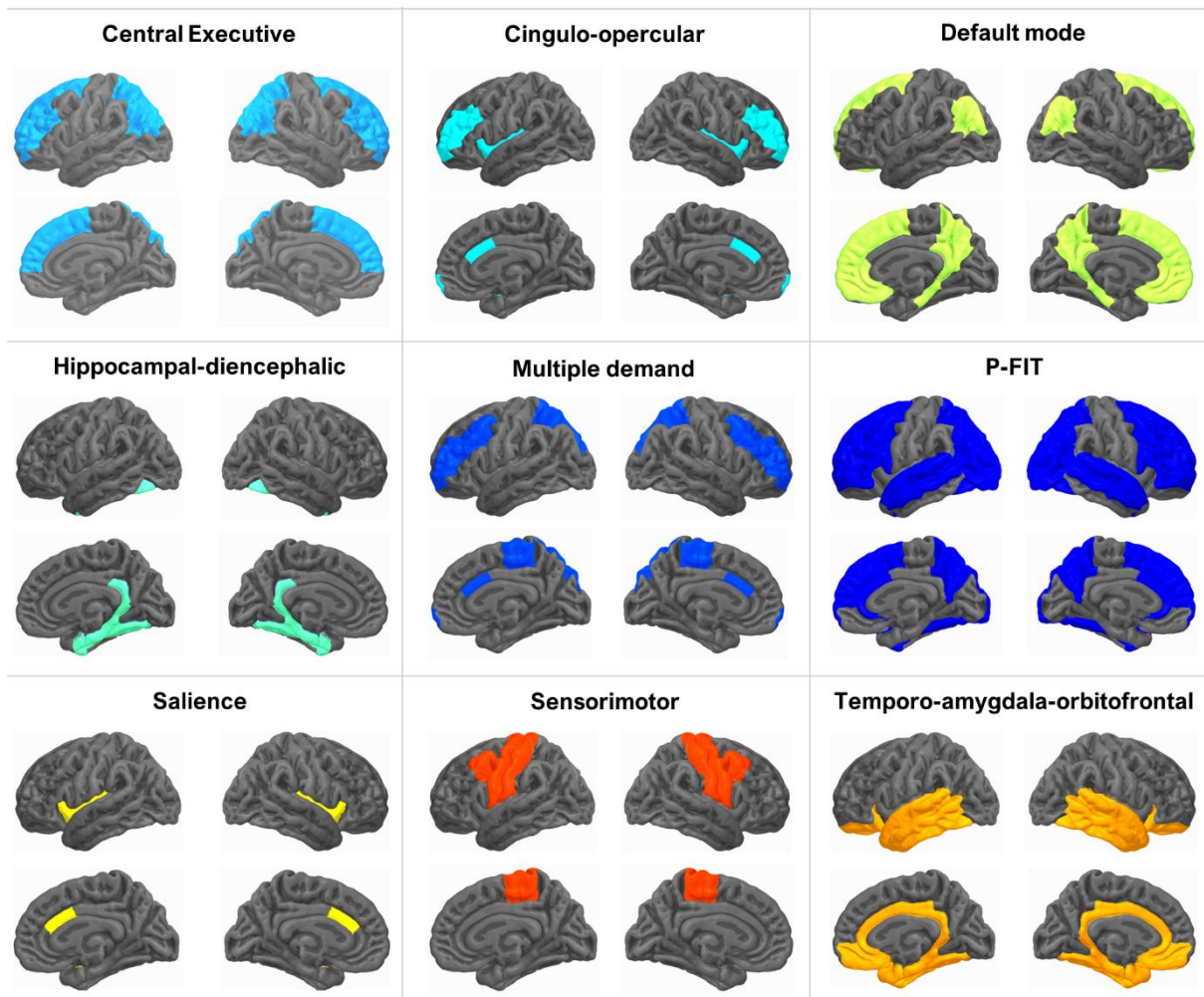
Fig.1. Canonical brain network definitions.

To scaffold the genetic architecture of human brain morphometry onto the canonical network organisation of the brain, we consider nine overlapping brain networks. Regional volumes thought to reside within these networks are represented through genome-wide association data of 83 grey-matter volumes (N = 36,778), and this Figure indicates which networks different volumes were allocated to. The network definitions were adopted from Madole et al. (2021), but are not indisputable. We used these theory-based network definitions to apply our novel dimensionality reduction technique *Genomic PCA*, to obtain genetic Principal Components underlying clearly labelled networks. Using these genetic PC1s, we tested whether different networks, or even the whole brain are genetically associated with cognitive aging.

The canonical brain networks examined in the present manuscript are based on a whole-brain perspective, considering the existing literature that describes synchronised (i.e., correlated) regional activity in functional MRI data (Madole et al., 2021), in addition to converging evidence from other modalities (i.e., structural MRI and lesion-based mapping (Bressler & Menon, 2010; Jung & Haier, 2007; Menon & Uddin, 2010). Among the most reported networks are the central executive (Sridharan et al., 2008), default mode (Buckner & DiNicola, 2019), salience (Downar et al., 2002), and multiple demand networks (Duncan, 2010). Our investigation focuses on brain volumes within these networks because they are highly heritable (Zhao, Ibrahim, et al., 2019), and are measured independently of mental processes during MRI scanning (compared with functional MRI). Grey matter volume is a robust predictor of general cognitive ability (Cox et al., 2019; Hilger et al., 2020), and it partly reflects age-related atrophy among middle-and-older adults; an important indicator of ageing and health outcomes (Cole et al., 2018).

There are substantial genetic links between brain structure and cognitive function in aging. For example, a recent investigation ran a GWAS on the *brain age* gap, which is an index of how much older (or younger) an individual's brain appears compared to their chronological age. Substantial genetic correlations were revealed between a dementia screening test (Mini Mental State Examination) and brain age in the whole brain (r_g = -0.3), as well as the four brain lobes (r_g = -0.15 to -0.22), suggesting that there is a genetic component to how quickly one's brain degrades with age.

Overall brain volume and cognitive ability are also genetically correlated (r_g = 0.24), implicating genes involved in regulating cell growth (Jansen et al., 2020). Biton et al. (2020) reported smaller genetic correlations between intelligence and seven

regional brain volumes (range r_g = 0.07-0.13), which is the only study we are aware of that considered regional volumes *not* normalised for global brain measures. Studies normalising for global measures report only small, or even negative associations between cognitive ability and regional brain structures (e.g., r_g = -0.13 between intelligence and frontal lobe) (see also Grasby et al., 2020; Zhao, Luo, et al., 2019), which is of secondary interest to our study because this only considers regional variance above and beyond variance that maps onto total brain size. Instead, we consider regional variance central to overall brain structure: rather than discarding it (and the regional information it carries; Reardon et al., 2018), we model interregional variance because cognitive ability and aging are brain-wide distributed phenomena (Cole et al., 2019; Hilger et al., 2020), that are more associated with shared between-region (rather than noisy region-specific) brain features (Cox et al., 2021).

The aims of this pre-registered study are twofold (https://osf.io/7n4qj). First, we link investigations of the genetic architecture of human brain morphometry with canonical brain networks, to test whether genetics operate on the same dimensions as are evident phenotypically. As Cheverud originally speculated, "If genetically and environmentally based phenotypic variations are produced by similar disruptions of developmental pathways, genetic and environmental correlations should be similar" (Cheverud, 1988). We therefore hypothesised a close correspondence of phenotypic and genetic morphometric correlations (as demonstrated across a range of traits in Biton et al., 2020; Sodini et al., 2018). A dissimilar organisation of phenotypic and genetic brain architecture would raise questions regarding the neurobiological validity of canonical brain networks in interindividual differences of structural grey matter. A similar organisation would be consistent with a measurable genetic foundation of structural brain networks.

Second, we investigate the extent to which genetic correlations among brain organisation, cognitive ability, and aging corroborate the magnitude and direction of well-established phenotypic associations. We hypothesised substantial genetic correlations of these variables with general morphometric dimensions across the whole brain, and nine overlapping structural brain networks. As implied by the phenotypic results of Madole et al. (2021), we expected the central executive network to play a disproportionate role in cognitive ability, which would confirm a more precise neurobiological foundation of cognitive ability.

2. Methods

The UK Biobank sample consisted of 36,778 unrelated White European participants (54% females) with available neuroimaging data. They had an average age of 63.3 years at neuroimaging visit (range from 40.0 to 81.8 years; SMethods 1.1). Standard quality checks were performed as described in SMethods 1.2-1.3. Then, we derived *Genomic Principal Components Analysis* (G*enomic PCA*; Fig.2) that follows three major steps to extract general dimensions of human brain morphometry underlying genetic covariance across multiple brain GWAS phenotypes (unlike ancestral PCs that index genetic similarity between participants).
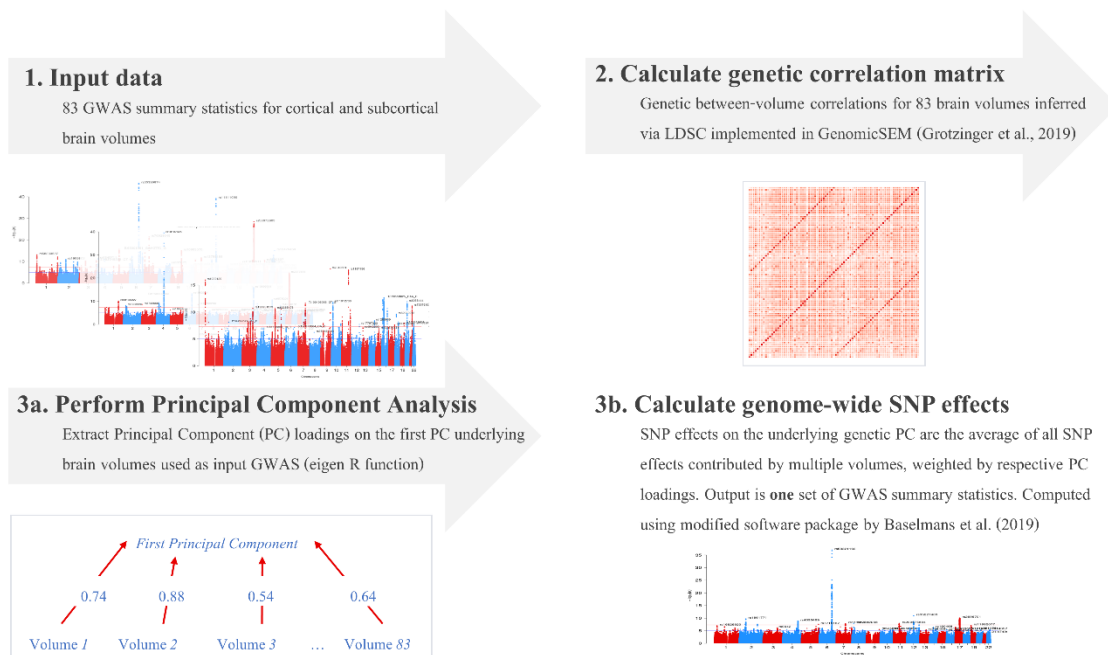
Fig.2. Genomic PCA pipeline.

(1) Input data: The pipeline takes GWAS summary statistics as input. Here, we calculated GWAS summary statistics for 83 cortical and subcortical grey-matter volumes, which were the input to the analyses presented throughout the manuscript. **(2) Calculate genetic correlation matrix:** We calculate interregional genetic correlations based on LDSC as implemented in GenomicSEM (Grotzinger et al., 2019). **(3a) Perform Principal Component Analysis:** We perform eigen decomposition of the genetic correlation matrix using the eigen function in R in order to extract PC1 loadings on the first PC underlying brain volumes for which we submitted GWAS summary statistics to the pipeline. Here we obtain PC1 loadings for each of the 83 brain volumes, and an estimate of R^2 quantifying how much genetic variance PC1 explained across all input volumes. **(3b) Calculate genome-wide SNP effects:** To obtain genome-wide SNP-wise effects on the underlying genetic PC1, we calculate each SNP effect as the average of all SNP effects contributed by the input volumes, weighted by respective volume-specific PC1 loadings. This creates one set of GWAS summary statistics representative of genetic correlates of an underlying genetic PC1. Individual SNP effects are computed with a modified function by Baselmans et al. (2019). We used the same procedure to also obtain PC1 underlying different brain networks, for which we submitted fewer volumes as input.

First, we calculated 83 GWAS summary statistics for 83 cortical and subcortical grey-matter volumes (33 cortical Desikan-Killiany (Desikan et al., 2006) regions in each hemisphere + 8 subcortical regions in each hemisphere + brain stem; Fig.2.1). UKB field IDs are listed in STable 1. GWAS effects were fitted in a linear mixed model using REGENIE (Mbatchou et al., 2021). SNP-heritability for each volume was comparable to those reported elsewhere (Zhao, Ibrahim, et al., 2019) (mean = 0.23, range = 0.07-0.42; Fig.3A).

Second, we calculated genetic correlation matrices indicating genetic overlap between the 83 volumes using linkage disequilibrium score regression (LDSC; Bulik-Sullivan, Finucane, et al., 2015) as implemented in the GenomicSEM software (Grotzinger et al., 2019) (Fig.2.2). Genetic between-volume correlations are displayed in SFig.1-10.

Third, we extracted the first genetic principal component (PC1) underlying genetic variance shared across multiple GWAS phenotypes (here we used 83 brain volumes as input), by which we reduced dimensionality from multiple to only one set of GWAS summary statistics. PC1 loadings and R^2 estimates were calculated with the eigen function in R (Fig.2.3.a). Genome-wide SNP effects are calculated as the average of SNP effects from multiple GWAS phenotypes weighted by (volume-specific) PC1 loadings. Standard errors are corrected for sample overlap by taking into account LDSC intercepts (Fig.2.3.b). In cases of complex and highly dimensional data (e.g., large numbers of variables, or complex loading structure making a factor model in GenomicSEM (Grotzinger et al., 2019) unfeasible), Genomic PCA permits a focus on the first dimension of maximal variation without assuming that there is only one dimension (which is what fitting a one-factor model would require). Genomic PCA is also computationally simpler given the large number of considered ROIs. It is a major

advantage that no access to individual-level phenotype data is needed to perform Genomic PCA, and we validated the approach by demonstrating that GWAS summary statistics produced by Genomic PCA are identical (r_g = .99) to GWAS summary statistics obtained from running GWAS analyses on a phenotypic PC1 (more details at https://annafurtjes.github.io/genomicPCA/, and in SMethods 2.5).

Using Genomic PCA, we performed theory-driven dimensionality reduction by extracting genetic PC1s from covariance structures across nine canonical brain networks (as well as the whole brain with 83 regions). That is, we submitted groups of brain volumes to Genomic PCA that are thought to be part of canonical brain networks (STable 2 lists volumes allocated to nine overlapping networks). Network definitions have been adopted from Madole et al. (2021), where networks were aligned with the structural, functional, and lesion-based literature (e.g., Bressler & Menon, 2010; Jung & Haier, 2007; Menon & Uddin, 2010).

The remainder of the Methods outlines analyses of genetic PC1s underlying multiple brain volumes (derived with Genomic PCA) and is structured according to the four major sub-sections of the Results:

1) First, we reported summary statistics (including volumetric PC1 loadings and variance explained by PC1; R^2) describing the genetic PC1s underlying the whole brain (83 regions), as well as nine canonical brain networks including fewer regions (Results 3.1).

2) Second, we tested whether genetic interregional covariance is similarly organised to phenotypic interregional covariance. To obtain comparable indices of phenotypic covariance, we ran a standard (phenotypic) PCA on a phenotypic correlation matrix obtained from the same brain volume variables used to calculate

GWAS. Phenotypic PCA is performed with the eigen function in R, which is also used in Genomic PCA. We quantified linear associations and the Tucker congruence coefficient (Lorenzo-Seva & Berge, 2006) to contrast genetic and phenotypic interregional correlations, as well as genetic and phenotypic PC1 loadings underlying brain-wide volumes (Results 3.2).

3) Third, to quantify the genetic relationship between general dimensions of brain morphometry and cognitive ability, we extracted a general factor of cognitive ability in GenomicSEM (Grotzinger et al., 2019) using factor analysis of seven cognitive traits as published by de la Fuente et al. (2021). The seven cognitive traits were *Matrix Pattern Completion task* for nonverbal reasoning, *Memory – Pairs Matching Test* for memory, *Reaction Time* for perceptual motor speed, *Symbol Digit Substitution Task* for information processing speed, *Trail Making Test – B* and *Tower Rearranging Task* for executive functioning, and *Verbal Numerical Reasoning Test* for verbal and numeric problem solving, or fluid intelligence. The main Results of this section are genetic correlations between general cognitive ability and genetic PC1s underlying the whole brain and nine different brain networks (Results 3.3). Additionally, we report genetic correlations with individual cognitive abilities, Q_{trait} analyses (Grotzinger et al., 2022), and we test whether the central executive network is particularly relevant for cognitive ability (SMethods 2.11).

4) Fourth, we tested for associations between general dimensions underlying the whole brain and age-related indices to understand whether generally more important regions for overall brain structure are also more susceptible to cognitive aging (which would indicate shared mechanisms). This fourth section is split into two parts: Firstly, we tested for a linear association between the genetic PC1 loadings of all 83 volumes (onto a PC1 underlying the whole brain) and a volume's cross-sectional

association with age ([Results 3.4.1](#)), which has previously been called its "age sensitivity" (Madole et al., 2021). This analysis was not repeated for smaller subnetworks, because the low degree of statistical power did not allow us to meaningfully estimate the correlation between the vectors.

In a second, non-registered analysis, we quantified a genetic correlation between a genetic PC1 underlying the whole brain and the brain age gap (the gap between chronological and biological brain age), for which we utilised GWAS summary statistics by Kaufmann et al. (2019). This brain age gap GWAS was based on the difference between an individual's chronological age and age predictions of how old (or young) an individual's brain appears from structural MRI measures (Results 3.4.2). This analysis was only performed for a genetic PC1 underlying the whole brain, but not PC1s underlying different networks, because they were so strongly associated that they indexed practically the same polygenic signal (as discussed in last paragraph [Section 3.1](#)). More details on Methods are in SMethods. Our analysis code is displayed at https://annafurtjes.github.io/Genetic_networks_project/.

3. Results

3.1 Descriptive statistics of genomic PC1s underlying whole brain and canonical brain networks

Genetic PC1s underlying volumes across the whole brain. In this section, we report variance explained (R^2) by the first underlying volumetric principal component, PC1 and corresponding PC1 loadings obtained from Genomic PCA of the whole brain (83 regions), as well as nine overlapping canonical brain networks. The PC1 underlying the whole brain explained 40% of the genetic variance across 83 regional volumes – larger than the 31% explained by the first phenotypic whole-brain PC1 (Fig.3F). For comparison, the second genetic PC2 accounted for a fraction of the variance that the first PC1 explained (R^2 = 6.7%), indicating that the first genetic PC1 accounted for the majority of systematic variance across structural networks. Genetic PC1 loadings onto the first PC1 underlying the whole brain ranged between 0.30 and 0.81 (mean = 0.62, SD = 0.13, median = 0.65; Fig.3E, STable 3).
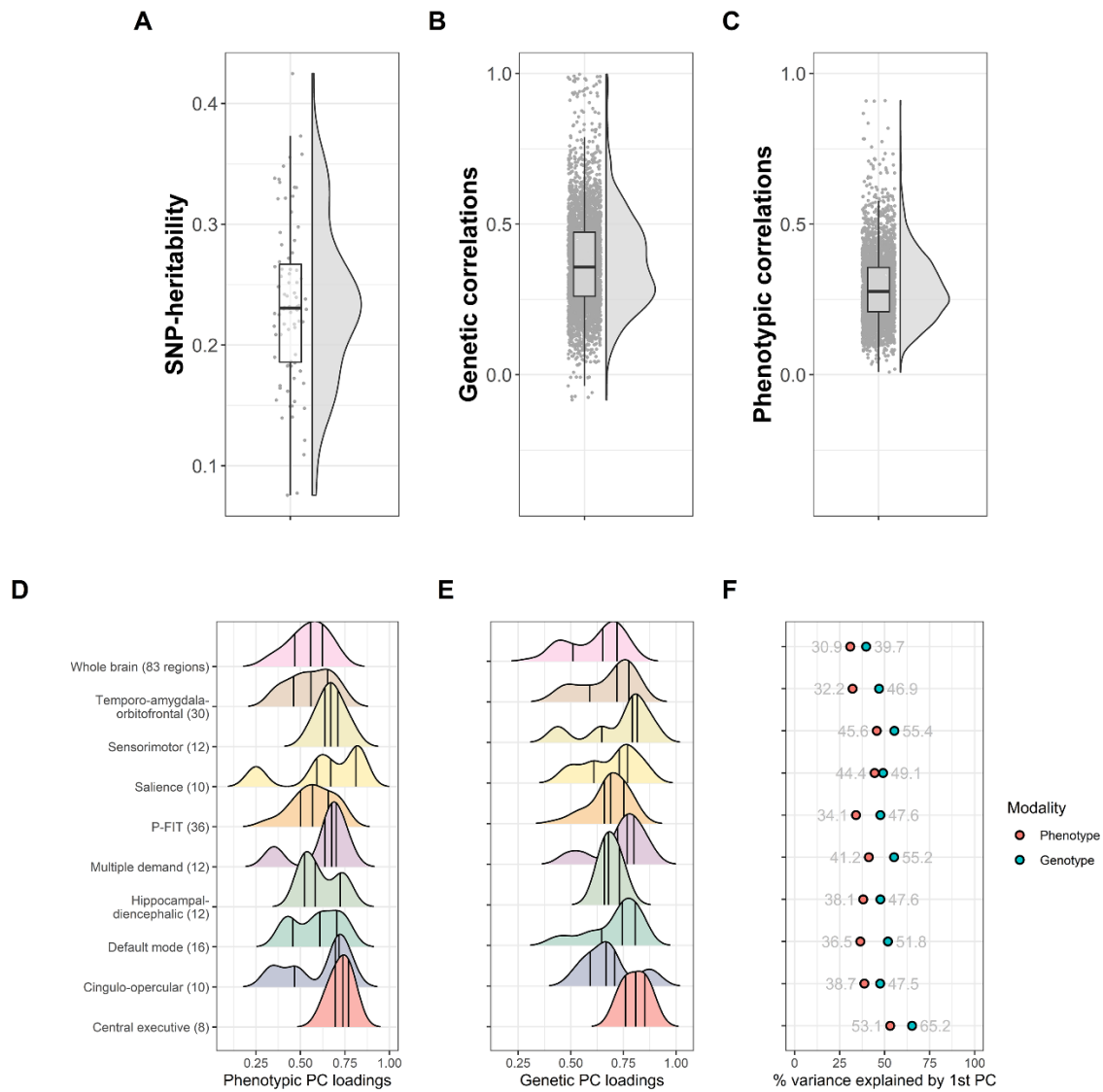
Fig.3. Descriptive statistics.

(A) Distribution of SNP-heritability estimates for 83 regional grey-matter volumes inferred through univariate LDSC. **(B)** Distribution of genetic correlations among 83 regional grey-matter volumes inferred through between-volume LDSC. **(C)** Distribution of phenotypic correlations among 83 regional grey-matter volumes inferred through Pearson's correlations. Raincloud plots were created based on code adapted from Allen et al. (2019). *Bottom row:* Density distributions of PC1 loadings on the first PC underlying volumes in **(D)** phenotypic and **(E)** genetic networks. Vertical lines indicate quantiles. Genetic PC1 loadings are plotted onto corresponding brain regions in SFig.24. **(F)** Variance explained by phenotypic and genetic first PC1 underlying volumes in each network.

Genetic PC1s underlying volumes in canonical networks. The first genetic PC1s underlying different brain networks accounted for greater R^2 than the genetic whole-brain PC1. R^2 ranged from 65% explained by the first genetic PC1 underlying the central executive network, to 47% accounted for by the first genetic PC1 underlying the temporo-amygdala-orbitofrontal network (Fig.3E). R² was larger for networks including fewer volumes, which tended to be more homogeneous, as indicated by PC1 loadings (e.g., range 0.74-0.88 for central executive, range 0.43-0.89 for sensorimotor). Parallel Analysis confirmed that genetic PC1s underlying all brain networks explained substantially more variance than expected by chance (Scree Plots SFig.11-20). Further simulations demonstrated that our theoretical grouping of volumes into networks resulted in more variance explained than expected by randomly grouping volumes (*STable 5*; *SMethods 2.7*).

To compare the polygenic signal captured by different brain networks, we calculated genetic correlations between them using Linkage Disequilibrium Score Regression (LDSC; Bulik-Sullivan, Finucane, et al., 2015). Those genetic correlations tended to be very high (mean r_g between networks 0.83, SD = 0.09; range = 0.63-0.97), suggesting different network PC1s captured roughly the same polygenic signal. For example, the central executive network was genetically associated with the whole brain at r_g = 0.91. That is, we obtained practically the same polygenic signal when extracting a genetic PC1 from the whole brain (83 volumes), as we obtained from extracting a genetic PC1 from fewer volumes (e.g., 8 volumes in the central executive).

3.2 Comparing genetic and phenotypic interregional covariance

To quantify how indices of genetic and phenotypic interregional covariance resemble each other, we calculated linear associations between phenotypic and genetic between-volume correlations, as well as linear associations and Tucker congruence coefficient between phenotypic and genetic PC1 loadings onto an underlying whole-brain PC1. The vectors of 3403 ($\frac{83(83-1)}{2}$) phenotypic and 3403 genetic interregional correlations were strongly positively associated (r = 0.84; b = 0.60; SE = 0.007, $p < 2$ x10^{-16}, R^2 = 70%) (Fig.4A), indicating that volumes that were strongly phenotypically correlated were also strongly genetically correlated. Magnitudes of genetic correlations tended to be slightly larger than phenotypic correlations (intercept = 0.06) which is consistent with previous reports (Biton et al., 2020).
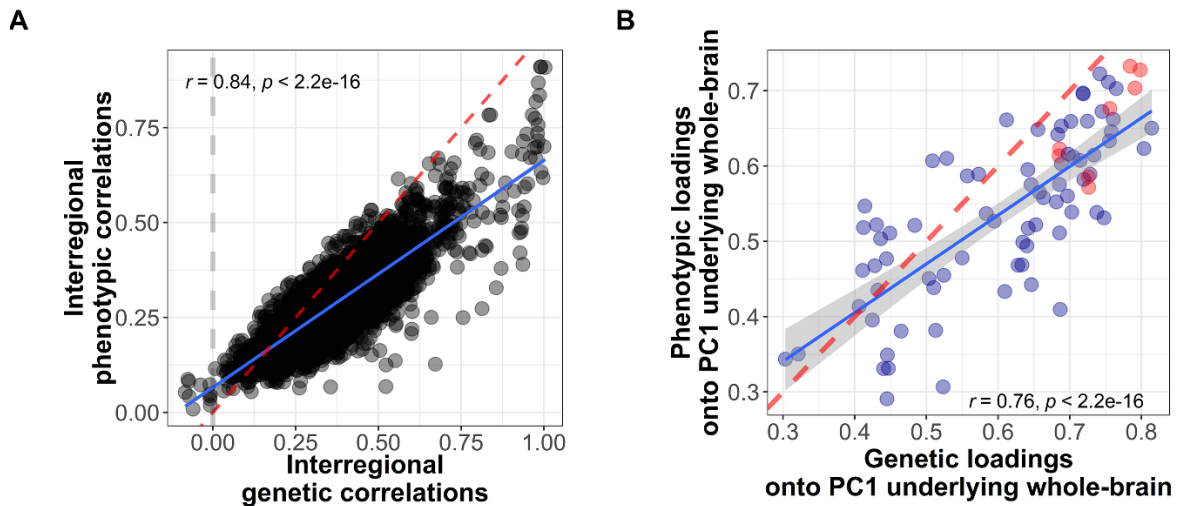
Fig.4. Quantitative comparison of phenotypic and genetic interregional covariance.

Figure **(A)** is contrasting 4303 between-volume correlations where the phenotypic correlations were obtained from phenotypic brain volumes, and the genetic correlations were obtained from LDSC of GWAS summary statistics of the same brain volumes. Figure **(B)** contrasts 83 phenotypic and genetic PC1 loadings onto an underlying whole-brain PC1. Regions coloured in red are regions allocated to the central executive network, which tend to be both phenotypically and genetically central to overall brain structure (i.e., high PC1 loadings).

The association between phenotypic PC1 loadings and genetic PC1 loadings was large and significant (b = 0.65, SE = 0.06, p = 5.07 x10^{-17}, R^2 = 58%, intercept = 0.15). The Tucker congruence coefficient was used to index the degree of similarity between genetic and phenotypic PC1 loadings, taking into account both their relative ordering and their absolute magnitudes (Lorenzo-Seva & Berge, 2006). It revealed very high congruence between phenotypic and genetic PC1 loadings for the 83 volumes (Tucker coefficient = 0.99). These results illustrate a close correspondence and an equivalent organisation of phenotypic and genetic dimensions of shared morphometry; a finding that aligns with Cheverud's Conjecture (see Discussion 4.2).

3.3 Genetic correlations between general cognitive ability and general dimensions of human brain morphometry

To quantify the genetic relationship between general dimensions underlying brain morphometry with cognitive ability, we fitted a general factor of cognitive ability indicated by seven cognitive test GWAS in GenomicSEM (Grotzinger et al., 2019) and calculated its genetic correlation with genetic PC1s underlying brain volumes in different brain networks (Fig.5). The whole-brain and all network-specific genetic PC1s were significantly genetically associated with general cognitive ability. Correlation magnitudes ranged between r_g = 0.17-0.21 (Table 1). According to commonly-used rules of thumb from Hu and Bentler (1998) (CFI > 0.95, RMSEA < 0.08), all models showed good model fit (STable 4).
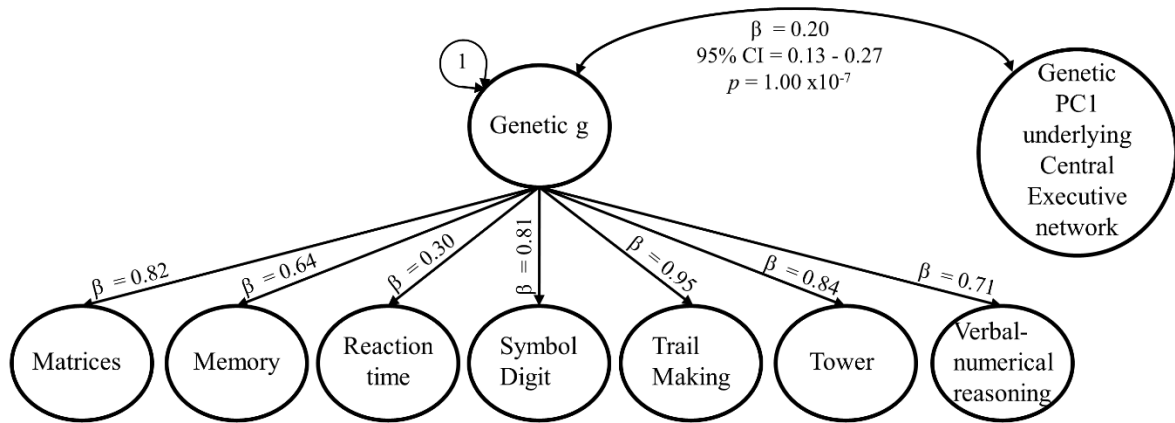
Fig 5. Genomic Structural Equation Model calculating genetic correlations between general cognitive ability and genetic PC1s.

We modelled a general *g* factor of general cognitive ability in GenomicSEM (Grotzinger et al., 2019) using cognitive ability GWAS summary statistics obtained from (de la Fuente et al., 2021). The genetic correlation between genetic *g* and general morphometric dimensions underlying the whole brain and nine canonical brain networks (modelled using Genomic PCA) are reported in Table 1. The seven cognitive traits and the networks are inferred through LDSC. Matrix = Matrix Pattern Completion task; Memory = Memory – Pairs Matching Test; RT = Reaction Time; Symbol Digit = Symbol Digit Substitution Task; Trails-B = Trail Making Test – B; Tower = Tower Rearranging Task; VNR = Verbal Numerical Reasoning Test. Model fit: χ^2 = 124.04, df = 20, *p*-value = 2.1 x10^{-20}, AIC = 174.04, CFI = 0.97, SRMR = 0.079

Table 1. Genetic correlations (r_g) between general cognitive ability and general dimensions of morphometry underlying the whole brain and nine canonical brain networks.

Network	Included volumes	r_g	95% CI	p-value	FDR q-value
Whole brain	83	0.21	0.13-0.29	1.00×10^{-7}	3.00×10^{-7}
Central executive	8	0.20	0.12-0.27	1.00×10^{-7}	3.00×10^{-7}
Cingulo-opercular	10	0.20	0.13-0.27	1.00×10^{-7}	3.00×10^{-7}
Default Mode	16	0.19	0.12-0.26	2.00×10^{-7}	3.00×10^{-7}
Hippocampal-Diencephalic	12	0.17	0.09-0.24	2.66×10^{-5}	2.66×10^{-5}
Multiple Demand	12	0.19	0.12-0.27	7.00×10^{-7}	9.00×10^{-7}
P-FIT	36	0.20	0.12-0.27	2.00×10^{-7}	3.00×10^{-7}
Salience	10	0.19	0.12-0.26	3.00×10^{-7}	4.00×10^{-7}
Sensorimotor	12	0.19	0.11-0.27	1.20×10^{-7}	1.30×10^{-6}
Temporo-amygdala-orbitofrontal	30	0.20	0.12-0.27	2.00×10^{-7}	4.00×10^{-7}

r_g = genetic correlation between genetic PC1s underlying nine canonical brain networks and a factor of general cognitive ability modelled from seven cognitive traits, SE = standard error, 95% CI = 95% confidence interval, *p*-value = original *p*-value as indicated by the GenomicSEM model, false discovery rate (FDR) *q*-value = *p*-value corrected using 5% false discovery rate.

We also report genetic correlations for three individual cognitive traits, because the available GWAS data (de la Fuente et al., 2021) did not warrant modelling separate cognitive domains. Each domain had a maximum of two traits only (e.g., logical reasoning is assessed by both Matrix Pattern Completion and Verbal Numerical Reasoning). Some cognitive tests are impure and contain various cognitive components (e.g., the Trail Making Test assesses executive and speed abilities). To reduce multiple testing burden, we pre-registered (https://osf.io/7n4qj) genetic correlations for three tests that represent relatively separate cognitive abilities: *Matrix Pattern Completion* consistently yielded the strongest genetic correlations with PCs underlying the brain networks (mean r_g across different networks = 0.18). Genetic correlations for *Symbol Digit Substitution Task* were slightly smaller (mean r_g = 0.12), followed by *Memory* which had the lowest average correlations (mean r_g = 0.09).

The significant genetic correlations – between general cognitive ability and genetic PC1s underlying different brain networks – seem to act through a factor of general cognitive ability, rather than through individual cognitive abilities, because individual cognitive traits had high loadings on the genetic cognitive ability factor (median = 0.81, range = 0.30-0.95; SFig.22). Also, Q_{trait} heterogeneity analyses (Grotzinger et al., 2022) demonstrated that the general cognitive ability factor accounted well for the patterns of association between specific cognitive abilities and brain network genetic PC1s (SFig.23). That is, models allowing independent associations for all individual cognitive traits did not yield better model fit than models forcing any association to go through the general cognitive ability factor ($\Delta \chi^2 \sim 0$; *df* = 6; STable 5).

Based on previous phenotypic findings that highlighted the importance of the central executive network to general cognitive function (Madole et al., 2021), we

hypothesised to find a stronger genetic correlation between general cognitive ability and volumetric PC1s underlying the central executive network, relative to other brain networks (see pre-registered plan https://osf.io/7n4qj). There was no evidence for differences in correlation magnitudes between the central executive network and general cognitive ability compared with other networks, even after accounting for network sizes (SFig.22; STable 6). Adjustments for network sizes were done by dividing effect sizes by the number of volumes contained in a network (SMethods 2.11).

3.4 Associations between aging and general dimensions of brain morphometry

3.4.1 Associations between genetic whole-brain PC1 loadings and age sensitivity

Previous phenotypic work demonstrated that brain volumes more central to overall brain structure – indexed by PC1 loadings onto a phenotypic PC1 underlying 83 brain-wide volumes – were most susceptible to aging. Aging was represented by cross-sectional Pearson's volume-age correlations (Madole et al., 2021), that are typically negative in adult populations. Here, we replicated this phenotypic association in a larger sample (r = -0.43, p = 4.4 x10^{-5}; Fig.6A), and we found a significant, though smaller association between *genetic* PC1 loadings and the same volume-age correlations (r = -0.27, p = 0.012; Fig.6B). This suggests that the more genetically central a region was to overall brain structure, the more sensitive that region also was to age-related shrinkage. Note that this association with age sensitivity emerged even though the PC1 loadings were extracted from brain volume GWAS residualised for age.
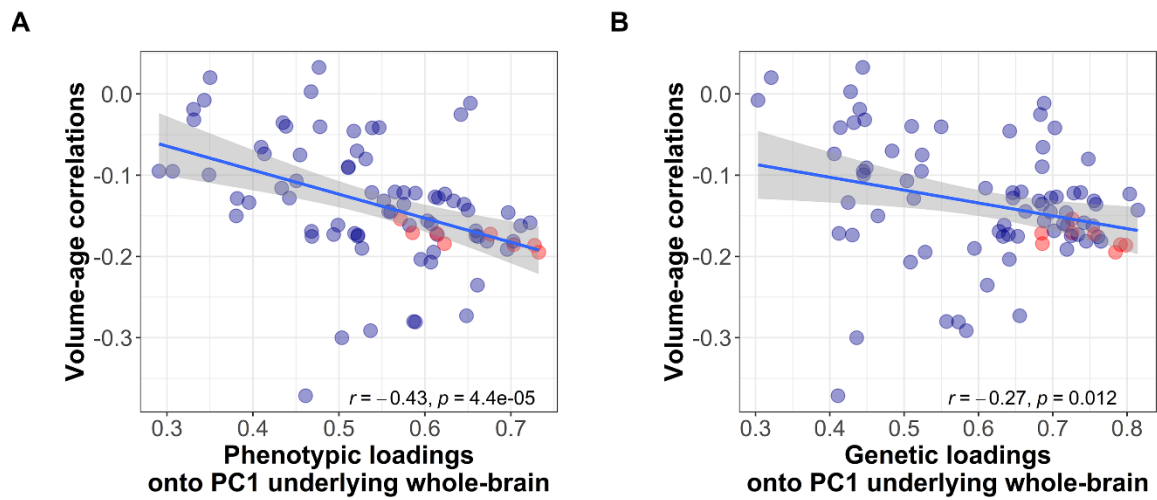
Fig.6. Association between **(A)** phenotypic, and **(B)** genetic PC1 loadings of all 83 volumes (onto a PC1 underlying the whole brain) and a volumes' cross-sectional association with age (Section 3.4.1), which is known as "age sensitivity" (Madole et al., 2021).

Volumes coloured in red are regions thought to reside in the central executive network, which tended to be both phenotypically and genetically central to overall brain structure (high PC1 loadings), and they tended to be more susceptible towards age (large volume-age correlation).

3.4.2 Genetic correlation between whole-brain genetic PC1 and brain age gap

Finally, we calculated a genetic correlation between a genetic PC1 underlying the whole brain and *brain age*, for which we used the brain age GWAS by Kaufmann et al. (2019). The genetic correlation was moderate and negative (r_g = -0.34; *SE* = 0.06), suggesting that there is a shared genetic basis to demonstrating younger brain age, and having consistently larger volumes across the whole brain (aging discussion in Section 4.3).

4. Discussion

Here, we have introduced a multivariate approach integrating covariance across both multiple brain regions and the genome (*Genomic PCA*) to help understand the links between the genetic architecture of human brain morphometry and the network organisation of the brain. In line with Cheverud's Conjecture (Cheverud, 1988), phenotypic and genetic brain organisation seemed to operate on the same major dimensions: phenotypic and genetic correlations were similar (Section 4.1). There was moderate genetic overlap between cognitive ability, aging, and global trends of morphometry underlying both the whole-brain and more parsimonious canonical brain networks (Sections 4.2 for cognitive ability, Section 4.3 for aging). To complement theory-driven perspectives like in this study, our method Genomic PCA may be used to identify regions most important to overall brain structure (e.g., volumes with largest PC loadings) to be prioritised in future investigations of the relationship between the brain and cognitive ability.

4.1. Analogous phenotypic and genetic interregional covariance across the brain

To our knowledge, this is the first genetically-informed study that corroborates the brain organisation observed in phenotypic studies - we demonstrated analogous interregional covariance across the whole brain derived from both phenotypic and genetic indices (i.e., highly corresponding interregional correlations and whole-brain PC1 loadings). Analogous to phenotypic findings in Madole et al. (2021), we found that some volumes were genetically more important for overall brain structure than others, indicated by high loadings onto the first principal component underlying the whole

brain. For example, frontal and parietal volumes, theorised to be part of the central executive network, had consistently high loadings, indicating their overall importance for overall brain structure

The close phenotypic and genetic correspondence in interregional covariance means that inferences from genetic to phenotypic dimensions are viable. This is in line with previous studies comparing phenotypic and genetic correlations between morphometric traits (Biton et al., 2020; Sodini et al., 2018). According to Cheverud's Conjecture, this indicates that genetics of brain organisation operate on the same dimensions as are evident phenotypically, and likely index the same developmental processes. More genetically-informed studies of brain organisation are needed to map those major dimensions onto the relevant biological pathways and mechanisms.

We suggest a similar organisation of phenotypic and genetic brain architecture is supporting evidence for the neurobiological validity of canonical brain networks considered in this study. The fact that our theoretical grouping of volumes into brain networks – informed by commonly-referenced studies of structural, functional, and lesion-based studies (Bressler & Menon, 2010; Jung & Haier, 2007; Madole et al., 2021; Menon & Uddin, 2010) – yielded networks that explained more variance than expected by randomly grouping volumes into networks, provides some evidence for the ontological reality of those networks.

However, it was surprising to find a lack of specificity between different networks at the level of their broad associated polygenic signal, which was quantified through very high genetic correlations between genetic PC1s underlying brain volumes in different canonical networks (range r_g = 0.63-0.97). This suggests that our Genomic PCA analyses picked up on general genes linked with global brain-wide features of

morphometric trends, which are practically the same across canonical networks and the whole brain. Future studies wishing to index the genetic correlates of these global features may focus on more parsimonious, and computationally more efficient, brain networks including a few volumes most representative of overall brain structure (e.g., 8 regions in central executive network), rather than modelling the whole brain.

4.2 Genetic correlations between general cognitive ability and general morphometry underlying canonical brain networks

Using a multivariate definition of general cognitive ability, we demonstrated PC1s underlying all nine brain networks, and the whole brain, were genetically associated with cognitive ability at small-to-moderate magnitudes (r_g = 0.17-0.21). The effect sizes were about the same magnitude as Jansen et al. (2020) found for a genetic correlation between total brain volume and cognitive ability (r_g = 0.24); this was even when some of our models considered only few brain regions (i.e., central executive included only 8 volumes and still yielded magnitudes as large as total brain size). Furthermore, our genetic network associations were numerically larger than genetic correlations obtained from individual brain volumes (range r_g = 0.07-0.13 in Biton et al. (2020)). This should encourage future studies to model general trends of morphometry underlying multiple brain regions, instead of considering individual regions only. This seems to distil less noisy genetic variance, more robustly relevant to cognitive ability.

In contrast to phenotypic findings (Madole et al., 2021), there was no evidence that genetic correlates underlying morphometry in the central executive network was any more strongly associated with cognitive ability than the other networks. This is compatible with the lack of specificity between different brain networks at the level of

135

their associated polygenic signal (discussed in <u>Section 4.1</u>): each network made a similar prediction of cognitive ability at the genetic level. The fact that a disproportionate role of the central executive network did not replicate in our genetically-informed design (even when accounting for network size), may suggest that genetics are more likely to predispose towards more general genes of global brain features shared across the brain. Tentatively, this would also suggest that instead of genes, environmental processes might drive phenotypically observed specialisations of brain networks, causing different morphometric structures to matter more (or less) for optimal cognitive performance.

4.3 Genetic associations between aging and general dimensions of brain morphometry

We demonstrated that regions genetically more important to overall brain structure (i.e., large whole-brain PC1 loadings) also tended to be more sensitive towards age-related shrinkage (i.e., cross-sectional volume-age correlations; $r = -0.27$). This may be due to more strenuous metabolic burden (or other functional stresses) on regions central to overall structure, possibly through more heavily-demanding cognitive processes. This could alter disproportionately the speed at which some regions atrophy with advancing age. Whereas this was previously described phenotypically, to our knowledge we present the first genetically-informed study to show this relation. However, we suggest it requires triangulation either by future longitudinal ageing studies, or cross-sectional studies modelling within-person atrophy by incorporating information on prior brain size (e.g., intracranial volume as a proxy for size at younger age).

We also found a substantial genetic correlation of general trends of morphometry across the whole brain with the brain age gap (r_g = -0.34), suggesting there is a shared genetic basis to brain age and general trends of brain organisation, even after residualising brain volume GWAS for age (which we had done prior to calculating interregional covariance and genetic PC1s). The genetics associated with younger appearing brains may act through overlapping biological processes that are also part of the mechanism of well-integrated global features of brain morphometry. That is, patterns of brain structural ageing may not just capture how quickly an individual's regional volumes decline compared to their peers, but rather, general healthy morphometry across the brain. This would be compatible with phenotypic research showing that younger brain age predicts better physical fitness, better fluid intelligence, and longer lifespan (Cole et al., 2018). Healthy brain morphometry could vary between people for many non-age-related reasons; our findings suggest it may, at least partly, be due to genetic predisposition, possibly towards better-integrated, more resilient brain biology.

4.4 Limitations

Analyses in this study come with limitations. Genetic correlations are representative for genetic associations across the entire genome, but do not give direct insight into specific DNA regions of sharing. As genetic correlations were calculated using LDSC, the limitations that apply to LDSC methodology are relevant to our study (discussion in Supplementary Note). We conclude based on heritability estimates, indexing signal-to-noise ratios in GWAS, that there was sufficient polygenic signal to warrant LDSC analysis (heritability ranged 7-42%). LDSC intercepts were perfectly associated with phenotypic correlations (R^2 = 0.99), indicating that the analyses

successfully separated confounding signal (including environmental factors) from the estimates of genetic correlations.

This study was conducted in the UK Biobank sample, which is not fully representative of the general population of the United Kingdom: its participants are more wealthy, healthy, and educated than average (Fry et al., 2017). Cohort effects may affect the degree to which differential brain-regional susceptibility to ageing can be inferred from cross-sectional data. It remains to be tested whether our results can be extrapolated to socio-economically poorer subpopulations, or outside European ancestry. Results were also dependent on the choice of brain parcellation to divide the cortex into separate regions.

4.5 Conclusion

To study the neurobiological bases of adult cognitive aging, we introduced a multivariate framework to integrate covariance across multiple brain regions and the genome (Genomic PCA), which allowed modelling general dimensions underlying brain-wide morphometry. In line with Cheverud's Conjecture, phenotypic and genetic brain organisation seemed to operate on the same major dimensions, and moderate genetic correlations supported that genes underlying general dimensions of brain morphometry are implicated in cognitive aging. Genetically more important regions to overall brain structure tended to be more susceptible towards age-related shrinkage. However, instead of uncovering localised brain network-specific genetic correlates, we only found evidence for general genetic correlates of brain-wide morphometric features. This may imply that environmental, or otherwise non-genetic, processes are more likely than genes to drive different morphometric structures to matter more (or less) for better cognitive performance. The evidence presented here brings us closer to

characterising the aetiology and robust neurobiological correlates of cognitive aging, and provides a foundation for future investigations ultimately working on interventions for cognitive decline.

Unpublished addition to Chapter 4:

Adjusting for brain size would have substantially altered study results

The results presented in Chapter 4 were obtained from sets of GWAS summary statistics that I calculated as part of this PhD. It may not be immediately apparent why this was necessary, considering several resources already provide GWAS summary statistics for regional Desikan-Killiany brain volumes in the UK Biobank (Smith, Douaud, et al., 2020). However, using these GWAS summary statistics produced irregular and unexpected results. This *Unpublished Addition to Chapter 4* aims to add to the conclusions drawn in Chapter 4 by contrasting the Chapter 4 results (that used GWAS unadjusted for brain size) with results that would have been obtained had I used GWAS summary statistics by Smith, Douaud, et al. (2020) that were adjusted for a volumetric scaling factor from the T1 head image (i.e., a proxy of brain size) and many other covariates. Their comparison also adds to the brain size adjustment discussion in Chapter 2 as it provides evidence that I argue is indicative of collider bias in GWAS summary statistics adjusted for brain size (or a proxy of brain size).

As pre-registered, I originally planned to use GWAS summary statistics by Smith, Douaud, et al. (2020) (for simplicity, below, I refer to this set of GWAS as *Smith GWAS* adjusted for brain size). When using the Smith GWAS for analyses outlined in Chapter 4, the results were unexpected and sparked my doubt about covariate adjustment in GWAS summary data: Interregional genetic correlations were dissociated from interregional phenotypic correlations whereby many genetic interregional correlations were negative (Fig.1 – same as Fig.4. in Chapter 4 but here extracted from GWAS adjusted for brain size). Also, brain networks modelled from

GWAS adjusted for brain size (Smith GWAS) were not significantly genetically correlated with cognitive ability (*g*-factor). It seemed implausible that Chapter 4 was the first study to discover that Cheverud's Conjecture did not apply to human brain morphometry.
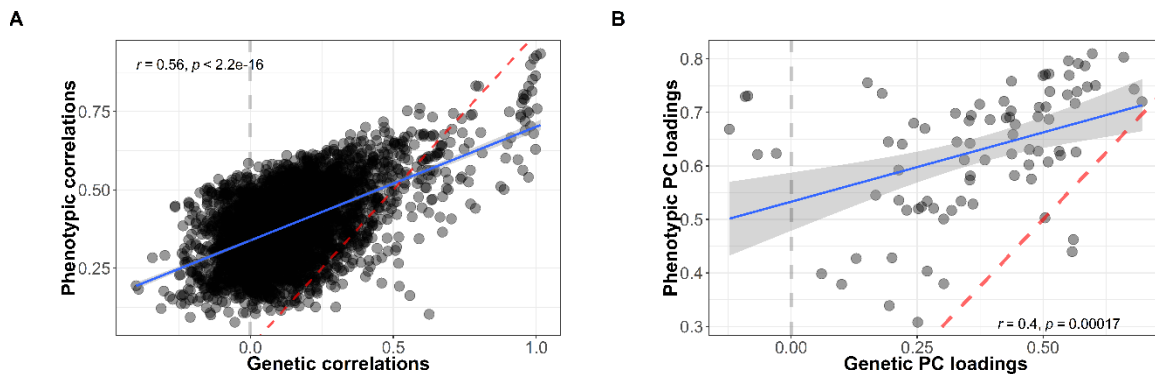
Fig.1. Interregional correlations and PC loadings from GWAS adjusted for brain size (Smith GWAS) as input.

(A) Correlation between phenotypic and genetic correlations of 83 regional grey-matter volumes. The dashed red line is the line of identity. **(B)** Correlation between phenotypic and genetic PC loadings on the first PC underlying 83 regional grey-matter volumes. The dashed red line is the line of identify, with a slope of 1 and an intercept of 0.

It subsequently became apparent that the GWAS adjusted for brain size (Smith GWAS) had been controlled for ~200 covariates which were listed in the Supplementary Material (Smith, Douaud, et al., 2020). That covariates were only reported in the Supplement indicates that the importance of covariate control is often deemed secondary. The main manuscript associated with Smith, Douaud, et al. (2020) only briefly mentions the inclusion of GWAS covariates with reference to the authors' *phenotypic* work in "*Confound modelling in UK Biobank brain imaging*" (Alfaro-Almagro et al., 2021). However, there is a difference in implied causal chains between phenotypic analyses and genetic analyses (as illustrated in Fig.9 in Chapter 2) which

suggests that different considerations apply in the covariate selection process of GWAS vs. phenotypic analyses.

Smith, Douaud, et al. (2020) used ~200 GWAS covariates including age, sex, acquisition site, head motion, and brain size. I would argue that variance accounted for by most of these variables is indeed nuisance variance because it can be reasonably assumed that head motion, for example, should not relate to the phenotype of interest or the genome and can therefore be discarded safely. This also applies to sex as only autosomal SNPs are considered in GWAS. On the contrary, brain size is strongly associated with brain regional volumes (Chapter 2) and the genome (Jansen et al., 2020).

In the context of Chapter 4, I aimed to capture genetic dimensions of variance *shared* between regions by modelling PCs underlying brain networks. In this context, residualizing brain region GWAS for brain size seemed inappropriate because brain size majorly overlaps with the variance captured by brain regions. I recalculated 83 brain volume GWAS summary statistics *unadjusted* for brain size in order to retain brain-size correlated variance and the regional information that it carries (Chapter 2), and to avoid potential collider bias by adjusting for heritable covariates (discussed in Chapter 2 and in Aschard et al. (2015)). More detail on the calculation of brain regional GWAS unadjusted for brain size are in the Appendix. These brain size unadjusted GWAS formed the basis for analyses presented in Chapter 4, and they rely on the same brain region phenotypes and approximately the same UKB sample as GWAS adjusted for brain size by Smith, Douaud, et al. (2020).

In my brain size unadjusted GWAS, I only included covariates about which I was certain that they exclusively tagged nuisance variance for my research question:

age, sex, genetic genotyping batch, and 40 ancestral principal components. I excluded time of year, scanner coordinates, and acquisition site as covariates because I found that they were empirically uncorrelated with brain region GWAS phenotypes (i.e., they explained less than an arbitrary, pre-registered 1% of brain region variance). I decided it was safe *not* to include those covariates because they unlikely act as a considerable confounder when they explain <1% of the variance, and it reduces the potential that I induce collider bias by adjusting for heritable covariates.

The remainder of this *Addition to Chapter 4* is dedicated to briefly compare the published results in Chapter 4 with those obtained from GWAS adjusted for brain size (Smith GWAS) to illustrate how adjusting for brain size would have altered the results. I felt this was important because the impact of brain size adjustment was substantial, and all existing brain regional GWAS – I am aware of – do indeed control for measures of brain size as a covariate (Grasby et al., 2020; Zhao, Luo, et al., 2019), which would mean that all publicly available brain regional GWAS are vulnerable to collider bias issues. The comparisons between results obtained from GWAS adjusted and unadjusted for brain size contrasted below deliver evidence that brain size GWAS adjustment may induce collider bias.

4a.1 Block jackknife comparisons between heritability and genetic overlap estimates

Block jackknife analyses aim to test significant differences between two heritability estimates (or genetic correlation estimates) through resampling. In essence, block jackknife enables significance testing by delivering standard errors of heritability (or genetic correlation) estimates. In addition to the full LDSC estimate, I calculated heritability (and genetic correlations) 200 times whereby a different

genomic LD block was excluded from the estimation to obtain pseudo heritability (or genetic overlap) estimates. To perform hypothesis tests, the mean and variance of the pseudovalues were used to derive z-statistics and corresponding confidence intervals (Hübel et al., 2019).

According to block jack-knife analyses, most brain regional volumes did not significantly differ in their heritability estimates when obtained from GWAS adjusted vs. unadjusted for brain size (Fig.2A). However, descriptively, GWAS adjusted for brain size (Smith GWAS) yielded somewhat smaller heritability estimates (median h^2 $_{adjusted\ for\ brain\ size}$ = 18%; $range$ = 7-43%), than those unadjusted for brain size (median h^2 $_{no\ brain\ size\ adjustment}$ = 23%; $range$ = 7-42%). To assess whether GWAS adjusted and unadjusted for brain size were identical, I performed block jack-knife analyses to test whether their genetic correlation was significantly different from one (Fig.2B). For most brain volumes, the genetic correlation between regional GWAS adjusted and unadjusted for brain size tended to be significantly different from one, indicating that brain size adjustment substantially altered GWAS signal associated with most brain regions.
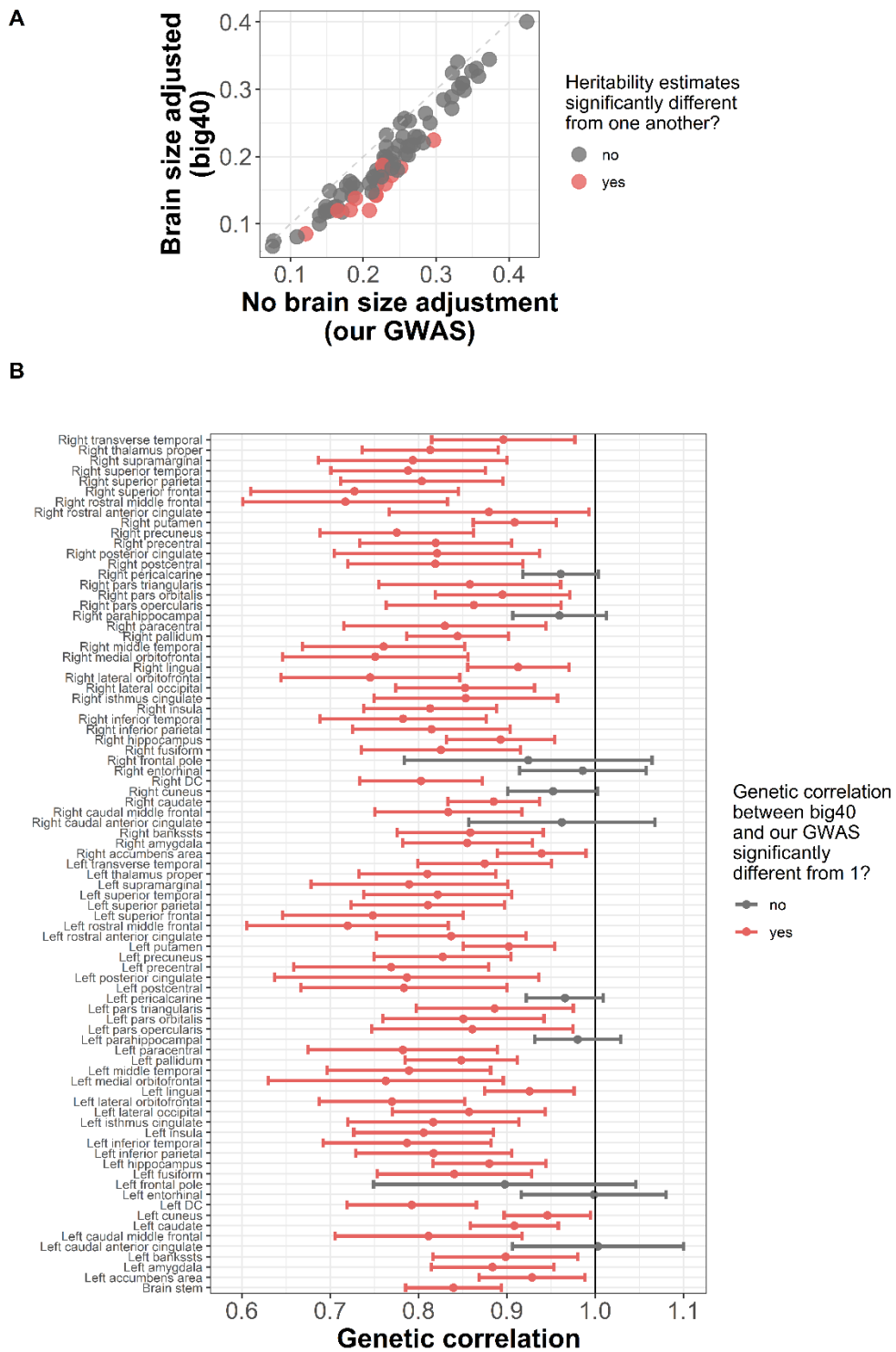
Fig.2. Block jack-knife analyses comparing GWAS adjusted and unadjusted for brain size in (A) heritability estimates, and (B) genetic correlations.

4a.2 Comparisons of genetic interregional correlations

Genetic interregional correlations were overall smaller when obtained from GWAS <u>adjusted</u> for brain size (median $r_{g \text{ adjusted for brain size}}$ = 0.12; $range$ = -0.4 − 1.0) than from GWAS <u>unadjusted</u> for brain size (median $r_{g \text{ no brain size adjustment}}$ = 0.35; $range$ = -0.08 − 1.0). GWAS <u>adjusted</u> for brain size had a wider spread in correlation magnitudes (Fig.3C vs. Fig.3D).
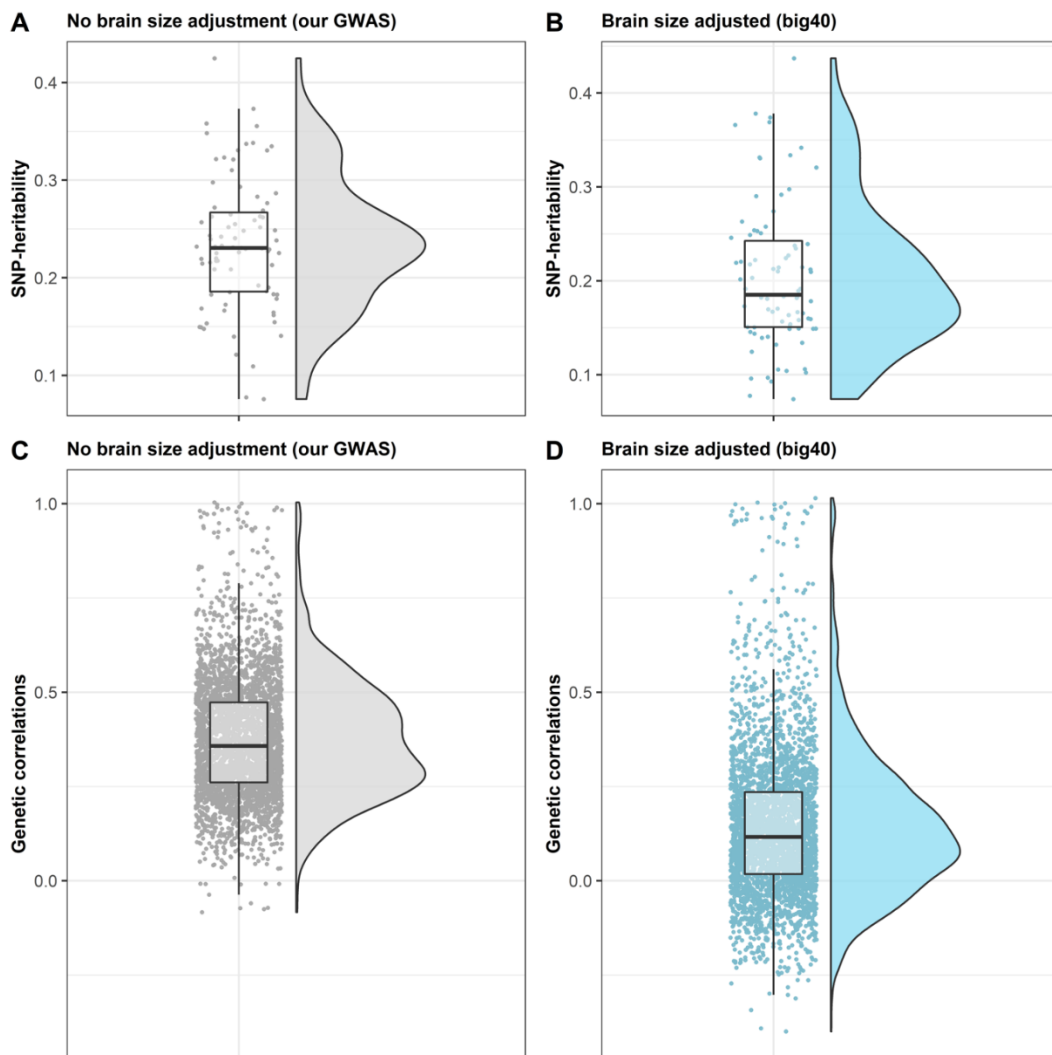
Fig.3. Comparison of brain regional heritability, and interregional genetic correlations obtained from GWAS unadjusted (Chapter 4 GWAS) and adjusted for brain size (Smith GWAS).

4a.3 Comparisons of genetic Principal Component loadings

PC loadings onto the first genetic PC underlying 83 regions were overall smaller when obtained from GWAS <u>adjusted</u> for brain size (Smith GWAS; *median PC loadings adjusted for brain size* = 0.39; *range* = -0.12 − 0.7) than from GWAS <u>unadjusted</u> for brain size (*median PC loadings no brain size adjustment* = 0.65; *range* = 0.30 − 0.81). Presumably, this reflects the fact that interregional correlations from GWAS <u>adjusted</u> for brain size (Smith GWAS) were overall smaller than those from <u>unadjusted</u> GWAS (shown in [Fig.3](#)C-D). The two sets of PC loadings were nevertheless substantially related (*r* = 0.74; [Fig.4](#)), suggesting that, even after brain size adjustment, regions retained approximately the same order in how well-representative they were of the rest of the brain.
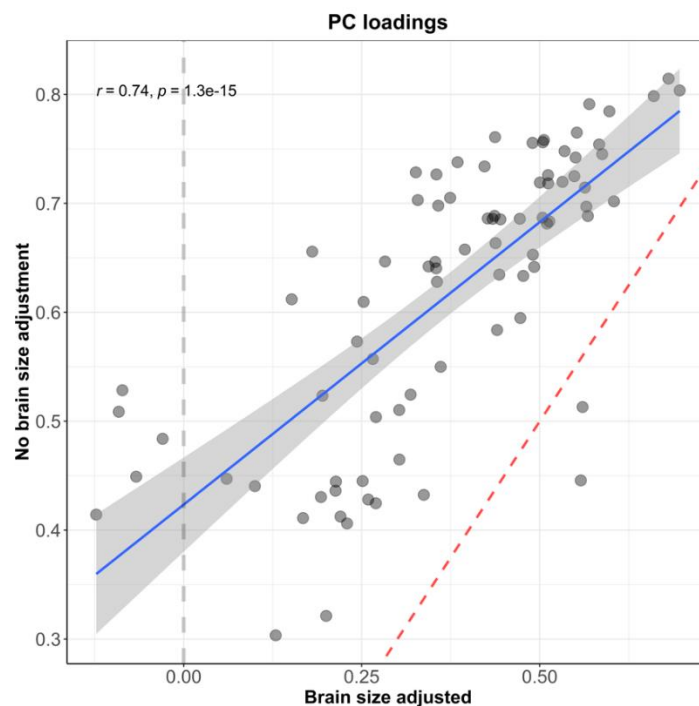
Fig.4. Relationship between PC loadings obtained from the Smith GWAS ("*Brain size adjusted*") and PC loadings obtained from GWAS used in Chapter 4 ("*No brain size adjustment*").

4a.4 Comparisons of genetic correlations between brain networks and cognitive ability

To test for significant genetic correlations between cognitive ability (*g*-factor) and brain networks, I performed Genomic Structural Equation Models (SEMs) in the GenomicSEM software (Grotzinger et al., 2019). In Chapter 4, Genomic SEMs were calculated based on genetic correlation matrices among seven individual cognitive abilities and nine canonical brain networks – underlined unadjusted for brain size (Fig.5A). For comparison, Fig.5B displays genetic correlation matrices among the same traits but adjusted for brain size (Smith GWAS). Brain networks unadjusted for brain size yielded overall larger genetic correlations with individual cognitive traits (median r_g $_{brain\ networks\ \&\ individual\ cognitive\ abilities}$ = 0.14; range = 0.01 – 0.30), than the brain networks adjusted for brain size (median r_g $_{brain\ networks\ \&\ individual\ cognitive\ abilities}$ = 0.06; range -0.13 – 0.36). It was unexpected that genetic brain networks obtained from GWAS adjusted for brain size (Smith GWAS) yielded overall more extreme genetic correlations with individual cognitive traits: For example, a PC underlying the default mode network – adjusted for brain size (Smith GWAS) – had a genetic correlation of r_g = -0.06 with the reaction time task, but r_g = 0.36 with the Tower Rearranging task (Fig.5).

A   Genetic correlations with cognitive ability:
no brain size adjustment

Brain networks	Matrix	Memory	RT	Symbol Digit	TMTB	Tower	VNR
whole_brain	0.23	0.1	0.05	0.14	0.15	0.27	0.25
temporo	0.24	0.1	0.03	0.11	0.13	0.3	0.24
sensori	0.11	0.09	0.05	0.12	0.15	0.17	0.23
salience	0.13	0.05	0.11	0.13	0.15	0.14	0.22
p_fit	0.23	0.11	0.03	0.12	0.13	0.27	0.24
multiple	0.17	0.11	0.02	0.11	0.14	0.21	0.23
hippocampal	0.18	0.03	0.07	0.13	0.11	0.2	0.19
default_mode	0.22	0.11	0.01	0.14	0.13	0.29	0.21
cingulo	0.19	0.09	0.05	0.12	0.15	0.22	0.25
central_exec	0.2	0.12	0.03	0.14	0.14	0.25	0.22

Cognitive tasks

B   Genetic correlations with cognitive ability:
brain size adjusted (big40)

Brain networks	Matrix	Memory	RT	Symbol Digit	TMTB	Tower	VNR
whole_brain	0.13	0.09	-0.03	0.07	0.07	0.27	0.11
temporo	0.17	0.04	-0.05	0.01	0.03	0.28	0.1
sensori	-0.13	0.08	-0.01	0.05	0.08	0.08	0.12
salience	-0.12	-0.01	0.09	0.05	0.01	-0.1	0.05
p_fit	0.16	0.1	-0.05	0.05	0.04	0.28	0.09
multiple	0.01	0.09	-0.06	0.02	0.06	0.21	0.09
hippocampal	0.14	-0.03	0.05	0.05	0.04	0.1	0.07
default_mode	0.1	0.08	-0.06	0.05	0.04	0.36	0.07
cingulo	0	0.04	-0.03	0.04	0.05	0.17	0.1
central_exec	0.06	0.12	-0.04	0.09	0.12	0.32	0.09

Cognitive tasks

Fig.5. Genetic correlations between structural brain networks obtained with Genomic PCA and individual cognitive ability traits.

These correlation matrices are the basis for GenomicSEM analyses. Cognitive Tasks: Matrix = Matrix Pattern Completion, Memory = Pairs Matching, RT = Reaction Time, Symbol Digit = Symbol Digit Substitution, TMTB = Trail Making − B, Tower = Tower Rearranging, VNR = Verbal Numerical Reasoning, temporo = temporo-amygdala-orbitofrontal network, sensori = sensori-motor network, salience = salience network, p_fit = Parieto-Frontal-Integration Theory network, multiple = multiple demand network, hippocampal = hippocampal-diencephalic, default mode = default mode network, cingulo = cingulo-opercular network, central exec = central executive network

149

Finally, I performed Genomic SEMs on the basis of genetic correlation matrices displayed in Fig.5 to quantify the genetic correlations between canonical brain networks and cognitive ability (*g*-factor). It was the aim of this to model the *g*-factor as the variance common between cognitive traits, which is then tested for a genetic correlation with each of the canonical brain networks. Genomic SEMs performed based on genetic correlations adjusted or unadjusted for brain size produced opposing conclusions: Brain networks unadjusted for brain size were significantly associated with the genetic *g*-factor, but brain networks adjusted for brain size (Smith GWAS) were *not* significantly associated with the genetic *g*-factor (Table 1; 99% confidence intervals multiple testing correction).

Table 1. Genetic correlations between structural correlation networks and a factor of general cognitive ability (*g*-factor).

Genetic correlation (rg) with g-factor	Brain size adjusted (big40)	99% CI	No brain size adjustment (our GWAS)	99% CI
Central Executive	0.13	0.01 - 0.25	0.20	0.1 - 0.29
Cingolo-opercular	0.07	-0.05 - 0.19	0.20	0.1 - 0.29
Default Mode	0.08	-0.03 - 0.19	0.19	0.1 - 0.29
Hippocampal	0.06	-0.05 - 0.18	0.17	0.07 - 0.27
Multiple Demand	0.08	-0.06 - 0.21	0.19	0.09 - 0.29
P-FIT	0.09	-0.02 - 0.21	0.20	0.1 - 0.29
Salience	0.04	-0.07 - 0.14	0.19	0.09 - 0.28
Sensorimotor	0.09	-0.02 - 0.21	0.19	0.09 - 0.29
Temoro-amygdala-orbitfrontal	0.07	-0.04 - 0.18	0.20	0.1 - 0.29
Whole brain	0.11	0 - 0.22	0.21	0.11 - 0.31

Abbreviations for the canonical brain networks: temporo = temporo-amygdala-orbitofrontal network, sensori = sensori-motor network, salience = salience network, p_fit = Parieto-Frontal-Integration Theory network, multiple = multiple demand network, hippocampal = hippocampal-diencephalic, default mode = default mode network, cingulo = cingulo-opercular network, central exec = central executive network

4a.5 Correlational patterns of results adjusted vs. unadjusted for brain size are consistent with the consequences of collider bias

I suggest that these overall correlational patterns are consistent with the idea that adjusting GWAS summary statistics for brain size induces collider bias which means that the GWAS likely captures (at least some) false positive SNPs (Aschard et al., 2015). Block jackknife comparisons (Section 4a.1) are consistent with this because GWAS adjusted and unadjusted for brain size did not differ in their heritability estimates (i.e., they had the same overall amount of captured polygenic signal; Fig.2A), but they captured different SNPs (i.e., significantly different polygenic signal; Fig.2B). This suggests that GWAS adjusted for brain size capture just as many – but different – SNPs compared with GWAS unadjusted for brain size. However, I argue that GWAS adjusted for brain size should have captured overall fewer SNPs because brain size adjustment substantially removes regional information (Chapter 2). That adjusted and unadjusted GWAS capture just as many, but different (and not simply fewer) SNPs is compatible with the idea false positive SNPs are captured in adjusted GWAS which falsely inflated heritability estimates.

When using GWAS summary statistics adjusted for brain size, it was unexpected to find widely differing genetic correlations between brain networks and individual cognitive abilities (e.g., PC underlying the default mode network widely ranged in its association with different cognitive traits from r_g = -0.06 to r_g = 0.36; Fig.5; Section 4a.4). This was unexpected because individual cognitive abilities tend to be phenotypically and genetically similar (de la Fuente et al., 2021). The widely differing genetic correlations between brain networks and individual cognitive abilities may reflect that false positive SNPs – induced by collider bias in the GWAS

adjusted for brain size – may have "accidentally" mapped onto individual cognitive abilities. This is conceivable considering individual cognitive abilities are highly polygenic. I suggest that this overlap between brain networks – adjusted for brain size – and individual cognitive abilities is likely accidental because it is so different between individual cognitive abilities that should be similar.

When modelling Genomic SEMs of the association between brain networks and cognitive ability (*g*-factor), brain networks adjusted for brain size were genetically uncorrelated with the *g*-factor, likely because the *g*-factor only indexes variance *systematically shared* between individual cognitive abilities: the *g*-factor would not have captured any of the polygenic signal associated with individual cognitive abilities that accidentally mapped onto brain networks. In sum, SNPs overlapping between brain networks adjusted for brain size and individual cognitive abilities could have mainly been the false positive, collider induced SNPs, whereas SNPs that would have been systematically correlated with the *g*-factor were not reliably captured by the regional GWAS adjusted for brain size. This is a powerful example how a multivariate approach such as the *g*-factor model helped sanity check correlation patterns of noisy morphometric variance.

To summarise, adjusting GWAS summary data for brain size prohibits capturing reliable associations between the *g*-factor and brain networks. This *Addition to Chapter 4* provides valuable context to the results presented in Chapter 4 because it is important to evaluate and build an intuition for the consequences of brain size adjustment. Brain size adjustment consequences likely affect all existing brain region GWAS summary statistics – I am aware of – as they all included a measure (or proxy measure) of brain size as a covariate (Grasby et al., 2020; Smith, Douaud, et al., 2020;

Zhao, Luo, et al., 2019). This also further underlines the importance of exploratory groundwork as presented in Chapter 2.

Chapter 5

A Quantified Comparison of Cortical Atlases on the Basis of Trait Morphometricty

Registered Report

A quantified comparison of cortical atlases on the basis of trait morphometricity

Anna E. Fürtjes [a,*], James H. Cole [b,c,d], Baptiste Couvy-Duchesne [e,f,1] and Stuart J. Ritchie [a,1]

[a] Social, Genetic and Developmental Psychiatry (SGDP) Centre, Institute of Psychiatry, Psychology & Neuroscience, King's College London, UK
[b] Department of Neuroimaging, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK
[c] Centre for Medical Image Computing, Department of Computer Science, University College London, London, UK
[d] Dementia Research Centre, Institute of Neurology, University College London, London, UK
[e] Paris Brain Institute (ICM), Inserm (U 1127), CNRS (UMR 7225), Sorbonne University, Inria Paris, Aramis Project-team, Paris, France
[f] Institute for Molecular Bioscience, The University of Queensland, St Lucia, Queensland, Australia

ARTICLE INFO

ABSTRACT

Background: Many different brain atlases exist that subdivide the human cortex into dozens or hundreds of regions-of-interest (ROIs). Inconsistency across studies using one or another cortical atlas may contribute to the replication crisis across the neurosciences.

Methods: Here, we provide a quantitative comparison between seven popular cortical atlases (Yeo, Desikan-Killiany, Destrieux, Jülich-Brain, Gordon, Glasser, Schaefer) and vertex-wise measures (thickness, surface area, and volume), to determine which parcellation retains the most information in the analysis of behavioural traits (incl. age, sex, body mass index, and cognitive ability) in the UK Biobank sample (N~40,000). We use linear mixed models to compare whole-brain morphometricity; the proportion of trait variance accounted for when using a given atlas.

Results: Commonly-used atlases resulted in a considerable loss of information compared to vertex-wise representations of cortical structure. Morphometricity increased linearly as a function of the log-number of ROIs included in an atlas, indicating atlas-based analyses miss many true associations and yield limited prediction accuracy. Likelihood ratio tests revealed that low-dimensional atlases accounted for unique trait variance rather than variance common between atlases, suggesting that previous studies likely returned atlas-specific findings. Finally, we found that the commonly-used atlases yielded brain-behaviour associations on par with those obtained with random parcellations, where specific region boundaries were randomly generated.

Discussion: Our findings motivate future structural neuroimaging studies to favour vertex-wise cortical representations over coarser atlases, or to consider repeating analyses across

Sex
Alcohol consumption
Age
Cigarette smoking

multiple atlases, should the use of low-dimensional atlases be necessary. The insights uncovered here imply that cortical atlas choices likely contribute to the lack of reproducibility in ROI-based studies.

1. Introduction

To better understand neuronal correlates of human behaviour, studies typically investigate the association between brain structural measurements in regions-of-interest (ROI) and behavioural traits. Recent efforts to collect large-scale neuroimaging data enable unprecedented opportunities for powerful statistical analyses. One of the most used software for structural brain analysis is *FreeSurfer* (Fischl et al., 2002) which is used to determine regional brain measures from a participant's brain image. Briefly summarised, FreeSurfer models the cortex in a two-dimensional mesh, with an arbitrary resolution of ~150,000 vertices per hemisphere (known as "fsaverage standard space"). Measurements are extracted at each vertex, and include cortical thickness, surface area and volume. With reference to a *cortical atlas* (Fig. 1), the software can subdivide the cortical mesh into ROIs and calculates their structural characteristics, such that the surface area—or grey-matter volume or cortical thickness.

Cortical atlases aim to outline structurally homogeneous and meaningful regions that reflect the organisation of the cortex, but there is no ground-truth parcellation. Different organisational characteristics are inferred based on distinct brain modalities, including anatomical landmarks, cytoarchitecture, and patterns of functional coactivation. Here, we consider seven available atlases: Desikan-Killiany (DK; Desikan et al., 2006), Destrieux (Destrieux, Fischl, Dale, & Halgren, 2010), Glasser (Glasser et al., 2016), Gordon (Gordon et al., 2014), Schaefer (Schaefer et al., 2017), Yeo (Yeo et al., 2011), Jülich-Brain (Amunts, Mohlberg, Bludau, & Zilles, 2020). Atlases differ in their anatomical boundaries and number of ROIs and were generated using different methods and samples.

Desikan-Killiany (68 ROIs) and Destrieux (148 ROIs) are landmark-based atlases, which means that the cortex is divided in a manner consistent with the macroscopic anatomy of gyri and sulci. ROIs in Desikan-Killiany were manually labelled on structural MRI scans from a sample of 40 participants between 19 and 87 years (mean age 55.95 years; Desikan et al., 2006), who had been originally recruited with a range of atrophy levels by the Washington University Alzheimer's Disease Research Center (Fotenos, Snyder, Girton, Morris, & Buckner, 2005). The Destrieux atlas follows what the authors described as "widely accepted anatomical conventions" (page 2; Destrieux et al., 2010) and was derived from 12 healthy participants between 18 and 33 years (mean age 21.67).

Yeo (34 ROIs), Gordon (333 ROIs) and Schaefer (500 ROIs) were derived from resting-state functional MRI data. Gordon et al. (2014) quantify gradients of functional activation across the cortex and use abrupt changes in these gradients as indicators of regional borders. Gordon's atlas was derived

from 120 healthy community-dwelling participants between the ages of 19–32 years (mean age 25). Schaefer et al. (2017) integrates this approach by maximising uniformity within regions while neighbouring voxels are only assigned to the same area if abrupt gradient changes do not separate them. Schaefer et al. used images of 1,489 brains from the Genomics Superstruct Project (participants aged between 18 and 35 years). Yeo et al. (2011) outline 17 macroscopic network configurations that were stably estimated using clustered functional connectivity data from 1,000 healthy participants (mean age 21.3 years).

Glasser et al. (2016, 360 ROIs) parcellate regions using a semi-automated approach where regional boundaries were defined from multiple indicators including cortical architecture, function, and connectivity measures. It was outlined from 200 participants between 22 and 35 years in the Human Connectome Project cohort. Finally, Amunts et al. (2020) present the 'Jülich-Brain', a microstructural parcellation of 137 cortical areas in each hemisphere, reflecting cytoarchitecture across the cortex, which was derived based on 23 postmortem brains (version 2.9). The mean age of this sample was 64 years (range 30–86).

Different atlases outline different regional boundaries with spatial discrepancies (Alexander-Bloch et al., 2018; Bohland, Bokil, Allen, & Mitra, 2009), which likely influences study results. A recent study showed that atlas choice affects estimates of network topology and functional brain connectivity (Revell et al., 2022). In another study, different atlases either induced or masked associations between ROIs and age, meaning that comparable regions were associated with age in one atlas but not in another (Yaakub et al., 2020). The fact that many different atlases are frequently used introduces uncertainty when comparing and attempting to reproduce study results; probably exacerbating the existing lack of consensus about associations between the brain and behaviour across neuroimaging studies (Kharabian Masouleh, Eickhoff, Hoffstaedter, Genon, & Alzheimer's Disease Neuroimaging Initiative, 2019; Marek et al., 2022). The optimal atlas may differ for any given study, and to our knowledge, there is no data-driven guideline about which atlas maximises brain-trait association and prediction of specific behavioural traits.

While variance captured by ROIs is reduced compared with vertex-wise measures, the main advantage of employing cortical atlases is to reduce dimensionality which facilitates interpretation. Recently, a novel statistical framework was presented (Sabuncu et al., 2016), allowing consideration of *all* brain vertices simultaneously to estimate whole-brain associations with behavioural traits. This approach accounts for correlations between vertices and negates the need to summarise vertex measures across ROIs. Using mixed linear models, this method quantifies the proportion of trait
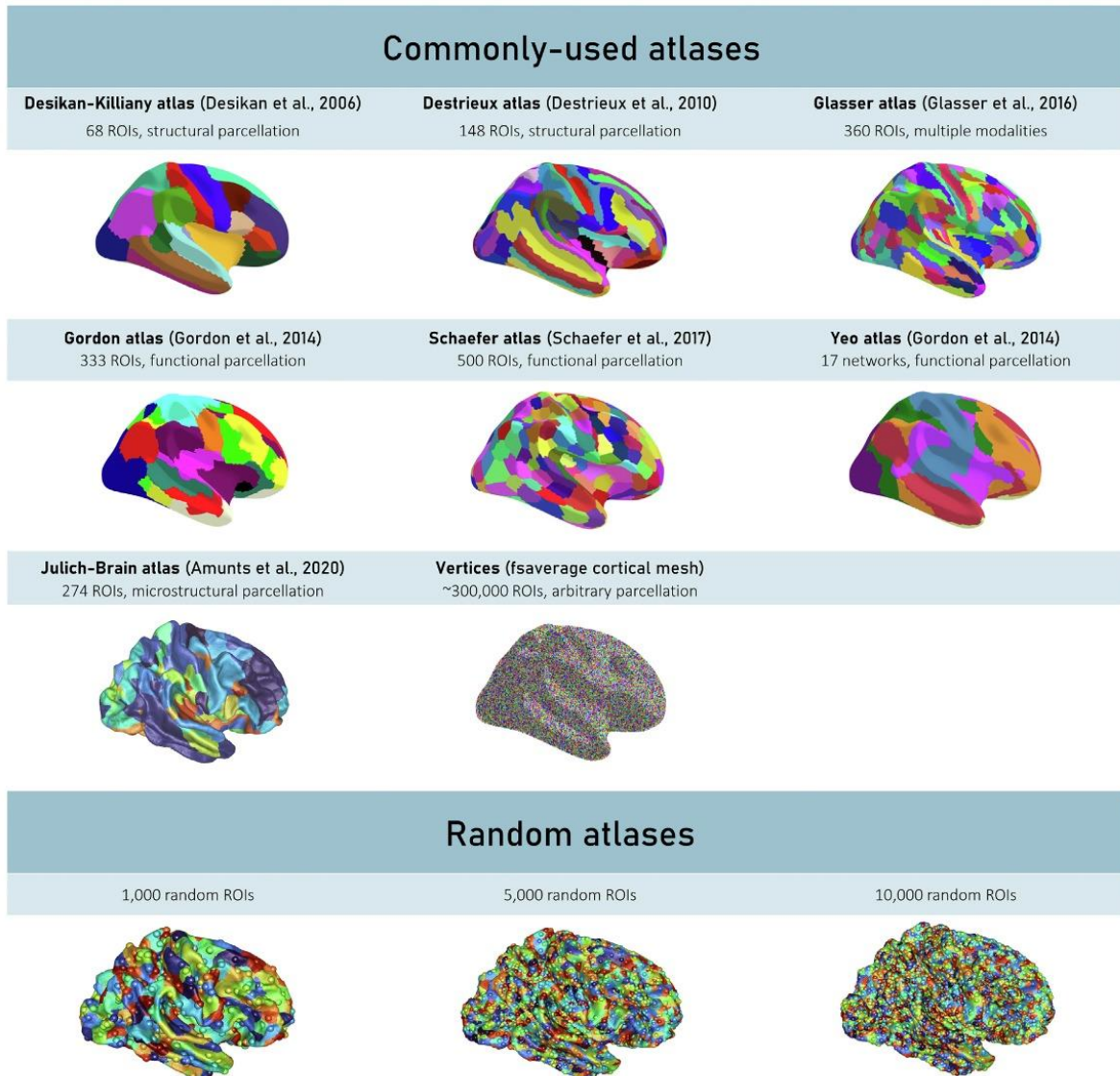
Fig. 1 — The surface-based cortical representations considered in this Registered Report (right hemisphere view). The first six commonly-used atlases were visualised with the ggseg package in R (Mowinckel & Vidal-Piñeiro, 2020). Note that the Yeo atlas annotation files dictate 17 networks in each hemisphere, resulting in 34 ROIs. Vertices were visualised based on 150,000 randomly simulated points. Some colours were generated with the *circulize* package in R (Gu, Gu, Eils, Schlesner, & Brors, 2014). The Julich-Brain atlas and the random parcellations were visualised from 3D coordinates using the *rgl* package in R (https://dmurdoch.github.io/rgl/). In random atlases, vertices with larger radius indicate the seeds from which random ROIs were grown. We also analysed a random atlas with 50,000 ROIs (i.e., 5 times the number of regions as displayed for the 10,000 ROIs), which is not shown in the figure.

variance (R^2) that can be attributed to brain morphology, or *morphometricity*. Conceptually, morphometricity is equivalent to *heritability* in genetics, which indicates the proportion of trait variance explained by many small polygenic effects.

A recent study estimated morphometricity for hundreds of traits, demonstrating that vertex-wise brain measures explained substantial proportions of trait variance (Couvy-Duchesne et al., 2020). In a supplementary analysis, DK ROI-summarised brain measures seemed to explain considerably less variance than the vertex-wise measures, though no statistical significance tests were performed to test this

difference. This effect was most pronounced for age, for which the proportion of variance explained by vertex-wise cortical measures was $R^2 \sim 65\%$, but only $R^2 \sim 25\%$ when derived from DK measures. This suggests that averaging across ROIs masks informative inter-individual variance which is retained in vertex-wise cortical representations. Vertex-wise morphometricity has been shown to be robust across samples, and it explained considerably more variance in traits such as fluid intelligence ($R^2 \sim 13.2\%$) compared with recent predictions models using different methods to maximise brain-based predictive ability ($R^2 \sim 7\%$) of behavioural

traits (Gong, Beckmann, & Smith, 2021). This illustrates that morphometricity estimates could be promising indicators for how well different representations of the brain (vertices or different atlases) can account for individual differences. This could help formulate a recommendation for which brain representation is most promising to capture structural brain associations with behavioural traits.

This Registered Report (https://osf.io/dkw9t/) is the first study to compare commonly-used cortical atlases of varying dimensionality by quantifying and contrasting morphometricity estimates of behavioural variables. We hypothesised that more fine-grained cortical representations would yield larger morphometricity estimates compared with coarser atlases, because explanatory variance is lost when averaging across ROIs. We expected at least a two-fold increase in morphometricity between the coarsest and most fine-grained atlas, which was formulated based on preliminary results in Couvy-Duchesne et al. (2020). It is reasonable to expect that some atlases may be detailed enough to capture maximal morphometric variance. Such an atlas should outperform randomly-generated ROIs and vertex-wise representations, as it summarises anatomically coherent ROIs that should be unaffected by registration imprecision and anatomical variability. We expect to find trait-dependent optimal levels of atlas dimensionality, as per recent work assessing atlas performances based on functional MRI data (Dadi et al., 2020). There, the highest performing predictions in an age variable were achieved by ~150 ROIs, while ~300 ROIs were optimal for intelligence.

The UK Biobank study currently provides FreeSurfer data from which we extracted cortical vertex-wise measurements, as well as ROI-summarised cortical representations in seven commonly-used atlases: Yeo (34 ROIs), Desikan-Killiany (68 ROIs), Destrieux (148 ROIs), Jülich-Brain (274 ROIs), Gordon (333 ROIs), Glasser (360 ROIs), and Schaefer (500 ROIs). To better assess a relationship between morphometricity and dimensionality of the atlas, we randomly generated atlases with 1,000; 5,000; 10,000; and 50,000 ROIs. Morphometricity estimates based on surface area, grey-matter volume, and cortical thickness are reported for seven non-brain traits that are well-measured in UKB and known to be morphometric, that is, robustly associated with grey-matter structure: age (Cole et al., 2018), sex (Ritchie et al., 2018), cognitive abilities (Cox, Ritchie, Fawns-Ritchie, Tucker-Drob, & Deary, 2019), body mass index (Couvy-Duchesne et al., 2020), education (Sabuncu et al., 2016), cigarette smoking, and alcohol consumption (Couvy-Duchesne et al., 2020). We fitted mixed linear models to estimate the morphometricity (i.e. total association) between behavioural traits as dependent variables and either vertex-wise or ROIs measurements as independent variables. We tested for significant differences in morphometricity using likelihood ratio tests.

2. Materials and methods

2.1. UKB neuroimaging data

The UK Biobank (UKB) study is a population-based cohort for health-related information from ~500,000 individuals across the United Kingdom (Sudlow et al., 2015). Among baseline characteristics, physical, and cognitive assessments, it provides pre-processed MRI data from around 40,000 participants (Littlejohns et al., 2020). The Research Ethics Committee ethically approved the UKB study and participants signed informed consent.

Researchers have access to the FreeSurfer outputs (UKB field ID 20263; https://biobank.ndph.ox.ac.uk/showcase/field. cgi?id=20263) from the first neuroimaging visit that have been pre-processed and quality controlled on behalf of UKB (Alfaro-Almagro et al., 2018). Detailed documentation on T1-weighted + T2 FLAIR processing using FreeSurfer 6.0 can be found online (Smith, Alfaro-Almagro, & Miller, 2020). Combining two structural images in processing is considered more precise and associated with improved cortical segmentation (Lindroth et al., 2019) compared to processing T1w images only. The Freesurfer outputs already contained individual-level data with reference to Desikan-Killiany (Desikan et al., 2006), Destrieux, and unsmoothed vertex-wise data. The vertex-wise measurements considered here correspond to the "fsaverage" cortical mesh representation (~300,000 vertices both hemispheres). We derived individual-level cortical measurements with reference to Yeo et al. (2011), Destrieux et al., (2010), Desikan et al. (2006), Glasser et al. (2016), Gordon et al. (2014), Schaefer et al. (2017), and Jülich-Brain (Amunts et al., 2020). Most fsaverage atlas annotation files were obtained from MultiAtlasTT (https://github. com/faskowit/multiAtlasTT. Click or tap if you trust this link.">https://github.com/faskowit/multiAtlasTT), which were in .gcs format and were easily mapped using FreeSurfer 6.0 commands (mris_ca_label & mris_anatomical_stats). From this resource, Schaefer atlas annotation files were only available for a maximum of 500 ROIs. Annotation files for Yeo et al. (2011) were downloaded as part of the FreeSurfer 6.0 software. Jülich-Brain atlas (Amunts et al., 2020) annotation files (v2.9) were obtained from Mangin, Rivière, and Amunts (2021). Mapping atlases from .annot files required a purpose-built pipeline, which we share on GitHub (https://annafurtjes. github.io/Comparing_atlases/). Surface area and volume ROIs are summarised as the sum of vertices, and cortical thickness is an average measure.

2.2. Random parcellations

We generated contiguous random parcellations by randomly selecting and summarising across vertices in the cortical mesh. The number of parcellations matched the number of ROIs contained in the commonly-used atlases (Fig. 1). We also generated an additional four, more fine-grained parcellations with 1,000; 5,000; 10,000 and 50,000 random ROIs. To ensure that parcellations resulted in contiguous ROIs, random seeds were selected from which we grew random ROIs. We used the vcgKDtree R clustering algorithm (Rvcg package) to iteratively grow the ROIs, adding at each iteration six spatially proximal vertices (when possible). This yielded ROIs of roughly the same size. We hypothesised that if random atlases yielded larger morphometricity compared with commonly-used atlases, it would indicate that morphometricity depends on dimensionality and not the exact ROI-specific boundaries outlined in atlases. In this manuscript, we refer to the

commonly-used and random brain parcellations as well as vertex-wise measures using the notion *atlases* that are composed of *ROIs*.

2.3. UKB traits of interest

Morphometricity was estimated for non-brain traits to quantify the proportion of variance accounted for when using cortical measures represented by ROIs from a given atlas. We selected the following dependent variables, because they are highly morphometric, and well-phenotyped in the UKB sample. We considered age (field ID 21003), sex (field ID 31) and body mass index (field ID 21001). We used measures of cognitive ability to construct a "g" factor of general cognitive ability using the lavaan package in R (Rosseel, 2012). We used a confirmatory factor analysis with one factor, based on the following cognitive ability tests: Verbal Numerical Reasoning (field ID 20016, 20191), Trail Making — B (field ID 6350, 20157), Matrix Pattern Completion (field ID 6373), Tower Rearranging (field ID 21004), Symbol Digit Substitution (field ID 23324 & 20159), Pairs Matching (field ID 399), and Reaction Time (field ID 200032). The tests have been described in detail elsewhere, and as in previous work, we considered the first measured occasion for each participant (de la Fuente, Davies, Grotzinger, Tucker-Drob, & Deary, 2021). Compared with standard reference tests, the cognitive assessments exhibit good concurrent validity ($r = .83$) and test-retest reliability (mean Pearson $r = .55$) (Fawns-Ritchie & Deary, 2020). Usually, this general cognitive ability factor explains about 40% of the total variance in the included cognitive tests (Deary, Penke, & Johnson, 2010); here it accounts for 34%. According to commonly used heuristics (Hu & Bentler, 1998), the factor demonstrated good model fit (CFI = .97, RMSEA = .05; SFig. 1).

Additionally, we estimated morphometricity for educational qualification (field ID 6138), number of cigarettes smoked daily (field ID 2887) and frequency of drinking alcohol (field ID 20414; excluding former drinkers field ID 20406). Outliers that fall beyond 4 standard deviations from the mean were removed from the sample. This criterion was only applied to non-brain traits, as this criterion applied to brain measures created too many missing data entries for LMMs to reliably converge. Covariates included in all models consisted of UKB acquisition site (field ID 54), and head positioning in the MRI scanner (X, Y, Z coordinates, field IDs 25756, 25757, 25758).

2.4. Statistical analyses

Morphometricity. Morphometricity quantifies the proportion of inter-individual differences in a non-brain trait (dependent variable) accounted for by cortical measurements (independent variables). The dependent variables in this study are listed above and include age, sex, body mass index, and cognitive ability. Independent variables include the following ROI measures within cortical atlases: either ~300,000 vertex-wise measures, or atlas-wide brain ROIs, which include different amounts of ROIs depending on the atlas. Additionally, we considered four covariates (outlined above in UKB traits of interest).

To estimate morphometricity, we fitted linear mixed models (LMMs, Fig. 2). This approach recognises the correlation structure between many independent variables, and was presented and validated elsewhere, using vertex-wise data (Couvy-Duchesne et al., 2020). Briefly, the LMM fits all cortical measurements as a vector of random effects that is constrained to a normal distribution and a structure of variance-covariance derived from the brain relatedness matrix (BRM, **B**). The BRM quantifies how similar participants are to one another based on cortical measures. We excluded 4 participants with outlying covariance (±8SD from mean brain-relatedness) from the analyses as this indicated oddly similar or dissimilar brains and could bias the LMM results. More detailed model definitions are outlined in the Supplementary Material. All covariates held constant, we used Restricted Maximum Likelihood (REML) implemented in the OSCA software (OmicS-data-based Complex trait Analysis; Zhang et al., 2019) to estimate the variance of the brain random effect (σ_b^2). σ_b^2 quantifies the total trait variance captured by all vertex-wise measurements, while σ_e^2 quantifies the residual variance accounted for by the error term. Morphometricity will be determined as follows (Couvy-Duchesne et al., 2020; Sabuncu et al., 2016):

$$R^2 = \sigma_b^2 / (\sigma_b^2 + \sigma_e^2)$$

As this is a Registered Report, we had introduced the detection of above-zero morphometricity as a "positive control", to test whether the proposed analyses allow for a fair test of the stated hypothesis. All morphometricity estimates presented in the results were significantly larger than zero, which confirms that there was enough systematic variance in the data to detect differences in atlas performance, should they exist. Though not pre-registered, we report in the Results that log-linear models described the relationship between morphometricity and atlas dimensionality well (it became obvious when plotting the data that there was a logarithmic relationship).

Comparison of atlases. To compare and test differences in morphometricity between two atlases (i.e., quantify the morphometricity specific to each atlas or shared between atlases), we extended the previous model by fitting two nested LMMs: the null model fits one variance component for all ROIs within one cortical atlas as random effects (non-brain trait ~ intercept + covariates + atlas1), and the alternative model includes an additional variance component for all ROIs within a second atlas (non-brain trait ~ intercept + covariates + atlas1 + atlas2). Likelihood ratio tests (LRT) allowed examination of whether morphometricity estimates (R^2) significantly differed between two commonly-used atlases (i.e., not the random atlases). The χ^2 distributed LRT statistic contrasts the improvement of fit from the null to the alternative model against the loss in degrees of freedom. To adjust for multiple testing, we considered Bonferroni corrected p-values below .05/ (588) = 8.5×10^{-5} as significant. Refer to Fig. 2 for an illustration. In practice, we performed 588 pair-wise comparisons between commonly-used atlases across all measurement types and traits (7 non-brain traits x 3 measurement types x 28 pairs of atlases). Each comparison contrasted morphometricity within each trait-brain measurement type configuration.

In each LRT, the loss in degrees of freedom equals one, because we model *one* variance component per atlas. This
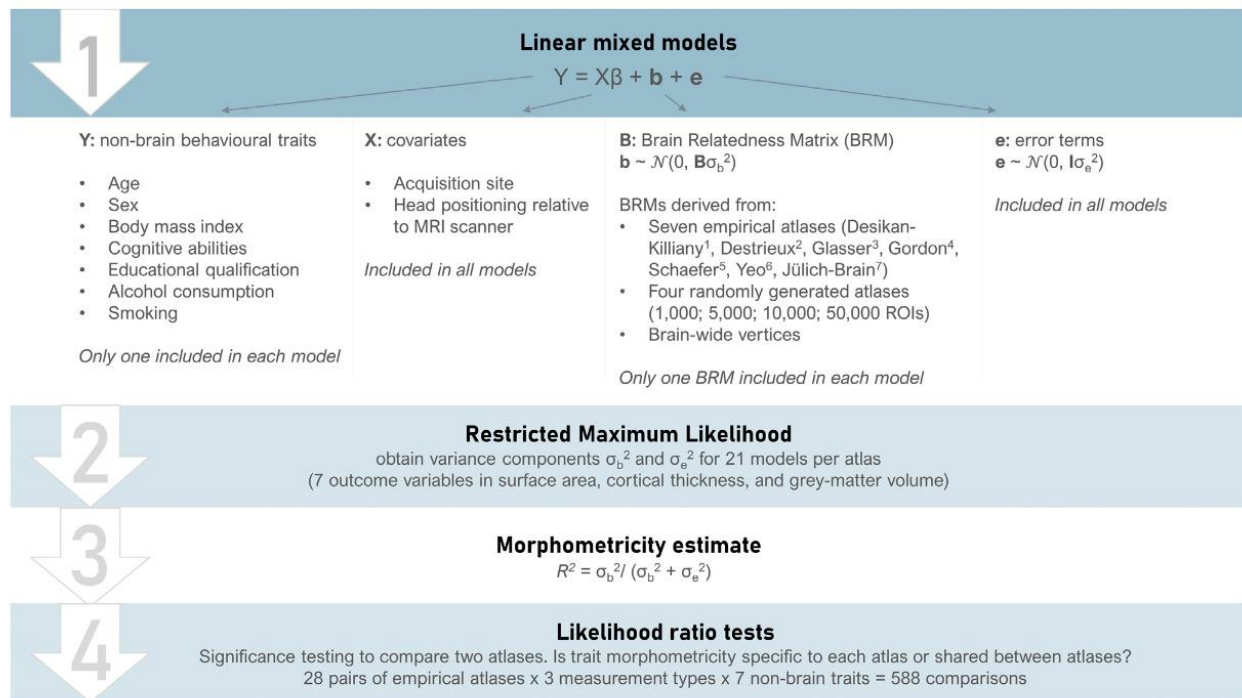
Fig. 2 — Step-by-step illustration of the statistical analyses described in this Registered Report. The linear mixed model equation in step 1 is shown to illustrate the variables considered. This equation will not be solved *per se*, instead we will use Restricted Maximum Likelihood on the basis of model assumptions outlined in the methods to obtain variance components (step 2), that allow calculation of the morphometricity estimate (step 3). Likelihood ratio tests will only compare commonly-used atlases and vertex-wise measures against one another, but not random parcellations. [1]Desikan et al. (2006); [2]Destrieux et al. (2010); [3]Glasser et al. (2016); [4]Gordon et al. (2014); [5]Schaefer et al. (2017); [6]Yeo et al. (2011); [7]Amunts et al. (2020).

means that testing does not depend on the numbers of cortical measurements contained in the different atlases, allowing a fair LRT comparison. To illustrate this lack of estimate inflation with more predictors, consider genome-wide complex trait analysis heritability estimates in which millions of genetic components are fit to a trait as random effects (Yang, Lee, Goddard, & Visscher, 2011). This is a well-established approach using LMMs equivalent to the one used here to estimate morphometricity. It is known that the heritability estimates do not alter when the number of genetic predictors considered in the analysis are increased or reduced, if they remain genome-wide representative. Refer to the supplementary material for pre-registered power calculations.

Best fitting model. In addition to the LRT, we calculated the Bayesian Information Criterion (BIC) (Schwarz, 1978) to compare and rank the models based on their penalised fit. The BIC is a well-established measure of fit that was calculated as the difference between model complexity ($\log(N_{participants})*p$; with p the number of parameters estimated in the model) and model fit ($2*LogLikelihood$), meaning that smaller values indicate better model fit.

Morphometricity of random parcellations. To understand whether commonly-used atlases explained more variance than expected using random atlas boundaries, we estimated the distributions of morphometricity estimates from 100 randomly generated atlases with the same number of ROIs as are included in the commonly-used atlases. For example, for comparison with Desikan-Killiany, we generated atlases with

68 random ROIs (across both hemispheres) and repeated the analysis 100 times, to obtain a distribution of morphometricity under random parcellations. We reported the percentage out of 147 estimates from commonly-used atlases (7 atlases x 7 traits x 3 measurement types) that were larger than 0%, 50%, and 99% of the 100 estimates from random parcellations.

2.5. Sensitivity analyses

Penalised regression. To test the robustness of the main results, we estimated how well grey-matter structure measures predicted non-brain traits by using LASSO models. LASSO is a "least absolute shrinkage and selection operator" (Tibshirani, 1996), that maximises predictive ability and can handle more predictors than observations. We estimated LASSO parameters on a training subset (random 80% sample subset, ~30,400 participants depending on number of total participants surviving data cleaning) with the big_spLinReg function in the bigstatsr package (Privé, Aschard, Ziyatdinov, & Blum, 2018). Using ten-fold cross-validation, the function separates the training set in ten folds, employs a Cross-Model Selection and Averaging procedure and outputs an averaged vector of coefficients.

We evaluated the prediction accuracy of the LASSO models in a validation sample (20%, ~7,600 participants) and report the prediction accuracy (i.e., R^2 between observed and predicted values), and mean absolute errors (i.e. mean absolute difference between observed and predicted measures of a
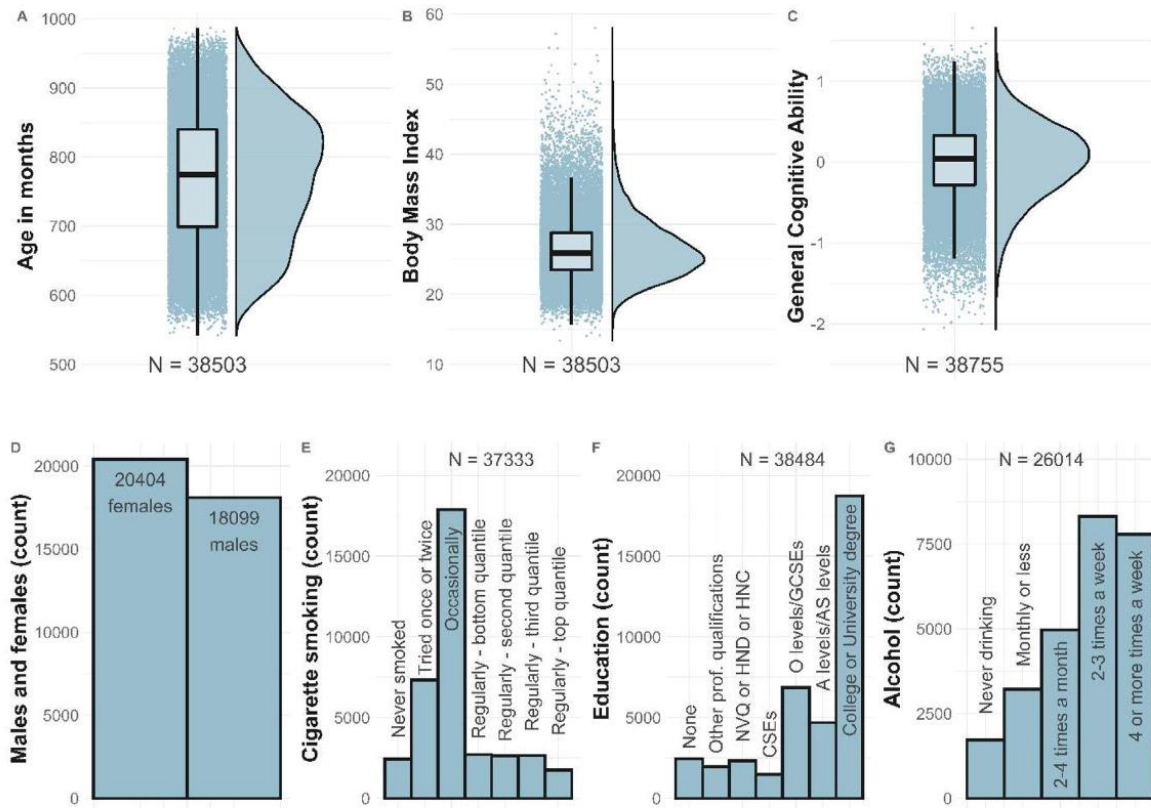
Fig. 3 — Descriptive statistics for each non-brain trait. Note the distribution for the education phenotype which is not representative of the general population but expected in the UKB sample (Fry et al., 2017).

non-brain trait). To statistically compare absolute errors between atlases, we contrasted the absolute errors obtained in models with different atlases using Wilcoxon's signed rank tests. To adjust for multiple testing, we considered Bonferroni corrected p-values below $.05/(588) = 8.5 \times 10^{-5}$ as significant.

3. Results

3.1. Descriptive statistics

Overall, we processed structural neuroimaging data for 42,957 participants (4 exclusions based on ±8SD BRM criterion). Participants who withdrew consent and those without T2-FLAIR measurements were removed from the phenotype data ($n_{removed} = 1,377$). Depending on data availability, differing numbers of participants were included in the morphometricity analysis of the seven non-brain traits. Final descriptive statistics are displayed by non-brain trait in Fig. 3.

3.2. Morphometricity estimates

First, we estimated the total association between each atlas and trait of interest, to assess if more complex atlases tend to retain more information. Morphometricity estimated from the commonly-used atlases (Yeo, Desikan-Killiany, Destrieux, Glasser, Gordon, Jülich-Brain, Schaefer), random parcellations with 1,000, 5,000, 10,000 and 50,000 ROIs, as well as from vertex-wise measurements are displayed in Fig. 4. All traits were morphometric as indicated by significant LRTs examining whether the estimate was larger than zero (STable 1). Estimates differed between non-brain traits, atlases, and brain measurement types, as did the BIC quantifying model performances (SFig. 2). Overall, age and sex yielded the highest morphometricity out of the seven non-brain traits.

As hypothesised, we observed that atlases with more ROIs produced larger morphometricity across all seven traits and all brain measurement types. This was more pronounced for surface area than cortical thickness measures. The largest estimates were always yielded by vertex-wise measures. More generally, we observed substantial increases in morphometricity from high-dimensional compared to low-dimensional atlases (Fig. 4; STable 1). We found a median 3.41-fold increase in morphometricity (range increase = 1.49-fold–15.17-fold) obtained from vertex-wise measures (300,000 ROIs) compared with the Yeo atlas (34 ROIs). For example, we found 25% morphometricity estimated for sex from surface area measures represented by the Yeo atlas, compared with 88% morphometricity by vertex-wise measures (all raw estimates in STable 2).
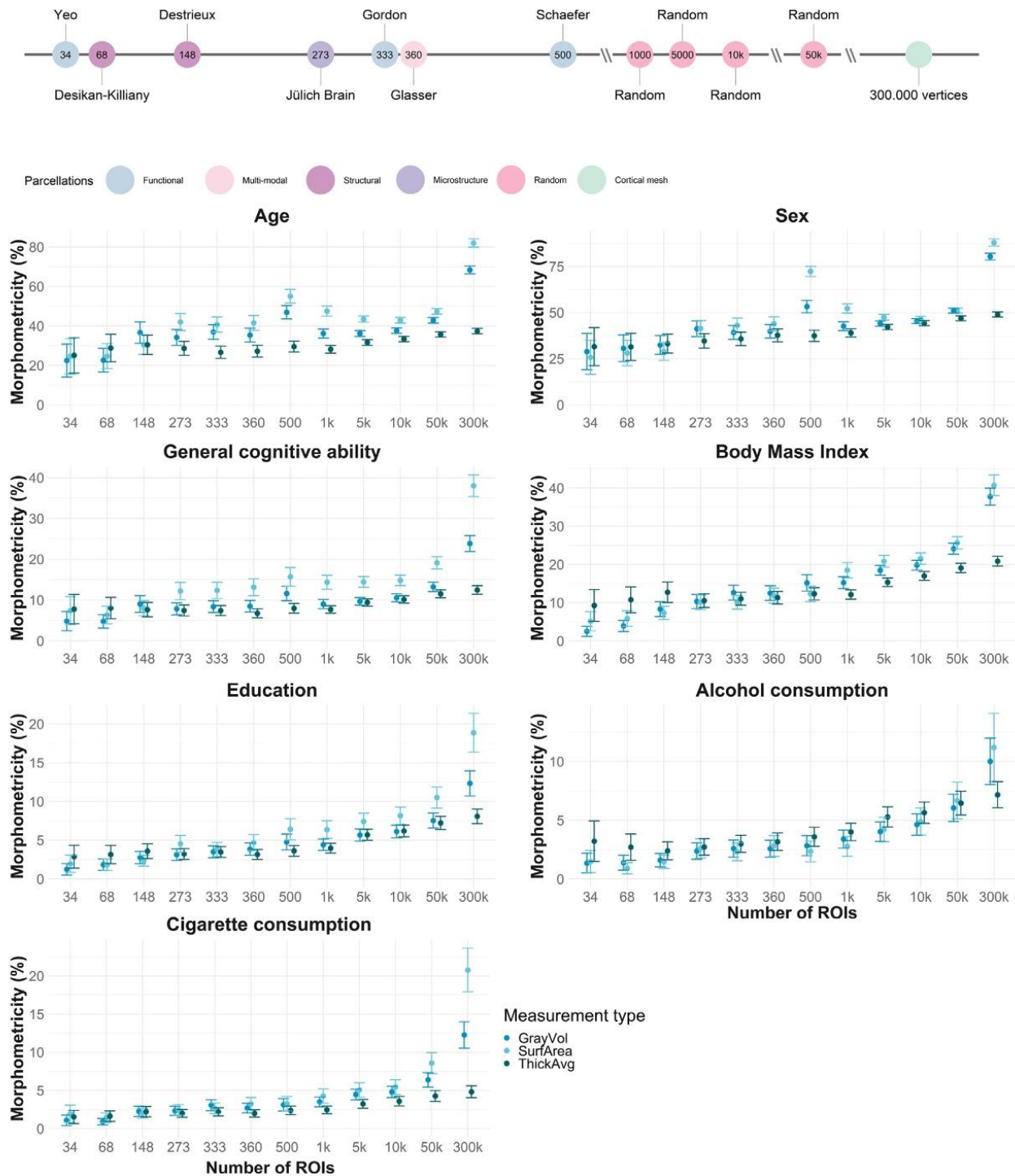
Fig. 4 – Morphometricity for seven behavioural traits estimated using different brain measurement types: GrayVol = cortical gray-matter volume, SurfArea = cortical surface area, ThickAvg = average cortical thickness.

3.3. Relationship between morphometricity and atlas dimensions

We observed strong linear-log relationships between morphometricity estimates and (logarithmically transformed) atlas dimensionality (SFig. 3), with R^2 estimates ranging between 56% and 98% (mean = 84.58%, SD = 11.89%). Table 1 reports intercepts (α) and coefficients (β) characterising this relationship for non-brain traits and brain measurement types separately. Based on information contained in Table 1, the expected morphometricity estimate (in %) may be inferred for each trait, as follows:

$$E(morphometricity) = \beta \times \log(number\ of\ ROIs) + a$$

With this formula, we can infer, for example, that the expected morphometricity for age using 200,000 surface area measurements is around 67% ($4.5 \times \log(200,000) + 12$). Table 1 shows that models of cortical thickness tended to yield smaller coefficients β than models of surface area, suggesting thickness-based morphometricity is less influenced by atlas dimensionality than surface area-based analyses.

3.4. Outlying estimates by Schaefer atlas

Morphometricity from the Schaefer atlas measures were unexpectedly large, they yielded worse model fit compared with other atlases (SFig. 2) and broke the continuity of increased estimates with larger atlas dimensionality (mainly associations with age and sex). According to our formula above, we expected the following morphometricity for an atlas with 500 ROIs: $age_{surface\ area} = 39.97\%$, $age_{volume} = 34.75\%$, $sex_{surface\ area} = 43.45\%$, $sex_{volume} = 40.73\%$. Instead, we found substantially larger estimates: $age_{surface\ area} = 55.01\%$, $age_{volume} = 46.90\%$, $sex_{surface\ area} = 72.30\%$, $sex_{volume} = 53.24\%$. Suspecting the Schaefer atlas measures may not be normally distributed, we re-estimated morphometricity using normalised Schaefer atlas data (Rank-based Inverse Normal Transformation) and obtained estimates as predicted by the formula: $age_{surface\ area} = 38.99\%$ [35−42%], $age_{volume} = 36.46\%$ [33−40%], $sex_{surface\ area} = 48.32\%$ [45−52%], $sex_{volume} = 44.46\%$ [41−48%]. Normalisation descriptively reduced morphometricity for all other traits too, but confidence intervals from adjusted and unadjusted estimates overlapped.

3.5. Atlases with random region boundaries

To understand whether commonly-used atlases explained more variance than expected using random atlas boundaries, we estimated the distributions of morphometricity estimates from 100 randomly generated atlases with the same number of ROIs as included in the commonly-used atlases. The distributions of estimates from random atlases are visualised in Fig. 5 alongside morphometricity estimated from the respective commonly-used atlas (red colours). In most cases, confidence intervals around the point morphometricity estimates from commonly-used atlases mapped well onto the spread in estimates from atlases with random parcellation (Fig. 5), suggesting that commonly-used atlases yielded estimates as expected (on average) by random parcellations. 69% of point estimates from commonly-used atlases (total of 147 estimates across seven atlases, seven traits and three brain measurement types) yielded morphometricity smaller than at least 50% of the point estimates from random atlases. 5% of commonly-used atlas estimates were larger than 99% of the null estimates, and 13% of the commonly-used estimates were smaller than any of the null estimates (mostly involving surface area; STable 4.1). The latter indicates that in some cases less variance is explained when using one of the established atlases (Desikan-Killiany, Destrieux, Gordon, Glasser, Schaefer) compared to using random parcellations when mapping surface area measures onto behavioural traits.

3.6. Atlas comparisons

We performed cross-atlas LRTs to quantify whether a more fine-grained atlas added explanatory variance in addition to

Table 1 − Linear regression results describing the relationship between morphometricity and logarithmic transformed atlas dimensionality.

Non-brain trait	Brain measurement type	a	β	β Standard error	p-value	R^2 (%)
Age	SurfArea	12	4.5	1	.0013	66
	GrayVol	13	3.5	.8	.0014	65
	ThickAvg	21	1.2	.17	2.8e-05	84
Alcohol consumption	SurfArea	−3.4	1	.1	2.4e-06	90
	GrayVol	−2.5	.86	.081	8.8e-07	92
	ThickAvg	.14	.56	.054	1.1e-06	92
Body Mass Index	SurfArea	−9.9	3.7	.27	8.5e-08	95
	GrayVol	−8.7	3.4	.26	1.3e-07	94
	ThickAvg	4.3	1.3	.11	2.3e-07	94
Cigarette consumption	SurfArea	−6.8	1.7	.3	.00022	76
	GrayVol	−3.4	1	.12	7.7e-06	88
	ThickAvg	.019	.38	.018	1.3e-09	98
Education	SurfArea	−5.1	1.6	.18	2.9e-06	90
	GrayVol	−2.6	1.1	.09	3.6e-07	93
	ThickAvg	.028	.64	.048	1e-07	95
General cognitive ability	SurfArea	−3.9	2.6	.46	.00019	77
	GrayVol	−1.1	1.6	.28	.00019	77
	ThickAvg	4.5	.6	.085	3.4e-05	83
Sex	SurfArea	13	4.9	1.4	.0054	56
	GrayVol	14	4.3	.8	.00032	74
	ThickAvg	24	2.1	.11	3.9e-09	97

GrayVol = cortical gray-matter volume, SurfArea = cortical surface area, ThickAvg = average cortical thickness. We observed linear-log relationships between morphometricity and atlas dimensionality which we describe through a linear regression analysis of logarithmically transformed atlas dimensionality. a is the intercept and β is the regression coefficient. Expected morphometricity estimates may be calculated as follows: $E(morphometricity) = \beta \times \log(number\ of\ ROIs) + a$

the variance already explained by a coarser atlas. For example, first, we modelled morphometricity using Desikan-Killiany (68 ROIs) and Destrieux (148 ROIs) brain measures in one model. Second, we dropped Destrieux from the model and recalculated morphometricity for Desikan-Killiany alone, which allowed quantification of the percentage of variance explained by Destrieux in addition to the variance already explained by Desikan-Killiany. The LRT quantifies whether the explanatory variance added by the higher dimensional atlas (i.e., Destrieux in this example) is larger than zero. To summarise the results, we display an index of the relative improvement made to the model by adding the higher dimensional atlas, which is calculated as the sum of the variance explained by both atlases together, divided by the morphometricity of the lower dimensional atlas alone (Fig. 6, SFig. 5). A ratio of 1 indicates no increase in morphometricity, and a ratio larger than 1 indicates the proportional increase gained by adding the higher dimensional atlas. None of the indices are below 1 because the combined morphometricity from two atlases cannot be smaller than the estimate yielded by one atlas alone.

A total of 576 LRTs (out of 588) were significant, and we found an average ratio of 2.73 for surface area, 2.70 for cortical volume, and 1.67 for cortical thickness (STable 5), indicating that the morphometricity roughly doubled when adding a second atlas to the equation. Vertex-wise measures added considerable explanatory variance in addition to all other atlases, but the greatest model improvements were gained when adding vertex-wise measures to low-dimensional atlases. Across the board, commonly-used atlases accounted for unique trait variance rather than variance shared between atlases suggesting they captured atlas-specific trait variance.

Surprisingly, some low-dimensional pairs of atlases surpassed vertex-wise estimates, for example, average thickness estimates for age were 29% [21—35%] for Desikan-Killiany alone, and 30% [25—36%] for Destrieux alone, but together Desikan-Killiany and Destrieux accounted for 63% [58—68%] which was substantially larger than the vertex-wise estimates of 37% [36—39%]. We re-calculated morphometricity and joint-atlas effects using simple linear regression (LM) (which was not pre-registered) to test whether these LMM estimates were biased, potentially due to the violation of LMM assumptions.

The rationale was that LMMs enforce a normal distribution on the ROI effects, which may be problematic for low-dimensional atlases with individual ROIs dominating and driving brain-wide effects. LM does not impose assumptions on the distribution of effects which we suggest allows to test for a violation of LMM assumptions (though LM will more likely produce inflated estimates, the more predictors are included). LM estimates are displayed in SFig. 7, confirming that LMMs are likely overestimating joint atlas effects by low-dimensional atlases. Overall, LMs validate most of the LMM results, but are likely overestimating effects by high-dimensional atlases.

Only twelve tests were not statistically significant ($p > .05/588$), all involving brain measures of surface area and alcohol and cigarette consumption. In these cases, adding the higher dimensional atlas to the equation did not significantly increase morphometricity. The low morphometricity estimates of the traits (<4%), probably resulted in a lower statistical power, which may, in part, explain these results.

3.7.   Sensitivity analysis: LASSO-based prediction

We trained LASSO models to investigate whether primary association results aligned with results from a machine learning approach. In other words, we sought to investigate if a larger morphometricity also led to a larger prediction accuracy. Indeed, we found that prediction accuracy improved with atlas dimensionality (and morphometricity), with the vertex-wise representation yielding the best prediction accuracy (i.e., R^2 between observed and LASSO-predicted values). Like morphometricity, prediction accuracy was substantial for most traits, but low for education, alcohol, and cigarette consumption (SFig. 6). We observed a gain in prediction accuracy by 1.95-fold (range 1.12-fold—9.46-fold) between the coarsest (34 ROIs, Yeo) and the most fine-grained cortical representations (vertices; prediction accuracies in STable 6). For example, predicted values of cognitive ability by volumetric Yeo measures explained 7% of the variance in observed values, and vertex-wise predictions explained 13% of observed cognitive ability values.

Wilcoxon Signed-Rank tests were used to test for statistical differences in absolute errors between two atlases. Out of 588 comparisons, only 3 reached a significance level of .05/588 (STable 6). Effects sizes were overall very small suggesting that we have no evidence to conclude that one atlas yielded smaller median prediction errors than another atlas. The three significant improvements in prediction were found between the Destrieux and Glasser measurements of ThickAvg in sex ($p = 1.62 \times 10^{-5}$; $r = .049$; $N_{evaluation} = 7701$), Destrieux and Schaefer measurements of ThickAvg in sex ($p = 6.51 \times 10^{-7}$; $r = .057$; $N_{evaluation} = 7701$), and Desikan-Killiany and Yeo measures of GrayVol in alcohol consumption ($p = 1.53 \times 10^{-5}$; $r = .056$; $N_{evaluation} = 5203$).

4.   Discussion

In this Registered Report, we provide a quantitative comparison of the information retained by commonly-used cortical atlases of varying dimensionality (i.e., number of considered ROIs) through reporting morphometricity estimates across seven behavioural traits. We calculated whole-brain morphometricity using linear mixed models (LMMs) and compared percentages of variance accounted for by cortical atlases using likelihood ratio tests (LRTs). As hypothesised, we found that using more fine-grained atlases to describe the cortical grey-matter structure resulted in larger morphometricity. This is consistent with our pre-registered hypothesis that lower dimensional cortical representations (i.e., fewer ROIs) tend to mask inter-individual variance, and that this variance can be retained when representing the brain using ~300,000 vertex-wise measures. We found, across all traits and types of measurements, that morphometricity increased linearly as a function of the log-atlas dimension. There was no evidence for a "tipping point" of atlas dimensionality beyond which morphometricity ceased to improve, nor was there evidence for trait-dependent optimal levels of atlas dimensionality.

Our findings should give studies of cortical structure reason to favour finer-grained cortical representations over
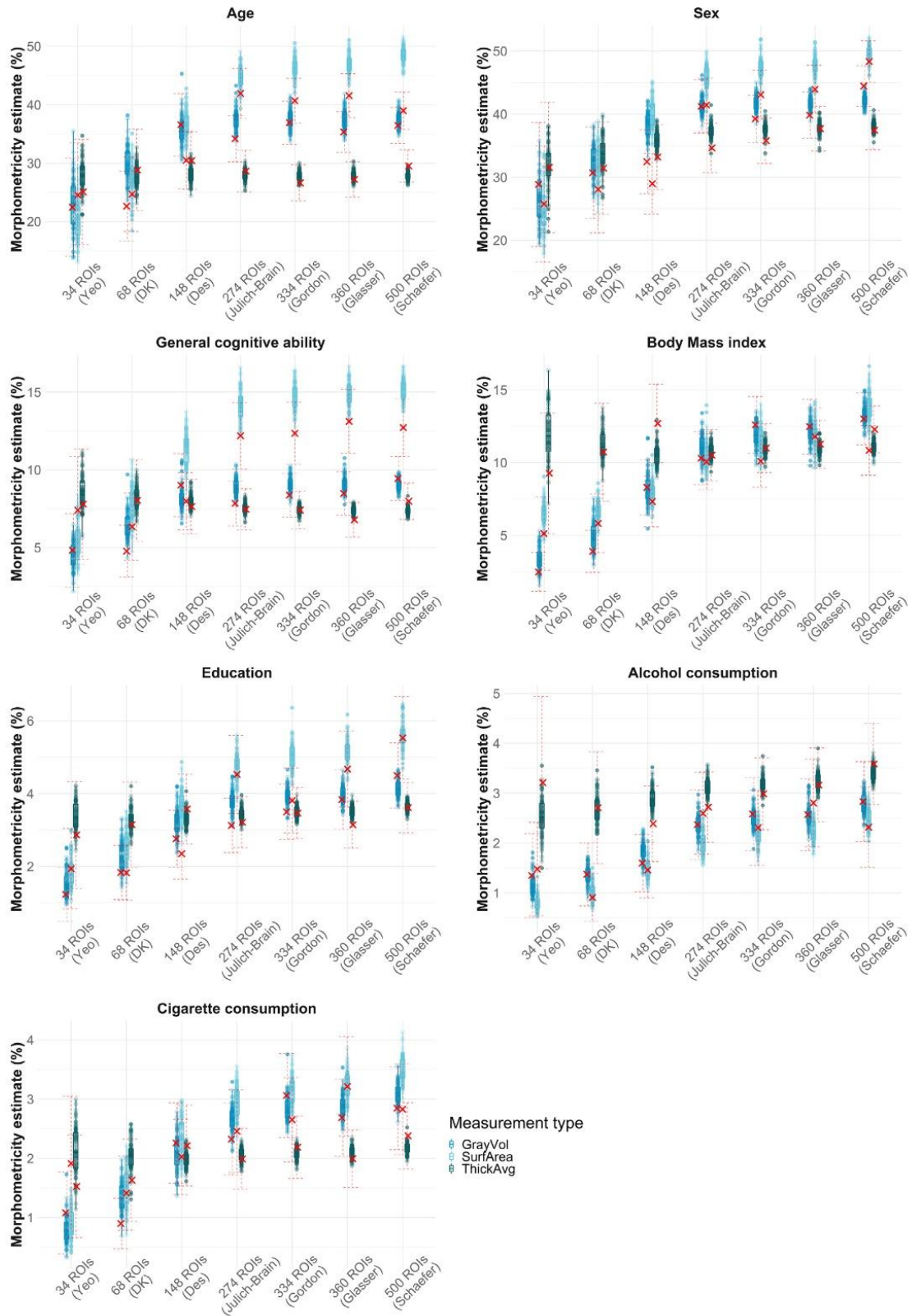
Fig. 5 — Distributions of morphometricity estimates from random parcellations. Each distribution was generated based on morphometricity estimates of random parcellations with matched numbers of ROIs. Red crosses indicate the corresponding estimate yielded by the commonly-used atlas; red lines indicate 95% confidence intervals around the point estimate. Random parcellations of 34 ROIs were compared to Yeo, 68 ROIs were compared to Desikan-Killiany, 148 ROIs were

coarser ones, which should enable better brain-behaviour mapping and prediction accuracy of behaviours from cortical measurements. We suggest that available neuro-imaging samples (such as the UKB) and computational resources are now large enough to accurately estimate associations between vertices and behavioural variables, promising future studies that more systematically account for brain-wide associations than can be obtained from ROI-behaviour associations (Smith & Nichols, 2018).

Should the use of lower-dimensional atlases be necessary in a study mapping interindividual differences in structural cortical measures onto behavioural traits, we would advise researchers, based on the findings presented here, to repeat analyses across multiple commonly-used atlases, or to iterate over multiple atlases with random region boundaries. This may complicate the interpretation of findings but promises more replicable brain-behaviour associations. While statistical power of these multidimensional approaches is limited by multiple testing correction, techniques to overcome this limitation have been derived. For example, one may group vertex-wise measures based on atlas ROIs and fit each set of vertices as a random effect, which permits to perform a single association test per ROI, while still modelling the fine-grained cortical structure (Couvy-Duchesne et al., 2020).

4.1. Greater atlas dimensionality yields substantially greater morphometricity estimates

Differences in morphometricity between the coarsest (Yeo, 34 ROIs) and the most fine-grained atlas (vertices, 300,000 ROIs) were considerable (median increase of 3.41-fold). One previous functional MRI study reported a doubling in prediction accuracy of cognitive ability between estimates represented by the Glasser atlas (360 ROIs) compared with vertex-wise representations (Feilong, Guntupalli, & Haxby, 2021). Though using structural neuroimaging data and an association framework, our results were comparable, showing 1.83-fold (thickness), 2.81-fold (volume) and 2.90-fold (surface area) improvements in morphometricity when using vertex-wise measurements relative to the Glasser atlas. Findings reported here are consistent with previous studies and together this indicates that selecting fine-grained cortical atlases substantially improves variance overlap between brain and behaviour. Future studies are needed to understand whether this applies to functional neuroimaging and diffusion tractography.

4.2. Commonly-used atlases yield the same morphometricity as expected from random parcellations

To assess whether the commonly-used atlases included here were superior to random parcellations, we generated 100 atlases with random ROI boundaries and, for better comparability, the same number of ROIs as in commonly-used atlases. Random parcellations were mapped onto participants' brain images and used to estimate morphometricity.

We observed that the morphometricity from commonly-used atlases was on par with the average morphometricity from random parcellations, which suggests commonly-used atlases do not maximise their potential in capturing specific trait variance. That random atlases yielded similar estimates as commonly-used atlases is in line with a recent study demonstrating that using random parcellations in predicting structure–function correlations resulted in similar power (Revell et al., 2022).

Surface area-based associations involving Desikan-Killiany, Destrieux et al. (2010), Gordon et al. (2014), Glasser et al. (2016) and Schaefer et al. (2017) produced estimates below any of the 100 random parcellations. This suggests for surface area-based associations that most random atlases would outperform the commonly-used ones. We suggest future studies should consider iterating over random parcellations to establish more robust association results. Note that this is referring to Schaefer morphometricity estimates that were corrected for non-normal distributions, and future studies should account for non-normal distributions when using Schaefer in FreeSurfer processing.

4.3. Morphometricity differs between surface area, cortical thickness, and volume

Morphometricity estimated from surface area measures were most affected by atlas dimensionality (largest linear-log regression coefficients; Table 1), which suggests that choosing fine-grained atlases will benefit studies of surface area most, compared with thickness measures, for example. The observation that atlas dimensionality has differential impact on estimates from different measurement types may fit with evidence that surface area and thickness measures have low phenotypic and genetic correlations (Panizzon et al., 2009; Wierenga, Langen, Oranje, & Durston, 2014; Winkler et al., 2010). Here, we did not specifically test for this, and future work is needed to understand whether different brain measurement types represent common or unique information. Morphometricity analyses in Couvy-Duchesne et al. (2020) suggested that surface area and thickness measures capture some common trait variance, but that they are mostly unique.

4.4. Low-dimensional atlases capture unique rather than common trait variance

We used LRTs to examine whether one of a pair of atlases outperformed the other in its overlap with behavioural traits. We found that most comparisons between atlases were significant, and that the variance accounted for was roughly doubled when jointly considering two low-dimensional atlases. This considerable improvement indicates that atlases explain unique, rather than shared trait variance. By selecting one coarse atlas to parcellate participants' brain images, researchers likely restrict their analyses to a specific dimension of variance that only partly overlaps with behavioural traits.

---

compared to Destrieux, 274 ROIs to Jülich-Brain, 334 to Gordon, 360 to Glasser, and 500 to Schaefer. Estimates for Schaefer were calculated based on Rank-based Inverse Normal Transformed data and are therefore not the same estimates as displayed in Fig. 4. GrayVol = gray matter volume, SurfArea = surface area, ThickAvg = average thickness.
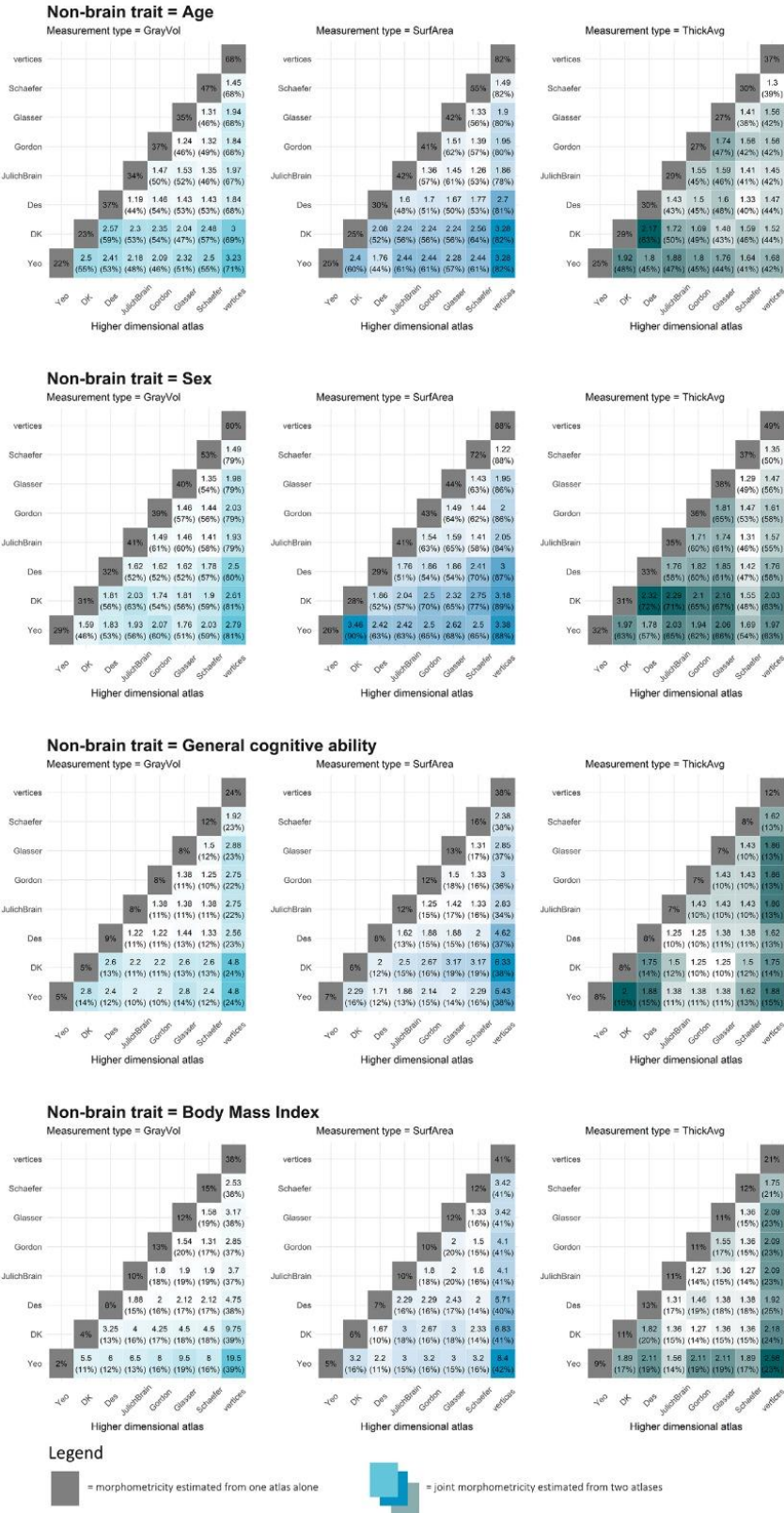
Fig. 6 − Atlas comparisons calculated from likelihood ratio tests for age, sex, cognitive ability, and body mass index. Find the equivalent visualisation for education, cigarette, and alcohol consumption in SFig. 5. Percentages displayed on the diagonal are morphometricity estimates for individual atlases, as indicated in Fig. 4. Indices on the off-diagonal show the relative improvement made to the model by adding the higher dimensional atlas, which we calculated as the sum of the

Alongside small sample sizes and large sampling variation (Genon, Eickhoff, & Kharabian, 2022; Marek et al., 2022), this is likely a contributor to highly heterogenous reports of brain-behaviour correlates and the lack of reproducibility across ROI-based studies. We demonstrate that two atlases can capture non-overlapping trait variance, which implies that results between two ROI-based studies cannot be translated if different atlases were used. Future studies by researchers who still wish to use low-dimensional atlases may yield more replicable results when repeating their analyses across multiple atlases, which may recover variance hidden by one, but not another low-dimensional atlas. This should yield more robust results but may complicate their interpretation. Future studies may parse the likely sources of brain-based signal using atlases derived from different principles (macro-anatomy, resting-state networks etc.).

4.5. LASSO-based predictions confirm morphometricity estimates

The general trend that atlas dimensionality was positively associated with increasing morphometricity was confirmed by LASSO-based predictions. This is in line with previous demonstrations that association and prediction results are related (Couvy-Duchesne et al., 2020). If unbiased and accurate, morphometricity estimated from LMMs represents the total linear association, that, in theory, LASSO-based predictions may reach if the algorithm was trained in an infinite sample allowing accurate estimates of predictor weights. Here, LASSO prediction accuracy is indeed smaller than LMMs in almost all cases. We observed few exceptions where cortical thickness-based predictions using LASSO surpassed LMM-based morphometricity (e.g., vertex-wise thickness-age: LASSO out-of-sample prediction = 59%; LMM association = 37%). This may suggest that the cortical thickness LMM estimates are downward biased, maybe due to violation of a normal distribution in effect sizes for traits like age and sex. It is known, for example, that different areas of the brain age unevenly across the life-span (Raz, Ghisletta, Rodrigue, Kennedy, & Lindenberger, 2010), which may introduce a skewed distribution of effect sizes.

4.6. Parallels between multi-dimensional neuroimaging and genetic analyses

The notion that morphometricity marks the maximum trait variance that surface-based structural brain measures could account for is analogous to narrow-range heritability. Both morphometricity and heritability have the same statistical definition. In parallel with the 'First Law' in behavioural genetics stating that all traits are heritable (Turkheimer, 2000), we found that all traits considered here were morphometric. Note that not all traits were morphometric in Couvy-Duchesne et al. (2020), but they used a considerably smaller sample (N~9,000). Using statistical genetics techniques to analyse neuroimaging data promises exciting avenues for future research that could generate innovative insights using statistical techniques that have already been derived and thoroughly tested using genetic data (Couvy-Duchesne et al., 2022).

We found exceptions in the analogy with heritability, where the morphometricity estimated by vertex-wise measures was surpassed by two joint low-dimensional atlases (prediction of age and sex by cortical thickness). Post-hoc analyses suggested that model assumptions imposed by LMMs were violated by low-dimensional atlases including individual ROIs with large effects. We suggest that the parallel between heritability and morphometricity is most consistent when considering vertex-wise data, as opposed to coarse cortical atlases, however, results must be interpreted in their specific genetic or neuroimaging context (e.g., genetic markers are stable across the lifespan, while brain structure changes in response to environments and behaviours).

4.7. Limitations

We emphasise that LMMs employed here fit all predictor variables as a single random effect, no matter how many predictor variables are included by an atlas. This makes the variance accounted for comparable between atlases, and it does not automatically increase with the number of predictors, as it would, for example, in a simple regression model. However, reported estimates are likely confounded (for example by disease or environmental influences), but this should be constant across atlases and should not impede a fair comparison between atlases. Some LMMs of joint morphometricity by low-dimensional atlases are likely overestimated due to individual ROIs driving brain-wide effects, but we demonstrate that we can test LMM assumptions using linear regression in this low-dimensional space. Furthermore, some categorical behavioural variables used in this study may be suboptimal. Cigarette and alcohol consumption categories were based on arbitrary cut-offs, and the education phenotype was dominated by the category capturing university education. Low levels of trait morphometricity may reflect these caveats and may require re-evaluation in other samples.

It is unknown whether our results are directly translatable to analyses using different software and pipelines. For reproducibility, we make all our analysis code available online (https://annafurtjes.github.io/Comparing_atlases/). Study results would likely differ across FreeSurfer versions (Gronenschild et al., 2012; we used v6.0.0), and between processing software, especially compared with other approaches than surface-based analyses (e.g., volumetric analyses in Statistical Parametric Mapping, SPM, or FMRIB Software Library, FSL). Future studies are needed to test whether results would replicate in functional MRI data, and

---

variance explained by two atlases together, divided by the morphometricity estimate of the lower dimensional atlas alone (joint$_{morphometricity}$/individual$_{morphometricity}$). **Squares are coloured according to this ratio, i.e., larger ratio, darker colour. The raw sum of variance explained is printed in brackets below the respective index. Non-significant results are marked with n.s.**

whether more elaborate means of summarising vertex-wise measures across ROIs could provide better cortical representation (e.g., taking the maximum thickness within an ROI).

5.   Conclusions

This Registered Report demonstrates the importance of appropriate cortical atlas choices across neuroimaging studies by demonstrating that using commonly-used cortical atlases resulted in considerable loss of information compared to vertex-wise representations of cortical structure. These atlases included Yeo et al. (2011), Desikan-Killiany (Desikan et al., 2006), Destrieux et al. (2010), Jülich-Brain (Amunts et al., 2020), Gordon et al. (2014), Glasser et al. (2016), and Schaefer et al. (2017), which all accounted for magnitudes of variance on par with random parcellations across seven behavioural traits using cortical thickness, volume, and surface area measures. In the interest of more replicable results, our findings should give researchers reason to leverage large-scale samples (like the UKB) to conduct association and prediction analyses using vertex-wise brain data which promises more systematic and localised accounts of the relationships between behaviours and structural measures of the cortex.

We further demonstrate that studies using one coarse atlas only (for example, Yeo, Desikan-Killiany, or Destrieux) tend to capture atlas-specific trait variance, implying that study results often cannot be translated between atlases. Hence, atlas choice is likely a contributor to the lack in reproducibility in the neuroimaging literature. We suggest that studies for which the use of coarse atlases is necessary, should either repeat analyses across multiple commonly-used atlases, or iterate over random atlases to produce more robust results.

5.1.   Statement of transparency for secondary analyses

AEF is the lead analysist on this project and has been granted access to the UKB data through UKB application 40933, of which JHC is the principal investigator. AEF previously worked with pre-processed Desikan-Killiany IDPs through another UKB application, but has not seen or worked with the bulk MRI data used in this project, or any of the behavioural traits. She therefore is naïve to any potential associations between variables considered in this study. The UKB application used here is unrelated to the one used in Couvy-Duchesne et al. (2020).

An application to download the required data was submitted to UKB on the 28th of March 2021. It has been approved on the 11th of May, neuroimaging data download occurred successively between November 2021 and January 2022 (our pipeline downloaded individual-by-individual, processed, and deleted data to keep cluster storage on King's College London's high performance computer *Rosalind* at a minimum). AEF had not downloaded or investigated any of the data until in-principle acceptance had been granted in November 2021. The analysis plan pre-registered as a Registered Report (https://osf.io/dkw9t), and analysis code was made available on GitHub (https://annafurtjes.github.io/Comparing_atlases/).

Credit author statement

Anna E. Fürtjes: Conceptualization, Formal analysis, Investigation, Project administration, Software, Visualization, Writing - original draft.

James H. Cole: Conceptualization, Resources, Supervision.

Baptiste Couvy-Duchesne: Conceptualization, Methodology, Supervision, Software.

Stuart J. Ritchie: Conceptualization, Resources, Supervision.

The final review and editing will be done by all authors.

Open practices

The study in this article earned Open Materials - Protected Access and Preregistered badges for transparent practices. Materials and data for the study are available at https://osf.io/dkw9t.

Declaration of competing interest

None.

Data availability

The authors do not have permission to share data.

Acknowledgements

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cortex.2022.11.001.

REFERENCES

Alexander-Bloch, A. F., Shou, H., Liu, S., Satterthwaite, T. D., Glahn, D. C., Shinohara, R. T., … Raznahan, A. (2018). On testing for spatial correspondence between maps of human brain structure and function. *Neuroimage, 178*, 540—551. https://doi.org/10.1016/j.neuroimage.2018.05.070

Alfaro-Almagro, F., Jenkinson, M., Bangerter, N. K., Andersson, J. L. R., Griffanti, L., Douaud, G., … Smith, S. M. (2018). Image processing and quality control for the first 10,000 brain imaging datasets from Uk Biobank. *Neuroimage, 166*, 400—424. https://doi.org/10.1016/j.neuroimage.2017.10.034

Amunts, K., Mohlberg, H., Bludau, S., & Zilles, K. (2020). Julich-brain: A 3d probabilistic atlas of the human brain's cytoarchitecture. *Science, 369*(6506), 988—992. https://doi.org/10.1126/science.abb4588

Bohland, J. W., Bokil, H., Allen, C. B., & Mitra, P. P. (2009). The brain atlas concordance problem: Quantitative comparison of anatomical parcellations. *Plos One, 4*(9), Article e7200. https://doi.org/10.1371/journal.pone.0007200

Cole, J. H., Ritchie, S. J., Bastin, M. E., Valdés Hernández, M. C., Muñoz Maniega, S., Royle, N., … Deary, I. J. (2018). Brain age predicts mortality. *Molecular Psychiatry, 23*(5), 1385—1392. https://doi.org/10.1038/mp.2017.62

Couvy-Duchesne, B., Strike, L. T., Zhang, F., Holtz, Y., Zheng, Z., Kemper, K. E., … Visscher, P. M. (2020). A unified framework for association and prediction from vertex-wise grey-matter structure. *Human Brain Mapping, 41*(14), 4062—4076. https://doi.org/10.1002/hbm.25109

Couvy-Duchesne, B., Zhang, F., Kemper, K., Sidorenko, J., Wray, N., Visscher, P., … Yang, J. (2022). Parsimonious model for mass-univariate vertexwise analysis. *The Journal of Medical Investigation: JMI, 9*(5), Article 052404. Retrieved from 10.1117/1.JMI.9.5.052404.

Cox, S. R., Ritchie, S. J., Fawns-Ritchie, C., Tucker-Drob, E. M., & Deary, I. J. (2019). Structural brain imaging correlates of general intelligence in Uk Biobank. *Intelligence, 76*, Article 101376. https://doi.org/10.1016/j.intell.2019.101376

Dadi, K., Varoquaux, G., Machlouzarides-Shalit, A., Gorgolewski, K. J., Wassermann, D., Thirion, B., et al. (2020). Fine-grain atlases of functional modes for fmri analysis. *Neuroimage, 221*, Article 117126. https://doi.org/10.1016/j.neuroimage.2020.117126

Deary, I. J., Penke, L., & Johnson, W. (2010). The neuroscience of human intelligence differences. *Nature Reviews Neuroscience, 11*(3), 201—211. https://doi.org/10.1038/nrn2793

de la Fuente, J., Davies, G., Grotzinger, A. D., Tucker-Drob, E. M., & Deary, I. J. (2021). A general dimension of genetic sharing across diverse cognitive traits inferred from molecular data. *Nature Human Behaviour, 5*(1), 49—58. https://doi.org/10.1038/s41562-020-00936-2

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., … Hyman, B. T. (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. https://doi.org/10.1016/j.neuroimage.2006.01.021

Destrieux, C., Fischl, B., Dale, A., & Halgren, E. (2010). Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *Neuroimage, 53*(1), 1—15. https://doi.org/10.1016/j.neuroimage.2010.06.010

Fawns-Ritchie, C., & Deary, I. J. (2020). Reliability and validity of the Uk Biobank cognitive tests. *Plos One, 15*(4), Article e0231627. https://doi.org/10.1371/journal.pone.0231627

Feilong, M., Guntupalli, J. S., & Haxby, J. V. (2021). The neural basis of intelligence in fine-grained cortical topographies. *eLife, 10*, Article e64058. https://doi.org/10.7554/eLife.64058

Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., … Dale, A. M. (2002). Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. *Neuron, 33*(3), 341—355. https://doi.org/10.1016/S0896-6273(02)00569-X

Fotenos, A. F., Snyder, A. Z., Girton, L. E., Morris, J. C., & Buckner, R. L. (2005). Normative estimates of cross-sectional and longitudinal brain volume decline in aging and ad. *Neurology, 64*(6), 1032—1039. https://doi.org/10.1212/01.Wnl.0000154530.72969.11

Fry, A., Littlejohns, T. J., Sudlow, C., Doherty, N., Adamska, L., Sprosen, T., … Allen, N. E. (2017). Comparison of sociodemographic and health-related characteristics of Uk Biobank participants with those of the general population. *American Journal of Epidemiology, 186*(9), 1026—1034. https://doi.org/10.1093/aje/kwx246

Genon, S., Eickhoff, S. B., & Kharabian, S. (2022). Linking interindividual variability in brain structure to behaviour. *Nature Reviews Neuroscience, 23*(5), 307—318. https://doi.org/10.1038/s41583-022-00584-7

Glasser, M. F., Coalson, T. S., Robinson, E. C., Hacker, C. D., Harwell, J., Yacoub, E., … Van Essen, D. C. (2016). A multi-modal parcellation of human cerebral cortex. *Nature, 536*(7615), 171—178. https://doi.org/10.1038/nature18933

Gong, W., Beckmann, C. F., & Smith, S. M. (2021). Phenotype discovery from population brain imaging. *Medical Image Analysis, 71*, Article 102050. https://doi.org/10.1016/j.media.2021.102050

Gordon, E. M., Laumann, T. O., Adeyemo, B., Huckins, J. F., Kelley, W. M., & Petersen, S. E. (2014). Generation and evaluation of a cortical area parcellation from resting-state correlations. *Cerebral Cortex, 26*(1), 288—303. https://doi.org/10.1093/cercor/bhu239

Gronenschild, E. H. B. M., Habets, P., Jacobs, H. I. L., Mengelers, R., Rozendaal, N., van Os, J., et al. (2012). The effects of freesurfer version, workstation type, and macintosh operating system version on anatomical volume and cortical thickness measurements. *Plos One, 7*(6), Article e38234. https://doi.org/10.1371/journal.pone.0038234

Gu, Z., Gu, L., Eils, R., Schlesner, M., & Brors, B. (2014). Circlize implements and enhances circular visualization in R. *Bioinformatics, 30*(19), 2811—2812. https://doi.org/10.1093/bioinformatics/btu393

Hu, L.-t., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological methods, 3*(4), 424. https://doi.org/10.1037/1082-989X.3.4.424

Kharabian Masouleh, S., Eickhoff, S. B., Hoffstaedter, F., Genon, S., & Alzheimer's Disease Neuroimaging Initiative. (2019). Empirical examination of the replicability of associations between brain structure and psychological variables. *eLife, 8*, Article e43464. https://doi.org/10.7554/eLife.43464

Lindroth, H., Nair, V. A., Stanfield, C., Casey, C., Mohanty, R., Wayer, D., … Sanders, R. D. (2019). Examining the identification of age-related atrophy between T1 and T1 + T2-flair cortical thickness measurements. *Scientific Reports, 9*(1), Article 11288. https://doi.org/10.1038/s41598-019-47294-2

Littlejohns, T. J., Holliday, J., Gibson, L. M., Garratt, S., Oesingmann, N., Alfaro-Almagro, F., … Allen, N. E. (2020). The Uk Biobank imaging enhancement of 100,000 participants: Rationale, data collection, management and future directions. *Nature Communications, 11*(1), 2624. https://doi.org/10.1038/s41467-020-15948-9

Mangin, J. F., Rivière, D., & Amunts, K. (2021). *Surface projections of julich-brain cytoarchitectonic maps*. EBRAINS. https://doi.org/10.25493/NZGY-6AS (V2.9) [Data Set].

Marek, S., Tervo-Clemmens, B., Calabro, F. J., Montez, D. F., Kay, B. P., Hatoum, A. S., … Dosenbach, N. U. F. (2022).

Reproducible brain-wide association studies require thousands of individuals. *Nature, 603*(7902), 654—660. https://doi.org/10.1038/s41586-022-04492-9

Mowinckel, A. M., & Vidal-Piñeiro, D. (2020). Visualization of brain statistics with R packages ggseg and Ggseg3d. *Advances in Methods and Practices in Psychological Science, 3*(4), 466—483. https://doi.org/10.1177/2515245920928009

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. *Cerebral Cortex, 19*(11), 2728—2735. https://doi.org/10.1093/cercor/bhp026

Prive, F., Aschard, H., Ziyatdinov, A., & Blum, M. G. B. (2018). Efficient analysis of large-scale genome-wide data with two R packages: Bigstatsr and bigsnpr. *Bioinformatics, 34*(16), 2781—2787. https://doi.org/10.1093/bioinformatics/bty185

Raz, N., Ghisletta, P., Rodrigue, K. M., Kennedy, K. M., & Lindenberger, U. (2010). Trajectories of brain aging in middle-aged and older adults: Regional and individual differences. *Neuroimage, 51*(2), 501—511. https://doi.org/10.1016/j.neuroimage.2010.03.020

Revell, A. Y., Silva, A. B., Arnold, T. C., Stein, J. M., Das, S. R., Shinohara, R. T., … Davis, K. A. (2022). A framework for brain atlases: Lessons from seizure dynamics. *Neuroimage, 254*, Article 118986. https://doi.org/10.1016/j.neuroimage.2022.118986

Ritchie, S. J., Cox, S. R., Shen, X., Lombardo, M. V., Reus, L. M., Alloza, C., … Deary, I. J. (2018). Sex differences in the adult human brain: Evidence from 5216 Uk Biobank participants. *Cerebral Cortex, 28*(8), 2959—2975. https://doi.org/10.1093/cercor/bhy109

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*(2), 1—36. https://doi.org/10.18637/jss.v048.i02

Sabuncu, M. R., Ge, T., Holmes, A. J., Smoller, J. W., Buckner, R. L., & Fischl, B. (2016). Morphometricity as a measure of the neuroanatomical signature of a trait. *Proceedings of the National Academy of Sciences, 113*(39), E5749—E5756. https://doi.org/10.1073/pnas.1604378113

Schaefer, A., Kong, R., Gordon, E. M., Laumann, T. O., Zuo, X.-N., Holmes, A. J., … Yeo, B. T. T. (2017). Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity mri. *Cerebral Cortex, 28*(9), 3095—3114. https://doi.org/10.1093/cercor/bhx179

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics, 6*(2), 461—464, 464. Retrieved from 10.1214/aos/1176344136.

Smith, S., Alfaro-Almagro, F., & Miller, K. (2020). *UK Biobank brain imaging documentation*. Retrieved from https://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/brain_mri.pdf.

Smith, S. M., & Nichols, T. E. (2018). Statistical challenges in "big data" human neuroimaging. *Neuron, 97*(2), 263—268. https://doi.org/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 medicine, 12*(3). https://doi.org/10.1371/journal.pmed.1001779. e1001779-e1001779.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological), 58*(1), 267—288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

Turkheimer, E. (2000). Three laws of behavior genetics and what they mean. *Current Directions in Psychological Science, 9*(5), 160—164. https://doi.org/10.1111/1467-8721.00084

Wierenga, L. M., Langen, M., Oranje, B., & Durston, S. (2014). Unique developmental trajectories of cortical thickness and surface area. *Neuroimage, 87*, 120—126. https://doi.org/10.1016/j.neuroimage.2013.11.010

Winkler, A. M., Kochunov, P., Blangero, J., Almasy, L., Zilles, K., Fox, P. T., … Glahn, D. C. (2010). Cortical thickness or grey matter volume? The importance of selecting the phenotype for imaging genetics studies. *Neuroimage, 53*(3), 1135—1146. https://doi.org/10.1016/j.neuroimage.2009.12.028

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. *Scientific Reports, 10*(1), 2837. https://doi.org/10.1038/s41598-020-57951-6

Yang, J., Lee, S. H., Goddard, M. E., & Visscher, P. M. (2011). Gcta: A tool for genome-wide complex trait analysis. *The American Journal of Human Genetics, 88*(1), 76—82. https://doi.org/10.1016/j.ajhg.2010.11.011

Yeo, B. T. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., … Buckner, R. L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology, 106*(3), 1125—1165. https://doi.org/10.1152/jn.00338.2011

Zhang, F., Chen, W., Zhu, Z., Zhang, Q., Nabais, M. F., Qi, T., … Yang, J. (2019). Osca: A tool for omic-data-based complex trait analysis. *Genome Biology, 20*(1), 107. https://doi.org/10.1186/s13059-019-1718-z

Chapter 6

General discussion

6.1 General purpose and contributions

This final discussion chapter summarises the key findings from the preceding empirical chapters, and it discusses their significance, as well as general limitations of the statistical methods. I argue that the presented findings encourage an integrated view of the brain and its many influencing factors that may be approximated most accurately using multidimensional study designs. Finally, I derive lessons that may help improve the reliability of future large-scale neuroimaging studies. Future directions emerging from specific findings are integrated throughout the discussion chapter. Implications and limitations specific to each empirical project are already discussed in the respective chapters and will not be reiterated here.

All in all, this thesis explored morphometric brain correlates of cognitive ability by modelling variance across thousands of individuals. The four empirical chapters employ statistical designs which aim to integrate behavioural, genetic, and neuroimaging data with an emphasis on robustly modelling interdependencies between multiple brain regions (i.e., brain networks). Genetic analysis designs provide an innovative way to triangulate previous phenotypic investigations from a biologically informed perspective. Overall, findings in this thesis help develop a more comprehensive understanding of the neuroanatomy and genetics underlying cognitive ability that may ultimately lay the foundations for society-wide interventions to mitigate old-age cognitive decline. Such interventions could help keep people healthier, wealthier, and more autonomous in old age which would considerably improve general quality of life.

This thesis contributes to the existing literature by addressing two overarching aims. First, I aimed to perform exploratory groundwork to help optimise analytical

decisions when testing the relationship between cognitive ability and brain morphometry. To my knowledge, Chapter 2 is the first hypothesis-free exploration of the impact brain size covariate adjustment has on the association between cognitive ability and brain regional volume. Chapter 3 derived an innovative multivariate genetic modelling technique, *Genomic PCA*, to address the limitations of existing multivariate genetics tools. Genomic PCA can incorporate a large number of GWAS traits while capturing dimensions of maximal genetic variance shared between these multiple GWAS traits. Chapter 5 is the first quantitative comparison of cortical atlases, I am aware of, that formulated a data-driven recommendation about which atlas may maximise brain-trait relationships in large-scale structural MRI studies.

Second, the thesis contributed to the literature by investigating the multivariate relationship between brain morphometry and cognitive ability leveraging genetic information (Chapter 4). By modelling structural brain networks using Genomic PCA (derived in Chapter 3), Chapter 4 is the first study that systematically integrated genetic data, age-related information, and multiple brain regions unadjusted for brain size (as recommended in Chapter 2). On this genetic level of analysis, Chapter 4 triangulated previous phenotypic studies by characterising the substantial role inherited factors play in the relationships between brain networks, cognitive ability, and ageing. Insights from the empirical studies are discussed below.

6.2 Overview of empirical findings

Studies in this thesis built on one another to optimise analytical decisions and to ensure meaningful hypothesis tests in Chapter 4. Chapter 2 provides evidence against adjusting for brain size in uni-regional associations between brain morphometry and cognitive ability. Regions' correlations with cognitive ability mostly reflected variance accounted for by brain size, as opposed to variance unique to each region. This shows that brain size does *not* index nuisance variance and that brain size should be modelled rather than discarded as a covariate. Strong empirical and theoretical evidence suggested that brain size adjusted GWAS are subject to collider bias. Hence, I calculated the first set of regional brain volume GWAS *unadjusted* for brain size.

Chapter 3 derived and validated Genomic PCA, a multivariate technique designed to integrate genome-wide data and large numbers of brain region phenotypes. Genomic PCA extracts genome-wide correlates of an underlying PC purely using GWAS summary statistics as input. I validated that resulting genome-wide PCs were reliable by demonstrating that the Genomic PCA extracted PCs are the same as are obtained by calculating GWAS of an underlying phenotypic PC based on individual-level data. The public availability of both GWAS summary statistics and the Genomic PCA analysis code contribute to the growing accessibility of genetic analyses.

Using Genomic PCA, Chapter 4 extracted genetic PCs from GWAS summary statistics of 83 brain volumes unadjusted for brain size. With these genetically informed analyses I demonstrate in Chapter 4 that there are moderate associations between morphometry across canonical brain networks, general cognitive ability, and cross-sectional indices of age-related neurodegeneration. The genetic links between

176

brain networks and cognitive ability suggest that brain network-cognition associations in the phenotypic literature (e.g., Madole et al., 2021) have shared genetic aetiology and are underpinned by inherited, possibly lifelong stable factors.

Genetic analyses in Chapter 4 revealed no differences between canonical brain networks in how correlated they were with cognitive ability. This implies that the most robust correlates of cognitive ability are distributed across the brain, rather than localised. Hence, sources of co-variation between brain morphometry and cognitive ability likely emerge from general brain-wide morphometric features which are *not* specific to individual regions or brain networks. Future studies should avoid uni-regional investigations of the brain and instead extract robust *general* morphometry shared across multiple brain regions. To achieve multivariate models using genetic information, future studies may model multiple brain phenotypes and brain modalities using techniques such as Genomic PCA (Chapter 3), GenomicSEM (Grotzinger et al., 2019), and Genomic ICA (Soheili-Nezhad et al., 2021).

6.2.1 General brain size genes hypothesis

When genetic studies deliver results that disagree with those from phenotypic studies, speculating about the reasons for their disagreement may help generate original hypotheses. Madole et al. (2021) showed that the central executive brain network was centrally important for cognitive ability in their phenotypic analysis. By contrast, the genetic analysis in Chapter 4 delivered no evidence that the central executive network was more important for cognitive ability than other brain networks or the whole brain.

This conflicting evidence between Madole et al. (2021) and Chapter 4 cannot have emerged from major methodological differences. Both studies had closely

matched designs: both used the Desikan-Killiany brain atlas, neither adjusted for brain size, they used PCA to reduce dimensionality across brain networks, and they both analysed structural brain volume data in the UK Biobank sample. The crucial difference between them was that Madole et al. (2021) considered phenotypic data, and Chapter 4 considered genome-wide data.

The fact that phenotypic analyses produced different conclusions to genetic analyses may inspire the following *general brain size genes* hypothesis: It hypothesises that the genetic correlation between brain morphometry and cognitive ability is largely underpinned by general genes affecting the size of the whole brain (rather than individual regions or networks) through general brain features such as cell growth factors (as shown in Jansen et al., 2020). This hypothesis could explain why my genetic analyses in Chapter 4 found the same correlations between cognitive ability and nine canonical brain networks. In line with this hypothesis, the genetic data only captured general brain features underlying brain-wide morphometry shared across all brain networks. Contrary to Chapter 4, Madole et al. (2021) may have found brain network differences because their phenotypic analyses would have been sensitive towards combined genetic *and* environmental factors. Environmental factors could have caused morphometric specialisation of the central executive network relevant to cognitive ability.

This *general brain size* hypothesis suggesting that genome-wide models of structural brain networks reflect general brain size genes was proposed in a 2020 review: Changeux et al. (2020) argued that brain size may be the single parsimonious factor accounting for the brain morphometry cognitive ability relationship. The *general brain size* hypothesis was formulated because it could explain the paradox that human genetics only marginally differ from primate genetics when human

cognitive ability greatly exceeds primate cognitive ability in language capabilities, abstract processing, working and long-term memory. It supports the *general brain size* hypothesis that the only consistent genetic differences found between primate and human genetics tend to affect brain size through genes involved in general brain growth, neuronal number, and neuronal maturation (Changeux et al., 2020).

Changeux et al. (2020) reason that larger brains do not only have more neurons, but they have features that are unique to humans. Such unique human features include selective growth in association areas, more differentiated neural regions, and more efficient connectivity that provide overall more workspace. Hence, genetic predisposition to brain size may parsimoniously account for complex human brain organisation that enables advanced human cognitive ability. While this hypothesis requires further scrutiny, it could explain why the genetic analyses in this thesis only found evidence for brain-wide and brain network-*un*specific correlates of cognitive ability. Future studies are needed to further test the general brain size hypothesis by integrating multi-dimensional biological information including neuroimaging, genetic, environmental, transcriptomic, epigenetic (and more) data, as well as their gene-brain-environment interplay.

6.2.2 Embracing brain complexity to better account for the relationship between cognitive ability and brain morphometry

Chapter 5 demonstrated that embracing neuroimaging data in its full vertex-wise complexity maximises the statistical relationship between cognitive ability and brain morphometry. Analysing vertex-wise neuroimaging data resulted in larger brain cognitive ability associations compared to analyses that reduced brain dimensionality via traditional brain atlases. Modelling hundreds of thousands of vertex-wise brain

volume measures simultaneously explained ~3-times more variance in cognitive ability (R^2 = 25%), than when the same brain measures were reduced into 68 regions in the Desikan-Killiany atlas (R^2 = 8%). The same vertex-wise measures explained 5-times more variance in cognitive ability than a single whole brain measure (R^2 = 5%) (Nave et al., 2018), which was, so far, considered the most reliable correlate of cognitive ability.

Going forward, structural MRI studies should embrace the complexity of hundreds of thousands of vertex-wise neuroimaging variables to obtain more informative models of the relationship between the brain and behavioural traits. Sample sizes and computational power are getting large enough to obtain sensible estimates from vertex-wise data (e.g., Chapter 5). To achieve more informative and complex brain models, future studies should adopt tools based in Statistical Genetics to analyse large-scale neuroimaging data computationally efficiently. In Chapter 5 I used the genome-wide complex trait analysis (GCTA) method to gain robust estimates of the relationship between structural MRI data, cognitive ability, and other behavioural traits. Future neuroimaging studies should use the wealth of statistical genetics tools to advance discovery without the need for novel methods development (further discussion of this in the Viewpoint below).

6.3 General limitations

6.3.1 Drawbacks of the cognitive ability (*g*-factor) definition

I defined cognitive ability as an abstract mathematical dimension (*g*-factor) extracted from variance shared between intercorrelated cognitive tests. This definition neglects specific cognitive abilities that do not map onto the *g*-factor but likely capture important variance. I focus on the *g*-factor as a reliable measure because it should not depend on the form, version, or wording of the specific cognitive tests used to index it. Therefore, the *g*-factor is objective, highly replicable, and (at least theoretically) error-free (Humphreys, 1992), which I argue is better suited to study human biology and lifelong health outcomes than other less reliable measures (self-report assessment for example). Additional studies investigating specific cognitive abilities are needed to comprehensively understand human behaviours and their lifelong consequences.

Cognitive ability as a measure is culturally loaded which means that an individual's performance in cognitive tests does not just reflect their cognitive ability, but also their exposure to academic activities through schooling, for example. Cognitive tests were developed in Western, educated, industrialised, rich and democratic (WEIRD) countries and tend to be delivered on a computer screen simulating artificial environments. Individuals raised in WEIRD countries likely received compulsory academic schooling and gained familiarity with computers which will mean they naturally perform better in cognitive tests than individuals raised in non-WEIRD countries who may have had less or no schooling. To widen the transferability of cognitive ability findings, novel cognitive tests are required to accommodate more diverse populations and their cultural environments.

To ensure scientific discoveries are not misused, cognitive ability researchers bear a responsibility to ensure their research is accurately interpreted by the wider population. The presence of differences in cognitive ability implies by no means that individuals with higher cognitive ability are worth more than those with lower cognitive ability. Researchers must carefully consider the potential impact of cognitive ability studies by framing the research using neutral language. Accurately outlining study limitations and realistic public health benefits are particularly important to avoid fuelling elitist arguments [14]. To better communicate the implications and intentions of cognitive ability studies, scientists must strengthen efforts and resources invested in public engagement.

I would like to explicitly distance my work from elitist agendas[14]: By considering within-group interindividual differences in cognitive ability, I researched a reliable and stable human trait that acts as a window into human biology and may lay foundations for interventions to mitigate old-age cognitive decline. I do not interpret the presence of genetic correlates of cognitive ability as causal which I hope helped portray a stochastic (rather than deterministic) view of the genetics of intelligence. I

---

[14] Here I outline one contemporary example illustrating how cognitive ability studies have been misused to propagate elitist agendas. Referencing cognitive ability studies, Boris Johnson implied in his 2013 Margaret Thatcher speech that people with greater cognitive ability may be worth more and deserve more privileges than people with lower cognitive ability. Below, I list three quotes from the speech that Johnson used to advocate for economic inequality:

- "Human beings [...] are already very far from equal in raw ability, if not spiritual worth".

- "Whatever you may think of the value of IQ tests, it is surely relevant to a conversation about equality that as many as 16% of our species have an IQ below 85, while about 2% have an IQ above 130. The harder you shake the pack, the easier it will be for some cornflakes to get to the top."

- "The income gap between the top cornflakes and the bottom cornflakes is getting wider than ever and I stress that I don't believe that economic equality is possible. Indeed, some measure of inequality is essential for the spirit of envy [...] that is a valuable spur of economic activity."

The quotes were taken from Gillborn (2016) and Johnson (2013).

would like to be explicit that none of my work is indicative of between-group differences (for example, between ancestries). further elaborates how genetics research can become more widely translatable to non-European ancestries.

6.3.2 Population-level inference of individual differences

Observational studies of healthy interindividual variability represent one of many approaches to investigate human traits. Compared with smaller interventional studies for example, interindividual studies isolate robust associations that generally apply to healthy humans but cannot be interpreted causally. Interindividual studies should be complemented by other approaches suited to uncover causal relationships (e.g., randomised control trials, Mendelian Randomisation, or natural experiments).

6.3.3 Cross-sectional data

Due to the cross-sectional nature of the used data, may have underestimated correlation magnitudes among brain regions. A 2021 longitudinal study showed that interregional correlations are considerably stronger among longitudinally measured brain volumes (average r = 0.81), than they are among cross-sectionally measured brain volumes (average r = 0.35) (Cox et al., 2021). It required longitudinal data to definitively show that brain volume and cognitive ability do indeed have a mutually dependent relationship (r = 0.43) and that they change over time in a coordinated fashion (Cox et al., 2021). Large-scale longitudinal collection of neuroimaging data is available in studies such as the Lothian Birth Cohort (LBC; Deary et al., 2012; Taylor et al., 2018), and the UK Biobank (Littlejohns et al., 2020). Increasing availability of longitudinal MRI data will allow future studies to characterise trait relationships more accurately.

6.3.4 Non-representative samples

Genetic and phenotypic analyses were conducted in White European ancestry. It remains unknown whether phenotypic and genetic correlations can be translated from Europeans to non-Europeans because ancestral groups differ in LD structures, as well as geographical, environmental and cultural differences that may be systematically correlated with ancestry (Abdellaoui et al., 2023; Mathieson & Scally, 2020). As genetic data of non-European ancestry – such as the African Genome Variation Project (Gurdasani et al., 2015), or the GenomeAsia 100K Project (Wall et al., 2019) – increase in size and availability, future multi-ancestral studies present opportunities to make findings more widely applicable.

Besides ancestry, analysed samples are not representative of the target population inferences are extrapolated to. The UKB sample is healthier, wealthier, and better educated than the age-matched general population in the UK: A strong volunteer participation bias means that healthier and well-off individuals were most likely to participate in the UKB data collection (Fry et al., 2017). Genetic analyses are significantly impacted by participation bias. A GWAS of study participation revealed 32 significantly associated SNPs (Tyrrell et al., 2021). According to a 2022 pre-print, participation bias affects most SNPs that reach genome-wide significance in education GWAS (Schoeler et al., 2022). To balance over- and under-represented individuals in GWAS, propensity score methods can be used to adjust SNP effect bias induced by imbalances of sample representativeness (Schoeler et al., 2022).

6.3.5 Indirect genetic effects

I used genetic correlations derived from GWAS summary statistics to replicate previous phenotypic studies more objectively. However, it would be inaccurate to

assume that genetic correlations simply captured genetic effects that directly affect traits under investigation. *In*direct genetic effects also contribute to estimates of heritability (e.g., Young et al., 2018) and genetic correlations (Bulik-Sullivan, Finucane, et al., 2015) and they index demographic effects such as assortative mating and population stratification (Abdellaoui et al., 2019), as well as gene-environment interplay (Abdellaoui et al., 2022). Hence, heritability and genetic correlations presented in this thesis likely reflect interrelated genetic and environmental factors.

GWAS designs that consider interindividual differences within families (e.g., between siblings) can help isolate direct genetic effects by mitigating indirect genetic effects (e.g., assortative mating, or gene-environment correlation; Howe et al., 2022). Compared with standard GWAS, family-based GWAS tend to substantially reduce genetic estimates in behavioural traits (~40% reduction in SNP-based heritability of cognitive ability; Howe et al., 2022). Future within-family studies could add to the thesis findings by characterising direct and indirect sources of genetic trait correlations to help unpack underlying genetic and environmental mechanisms .

6.3.6 Putative brain network characterisations

Finally, brain network characterisations used in Chapter 4 are putative as I assigned individual brain regions to canonical brain networks based on findings in some previous publications. The quality of theory-driven brain network models can only be as good as the evidence that informs them. Exactly which regions reside in different macro-scale brain networks remains an active area of research (Uddin et al., 2019) and future research may encourage refinement of the brain networks modelled in Chapter 4. Future data-driven work is needed to refine brain network characterisations through different sources of brain-based information, for example,

cellular micro-organisation of the brain (Amunts et al., 2020) or neurotransmitter systems (Hansen et al., 2022).

6.4 Significance of these findings for the design of neuroimaging studies

This thesis finds evidence that cognitive ability is anchored in more general brain-wide features (rather than localised ones) which emphasises a more integrated view of the brain. Such an integrated view would require holistic study designs integrating multiple brain regions, genetic data, environmental information and more. This is in line with Westlin et al. (2023) calling for a reconsideration of fundamental study assumptions about the brain: those fundamental assumptions tend to be misleading as they can dictate imprecise statistical designs in neuroimaging studies. Westlin et al. (2023) outline that many studies adopt a typological view of the brain that assumes that human cognition and behaviour emerges from (1) a simple region-to-behaviour mapping of (2) localised neuronal signal, that is thought (3) to be independent of other signals, including the rest of brain, the body, and the physical environment. These assumptions justify simplistic statistical models mapping individual brain regions onto complex traits, at the same time as adjusting for brain size, as well as other associated factors.

Westlin et al. (2023) elaborate that this view of the brain stands in contrast to empirical evidence (e.g., functional MRI brain networks or animal studies). They argue that brain-behaviour relationships can become more robust and replicable when studies are designed under the following alternative assumptions, more in line with empirical evidence: Behavioural traits emerge from (1) multiple coordinated regions,

(2) whose activity may spread throughout the whole brain, (3) which is directly dependent upon a complex web of interacting signals from the brain, the body, and the environment. Individually, each factor that is correlated with a behavioural trait (e.g., cognitive ability) may have weak influences on underlying brain function, but together these factors give rise to cognitive ability collectively, and can therefore not be reasonably isolated from one another in statistical analyses.

Although these alternative assumptions in Westlin et al. (2023) were formulated mainly with reference to functional MRI studies, the alternative assumptions are also supported by structural MRI studies in this thesis. My thesis demonstrated that genetic correlates of brain morphometry relate to cognitive ability through (1) general features shared between multiple regions (2) that are spread across the whole brain and that are not specific to brain regions or brain networks. (3) General dimensions of brain morphometry are not independent of global brain features (e.g., total brain volume), and they cannot be reliably modelled when made statistically independent of brain size.

In line with the alternative more holistic view of the brain suggested by Westlin et al. (2023), this thesis identified four key methodological choices that may help produce more robust and biologically plausible brain trait correlations: First, do not include brain size as a covariate especially in GWAS. Second, use multivariate models of the brain accounting for correlations among brain regions. Third, integrate genetic information. Fourth, use vertex-wise cortical representations instead of low-dimensional brain atlases. Adopting these methodological choices should support future studies to capture more reliable, more complex, and biologically more plausible correlations between brain morphometry and cognitive ability, as well as other behavioural traits.

6.5 Transferring lessons across research disciplines

The process of conducting the brain atlas research in Chapter 5 led me to appreciate the value of working between disciplines. It made me aware of the strong parallels between statistical genetics and neuroimaging research in the challenges and opportunities they face. I suggest that practices from statistical genetics – which is an inherently multivariate and computational discipline – can teach valuable lessons relevant to neuroimaging research which has a tradition of much smaller samples. Neuroimaging studies tend to focus on less computationally intensive analyses of individual brain regions in clinical settings, for example. By applying genetics analysis techniques to neuroimaging data (as done in Chapter 5), we may be able to drive innovation and accelerate the evolution of large-scale neuroimaging research without needing to develop new neuroimaging-specific techniques. In a Viewpoint that is currently under review with *Cortex* I present five major lessons derived from genetics research that could help address many issues of replicability in neuroimaging research.

First, the failure of candidate gene studies motivates neuroimaging studies to avoid overly simplistic studies mapping individual brain regions onto complex traits.

Second, genetics research findings demonstrate that the robustness of study results is greatly improved by increasing sample sizes.

Third and fourth, the success of GWAS inspires neuroimaging studies to adopt hypothesis-free mass-univariate testing and sharing of summary-level (vertex-wise) association data to aid cross-sample meta-analyses and large-scale collaboration that is otherwise hindered by data privacy concerns.

Finally, genetics methods can account for complex data structures and are computationally efficiently implemented. Applying genetics analysis techniques to vertex-wise neuroimaging data may boost scientific discovery without the need for novel neuroimaging-specific methods.

In the Viewpoint I argue that these five practices should either be further endorsed, or newly adopted by the neuroimaging community which may accelerate the evolution of neuroimaging as the field embarks on a big-data journey of replicable studies of the brain. This Viewpoint summary was partially adopted from its abstract, and a word-by-word copy of the version submitted to *Cortex* is presented in the Appendix.

6.6 Concluding remarks

The overarching aim of this thesis was to investigate brain morphometric correlates of general cognitive ability using novel, partly genetically informed analyses in well-powered samples. It contributed empirical investigations that aimed to improve methodological decisions in neuro-cognitive studies. Specifically, this thesis gathered evidence weighing against covariate adjustment for brain size, especially in GWAS. My work also demonstrated that vertex-wise cortical representations can help maximise brain-trait associations in comparison with traditional cortical atlases. Informed by this methodological groundwork, this thesis used structural MRI data and a novel multivariate and genetically informed approach to revisit the longstanding question of where in the brain cognitive abilities may be located. Results proposed that morphometric brain organisation is moderately linked with general cognitive ability, whereby sources of their covariation seem to emerge from brain-wide, rather than region-specific or network-specific features, most likely related to overall brain size

and cell growth. As neuroimaging research embarks on a big-data journey of reproducible studies, work presented here inspires interdisciplinary and multidimensional approaches that aim to further our understanding of the biology of cognitive ability by embracing the full complexity of the brain, and its genetic, environmental, and developmental influences.

References

Abdellaoui, A., Dolan, C. V., Verweij, K. J. H., & Nivard, M. G. (2022). Gene–Environment Correlations across Geographic Regions Affect Genome-Wide Association Studies. *Nature Genetics, 54*(9), 1345-1354. doi:10.1038/s41588-022-01158-0

Abdellaoui, A., Hugh-Jones, D., Yengo, L., Kemper, K. E., Nivard, M. G., Veul, L., . . . Visscher, P. M. (2019). Genetic Correlates of Social Stratification in Great Britain. *Nature Human Behaviour, 3*(12), 1332-1342. doi:10.1038/s41562-019-0757-5

Abdellaoui, A., Yengo, L., Verweij, K. J. H., & Visscher, P. M. (2023). 15 Years of Gwas Discovery: Realizing the Promise. *The American Journal of Human Genetics, 110*(2), 179-194. doi:10.1016/j.ajhg.2022.12.011

Abraham, G., Qiu, Y., & Inouye, M. (2017). Flashpca2: Principal Component Analysis of Biobank-Scale Genotype Datasets. *Bioinformatics, 33*(17), 2776-2778. doi:10.1093/bioinformatics/btx299

Acheson, D. J., & Hagoort, P. (2013). Stimulating the Brain's Language Network: Syntactic Ambiguity Resolution after Tms to the Inferior Frontal Gyrus and Middle Temporal Gyrus. *Journal of Cognitive Neuroscience, 25*(10), 1664-1677. doi:10.1162/jocn_a_00430

Akshoomoff, N., Beaumont, J. L., Bauer, P. J., Dikmen, S. S., Gershon, R. C., Mungas, D., . . . Heaton, R. K. (2013). Viii. Nih Toolbox Cognition Battery (Cb): Composite Scores of Crystallized, Fluid, and Overall Cognition. *Monographs of the Society for Research in Child Development, 78*(4), 119-132. doi:10.1111/mono.12038

Alexander-Bloch, A. F., Mathias, S. R., Fox, P. T., Olvera, R. L., Göring, H. H. H., Duggirala, R., . . . Glahn, D. C. (2017). Human Cortical Thickness Organized into Genetically-Determined Communities across Spatial Resolutions. *Cerebral Cortex, 29*(1), 106-118. doi:10.1093/cercor/bhx309

Alfaro-Almagro, F., Jenkinson, M., Bangerter, N. K., Andersson, J. L. R., Griffanti, L., Douaud, G., . . . Smith, S. M. (2018). Image Processing and Quality Control for the First 10,000 Brain Imaging Datasets from Uk Biobank. *NeuroImage, 166*, 400-424. doi:10.1016/j.neuroimage.2017.10.034

Alfaro-Almagro, F., McCarthy, P., Afyouni, S., Andersson, J. L. R., Bastiani, M., Miller, K. L., . . . Smith, S. M. (2021). Confound Modelling in Uk Biobank Brain Imaging. *NeuroImage, 224*, 117002. doi:10.1016/j.neuroimage.2020.117002

Allen, M., Poggiali, D., Whitaker, K., Marshall, T., & Kievit, R. (2019). Raincloud Plots: A Multi-Platform Tool for Robust Data Visualization [Version 1; Peer Review: 2 Approved]. *Wellcome Open Research, 4*(63). doi:10.12688/wellcomeopenres.15191.1

Amunts, K., Mohlberg, H., Bludau, S., & Zilles, K. (2020). Julich-Brain: A 3d Probabilistic Atlas of the Human Brain's Cytoarchitecture. *Science, 369*(6506), 988-992. doi:10.1126/science.abb4588

Anderson, K. M., Ge, T., Kong, R., Patrick, L. M., Spreng, R. N., Sabuncu, M. R., . . . Holmes, A. J. (2021). Heritability of Individualized Cortical Network Topography. *Proceedings of the National Academy of Sciences, 118*(9), e2016271118. doi:10.1073/pnas.2016271118

Anvari, F., Kievit, R., Lakens, D., Pennington, C. R., Przybylski, A. K., Tiokhin, L., . . . Orben, A. (2022). Not All Effects Are Indispensable: Psychological Science Requires Verifiable Lines of Reasoning for Whether an Effect Matters. *Perspectives on Psychological Science, 0*(0). doi:10.1177/17456916221091565

Arnatkevičiūtė, A., Fulcher, B. D., Oldham, S., Tiego, J., Paquola, C., Gerring, Z., . . . Fornito, A. (2021). Genetic Influences on Hub Connectivity of the Human Connectome. *Nature Communications, 12*(1), 4237. doi:10.1038/s41467-021-24306-2

Aschard, H., Vilhjálmsson, Bjarni J., Joshi, Amit D., Price, Alkes L., & Kraft, P. (2015). Adjusting for Heritable Covariates Can Bias Effect Estimates in Genome-Wide Association Studies. *The American Journal of Human Genetics, 96*(2), 329-339. doi:10.1016/j.ajhg.2014.12.021

Ashburner, J., & Friston, K. J. (2000). Voxel-Based Morphometry—the Methods. *NeuroImage, 11*(6), 805-821. doi:10.1006/nimg.2000.0582

Barbey, A. K. (2018). Network Neuroscience Theory of Human Intelligence. *Trends in Cognitive Sciences, 22*(1), 8-20. doi:10.1016/j.tics.2017.10.001

Baselmans, B. M. L., Jansen, R., Ip, H. F., van Dongen, J., Abdellaoui, A., van de Weijer, M. P., . . . Social Science Genetic Association, C. (2019). Multivariate Genome-Wide Analyses of the Well-Being Spectrum. *Nature Genetics, 51*(3), 445-451. doi:10.1038/s41588-018-0320-8

Basten, U., Hilger, K., & Fiebach, C. J. (2015). Where Smart Brains Are Different: A Quantitative Meta-Analysis of Functional and Structural Brain Imaging Studies on Intelligence. *Intelligence, 51*, 10-27. doi:10.1016/j.intell.2015.04.009

Batty, G. D., Deary, I. J., & Gottfredson, L. S. (2007). Premorbid (Early Life) Iq and Later Mortality Risk: Systematic Review. *Annals of Epidemiology, 17*(4), 278-288. doi:10.1016/j.annepidem.2006.07.010

Baurley, J. W., Edlund, C. K., Pardamean, C. I., Conti, D. V., & Bergen, A. W. (2016). Smokescreen: A Targeted Genotyping Array for Addiction Research. *BMC genomics, 17*, 145-145. doi:10.1186/s12864-016-2495-7

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. doi:10.1111/j.2517-6161.1995.tb02031.x

Bethlehem, R. A. I., Seidlitz, J., White, S. R., Vogel, J. W., Anderson, K. M., Adamson, C., . . . Vetsa. (2022). Brain Charts for the Human Lifespan. *Nature, 604*(7906), 525-533. doi:10.1038/s41586-022-04554-y

Bilker, W. B., Hansen, J. A., Brensinger, C. M., Richard, J., Gur, R. E., & Gur, R. C. (2012). Development of Abbreviated Nine-Item Forms of the Raven's Standard Progressive Matrices Test. *Assessment, 19*(3), 354-369. doi:10.1177/1073191112446655

Biton, A., Traut, N., Poline, J.-B., Aribisala, B. S., Bastin, M. E., Bülow, R., . . . Toro, R. (2020). Polygenic Architecture of Human Neuroanatomical Diversity. *Cerebral Cortex, 30*(4), 2307-2320. doi:10.1093/cercor/bhz241

Blokland, G. A. M., de Zubicaray, G. I., McMahon, K. L., & Wright, M. J. (2012). Genetic and Environmental Influences on Neuroimaging Phenotypes: A Meta-Analytical Perspective on Twin Imaging Studies. *Twin Research and Human Genetics, 15*(3), 351-371. doi:10.1017/thg.2012.11

Border, R., Johnson, E. C., Evans, L. M., Smolen, A., Berley, N., Sullivan, P. F., & Keller, M. C. (2019). No Support for Historical Candidate Gene or Candidate Gene-by-Interaction Hypotheses for Major Depression across Multiple Large Samples. *American Journal of Psychiatry, 176*(5), 376-387. doi:10.1176/appi.ajp.2018.18070881

Botvinik-Nezer, R., Holzmeister, F., Camerer, C. F., Dreber, A., Huber, J., Johannesson, M., . . . Schonberg, T. (2020). Variability in the Analysis of a Single Neuroimaging Dataset by Many Teams. *Nature, 582*(7810), 84-88. doi:10.1038/s41586-020-2314-9

Bressler, S. L., & Menon, V. (2010). Large-Scale Brain Networks in Cognition: Emerging Methods and Principles. *Trends in Cognitive Sciences, 14*(6), 277-290. doi:10.1016/j.tics.2010.04.004

Bromet, E., Andrade, L. H., Hwang, I., Sampson, N. A., Alonso, J., de Girolamo, G., . . . Kessler, R. C. (2011). Cross-National Epidemiology of Dsm-Iv Major Depressive Episode. *BMC Medicine, 9*(1), 90. doi:10.1186/1741-7015-9-90

Buckner, R. L., & DiNicola, L. M. (2019). The Brain's Default Network: Updated Anatomy, Physiology and Evolving Insights. *Nature Reviews Neuroscience, 20*(10), 593-608. doi:10.1038/s41583-019-0212-7

Bulik-Sullivan, B., Finucane, H. K., Anttila, V., Gusev, A., Day, F. R., Loh, P.-R., . . . Robinson, E. B. (2015). An Atlas of Genetic Correlations across Human Diseases and Traits. *Nature Genetics, 47*(11), 1236. doi:10.1038/ng.3406

Bulik-Sullivan, B. K., Loh, P.-R., Finucane, H. K., Ripke, S., Yang, J., Patterson, N., . . . Schizophrenia Working Group of the Psychiatric Genomics, C. (2015). Ld Score Regression Distinguishes Confounding from Polygenicity in Genome-Wide Association Studies. *Nature Genetics, 47*(3), 291-295. doi:10.1038/ng.3211

Bullmore, E. T., Woodruff, P. W. R., Wright, I. C., Rabe-Hesketh, S., Howard, R. J., Shuriquie, N., & Murray, R. M. (1998). Does Dysplasia Cause Anatomical Dysconnectivity in Schizophrenia? *Schizophrenia Research, 30*(2), 127-135. doi:10.1016/S0920-9964(97)00141-2

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

Casey, B. J., Cannonier, T., Conley, M. I., Cohen, A. O., Barch, D. M., Heitzeg, M. M., . . . Dale, A. M. (2018). The Adolescent Brain Cognitive Development (Abcd) Study: Imaging Acquisition across 21 Sites. *Developmental Cognitive Neuroscience, 32*, 43-54. doi:10.1016/j.dcn.2018.03.001

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*(1). doi:10.1186/s13742-015-0047-8

Changeux, J.-P., Goulas, A., & Hilgetag, C. C. (2020). A Connectomic Hypothesis for the Hominization of the Brain. *Cerebral Cortex, 31*(5), 2425-2449. doi:10.1093/cercor/bhaa365

Chen, C.-H., Panizzon, Matthew S., Eyler, Lisa T., Jernigan, Terry L., Thompson, W., Fennema-Notestine, C., . . . Dale, Anders M. (2011). Genetic Influences on Cortical Regionalization in the Human Brain. *Neuron, 72*(4), 537-544. doi:10.1016/j.neuron.2011.08.021

Cheverud, J. M. (1988). A Comparison of Genetic and Phenotypic Correlations. *Evolution, 42*(5), 958-968. doi:10.2307/2408911

Cole, J. H., & Franke, K. (2017). Predicting Age Using Neuroimaging: Innovative Brain Ageing Biomarkers. *Trends in Neurosciences, 40*(12), 681-690. doi:10.1016/j.tins.2017.10.001

Cole, J. H., Marioni, R. E., Harris, S. E., & Deary, I. J. (2019). Brain Age and Other Bodily 'Ages': Implications for Neuropsychiatry. *Molecular Psychiatry, 24*(2), 266-281. doi:10.1038/s41380-018-0098-1

Cole, J. H., Ritchie, S. J., Bastin, M. E., Valdés Hernández, M. C., Muñoz Maniega, S., Royle, N., . . . Deary, I. J. (2018). Brain Age Predicts Mortality. *Molecular Psychiatry, 23*(5), 1385-1392. doi:10.1038/mp.2017.62

Coleman, J. R. I., Peyrot, W. J., Purves, K. L., Davis, K. A. S., Rayner, C., Choi, S. W., . . . on the behalf of Major Depressive Disorder Working Group of the Psychiatric Genomics, C. (2020). Genome-Wide Gene-Environment Analyses of Major Depressive Disorder and Reported Lifetime Traumatic Experiences in Uk Biobank. *Molecular Psychiatry, 25*(7), 1430-1446. doi:10.1038/s41380-019-0546-6

Coleman, M. D., Grainger, A. I., Parri, H. R., & Hill, E. J. (2021). Chapter 1 - In vitro Human Stem Cell–Mediated Central Nervous System Platforms: Progress and Challenges∗∗Dedicated to the Memory of Mark J. Winn, Ph.D., (1960-1993). In Birbrair, A. (Ed.), *Ipscs for Modeling Central Nervous System Disorders* (Vol. 6, pp. 1-19): Academic Press.

Couvy-Duchesne, B., Strike, L. T., Zhang, F., Holtz, Y., Zheng, Z., Kemper, K. E., . . . Visscher, P. M. (2020). A Unified Framework for Association and Prediction from Vertex-Wise Grey-Matter Structure. *Human Brain Mapping, 41*(14), 4062-4076. doi:10.1002/hbm.25109

Cox, S. R., Harris, M. A., Ritchie, S. J., Buchanan, C. R., Valdés Hernández, M. C., Corley, J., . . . Tucker-Drob, E. M. (2021). Three Major Dimensions of Human Brain Cortical Ageing in Relation to Cognitive Decline across the Eighth Decade of Life. *Mol Psychiatry, 26*(6), 2651-2662. doi:10.1038/s41380-020-00975-1

Cox, S. R., Ritchie, S. J., Fawns-Ritchie, C., Tucker-Drob, E. M., & Deary, I. J. (2019). Structural Brain Imaging Correlates of General Intelligence in Uk Biobank. *Intelligence, 76*, 101376. doi:10.1016/j.intell.2019.101376

Cox, S. R., Ritchie, S. J., Tucker-Drob, E. M., Liewald, D. C., Hagenaars, S. P., Davies, G., . . . Deary, I. J. (2016). Ageing and Brain White Matter Structure in 3,513 Uk Biobank Participants. *Nature Communications, 7*(1), 13629. doi:10.1038/ncomms13629

Davies, G., Armstrong, N., Bis, J. C., Bressler, J., Chouraki, V., Giddaluru, S., . . . Deary, I. J. (2015). Genetic Contributions to Variation in General Cognitive Function: A Meta-Analysis of Genome-Wide Association Studies in the Charge Consortium (N=53 949). *Molecular Psychiatry, 20*(2), 183-192. doi:10.1038/mp.2014.188

Davies, G., Lam, M., Harris, S. E., Trampush, J. W., Luciano, M., Hill, W. D., . . . Deary, I. J. (2018). Study of 300,486 Individuals Identifies 148 Independent Genetic Loci Influencing General Cognitive Function. *Nature Communications, 9*(1), 2098. doi:10.1038/s41467-018-04362-x

Davies, N. M., Hill, W. D., Anderson, E. L., Sanderson, E., Deary, I. J., & Davey Smith, G. (2019). Multivariable Two-Sample Mendelian Randomization Estimates of the Effects of Intelligence and Education on Health. *eLife, 8*, e43990. doi:10.7554/eLife.43990

de la Fuente, J., Davies, G., Grotzinger, A. D., Tucker-Drob, E. M., & Deary, I. J. (2021). A General Dimension of Genetic Sharing across Diverse Cognitive Traits Inferred from Molecular Data. *Nature Human Behaviour, 5*(1), 49-58. doi:10.1038/s41562-020-00936-2

de Leeuw, C. A., Mooij, J. M., Heskes, T., & Posthuma, D. (2015). Magma: Generalized Gene-Set Analysis of Gwas Data. *PLOS Computational Biology, 11*(4), e1004219. doi:10.1371/journal.pcbi.1004219

de Vlaming, R., Slob, E. A. W., Jansen, P. R., Dagher, A., Koellinger, P. D., Groenen, P. J. F., & Rietveld, C. A. (2021). Multivariate Analysis Reveals Shared Genetic Architecture of Brain Morphology and Human Behavior. *Communications Biology, 4*(1), 1180. doi:10.1038/s42003-021-02712-y

Deary, I. J., Gow, A. J., Pattie, A., & Starr, J. M. (2012). Cohort Profile: The Lothian Birth Cohorts of 1921 and 1936. *International journal of epidemiology, 41*(6), 1576-1584. doi:10.1093/ije/dyr197

Deary, I. J., Johnson, W., & Houlihan, L. M. (2009). Genetic Foundations of Human Intelligence. *Human Genetics, 126*(1), 215-232. doi:10.1007/s00439-009-0655-4

Deary, I. J., Liewald, D., & Nissan, J. (2011). A Free, Easy-to-Use, Computer-Based Simple and Four-Choice Reaction Time Programme: The Deary-Liewald Reaction Time Task. *Behavior Research Methods, 43*(1), 258-268. doi:10.3758/s13428-010-0024-1

Deary, I. J., Penke, L., & Johnson, W. (2010). The Neuroscience of Human Intelligence Differences. *Nature Reviews Neuroscience, 11*(3), 201-211. doi:10.1038/nrn2793

Deary, I. J., Weiss, A., & Batty, G. D. (2010). Intelligence and Personality as Predictors of Illness and Death:How Researchers in Differential Psychology and Chronic Disease Epidemiology Are Collaborating to Understand and Address Health Inequalities. *Psychological Science in the Public Interest, 11*(2), 53-79. doi:10.1177/1529100610387081

Deng, W., Rolls, E. T., Ji, X., Robbins, T. W., Banaschewski, T., Bokde, A. L. W., . . . Feng, J. (2017). Separate Neural Systems for Behavioral Change and for Emotional Responses to Failure During Behavioral Inhibition. *Human Brain Mapping, 38*(7), 3527-3537. doi:10.1002/hbm.23607

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., . . . Hyman, B. T. (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

Downar, J., Crawley, A. P., Mikulis, D. J., & Davis, K. D. (2002). A Cortical Network Sensitive to Stimulus Salience in a Neutral Behavioral Context across Multiple Sensory Modalities. *Journal of Neurophysiology, 87*(1), 615-620. doi:10.1152/jn.00636.2001

Dudbridge, F. (2013). Power and Predictive Accuracy of Polygenic Risk Scores. *PLOS Genetics, 9*(3), e1003348. doi:10.1371/journal.pgen.1003348

Duncan, J. (2010). The Multiple-Demand (Md) System of the Primate Brain: Mental Programs for Intelligent Behaviour. *Trends in Cognitive Sciences, 14*(4), 172-179. doi:10.1016/j.tics.2010.01.004

Dunn, L. M., & Dunn, L. M. (1965). *Peabody Picture Vocabulary Test*: Third Edition (PPVT-III) [Database record]. APA PsycTests.

Elliott, M. L., Knodt, A. R., Ireland, D., Morris, M. L., Poulton, R., Ramrakha, S., . . . Hariri, A. R. (2020). What Is the Test-Retest Reliability of Common Task-Functional Mri Measures? New Empirical Evidence and a Meta-Analysis. *Psychological Science, 31*(7), 792-806. doi:10.1177/0956797620916786

Fawns-Ritchie, C., & Deary, I. J. (2020). Reliability and Validity of the Uk Biobank Cognitive Tests. *PLOS ONE, 15*(4), e0231627. doi:10.1371/journal.pone.0231627

Ferreira, L. K., & Busatto, G. F. (2013). Resting-State Functional Connectivity in Normal Brain Aging. *Neurosci Biobehav Rev, 37*(3), 384-400. doi:10.1016/j.neubiorev.2013.01.017

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

Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., . . . Dale, A. M. (2002). Whole Brain Segmentation: Automated Labeling of Neuroanatomical Structures in the Human Brain. *Neuron, 33*(3), 341-355. doi:10.1016/S0896-6273(02)00569-X

Fjell, A. M., & Walhovd, K. B. (2010). Structural Brain Changes in Aging: Courses, Causes and Cognitive Consequences. *Reviews in the Neurosciences, 21*(3), 187-222. doi:10.1515/REVNEURO.2010.21.3.187

Fry, A., Littlejohns, T. J., Sudlow, C., Doherty, N., Adamska, L., Sprosen, T., . . . Allen, N. E. (2017). Comparison of Sociodemographic and Health-Related Characteristics of Uk Biobank Participants with Those of the General Population. *American Journal of Epidemiology, 186*(9), 1026-1034. doi:10.1093/aje/kwx246

Furberg, H., Kim, Y., Dackor, J., Boerwinkle, E., Franceschini, N., Ardissino, D., . . . Genetics, C. (2010). Genome-Wide Meta-Analyses Identify Multiple Loci Associated with Smoking Behavior. *Nature Genetics, 42*(5), 441-447. doi:10.1038/ng.571

Furnham, A., & Cheng, H. (2017). Childhood Cognitive Ability Predicts Adult Financial Well-Being. *Journal of Intelligence, 5*(1), 3. doi:10.3390/jintelligence5010003

Fürtjes, A. E., Arathimos, R., Coleman, J. R., Cole, J. H., Cox, S. R., Deary, I. J., . . . Ritchie, S. J. (2021). General Dimensions of Human Brain Morphometry Inferred from Genome-Wide Association Data. *bioRxiv*, 2021.2010.2022.465437. doi:10.1101/2021.10.22.465437

Fürtjes, A. E., Cole, J. H., Couvy-Duchesne, B., & Ritchie, S. J. (2023). A Quantified Comparison of Cortical Atlases on the Basis of Trait Morphometricity. *Cortex, 158*, 110-126. doi:10.1016/j.cortex.2022.11.001

Garavan, H., Bartsch, H., Conway, K., Decastro, A., Goldstein, R. Z., Heeringa, S., . . . Zahs, D. (2018). Recruiting the Abcd Sample: Design Considerations and Procedures. *Developmental Cognitive Neuroscience, 32*, 16-22. doi:10.1016/j.dcn.2018.04.004

Gareth, J., Daniela, W., Trevor, H., & Robert, T. (2013). *An Introduction to Statistical Learning: With Applications in R*: Spinger, New York.

Genç, E., Fraenz, C., Schlüter, C., Friedrich, P., Hossiep, R., Voelkle, M. C., . . . Jung, R. E. (2018). Diffusion Markers of Dendritic Density and Arborization in Gray Matter Predict Differences in Intelligence. *Nat Commun, 9*(1), 1905. doi:10.1038/s41467-018-04268-8

Genon, S., Eickhoff, S. B., & Kharabian, S. (2022). Linking Interindividual Variability in Brain Structure to Behaviour. *Nature Reviews Neuroscience, 23*(5), 307-318. doi:10.1038/s41583-022-00584-7

Gillborn, D. (2016). Softly, Softly: Genetics, Intelligence and the Hidden Racism of the New Geneism. *Journal of Education Policy, 31*(4), 365-388. doi:10.1080/02680939.2016.1139189

Gläscher, J., Tranel, D., Paul, L. K., Rudrauf, D., Rorden, C., Hornaday, A., . . . Adolphs, R. (2009). Lesion Mapping of Cognitive Abilities Linked to Intelligence. *Neuron, 61*(5), 681-691. doi:10.1016/j.neuron.2009.01.026

Gottfredson, L. S., & Deary, I. J. (2004). Intelligence Predicts Health and Longevity, but Why? *Current Directions in Psychological Science, 13*(1), 1-4. doi:10.1111/j.0963-7214.2004.01301001.x

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), eaay6690. doi:10.1126/science.aay6690

Grotzinger, A. D., Mallard, T. T., Akingbuwa, W. A., Ip, H. F., Adams, M. J., Lewis, C. M., . . . Nivard, M. G. (2022). Genetic Architecture of 11 Major Psychiatric Disorders at Biobehavioral, Functional Genomic and Molecular Genetic Levels of Analysis. *Nature Genetics, 54*, 548–559. doi:10.1038/s41588-022-01057-4

Grotzinger, A. D., Rhemtulla, M., de Vlaming, R., Ritchie, S. J., Mallard, T. T., Hill, W. D., . . . Tucker-Drob, E. M. (2019). Genomic Structural Equation Modelling Provides Insights into the Multivariate Genetic Architecture of Complex Traits. *Nature Human Behaviour, 3*(5), 513-525. doi:10.1038/s41562-019-0566-x

Gurdasani, D., Carstensen, T., Tekola-Ayele, F., Pagani, L., Tachmazidou, I., Hatzikotoulas, K., . . . Sandhu, M. S. (2015). The African Genome Variation Project Shapes Medical Genetics in Africa. *Nature, 517*(7534), 327-332. doi:10.1038/nature13997

Hansen, J. Y., Shafiei, G., Markello, R. D., Smart, K., Cox, S. M. L., Nørgaard, M., . . . Misic, B. (2022). Mapping Neurotransmitter Systems to the Structural and Functional Organization of the Human Neocortex. *Nature Neuroscience, 25*, 1569–1581. doi:10.1038/s41593-022-01186-3

Hayes, S. M., Nadel, L., & Ryan, L. (2007). The Effect of Scene Context on Episodic Object Recognition: Parahippocampal Cortex Mediates Memory Encoding and Retrieval Success. *Hippocampus, 17*(9), 873-889. doi:10.1002/hipo.20319

Herculano-Houzel, S. (2009). The Human Brain in Numbers: A Linearly Scaled-up Primate Brain. *Frontiers in Human Neuroscience, 3*, 31. doi:10.3389/neuro.09.031.2009

Herculano-Houzel, S. (2016). *The Human Advantage: A New Understanding of How Our Brain Became Remarkable*: The MIT Press, Cambridge, MA.

Herculano-Houzel, S. (2017). Numbers of Neurons as Biological Correlates of Cognitive Capability. *Current Opinion in Behavioral Sciences, 16*, 1-7. doi:10.1016/j.cobeha.2017.02.004

Hilger, K., Winter, N. R., Leenings, R., Sassenhagen, J., Hahn, T., Basten, U., & Fiebach, C. J. (2020). Predicting Intelligence from Brain Gray Matter Volume. *Brain Structure and Function, 225*(7), 2111-2129. doi:10.1007/s00429-020-02113-7

Hill, W. D., Marioni, R. E., Maghzian, O., Ritchie, S. J., Hagenaars, S. P., McIntosh, A. M., . . . Deary, I. J. (2019). A Combined Analysis of Genetically Correlated Traits Identifies 187 Loci and a Role for Neurogenesis and Myelination in Intelligence. *Molecular Psychiatry, 24*(2), 169-181. doi:10.1038/s41380-017-0001-5

Howe, L. J., Nivard, M. G., Morris, T. T., Hansen, A. F., Rasheed, H., Cho, Y., . . . Within Family, C. (2022). Within-Sibship Genome-Wide Association Analyses Decrease Bias in Estimates of Direct Genetic Effects. *Nature Genetics, 54*(5), 581-592. doi:10.1038/s41588-022-01062-7

Hu, L.-t., & Bentler, P. M. (1998). Fit Indices in Covariance Structure Modeling: Sensitivity to Underparameterized Model Misspecification. *Psychological methods, 3*(4), 424. doi:10.1037/1082-989X.3.4.424

Hübel, C., Gaspar, H. A., Coleman, J. R. I., Finucane, H., Purves, K. L., Hanscombe, K. B., . . . Breen, G. (2019). Genomics of Body Fat Percentage May Contribute to Sex Bias in Anorexia Nervosa. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics, 180*(6), 428-438. doi:10.1002/ajmg.b.32709

Humphreys, L. G. (1992). Commentary: What Both Critics and Users of Ability Tests Need to Know. *Psychological Science, 3*(5), 271-275.

Huntenburg, J. M., Bazin, P.-L., & Margulies, D. S. (2018). Large-Scale Gradients in Human Cortical Organization. *Trends in Cognitive Sciences, 22*(1), 21-31. doi:10.1016/j.tics.2017.11.002

Jansen, P. R., Nagel, M., Watanabe, K., Wei, Y., Savage, J. E., de Leeuw, C. A., . . . Posthuma, D. (2020). Genome-Wide Meta-Analysis of Brain Volume Identifies Genomic Loci and Genes Shared with Intelligence. *Nature Communications, 11*(1), 5606. doi:10.1038/s41467-020-19378-5

Jerison, H. (1973). *Evolution of the Brain and Intelligence*: Elsevier, Academic Press, New York.

Jiang, L., Cao, X., Li, T., Tang, Y., Li, W., Wang, J., . . . Li, C. (2016). Cortical Thickness Changes Correlate with Cognition Changes after Cognitive Training: Evidence from a Chinese Community Study. *Frontiers in aging neuroscience, 8*, 118. doi:10.3389/fnagi.2016.00118

Johnson, B. (2013). Third Annual Margaret Thatcher Lecture at the Centre for Policy Studies, London, November. Retrieved from https://www.theguardian.com/politics/video/2013/nov/28/boris-johnson-wealth-envy-inequality-margaret-thatcher-video

Johnston, R., Jones, K., & Manley, D. (2018). Confounding and Collinearity in Regression Analysis: A Cautionary Tale and an Alternative Procedure, Illustrated by Studies of British Voting Behaviour. *Quality & Quantity, 52*(4), 1957-1976. doi:10.1007/s11135-017-0584-6

Jung, R. E., & Haier, R. J. (2007). The Parieto-Frontal Integration Theory (P-Fit) of Intelligence: Converging Neuroimaging Evidence. *Behavioral and Brain Sciences, 30*(2), 135-154. doi:10.1017/S0140525X07001185

Kaufmann, T., van der Meer, D., Doan, N. T., Schwarz, E., Lund, M. J., Agartz, I., . . . Karolinska Schizophrenia, P. (2019). Common Brain Disorders Are Associated with Heritable Patterns of Apparent Aging of the Brain. *Nature Neuroscience, 22*(10), 1617-1623. doi:10.1038/s41593-019-0471-7

Klein, A., & Tourville, J. (2012). 101 Labeled Brain Images and a Consistent Human Cortical Labeling Protocol. *Frontiers in Neuroscience, 6*(171). doi:10.3389/fnins.2012.00171

Lam, M., Awasthi, S., Watson, H. J., Goldstein, J., Panagiotaropoulou, G., Trubetskoy, V., . . . Ripke, S. (2019). Ricopili: Rapid Imputation for Consortias Pipeline. *Bioinformatics, 36*(3), 930-933. doi:10.1093/bioinformatics/btz633

Lee, J. J., McGue, M., Iacono, W. G., & Chow, C. C. (2018). The Accuracy of Ld Score Regression as an Estimator of Confounding and Genetic Correlations in Genome-Wide Association Studies. *Genetic Epidemiology, 42*(8), 783-795. doi:10.1002/gepi.22161

Lee, J. J., McGue, M., Iacono, W. G., Michael, A. M., & Chabris, C. F. (2019). The Causal Influence of Brain Size on Human Intelligence: Evidence from within-Family Phenotypic Associations and Gwas Modeling. *Intelligence, 75*, 48-58. doi:10.1016/j.intell.2019.01.011

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

Lerch, J. P., van der Kouwe, A. J. W., Raznahan, A., Paus, T., Johansen-Berg, H., Miller, K. L., . . . Sotiropoulos, S. N. (2017). Studying Neuroanatomy Using Mri. *Nature Neuroscience, 20*(3), 314-326. doi:10.1038/nn.4501

Lewis, C. M., & Vassos, E. (2022). Polygenic Scores in Psychiatry: On the Road from Discovery to Implementation. *American Journal of Psychiatry, 179*(11), 800-806. doi:10.1176/appi.ajp.20220795

Li, R., Zhang, S., Yin, S., Ren, W., He, R., & Li, J. (2018). The Fronto-Insular Cortex Causally Mediates the Default-Mode and Central-Executive Networks to Contribute to Individual Cognitive Performance in Healthy Elderly. *Human Brain Mapping, 39*(11), 4302-4311. doi:10.1002/hbm.24247

Lindroth, H., Nair, V. A., Stanfield, C., Casey, C., Mohanty, R., Wayer, D., . . . Sanders, R. D. (2019). Examining the Identification of Age-Related Atrophy between T1 and T1 + T2-Flair Cortical Thickness Measurements. *Scientific Reports, 9*(1), 11288. doi:10.1038/s41598-019-47294-2

Linnér, R. K., Biroli, P., Kong, E., Meddens, S. F. W., Wedow, R., Fontana, M. A., . . . Social Science Genetic Association, C. (2019). Genome-Wide Association Analyses of Risk Tolerance and Risky Behaviors in over 1 Million Individuals Identify Hundreds of Loci and Shared Genetic Influences. *Nature Genetics, 51*(2), 245-257. doi:10.1038/s41588-018-0309-3

Littlejohns, T. J., Holliday, J., Gibson, L. M., Garratt, S., Oesingmann, N., Alfaro-Almagro, F., . . . Allen, N. E. (2020). The Uk Biobank Imaging Enhancement of 100,000 Participants: Rationale, Data Collection, Management and Future Directions. *Nature Communications, 11*(1), 2624. doi:10.1038/s41467-020-15948-9

Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological Investigation of the Basis of the Fmri Signal. *Nature, 412*(6843), 150-157. doi:10.1038/35084005

Lorenzo-Seva, U., & Berge, J. M. F. t. (2006). Tucker's Congruence Coefficient as a Meaningful Index of Factor Similarity. *Methodology, 2*(2), 57-64. doi:10.1027/1614-2241.2.2.57

Madole, J. W., Ritchie, S. J., Cox, S. R., Buchanan, C. R., Hernández, M. V., Maniega, S. M., . . . Tucker-Drob, E. M. (2021). Aging-Sensitive Networks within the Human Structural Connectome Are Implicated in Late-Life Cognitive Declines. *Biological Psychiatry, 89*(8), 795-806. doi:10.1016/j.biopsych.2020.06.010

Mak, T. S. H., Porsch, R. M., Choi, S. W., & Sham, P. C. (2018). Polygenic Scores for Uk Biobank Scale Data. *bioRxiv*, 252270. doi:10.1101/252270

Mak, T. S. H., Porsch, R. M., Choi, S. W., Zhou, X., & Sham, P. C. (2017). Polygenic Scores Via Penalized Regression on Summary Statistics. *Genetic Epidemiology, 41*(6), 469-480. doi:10.1002/gepi.22050

Manichaikul, A., Mychaleckyj, J. C., Rich, S. S., Daly, K., Sale, M., & Chen, W.-M. (2010). Robust Relationship Inference in Genome-Wide Association Studies. *Bioinformatics, 26*(22), 2867-2873. doi:10.1093/bioinformatics/btq559

Marek, S., Tervo-Clemmens, B., Calabro, F. J., Montez, D. F., Kay, B. P., Hatoum, A. S., . . . Dosenbach, N. U. F. (2022). Reproducible Brain-Wide Association Studies Require Thousands of Individuals. *Nature, 603*(7902), 654-660. doi:10.1038/s41586-022-04492-9

Mathieson, I., & Scally, A. (2020). What Is Ancestry? *PLOS Genetics, 16*(3), e1008624. doi:10.1371/journal.pgen.1008624

Mbatchou, J., Barnard, L., Backman, J., Marketta, A., Kosmicki, J. A., Ziyatdinov, A., . . . Marchini, J. (2021). Computationally Efficient Whole-Genome Regression for Quantitative and Binary Traits. *Nature Genetics, 53*, 1097–1103. doi:10.1038/s41588-021-00870-7

McCarthy, S., Das, S., Kretzschmar, W., Delaneau, O., Wood, A. R., Teumer, A., . . . the Haplotype Reference, C. (2016). A Reference Panel of 64,976 Haplotypes for Genotype Imputation. *Nature Genetics, 48*(10), 1279-1283. doi:10.1038/ng.3643

Meer, D. v. d., Kaufmann, T., Shadrin, A. A., Makowski, C., Frei, O., Roelfs, D., . . . Dale, A. M. (2021). The Genetic Architecture of Human Cortical Folding. *Science Advances, 7*(51), eabj9446. doi:doi:10.1126/sciadv.abj9446

Menon, V., & Uddin, L. Q. (2010). Saliency, Switching, Attention and Control: A Network Model of Insula Function. *Brain Structure and Function, 214*(5), 655-667. doi:10.1007/s00429-010-0262-0

Mills, M. C., & Rahal, C. (2019). A Scientometric Review of Genome-Wide Association Studies. *Communications Biology, 2*(1), 9. doi:10.1038/s42003-018-0261-x

Nave, G., Jung, W. H., Karlsson Linnér, R., Kable, J. W., & Koellinger, P. D. (2018). Are Bigger Brains Smarter? Evidence from a Large-Scale Preregistered Study. *Psychological Science, 30*(1), 43-54. doi:10.1177/0956797618808470

Nguyen, L., Murphy, K., & Andrews, G. (2019). Cognitive and Neural Plasticity in Old Age: A Systematic Review of Evidence from Executive Functions Cognitive Training. *Ageing Research Reviews, 53*, 100912. doi:10.1016/j.arr.2019.100912

Nieuwboer, H. A., Pool, R., Dolan, C. V., Boomsma, D. I., & Nivard, M. G. (2016). Gwis: Genome-Wide Inferred Statistics for Functions of Multiple Phenotypes. *The American Journal of Human Genetics, 99*(4), 917-927. doi:10.1016/j.ajhg.2016.07.020

Northey, J. M., Cherbuin, N., Pumpa, K. L., Smee, D. J., & Rattray, B. (2018). Exercise Interventions for Cognitive Function in Adults Older Than 50: A Systematic Review with Meta-Analysis. *British Journal of Sports Medicine, 52*(3), 154-160. doi:10.1136/bjsports-2016-096587

O'Brien, L. M., Ziegler, D. A., Deutsch, C. K., Frazier, J. A., Herbert, M. R., & Locascio, J. J. (2011). Statistical Adjustments for Brain Size in Volumetric Neuroimaging Studies: Some Practical Implications in Methods. *Psychiatry Research: Neuroimaging, 193*(2), 113-122. doi:10.1016/j.pscychresns.2011.01.007

Pain, O., Glanville, K. P., Hagenaars, S. P., Selzam, S., Fürtjes, A. E., Gaspar, H. A., . . . Lewis, C. M. (2021). Evaluation of Polygenic Prediction Methodology within a Reference-Standardized Framework. *PLOS Genetics, 17*(5), e1009021. doi:10.1371/journal.pgen.1009021

Pandya, D. N., & Yeterian, E. H. (1985). Architecture and Connections of Cortical Association Areas. In Peters, A. & Jones, E. G. (Eds.), *Association and Auditory Cortices* (pp. 3-61). Boston, MA: Springer US.

Peelle, J. E., Cusack, R., & Henson, R. N. A. (2012). Adjusting for Global Effects in Voxel-Based Morphometry: Gray Matter Decline in Normal Aging. *NeuroImage, 60*(2), 1503-1516. doi:10.1016/j.neuroimage.2011.12.086

Pietschnig, J., Penke, L., Wicherts, J. M., Zeiler, M., & Voracek, M. (2015). Meta-Analysis of Associations between Human Brain Volume and Intelligence Differences: How Strong Are They and What Do They Mean? *Neuroscience & Biobehavioral Reviews, 57*, 411-432. doi:10.1016/j.neubiorev.2015.09.017

Plomin, R., Haworth, C. M. A., Meaburn, E. L., Price, T. S., & Davis, O. S. P. (2013). Common DNA Markers Can Account for More Than Half of the Genetic Influence on Cognitive Abilities. *Psychological Science, 24*(4), 562-568. doi:10.1177/0956797612457952

Power, Jonathan D., Cohen, Alexander L., Nelson, Steven M., Wig, Gagan S., Barnes, Kelly A., Church, Jessica A., . . . Petersen, Steven E. (2011). Functional Network Organization of the Human Brain. *Neuron, 72*(4), 665-678. doi:10.1016/j.neuron.2011.09.006

Purves, D., Augustine, G., & Fitzpatrick, D. (2001). *'The Premotor Cortex', in Neuroscience* (2nd edition ed. Vol. 682): Sinauer Associates, New York.

Raz, N., Ghisletta, P., Rodrigue, K. M., Kennedy, K. M., & Lindenberger, U. (2010). Trajectories of Brain Aging in Middle-Aged and Older Adults: Regional and Individual Differences. *NeuroImage, 51*(2), 501-511. doi:10.1016/j.neuroimage.2010.03.020

Reardon, P. K., Seidlitz, J., Vandekar, S., Liu, S., Patel, R., Park, M. T. M., . . . Raznahan, A. (2018). Normative Brain Size Variation and Brain Shape Diversity in Humans. *Science, 360*(6394), 1222-1227. doi:10.1126/science.aar2578

Resnick, S. M., Pham, D. L., Kraut, M. A., Zonderman, A. B., & Davatzikos, C. (2003). Longitudinal Magnetic Resonance Imaging Studies of Older Adults: A Shrinking Brain. *The Journal of Neuroscience, 23*(8), 3295-3301. doi:10.1523/jneurosci.23-08-03295.2003

Ritchie, S. J., & Tucker-Drob, E. M. (2018). How Much Does Education Improve Intelligence? A Meta-Analysis. *Psychological Science, 29*(8), 1358-1369. doi:10.1177/0956797618774253

Rosseel, Y. (2012). Lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software, 48*(2), 1-36. doi:10.18637/jss.v048.i02

Salthouse, T. A. (2010). Selective Review of Cognitive Aging. *Journal of the International Neuropsychological Society, 16*(5), 754-760. doi:10.1017/S1355617710000706

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. *Nature Genetics, 50*(7), 912-919. doi:10.1038/s41588-018-0152-6

Schaefer, A., Kong, R., Gordon, E. M., Laumann, T. O., Zuo, X.-N., Holmes, A. J., . . . Yeo, B. T. T. (2017). Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity Mri. *Cerebral Cortex, 28*(9), 3095-3114. doi:10.1093/cercor/bhx179

Scheel, A. M., Tiokhin, L., Isager, P. M., & Lakens, D. (2021). Why Hypothesis Testers Should Spend Less Time Testing Hypotheses. *Perspectives on Psychological Science, 16*(4), 744-755. doi:10.1177/1745691620966795

Schoeler, T., Speed, D., Porcu, E., Pirastu, N., Pingault, J.-B., & Kutalik, Z. (2022). Correction for Participation Bias in the Uk Biobank Reveals Non-Negligible Impact on Genetic Associations and Downstream Analyses. *bioRxiv*, 2022.2009.2028.509845. doi:10.1101/2022.09.28.509845

Schwarzer, G., Carpenter, J., & Rücker, G. (2015). *Meta-Analysis with R (Use-R!)*: Springer International Publishing, Switzerland.

Scoville, W. B., & Milner, B. (1957). Loss of Recent Memory after Bilateral Hippocampal Lesions. *J Neurol Neurosurg Psychiatry, 20*(1), 11-21. doi:10.1136/jnnp.20.1.11

Sham, P. C., & Purcell, S. (2001). Equivalence between Haseman-Elston and Variance-Components Linkage Analyses for Sib Pairs. *The American Journal of Human Genetics, 68*(6), 1527-1532. doi:10.1086/320593

Shen, E. H., Overly, C. C., & Jones, A. R. (2012). The Allen Human Brain Atlas: Comprehensive Gene Expression Mapping of the Human Brain. *Trends Neurosci, 35*(12), 711-714. doi:10.1016/j.tins.2012.09.005

Shi, H., Burch, K. S., Johnson, R., Freund, M. K., Kichaev, G., Mancuso, N., . . . Pasaniuc, B. (2020). Localizing Components of Shared Transethnic Genetic Architecture of Complex Traits from Gwas Summary Data. *The American Journal of Human Genetics, 106*(6), 805-817. doi:10.1016/j.ajhg.2020.04.012

Shin, J., Ma, S., Hofer, E., Patel, Y., Vosberg, D. E., Tilley, S., . . . Gottesman, R. (2020). Global and Regional Development of the Human Cerebral Cortex: Molecular Architecture and Occupational Aptitudes. *Cerebral Cortex, 30*(7), 4121-4139. doi:10.1093/cercor/bhaa035

Singhal, I., Soni, A. K., & Srinivasan, N. (2020). Default(Y) Mode Network: Important Regions of Dmn Do Not Survive Alterations in Flip Angles. *bioRxiv*, 2020.2007.2009.196568. doi:10.1101/2020.07.09.196568

Smith, S., Alfaro-Almagro, F., & Miller, K. (2020). Uk Biobank Brain Imaging Documentation. Retrieved from https://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/brain_mri.pdf

Smith, S. M., Douaud, G., Chen, W., Hanayik, T., Alfaro-Almagro, F., Sharp, K., & Elliott, L. T. (2020). Enhanced Brain Imaging Genetics in Uk Biobank. *bioRxiv*, 2020.2007.2027.223545. doi:10.1101/2020.07.27.223545

Sodini, S. M., Kemper, K. E., Wray, N. R., & Trzaskowski, M. (2018). Comparison of Genotypic and Phenotypic Correlations: Cheverud's Conjecture in Humans. *Genetics, 209*(3), 941-948. doi:10.1534/genetics.117.300630

Soheili-Nezhad, S., Beckmann, C. F., & Sprooten, E. (2021). Independent Genomic Sources of Brain Structure and Function. *bioRxiv*, 2021.2001.2006.425535. doi:10.1101/2021.01.06.425535

Spearman, C. (1904). "General Intelligence," Objectively Determined and Measured. *American Journal of Psychology, 15*, 201-292.

Sporns, O. (2011). The Human Connectome: A Complex Network. *Annals of the New York Academy of Sciences, 1224*(1), 109-125. doi:10.1111/j.1749-6632.2010.05888.x

Sridharan, D., Levitin, D. J., & Menon, V. (2008). A Critical Role for the Right Fronto-Insular Cortex in Switching between Central-Executive and Default-Mode Networks. *Proceedings of the National Academy of Sciences, 105*(34), 12569-12574. doi:10.1073/pnas.0800005105

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 medicine, 12*(3), e1001779-e1001779. doi:10.1371/journal.pmed.1001779

Taylor, A. M., Pattie, A., & Deary, I. J. (2018). Cohort Profile Update: The Lothian Birth Cohorts of 1921 and 1936. *International journal of epidemiology, 47*(4), 1042-1042r. doi:10.1093/ije/dyy022

Tucker-Drob, E. M. (2011). Neurocognitive Functions and Everyday Functions Change Together in Old Age. *Neuropsychology, 25*(3), 368-377. doi:10.1037/a0022348

Tucker-Drob, E. M. (2019). Cognitive Aging and Dementia: A Life-Span Perspective. *Annual Review of Developmental Psychology, 1*(1), 177-196. doi:10.1146/annurev-devpsych-121318-085204

Tucker, L. R. (1951). *A Method for Synthesis of Factor Analysis Studies*. Retrieved from

Turkheimer, E. (2000). Three Laws of Behavior Genetics and What They Mean. *Current Directions in Psychological Science, 9*(5), 160-164. doi:10.1111/1467-8721.00084

Tyrrell, J., Zheng, J., Beaumont, R., Hinton, K., Richardson, T. G., Wood, A. R., . . . Tilling, K. (2021). Genetic Predictors of Participation in Optional Components of Uk Biobank. *Nature Communications, 12*(1), 886. doi:10.1038/s41467-021-21073-y

Uddin, L. Q., Yeo, B. T. T., & Spreng, R. N. (2019). Towards a Universal Taxonomy of Macro-Scale Functional Human Brain Networks. *Brain Topography, 32*(6), 926-942. doi:10.1007/s10548-019-00744-6

van der Meer, D., Frei, O., Kaufmann, T., Shadrin, A. A., Devor, A., Smeland, O. B., . . . Dale, A. M. (2020). Understanding the Genetic Determinants of the Brain with Mostest. *Nature Communications, 11*(1), 3512. doi:10.1038/s41467-020-17368-1

van Rheenen, W., Peyrot, W. J., Schork, A. J., Lee, S. H., & Wray, N. R. (2019). Genetic Correlations of Polygenic Disease Traits: From Theory to Practice. *Nature Reviews Genetics, 20*(10), 567-581. doi:10.1038/s41576-019-0137-z

Visscher, P. M., Hemani, G., Vinkhuyzen, A. A. E., Chen, G.-B., Lee, S. H., Wray, N. R., . . . Yang, J. (2014). Statistical Power to Detect Genetic (Co)Variance of Complex Traits Using Snp Data in Unrelated Samples. *PLOS Genetics, 10*(4), e1004269. doi:10.1371/journal.pgen.1004269

Visscher, P. M., Wray, N. R., Zhang, Q., Sklar, P., McCarthy, M. I., Brown, M. A., & Yang, J. (2017). 10 Years of Gwas Discovery: Biology, Function, and Translation. *The American Journal of Human Genetics, 101*(1), 5-22. doi:10.1016/j.ajhg.2017.06.005

Vittinghoff, E., Glidden, D. V., Shiboski, S. C., & McCulloch, C. E. (2006). *Regression Methods in Biostatistics: Linear, Logistic, Survival, and Repeated Measures Models*: 2nd ed. New York: Springer; 2012.

Wall, J. D., Stawiski, E. W., Ratan, A., Kim, H. L., Kim, C., Gupta, R., . . . GenomeAsia, K. C. (2019). The Genomeasia 100k Project Enables Genetic Discoveries across Asia. *Nature, 576*(7785), 106-111. doi:10.1038/s41586-019-1793-z

Warren, H. R., Evangelou, E., Cabrera, C. P., Gao, H., Ren, M., Mifsud, B., . . . Understanding Society Scientific, g. (2017). Genome-Wide Association Analysis Identifies Novel Blood Pressure Loci and Offers Biological Insights into Cardiovascular Risk. *Nature Genetics, 49*(3), 403-415. doi:10.1038/ng.3768

Wechsler, D. (2010). *Wms-Iv Administration and Scoring Manual*: London: Pearson.

Westlin, C., Theriault, J. E., Katsumi, Y., Nieto-Castanon, A., Kucyi, A., Ruf, S. F., . . . Brooks, D. H. (2023). Improving the Study of Brain-Behavior Relationships by Revisiting Basic Assumptions. *Trends in Cognitive Sciences*. doi:10.1016/j.tics.2022.12.015

Williams, C. M., Peyre, H., & Ramus, F. (2022). Brain Volumes, Thicknesses, and Surface Areas as Mediators of Genetic Factors and Childhood Adversity on Intelligence. *bioRxiv*, 2022.2009.2008.507068. doi:10.1101/2022.09.08.507068

Winkler, A. M., Kochunov, P., Blangero, J., Almasy, L., Zilles, K., Fox, P. T., . . . Glahn, D. C. (2010). Cortical Thickness or Grey Matter Volume? The Importance of Selecting the Phenotype for Imaging Genetics Studies. *NeuroImage, 53*(3), 1135-1146. doi:10.1016/j.neuroimage.2009.12.028

Wray, N. R. (2005). Allele Frequencies and the R2 Measure of Linkage Disequilibrium: Impact on Design and Interpretation of Association Studies. *Twin Research and Human Genetics, 8*(2), 87-94. doi:10.1375/twin.8.2.87

Wu, K., Taki, Y., Sato, K., Qi, H., Kawashima, R., & Fukuda, H. (2013). A Longitudinal Study of Structural Brain Network Changes with Normal Aging. *Frontiers in Human Neuroscience, 7*. doi:10.3389/fnhum.2013.00113

Wysocki, A. C., Lawson, K. M., & Rhemtulla, M. (2022). Statistical Control Requires Causal Justification. *Advances in Methods and Practices in Psychological Science, 5*(2), 25152459221095823. doi:10.1177/25152459221095823

Xue, G., Chen, C., Lu, Z. L., & Dong, Q. (2010). Brain Imaging Techniques and Their Applications in Decision-Making Research. *Xin Li Xue Bao, 42*(1), 120-137. doi:10.3724/sp.J.1041.2010.00120

Yang, J., Lee, S. H., Goddard, M. E., & Visscher, P. M. (2011). Gcta: A Tool for Genome-Wide Complex Trait Analysis. *American journal of human genetics, 88*(1), 76-82. doi:10.1016/j.ajhg.2010.11.011

Yengo, L., Vedantam, S., Marouli, E., Sidorenko, J., Bartell, E., Sakaue, S., . . . Bisgaard, H. (2022). A Saturated Map of Common Genetic Variants Associated with Human Height. *Nature, 610*(7933), 704-712. doi:10.1038/s41586-022-05275-y

Yengo, L., Yang, J., & Visscher, P. M. (2018). Expectation of the Intercept from Bivariate Ld Score Regression in the Presence of Population Stratification. *bioRxiv*, 310565. doi:10.1101/310565

Yeo, B. T. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., . . . Buckner, R. L. (2011). The Organization of the Human Cerebral Cortex Estimated by Intrinsic Functional Connectivity. *Journal of Neurophysiology, 106*(3), 1125-1165. doi:10.1152/jn.00338.2011

Young, A. I., Frigge, M. L., Gudbjartsson, D. F., Thorleifsson, G., Bjornsdottir, G., Sulem, P., . . . Kong, A. (2018). Relatedness Disequilibrium Regression Estimates Heritability without Environmental Bias. *Nature Genetics, 50*(9), 1304-1310. doi:10.1038/s41588-018-0178-9

Zhao, B., Ibrahim, J. G., Li, Y., Li, T., Wang, Y., Shan, Y., . . . Zhu, H. (2019). Heritability of Regional Brain Volumes in Large-Scale Neuroimaging and Genetic Studies. *Cerebral Cortex, 29*(7), 2904-2914. doi:10.1093/cercor/bhy157

Zhao, B., Luo, T., Li, T., Li, Y., Zhang, J., Shan, Y., . . . Neurocognition Genetics. (2019). Genome-Wide Association Analysis of 19,629 Individuals Identifies Variants Influencing Regional Brain Volumes and Refines Their Genetic Co-Architecture with Cognitive and Mental Health Traits. *Nature Genetics, 51*(11), 1637-1644. doi:10.1038/s41588-019-0516-6

Zhao, Q., Voon, V., Zhang, L., Shen, C., Zhang, J., & Feng, J. (2022). The Abcd Study: Brain Heterogeneity in Intelligence During a Neurodevelopmental Transition Stage. *Cerebral Cortex, 32*(14), 3098-3109. doi:10.1093/cercor/bhab403

Appendices

Contents of Appendices

Supplementary Material for Chapter 4

Supplementary Methods

1. Study Design

1.1. UK Biobank data

The UKB study received ethical approval by the Research Ethics Committee. All participants signed informed consent for their data to be analysed. Access to phenotypic and genetic UK Biobank data was granted through the approved application 18177. Magnetic resonance imaging (MRI) data was collected by the UK Biobank study with identical hardware and software in Manchester, Newcastle, and Reading. Brain volumetric phenotypes were pre-processed by an imaging-pipeline developed and executed on behalf of UK Biobank (Alfaro-Almagro et al., 2018). More information on T1 processing can be found in the UK Biobank online documentation (Smith, Alfaro-Almagro, et al., 2020). Briefly, cortical surfaces were modelled using FreeSurfer, and 33 volumes were extracted based on Desikan-Killiany surface templates (Desikan et al., 2006) (UKB does not provide frontal pole measures due to high missingness); 8 subcortical areas were derived using FreeSurfer aeseg tools (Fischl et al., 2002). Volumetric measures (mm^3) have been generated in each participant's native space. We used 83 available imaging-derived phenotypes (IDPs) of cortical and subcortical grey-matter volumes in regions of interest spanning the whole brain (UK Biobank category 192 & 190; STable 1). We assume the IDPs to be normally-distributed.

1.2. Phenotypic quality control

Excluding participants who withdrew consent, we considered 41,776 participants with non-missing T1-weighted IDPs that had been processed in conjunction with T2-weighted FLAIR (UK Biobank field ID 26500) where available. Using both T1 and T2 measures ensures more precise cortical segmentation (Lindroth et al., 2019). Extreme outliers outside of 4 standard deviations from the mean were excluded, which ranged between 41,686 to 41,769 available participants between different IDP. 381 participants were excluded as they self-reported non-European ethnicity. (In GWAS it is necessary to include unrelated, ethnically homogeneous participants. Different ethnicities display different linkage disequilibrium structures across the genome, and the analyses conducted here would be uninterpretable if conducted across ethnicities. To maximise sample size, we analyse European participants, as they represent the largest ethnically homogeneous group in the UKB cohort). Across the 83 brain volumes and the covariates, this phenotypic quality control resulted in 39,947 complete cases, for whom the following genetic quality control steps were performed.

1.3. Genetic quality control

Out of the 39,947 UK Biobank participants, genetic data were available for 38,957 participants. Genetic data was quality controlled by UK Biobank and were downloaded from the full release (Bycroft et al., 2018). We applied additional quality control as previously described in Coleman et al. (2020) using PLINK2 (Chang et al., 2015). 38,038 participants were of European ancestry according to 4-means clustering on the first two genetic principal components available through UK Biobank (Warren

et al., 2017). Of those participants, we removed 72 due to quality assurance issues reported by UK Biobank and 204 participants due to high rates of missingness (2% missingness). To obtain a sample of unrelated individuals, 956 participants were removed using the greedyRelated algorithm (KING $r < 0.044$ (Manichaikul et al., 2010)). The algorithm is "greedy" because it maximises sample size; for example, it removes the child in a parent-child-trio. Finally, 28 participants were removed because genetic sex did not align with self-reported sex, resulting in a total of 36,778 participants (STable 10). Genetic sex was identified based on measures of X-chromosome homozygosity (F_X ; removal of participants with $F_X < 0.9$ for phenotypic males, $F_X > 0.5$ for phenotypic females). The final sample ($N = 36,778$) included 19,888 females (54 %) and had an average age of 63.3 years at the neuroimaging visit (range from 40.0 to 81.8 years).

Out of 805,426 available directly genotyped variants, 104,771 were removed for high rates of missing genotype data (> 98%). 103,137 variants were removed due to a minimum allele frequency of 0.01, and 9,935 variants were removed as they failed the Hardy-Weinberg exact test (p-value = 10^{-8}). After excluding 16,326 variants on the sex chromosomes and those with chromosome labels larger than 22, we obtained a final sample of 571,257 directly genotyped SNPs. Imputed genotype data was obtained by UK Biobank with reference to the Haplotype Reference Consortium (McCarthy et al., 2016), and we filtered them for a minor allele frequency of above 0.01 and an IMPUTE INFO metric of above 0.4.

1.4. Measures of cognitive performance

In this study, we considered GWAS summary statistics of performance in seven cognitive tests by de la Fuente et al. (2021) that were calculated with between 11,263

and 331,679 participants for each test. We considered the HapMap 3 reference SNPs with the MHC regions removed. These seven cognitive performance tests were assessed on a touchscreen computer: *Matrix Pattern Completion task* for nonverbal reasoning, *Memory – Pairs Matching Test* for memory, *Reaction Time* for perceptual motor speed, *Symbol Digit Substitution Task* for information processing speed, *Trail Making Test – B* and *Tower Rearranging Task* for executive functioning, and *Verbal Numerical Reasoning Test* for verbal and numeric problem solving, or fluid intelligence. Despite the non-standard and unsupervised delivery of assessment, these cognitive tests demonstrate strong concurrent validity compared with standard reference tests (r = .83) and good test-retest reliability (Pearson r range for different cognitive tests = 0.4–0.78) (Fawns-Ritchie & Deary, 2020).

2. *Statistical analysis*

2.1. GWAS summary statistics calculation

The 83 regional brain volumes described in Supplementary Methods 1.1. were quality controlled as described in Supplementary Methods 1.2-1.3. GWAS summary statistics for these 83 regional brain volumes (continuous variables) were calculated using REGENIE (Mbatchou et al., 2021), which fits polygenic effects in a linear mixed model using Ridge regression. The REGENIE pipeline is split into two steps: First, blocks of directly genotyped SNPs are used to fit a cross-validated whole-genome regression model using Ridge regression, to determine the amount of phenotypic variance explained by genetic effects. Second, the association between the phenotype and imputed genetic variants is calculated conditional upon Ridge regression

predictions from the first step. Proximal contamination is circumvented by using a leave-one-chromosome-out scheme.

Covariates included in the GWAS analyses were *age at neuroimaging visit*, *sex*, *genotyping batch*, and *40 genetic principal components* as provided by UK Biobank. We also derived the variables *time of year*, *head position*, and *acquisition site,* but excluded them from our set of GWAS covariates because they were not associated with the brain volumes at the pre-registered arbitrary cut-off of $r \leq .10$ (STable 9), and therefore explained less than 1% of the phenotype variance. Note that, in contrast to other existing brain-volume GWAS in UK Biobank (e.g., Smith, Douaud, et al., 2020), our analyses were conducted *without* controlling for brain size (or any other global brain measure such as total grey-matter volume or intracranial volume). This is because we wanted to represented total variance associated with regional volumes, rather than capturing variance that persisted above and beyond variance that mapped onto global measures (also discussed in the introduction of the manuscript). Genetic correlations calculated relative to such global measures are known to attenuate genetic correlations among volumes, as well as with other traits such as cognitive abilities (de Vlaming et al., 2021). In the context of this study, we aim to model general dimensions of variance shared between brain volumes which will probably closely covary with overall brain size. Attenuated genetic correlations would hide major dimensions of variance across genetic brain networks, because much of the variance shared between volumes overlaps with variance indexed by brain size and would therefore not tag general dimensions of shared genetic variance between brain volumes. This variance is of interest because general intelligence yields global rather than a region-specific associations with grey matter volume (Hilger et al., 2020). Equally, aging affects the whole brain rather than individual regions (Cole et al., 2019).

2.2. Genetic and phenotypic correlation matrices between brain volumes.

To derive dimensions of shared morphometry across brain volumes, we calculated both a phenotypic and a genetic correlation matrix from 83 grey-matter volumes. Phenotypic regional brain volumes were residualised for age at neuroimaging visit and sex, and then used to estimate a phenotypic correlation matrix through Pearson's correlations with complete pairwise observations. The genetic correlation matrix was inferred through LDSC (Bulik-Sullivan, Finucane, et al., 2015), a technique quantifying shared genome-wide polygenic effects between traits using GWAS summary statistics. Cross-trait LDSC regresses the product of effect sizes in two GWAS onto linkage disequilibrium (LD) scores, indicating how correlated a genetic variant is with its neighbouring variants (Bulik-Sullivan, Finucane, et al., 2015). The slope indexes the genetic correlation, while the intercept captures signal uncorrelated with LD, such as population stratification, environmental confounding, and sample overlap (Lee, McGue, et al., 2018). Interregional genetic correlations ranged from $r_g =$ -0.08 (SE = 0.013) between right frontal pole and left pallidum, to r_g = 0.87 (SE = 0.017) between left middle temporal and left inferior temporal (*Fig.3B, SFig.1*). Corresponding standard errors ranged between 0.01 and 0.03 (mean = 0.014; SD = 0.002).

2.3. Comparing phenotypic and genetic correlations

To quantify the relationship between phenotypic and genetic correlations, we estimated the correlation between 3403 phenotypic and 3403 genetic interregional brain volume correlations ($\frac{83(83-1)}{2}$ = 3403 correlations between 83 volumes). We calculated genetic correlation matrices indicating genetic overlap between the 83 volumes using linkage disequilibrium score regression (LDSC; Bulik-Sullivan,

Finucane, et al., 2015) as implemented in the GenomicSEM software (Grotzinger et al., 2019; *Fig.2.2*). Genetic between-volume correlations are displayed in *SFig.1-10*. To obtain comparable indices of phenotypic covariance, we ran PCA on a phenotypic correlation matrix obtained from the same brain volume variables used to calculate GWAS.

2.4. Principal component analysis (PCA) of genetic and phenotypic correlation matrices

PCA was applied to the phenotypic and genetic correlation matrices indicating genetic overlap between brain volumes described above to obtain their respective first principal component (PC). The first PC1 represents an underlying dimension of common structural sharing across regional volumes, which we refer to as general dimensions of shared morphometry throughout this manuscript. PC1 loadings were calculated for all volumes in the whole brain, as well as theoretical grouping of fewer volumes thought to reside in smaller canonical networks to quantify contributions of regional volumes to this either brain-wide, or network-specific dimension of shared morphometry. Phenotypically, this was done using the eigen function in R, and genetically this was done following the pipeline outlined in detail in the main Methods section (*Fig.2*).

2.5. Genome-wide PCs of morphometry across the whole brain and canonical networks

To statistically represent volumetric PCs on a genome-wide level, we averaged genome-wide SNP effects contributed by multiple grey-matter volume GWAS summary statistics, weighted by their respective (region-specific) PC1 loadings. We obtained one set of GWAS summary statistics showing SNP associations of a genetic

principal component underlying multiple GWAS phenotypes derived from samples of unknown degrees of overlap. SNP effects were calculated by adapting existing software for genome-wide multivariate meta-analysis by Baselmans et al. (2019). More details are in the main Methods section (*Fig.2*).

We tested and validated this novel approach in an independent set of GWAS summary statistics of four risky behaviours (Linnér et al., 2019). In addition to the risky behaviour GWAS, another set of summary statistics is available for a phenotypic PC1 underlying these risky behaviour phenotypes that the authors had calculated phenotypically before running GWAS analyses. We compared these phenotypic PC1 GWAS summary statistics provided by Linnér et al. (2019) with summary statistics for a *genetic* PC1 underlying the four risky behaviours GWAS that we calculated using Genomic PCA. They were genetically correlated at r_g = 0.99 (*SE* = 0.037) confirming that our method captures the same signal as can be obtained from phenotypic PCs, by simply relying on publicly available GWAS data. For details of the analysis and code refer to: https://annafurtjes.github.io/genomicPCA/ .

2.6. Parallel analysis

We tested whether genetic PCs underlying multiple volumes in the whole brain, and canonical brain networks explained more variance than expected by chance, that is, whether they explained more than 95% of their corresponding PCs generated under a simulated null correlation matrix. We developed a version of parallel analysis to generate null distributions of eigenvalues by simulating null correlation matrices sampled from a diagonal population correlation matrix, where the multivariate sampling distribution is specified to take the form of the sampling distribution of the

standardised empirical genetic correlation matrix (the V_{STD} matrix, as estimated using GenomicSEM (Grotzinger et al., 2019)). This sampling correlation matrix serves as an index of the precision of the elements in the empirical genetic covariance matrix (i.e., heritabilities and co-heritabilities across traits) and the sampling dependencies among these when generating the random null models. We specified 1,000 replications to simulate the null correlation matrices and used a 95% threshold for distinguishing true eigenvalues from noise.

2.7. Simulation of networks with randomly included brain volumes.

We performed an additional sensitivity analysis simulating networks with randomly included brain volumes, to determine whether shared structural variance relied on network membership, or arose through phenotypic properties common to all regional brain volumes. To compare explained variances between canonical networks and random networks, we quantified the expected explained variance in random networks by randomly sampling regions 800 times each, for different numbers of included volumes (because networks including fewer volumes generally tend to explain a larger percentage of variance, as larger networks are more heterogeneous). That is, simulations were run for 8, 10, 12, 16, 30, and 36 included regions, to obtain a distribution for each networks size to compare the corresponding network's explained variance to. We reported the mean explained variance by PCs for networks with randomly included volumes and a 95% confidence interval. Comparisons between explained variances for random and empirical networks were done for the same number of included volumes.

2.8. Comparing phenotypic and genetic loadings onto PC1 underlying 83 brain-wide volumes

To compare whether genetic correlation structures of regional brain morphometry resembled the phenotypic correlation structure of the same regions, we calculated an un-standardised linear regression with a vector of 83 phenotypic whole-brain PC1 loadings as the dependent variable, and a vector containing 83 genetic whole-brain PC1 loadings as the independent variable. We calculated the Tucker congruence coefficient to quantify the relative similarity between the two sets of PC1 loadings independent of their absolute magnitude. The coefficient is insensitive to scalar multiplication (Tucker, 1951).

2.9. Correlation between genetic loadings onto PC1 underlying 83 brain-wide volumes with age sensitivity

Pearson's correlations between 83 phenotypic grey-matter volumes and age at neuroimaging visit were calculated to quantify cross-sectional age-volume-correlations for each of the 83 brain volumes. These age-volume correlations are referred to as *age sensitivity* throughout the manuscript. We estimated the correlation between a vector containing indices of age sensitivity and (1) a vector of *genetic* whole-brain PC1 loadings, and for comparison (2) a vector of *phenotypic* whole-brain PC1 loadings.

2.10. Genetic correlation between general dimensions of shared morphometry across the whole-brain and brain age

Using cross-trait LDSC (Bulik-Sullivan, Finucane, et al., 2015), we calculated a genetic correlation between genetic PC1 underlying brain-wide volumes (as created by

Genomic PCA) and the *brain age* gap. We used the brain age GWAS summary statistics by Kaufmann et al. (2019). Brain age is a phenotype based on individual-level predictions of how much older (or younger) an individual's brain appears, relative to their chronological age. It is estimated using parameters characterising the relationship between age and structural neuroimaging measures (volume, thickness, and surface area) that were tuned using machine learning in an independent sample. The final brain age phenotype indexed in the GWAS was calculated as the difference between participants chronological age and their age as predicted based on structural brain characteristics.

2.11. Genomic SEMs of genetic correlations between structural brain networks and a factor of general cognitive ability

We assessed genetic correlations between genetic PC1s underlying canonical brain networks and general cognitive ability using GenomicSEM (Grotzinger et al., 2019) (*Fig.5*). In GenomicSEM (Grotzinger et al., 2019), the genetic general cognitive ability factor was modelled from seven cognitive ability GWAS summary statistics (described in Supplementary Methods 1.4.), and the genetic correlation between genetic general cognitive ability and genetic brain PCs was estimated with diagonally weighted least squares. To quantify model fit, we reported default fit indices calculated by the GenomicSEM package: $\chi 2$ values, the Akaike Information Criterion (AIC), the Comparative Fit Index (CFI) and the Standardised Root Mean Square Residuals (SRMR). The multiple testing burden was addressed by correcting *p*-values from the genetic correlations for multiple testing with a false-positive discovery rate of 5% (Benjamini & Hochberg, 1995).

We preregistered that we would test for significant differences in correlation magnitudes between the networks whose underlying PC1 yielded a significant association with general cognitive abilities. Because we hypothesised a particularly strong association for the central executive network, we planned to perform this comparison between the central executive and all other networks, to reduce the multiple testing burden. We fitted two GenomicSEM models in which correlation magnitudes between general cognitive ability and both the central executive and another network were either freely estimated, or they were forced to be the same. A significant decrease in model fit (as indicated as the difference in $\chi 2$) between the freely estimated model and the constrained model (df = 1) would indicate that there likely are differences in correlation magnitudes between the networks in how strongly they correlate with general cognitive ability (*SFig.22*). We found no evidence that the PC1 underlying the central executive network was any more genetically correlated with cognitive ability than the PC1 underlying any other network (STable 5).

Additionally, we assessed whether the PC1 underlying the central executive network was disproportionately genetically correlated with general cognitive ability considering its small size (i.e., few included volumes). Similar to the approach described above, we fitted two models: One, in which we freely estimate the correlation between the genetic PC1 underlying the central executive and general cognitive ability, and the correlation between genetic PC1 underlying another network and general cognitive ability. We then divided the correlation magnitude by the number of regions included in the network (i.e., magnitude was divided by 8 for the central executive network, it was divided by 16 for the default mode, by 36 for the P-FIT etc.). The second model had the same set up, but we forced the adjusted correlations of the two networks with cognitive ability to be equal (e.g., $r_{central\ executive}$ / 8 == $r_{default}$ / 16). We assessed

whether there was a significant difference in χ^2 model fit between these two models. As above, a significant decrease in model fit between the freely estimated model and the constrained model (df = 1) would indicate that there likely are differences in relative correlation magnitudes (i.e., magnitudes adjusted for network sizes). Based on previous findings, we expected the relative magnitude for the central executive network to be significantly larger than the relative magnitude for any other network. Even when accounting for network size, we found no evidence that the genetic PC1 underlying the central executive network was any more genetically correlated with cognitive ability than a genetic PC1 underlying any other canonical network (STable 6).

To probe whether any specific cognitive ability might have driven the genetic correlation between PCs underlying brain networks and general cognitive ability, we reported genetic correlations between the significant networks and three specific cognitive abilities: (1) *Matrix Pattern Completion task* to represent nonverbal reasoning, (2) *Memory – Pairs Matching Test* to represent memory, and (3) *Symbol Digit Substitution Task* to represent information processing speed. Reducing the analyses to only three consistent and representative cognitive measures reduced the burden of multiple testing. Matrix Pattern Completion consistently yielded the strongest genetic correlations with PCs underlying all the brain networks (mean r_g across different networks = 0.18). Genetic correlations for Symbol Digit Substitution Task were slightly smaller (mean r_g = 0.12), followed by Memory which had the lowest average correlations (mean r_g = 0.09; STable 7).

We also calculated Q_{trait} heterogeneity indices (Grotzinger et al., 2022) to evaluate whether the general cognitive ability factor that we fit in the models above accounts well for the specific cognitive abilities. To this end, we compared the fit of

two models for each network as displayed in *SFig.23*. One model allows for independent associations between the seven cognitive traits, and both general cognitive ability and the PC1 underlying the considered brain network. The second model forces the association between the seven cognitive traits and the PC1 underlying the brain network to go through the general cognitive ability factor. We obtained χ^2 fit statistics for both models and tested their difference for statistical significance ($\Delta \chi^2 \neq 0$; df = 6). Non-significant results ($p > 0.05/10$) would suggest that genetic associations between cognitive abilities and brain networks are likely general and act through a factor of general cognitive ability. We obtained non-significant Q_{trait} heterogeneity indices (Grotzinger et al., 2022) for all brain networks, demonstrating that the general cognitive ability factor accounted well for the patterns of association between specific cognitive abilities and the brain networks (*SFig.23*).

Data and code availability. Access to phenotypic and genetic UK Biobank data was granted through the approved application 18177. We have made the 83 GWAS summary statistics of regional volumes available at the GWAS catalogue (https://www.ebi.ac.uk/gwas/).

GWAS summary statistics for the seven cognitive traits by (de la Fuente et al., 2021) were downloaded at https://datashare.ed.ac.uk/handle/10283/3756.

The pre-registration for this analysis can be found online (https://osf.io/7n4qj).

Full analysis code including results for this study are available at https://annafurtjes.github.io/Genetic_networks_project.

Further discussion of study limitations

All limitations that apply to LDSC methodology are relevant to this study. It is known that LDSC estimates can be downwardly biased (Lee, McGue, et al., 2018) with larger standard errors as compared with genome-wide complex trait analysis, for example (van Rheenen et al., 2019). However, there is no available software allowing genomic structural equation modelling using any of these other methods.

LDSC intercepts should capture some shared phenotypic variation, which could confound the genetic correlation estimates. We found a perfect correlation between intercepts and volume-by-volume phenotypic correlations ($b = 0.98$, SE $= 0.009$, p-value < 2 x10^{-16}, $R^2 = 0.997$), indicating that LDSC produced larger intercepts for highly phenotypically correlated traits. This suggests that genetic correlations found in this study are reliable and probably do not rely on methodological artefacts.

LDSC regresses effect sizes from GWAS summary statistics on LD Scores which separates the polygenic signal into signal that correlates with LD (slope), and signal that does not correlate with LD (intercept). Variation captured with the intercept characterises confounding on the cross-trait associations such as sample overlap, population stratification and environmental influences which should be uncorrelated with LD (Lee, McGue, et al., 2018; Yengo et al., 2018). The correlation between phenotypic associations and intercepts is plausible as it indicates that the intercepts successfully captured and removed confounding effects from the estimate of genetic correlations (slope). We probably obtained genome-wide representative genetic correlations in this study because we discovered pronounced bilateral symmetry of genetic influences, that is near perfect correlations between areas and their homologous counterpart in the opposite hemisphere. This bilateral symmetry was

previously reported in twin studies (Alexander-Bloch et al., 2017; Chen et al., 2011) that are sensitive towards overall additive genetic effects and can help inform whether genetic correlations are typical for the whole allelic spectrum.

Another limitation is that precision of genetic correlations depends on the polygenic signal contained in GWAS summary statistics. Low precision and limited systematic variance were linked with unstable genetic correlations (Lee, McGue, et al., 2018). SNP-heritability can be used as a signal-to-noise ratio, to inform whether sufficient polygenic signal was present in a GWAS, and a previous study suggested to exclude traits with SNP-heritability estimates below 0.05 (Sodini et al., 2018). SNP-heritability estimates of our brain-volumetric traits ranged between 7.6% (SE = 0.01) and 42% (SE = 0.037; median h^2 across all volumes = 0.23). The left and right frontal poles had the lowest SNP-heritability estimates, probably because the boundaries of this region are ambiguous (Klein & Tourville, 2012) resulting in limited systematic variation.

Supplementary Figures
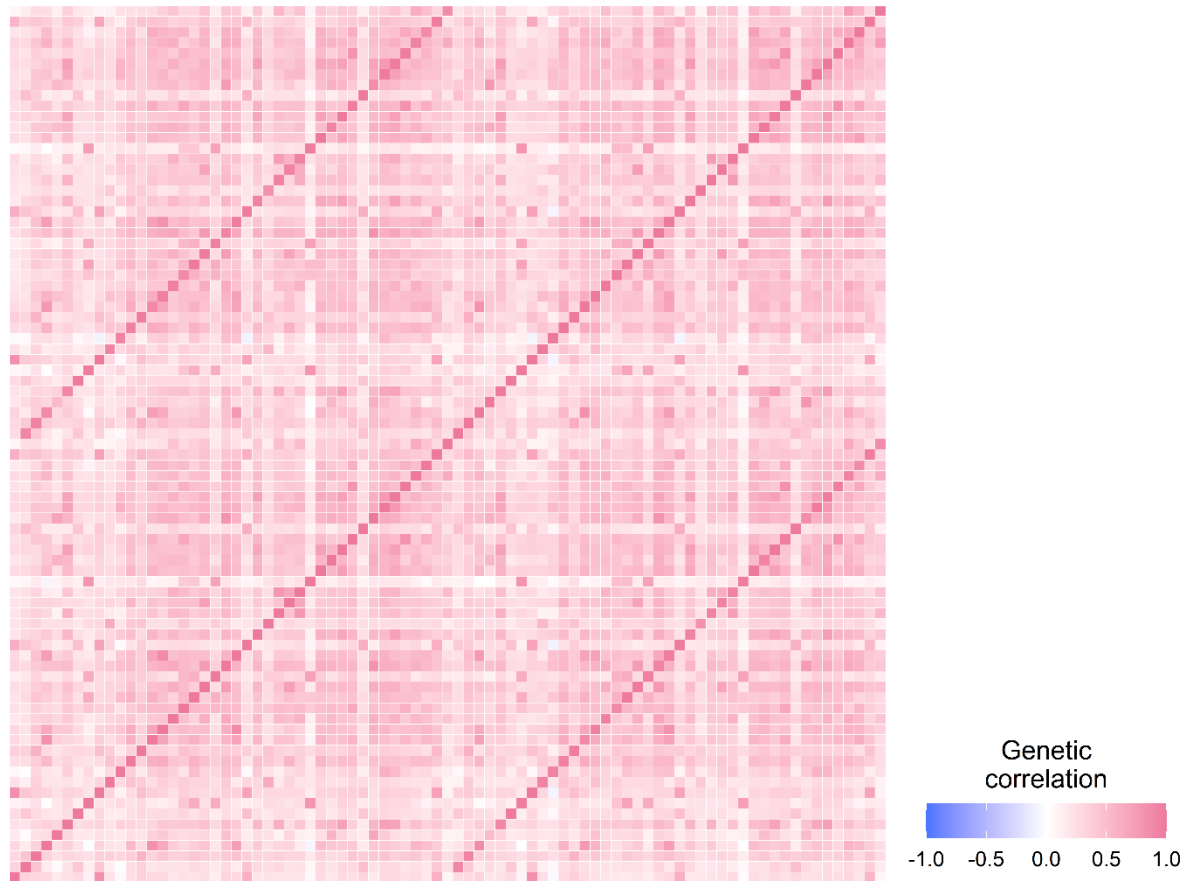
Supplementary Figure 1. Genetic correlation matrix inferred through bivariate LDSC across the whole brain (83 volumes). High correlations along the off-diagonal lines, parallel to diagonal, indicate nearly perfect correlations between regions and their homologous counterparts in the opposite hemisphere (brain stem is an exception as it has no counterpart in the opposite hemisphere).
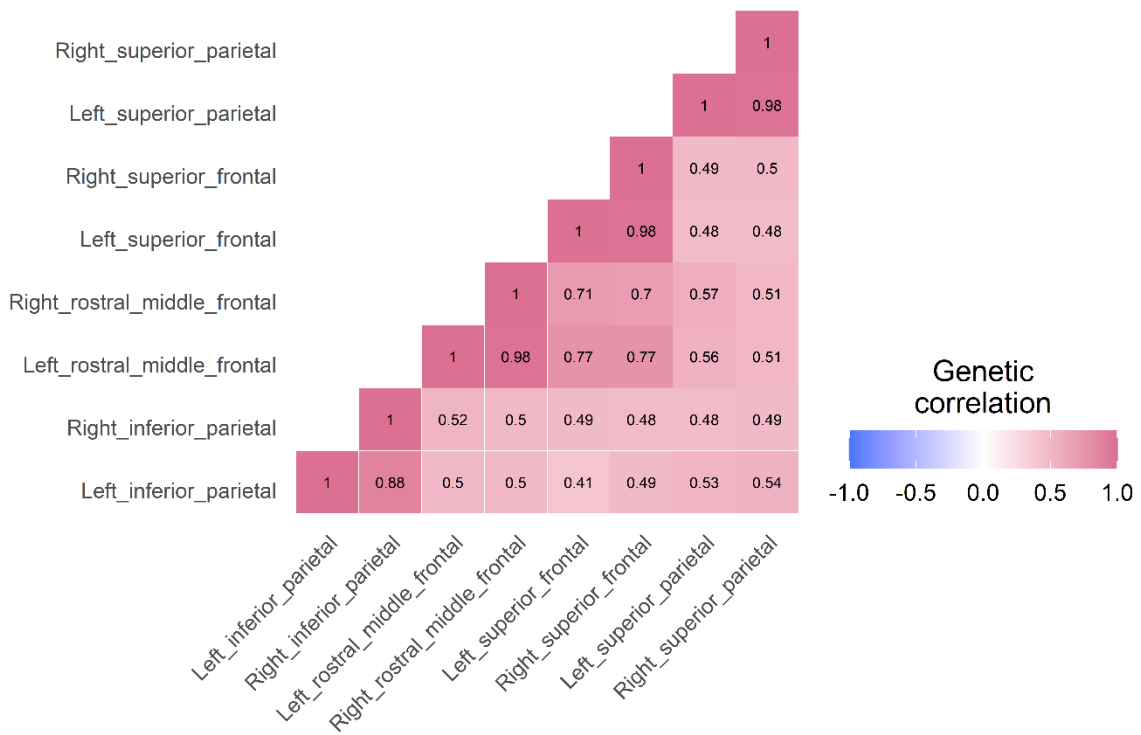
Supplementary Figure 2. Genetic correlations inferred through LDSC among the central executive network (8 volumes).
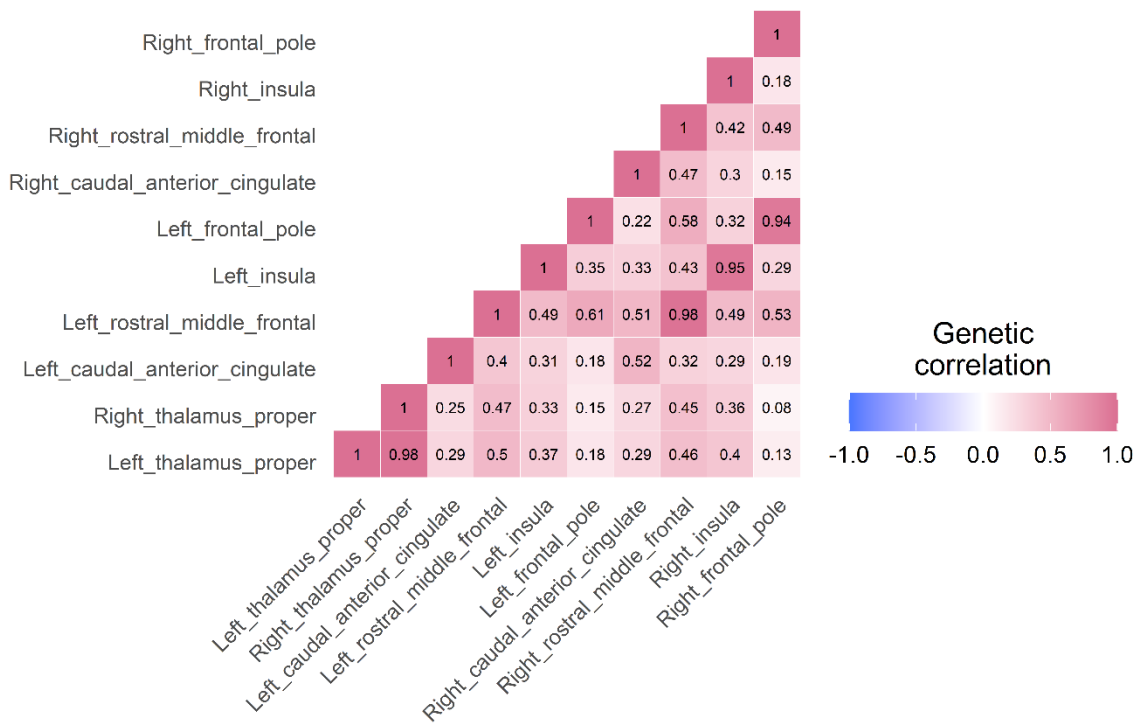
Supplementary Figure 3. Genetic correlations inferred through LDSC among the cingulo-opercular network (10 volumes).
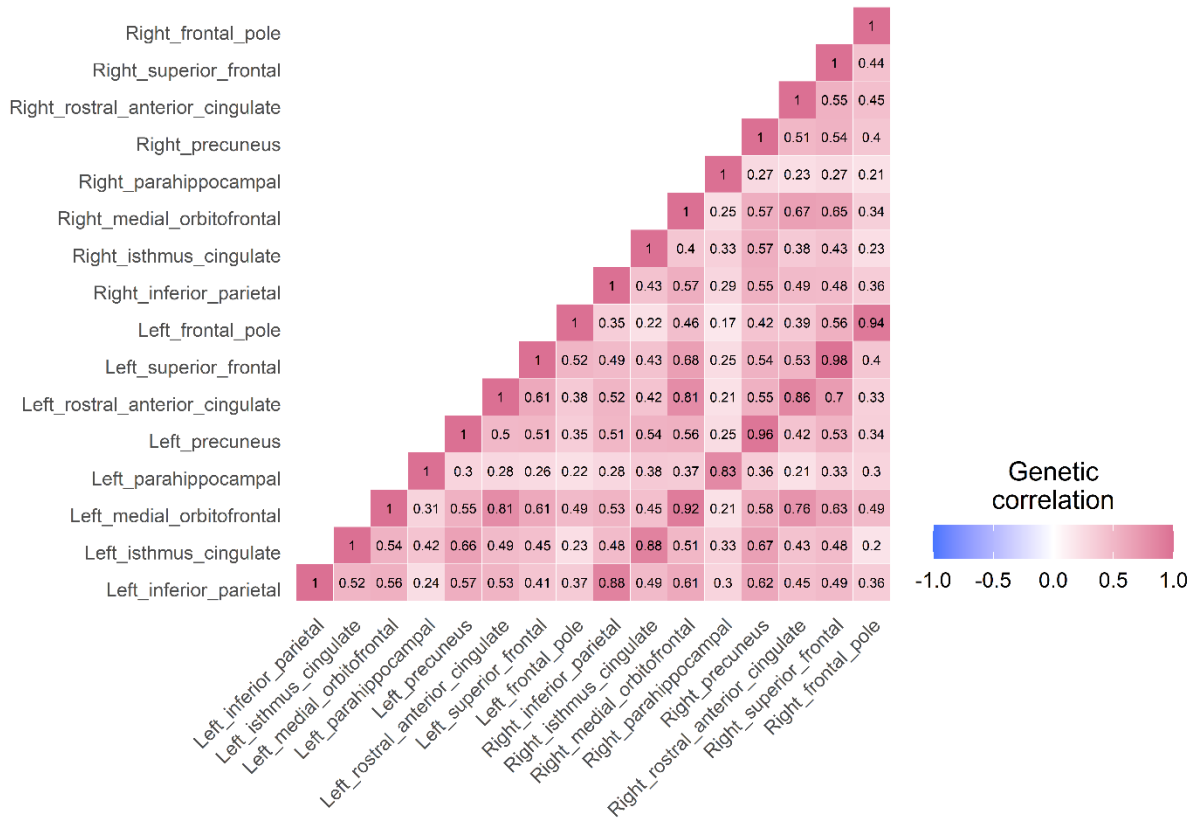
Supplementary Figure 4. Genetic correlations inferred through LDSC among the default mode network (16 volumes).
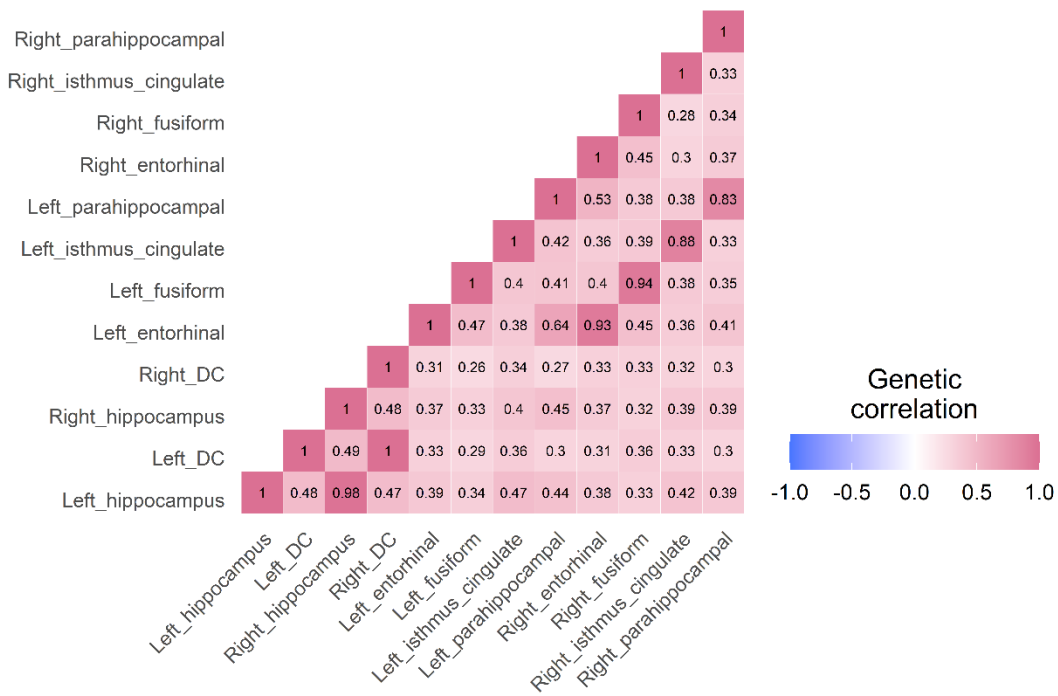
Supplementary Figure 5. Genetic correlations inferred through LDSC among the hippocampal-diencephalic network (12 volumes).
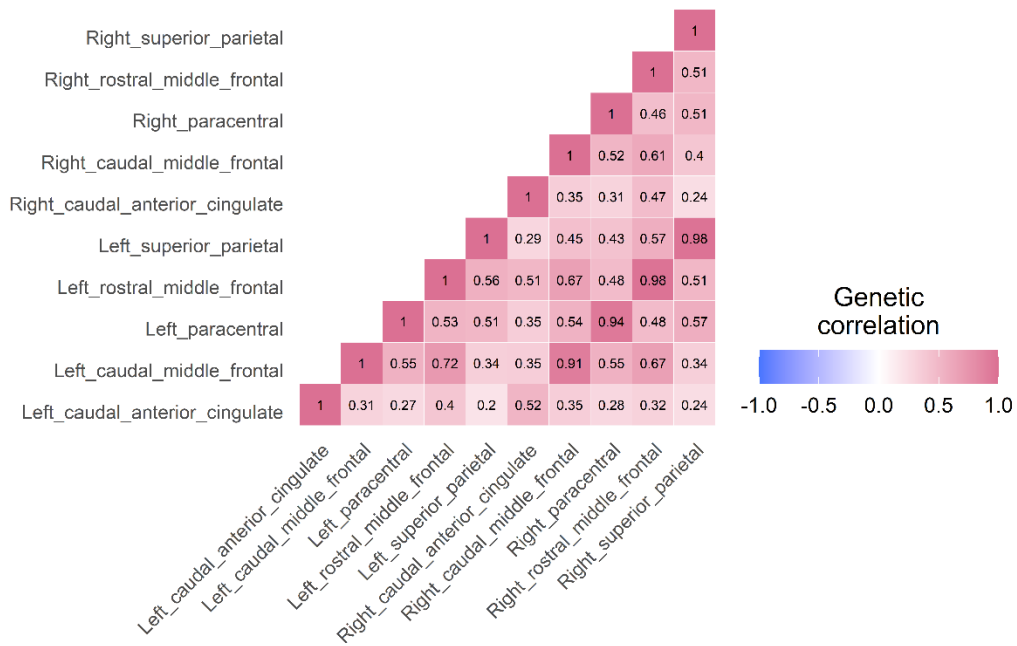
Supplementary Figure 6. Genetic correlations inferred through LDSC among the multiple demand network (12 volumes).

Supplementary Figure 7. Genetic correlations inferred through LDSC among the P-FIT network (36 volumes).
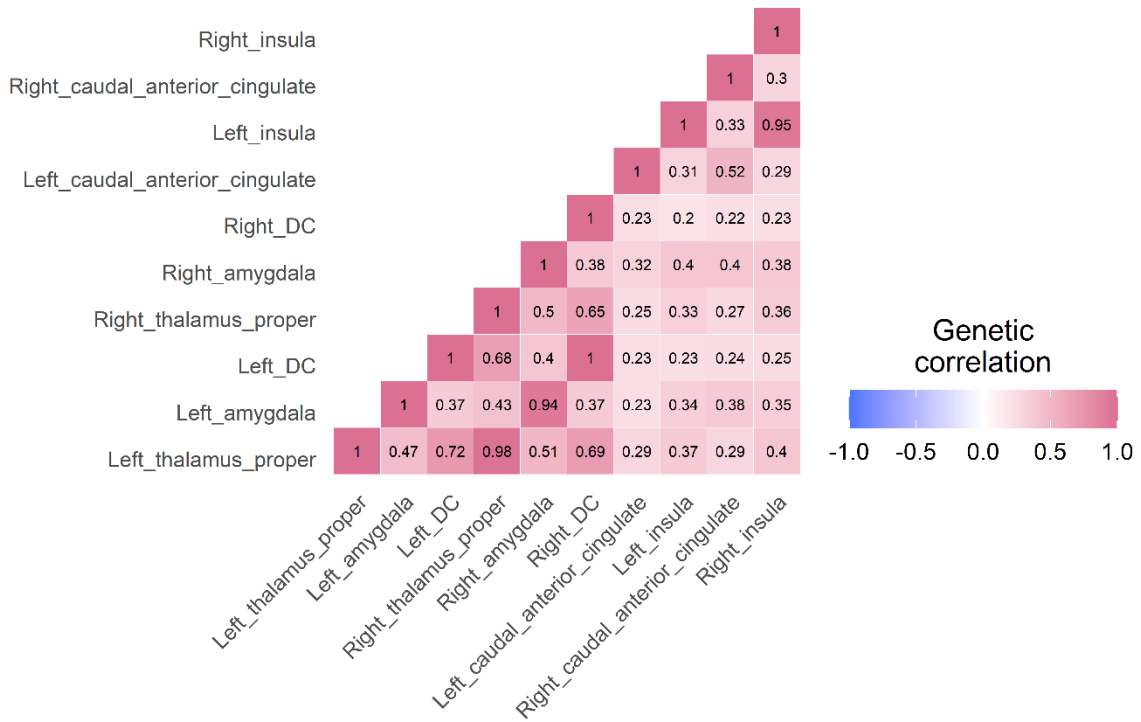
Supplementary Figure 8. Genetic correlations inferred through LDSC among the salience network (10 volumes).
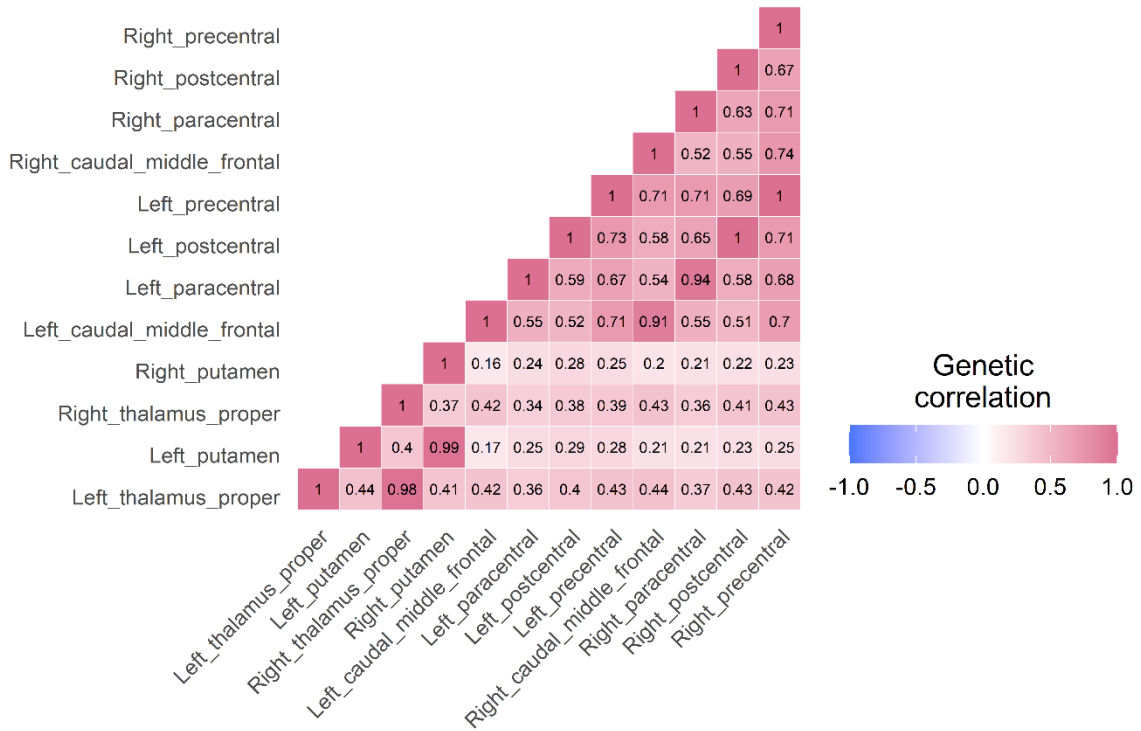
Supplementary Figure 9. Genetic correlations inferred through LDSC among the sensorimotor network (12 volumes).
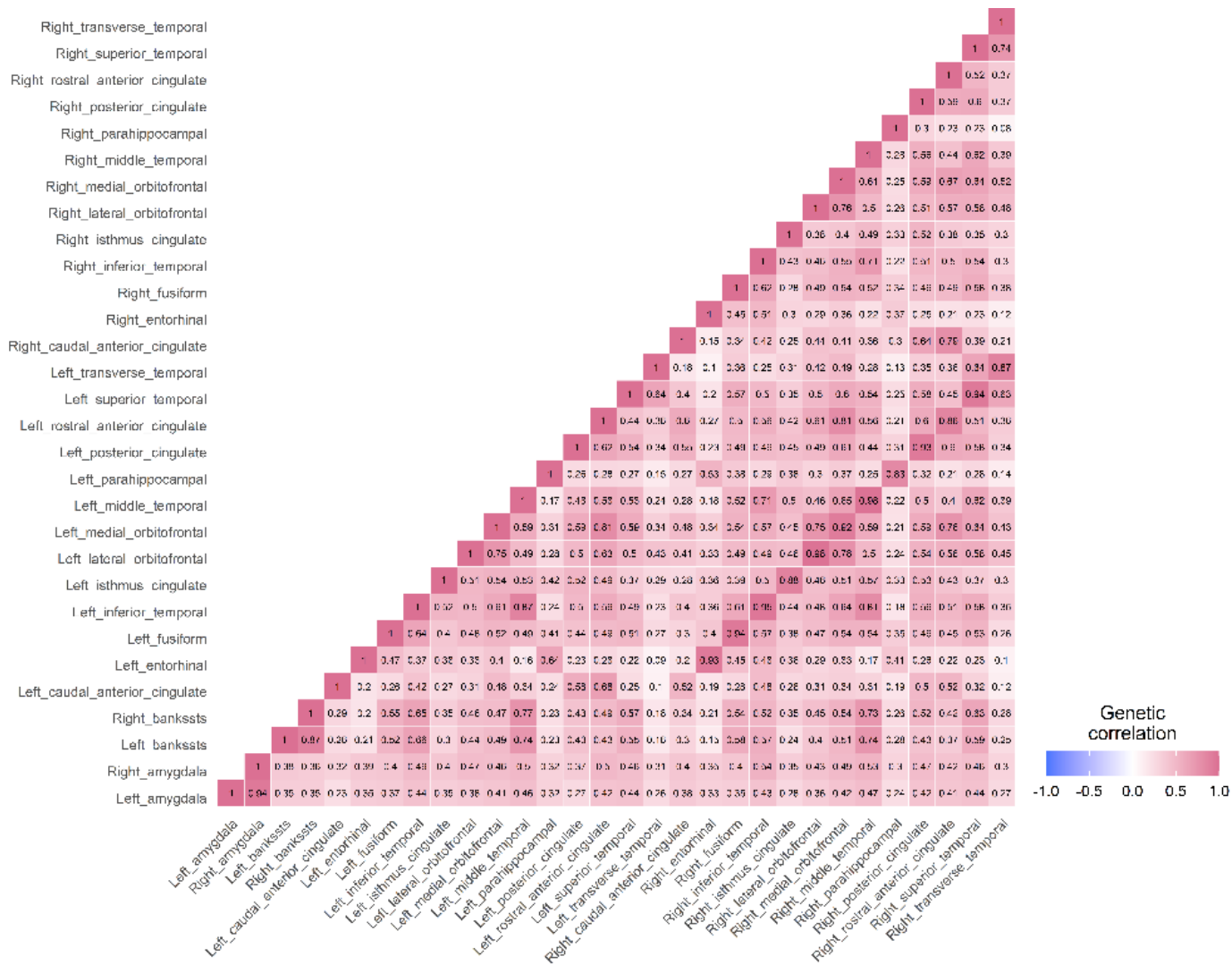
Supplementary Figure 10. Genetic correlations inferred through LDSC among the temporo-amygdala-orbitofrontal network (30 volumes).
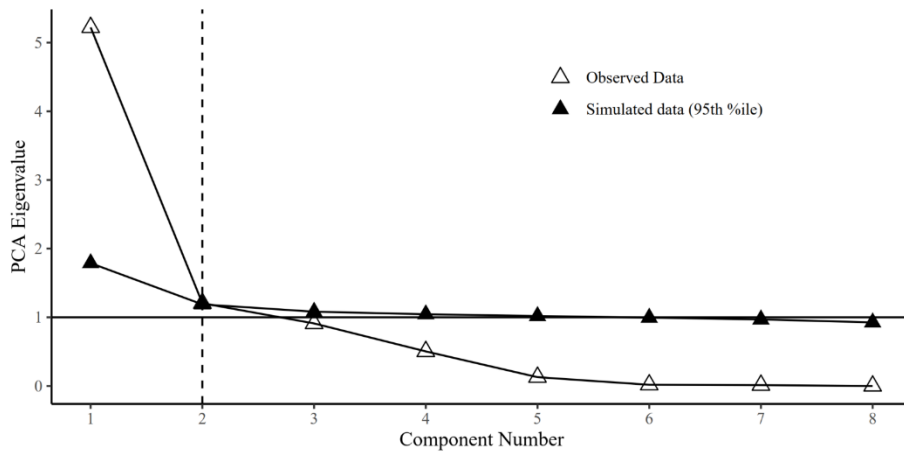
Supplementary Figure 11. Parallel analysis in the central executive network
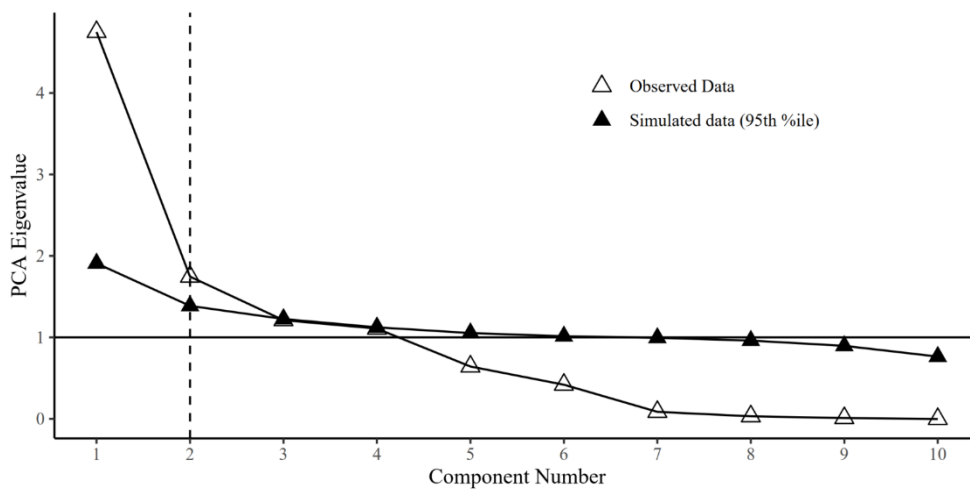
Supplementary Figure 12. Parallel analysis in the cingulo-operular network
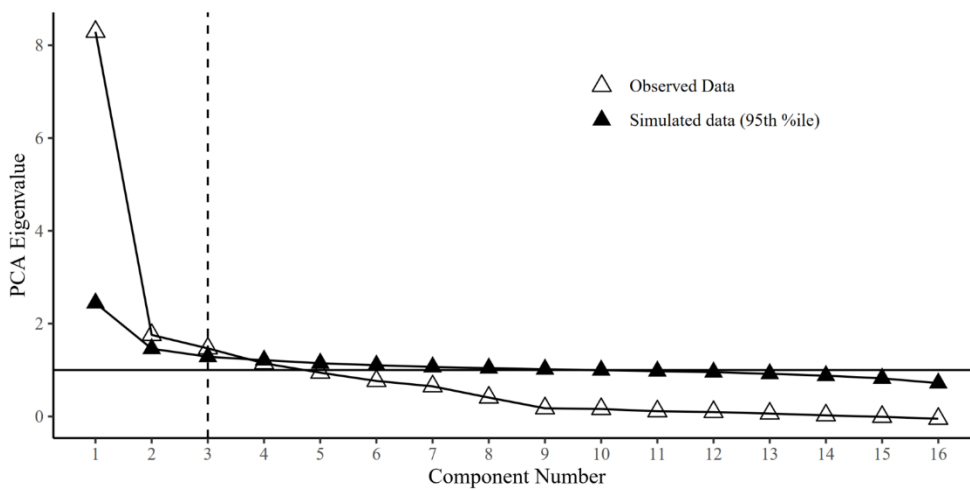
Supplementary Figure 13. Parallel analysis in the default mode network

232
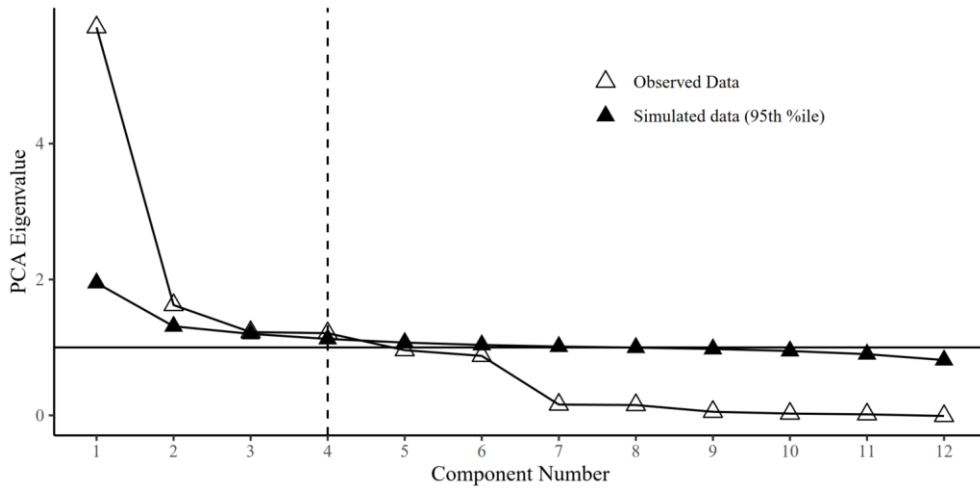
Supplementary Figure 14. Parallel analysis in the hippocampal-diencephalic network
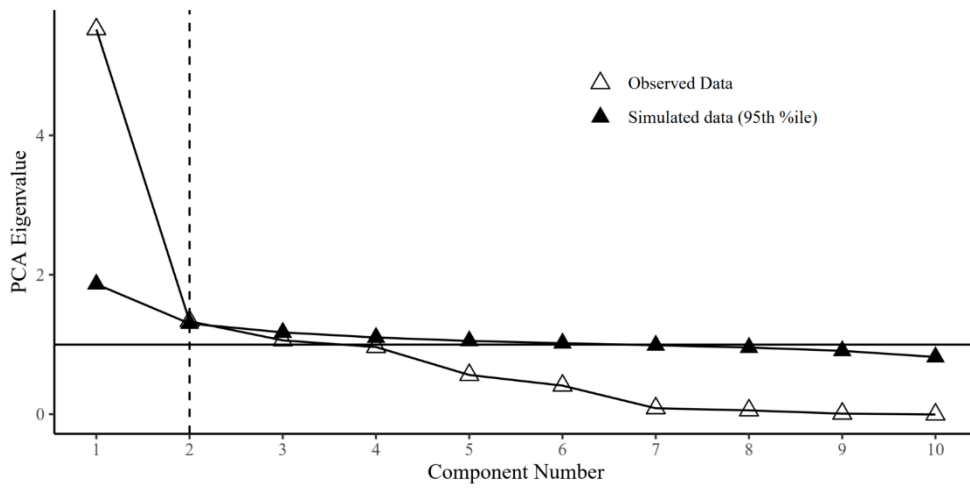
Supplementary Figure 15. Parallel analysis in the multiple demand network
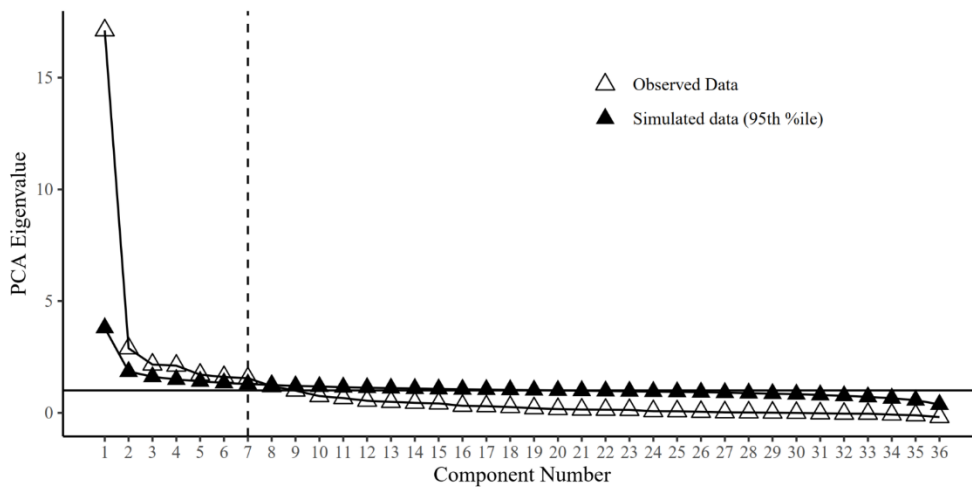
Supplementary Figure 16. Parallel analysis in the P-FIT network
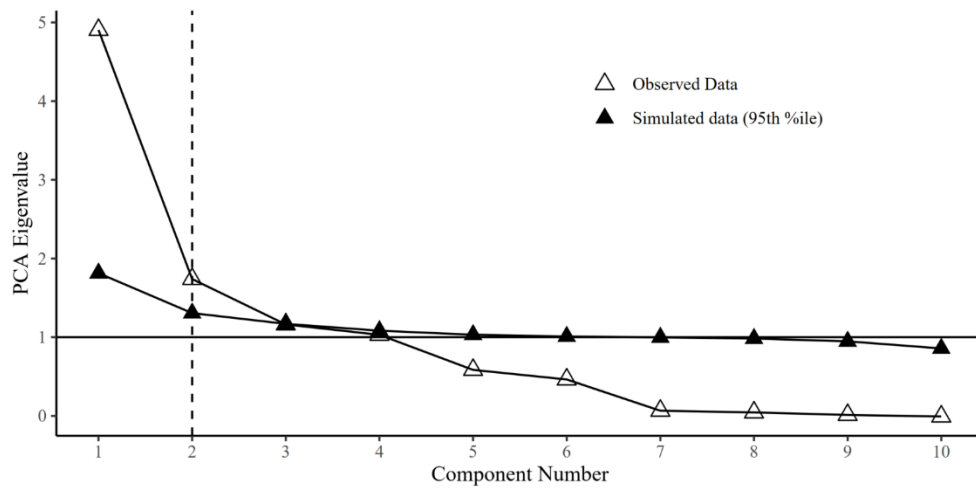
Supplementary Figure 17. Parallel analysis in the salience network
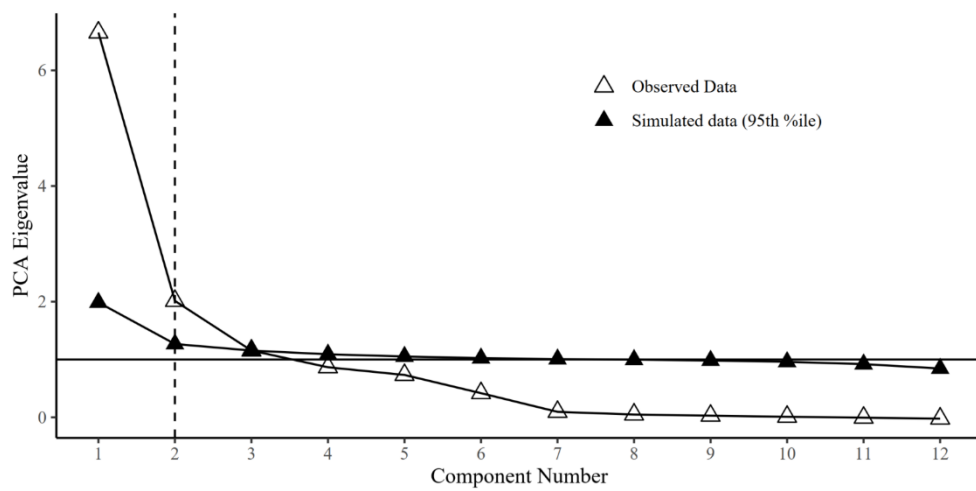
Supplementary Figure 18. Parallel analysis in the sensorimotor network
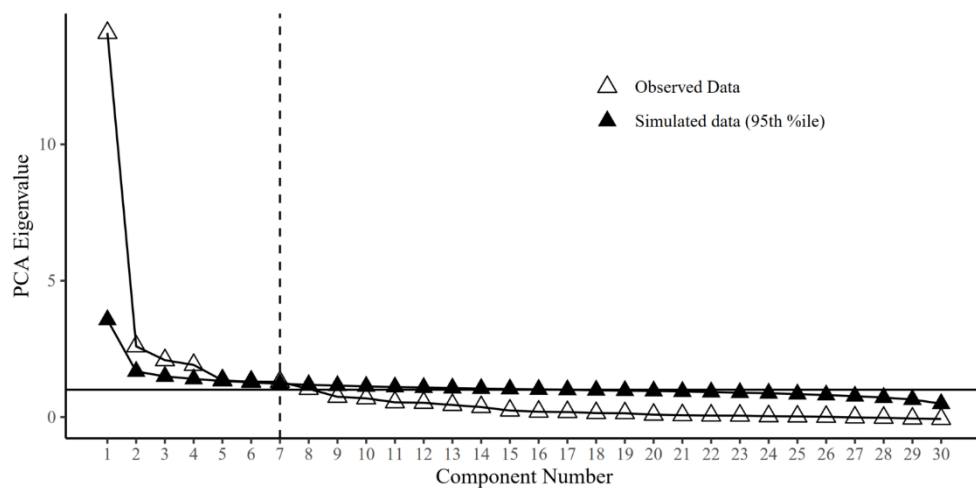
Supplementary Figure 19. Parallel analysis in the temporo-amygdala-orbitofrontal network
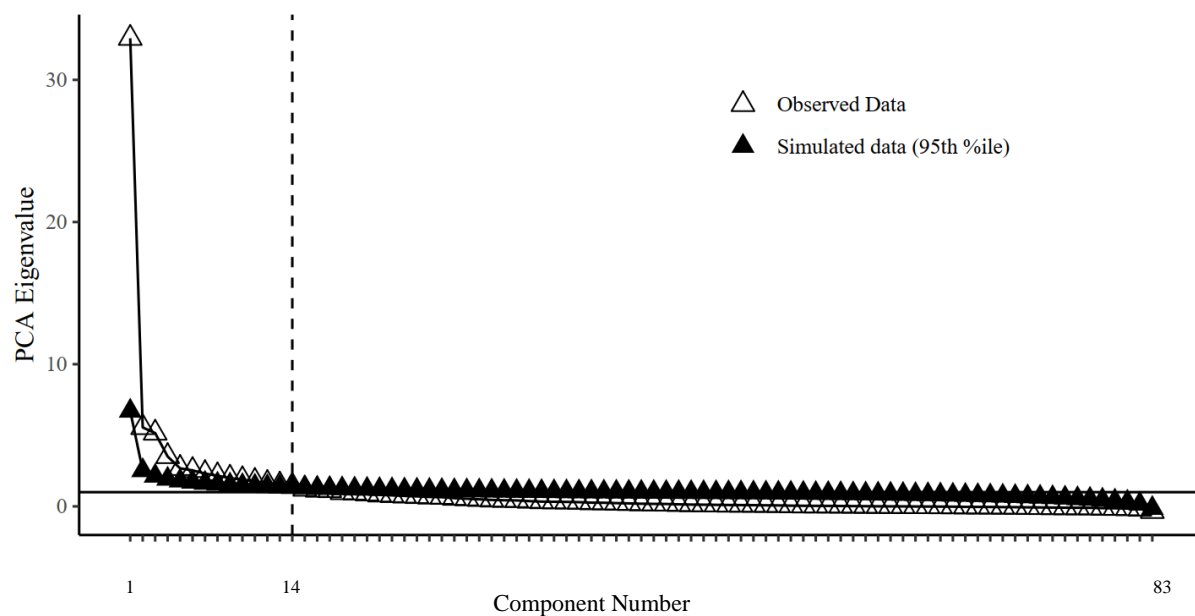
Supplementary Figure 20. Parallel analysis in the whole brain with 83 nodes.
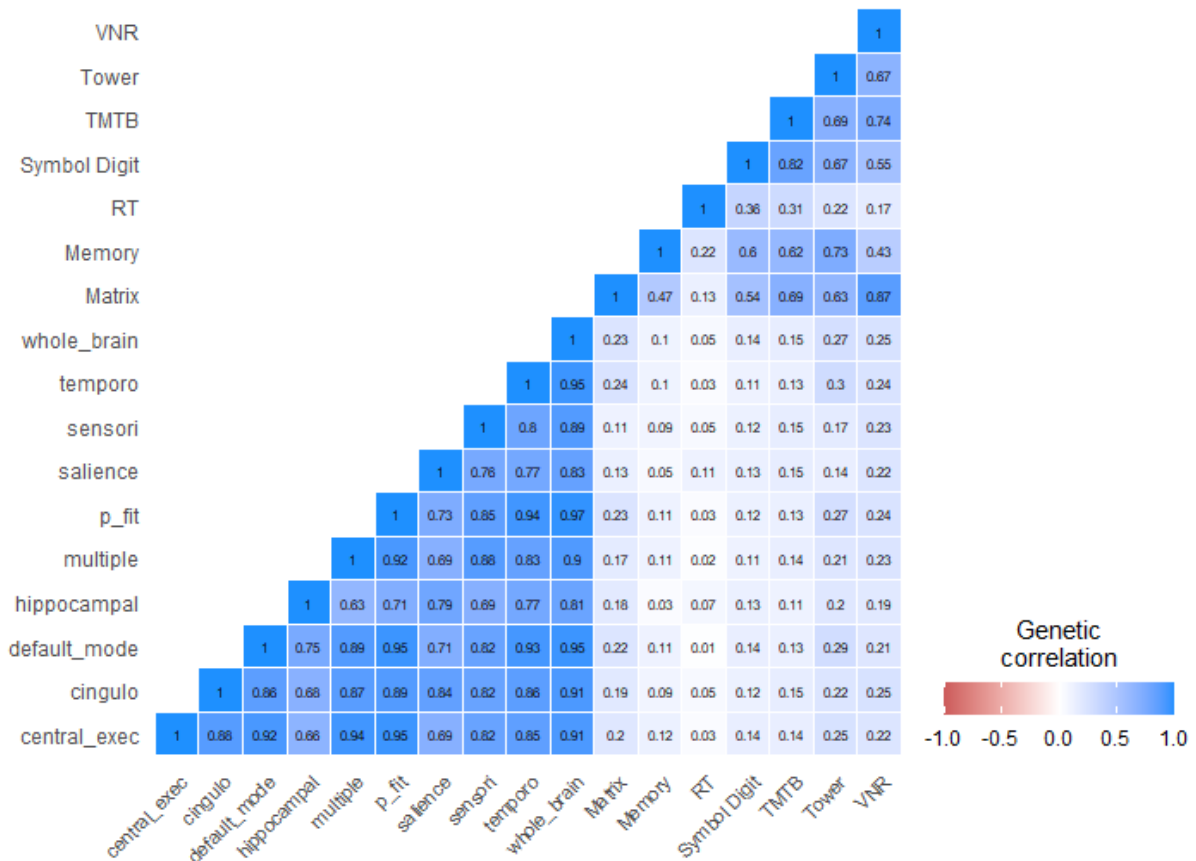
Supplementary Figure 21. Genetic correlations between seven cognitive traits and genetic PC1s underlying structural covariance networks. Descriptively, performance in the Tower Rearranging Task has the largest association with brain networks in comparison with other cognitive tasks. Abbreviations: Matrix = Matrix Pattern Completion task; Memory = Memory – Pairs Matching Test; RT = Reaction Time; Symbol Digit = Symbol Digit Substitution Task; Trails-B = Trail Making Test – B; Tower = Tower Rearranging Task; VNR = Verbal Numerical Reasoning Test; central exec = central executive; cingulo = cingulo-opercular; hippocampal = hippocampal-diencephalic; multiple = multiple demand; p fit = parieto-frontal integration theory; sensori = sensorimotor; temporo = temporo-amygdala-orbitofrontal
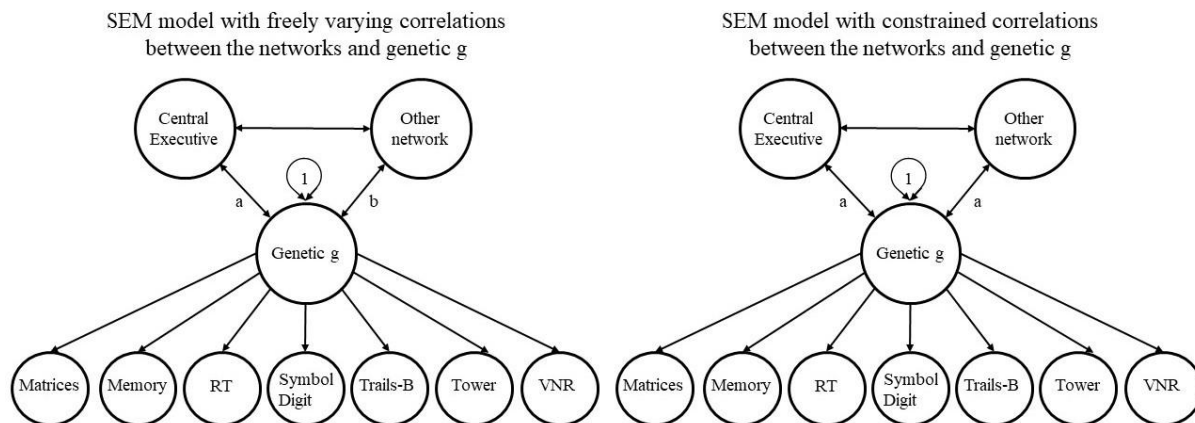
Supplementary Figure 22. Illustration of the genomic structural equation models used to test whether correlation magnitudes with genetic general cognitive ability differ between the central executive network and other significantly associated networks. The model on the right freely estimates correlation parameters between two networks and genetic g while allowing for correlations between the networks. In the left model, we force the correlation magnitudes to be the same, and assess whether model fit deteriorates significantly, to conclude whether correlation magnitudes between networks are likely different from each other. The difference between these two models is in the paths leading from the networks to genetic g. In the left model, the paths can take a different value which is why they are labelled with a and b. In the right model, the two paths are specified to be the same and are therefore both labelled with a.

Independent pathways model
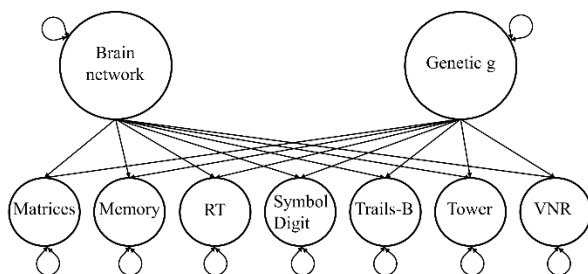
Common pathway model
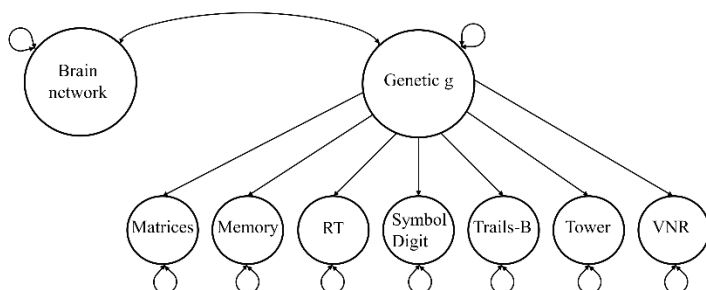
Supplementary Figure 23. Structural equation models to calculate Q_{trait} heterogeneity indices
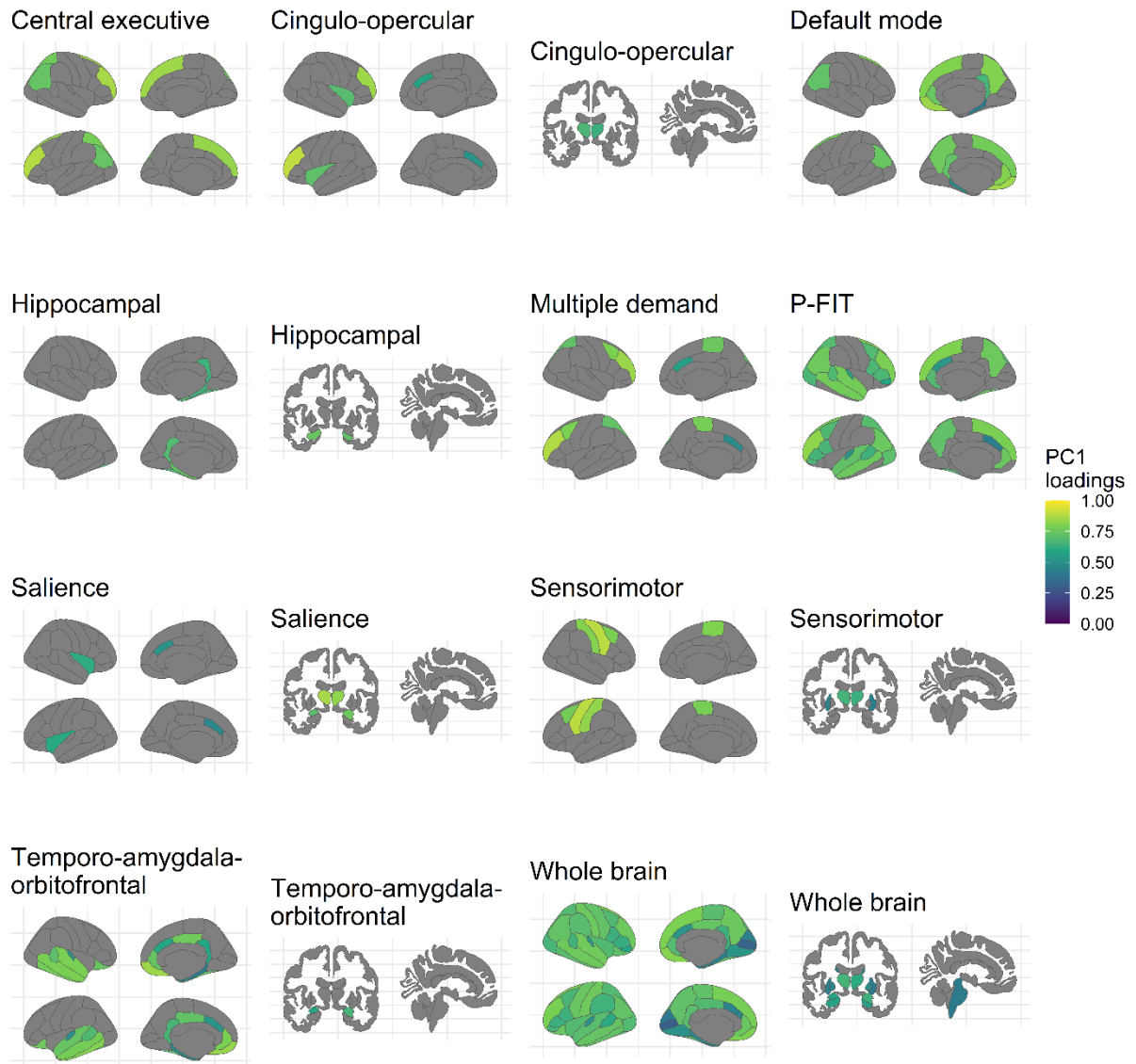
Supplementary Figure 24. PC1 loadings displayed in *Fig.3* plotted onto brain regions.

Supplementary Tables

Supplementary Table 1. 83 cortical and subcortical grey-matter regions of interest

No. volumes	Region of interest	UK Biobank field ID
1.	Left banks of superior temporal sulcus	26789
2.	Left caudal anterior cingulate	26790
3.	Left caudal middle frontal	26791
4.	Left cuneus	26792
5.	Left entorhinal	26793
6.	Left fusiform	26794
7.	Left inferior parietal	26795
8.	Left inferior temporal	26796
9.	Left isthmus cingulate	26797
10.	Left lateral occipital	26798
11.	Left lateral orbitofrontal	26799
12.	Left lingual	26800
13.	Left medial orbitofrontal	26801
14.	Left middle temporal	26802
15.	Left parahippocampal	26803
16.	Left paracentral	26804
17.	Left pars opercularis	26805
18.	Left pars orbitalis	26806
19.	Left pars triangularis	26807
20.	Left pericalcarine	26808
21.	Left postcentral	26809
22.	Left posterior cingulate	26810
23.	Left precentral	26811
24.	Left precuneus	26812
25.	Left rostral anterior cingulate	26813
26.	Left rostral middle frontal	26814
27.	Left superior frontal	26815
28.	Left superior parietal	26816
29.	Left superior temporal	26817
30.	Left supramarginal	26818
31.	Left frontal pole	26819
32.	Left transverse temporal	26820
33.	Left insula	26821
34.	Right banks of superior temporal sulcus	26890
35.	Right caudal anterior cingulate	26891
36.	Right caudal middle frontal	26892

37.	Right cuneus	26893
38.	Right entorhinal	26894
39.	Right fusiform	26895
40.	Right inferior parietal	26896
41.	Right inferior temporal	26897
42.	Right isthmus cingulate	26898
43.	Right lateral occipital	26899
44.	Right lateral orbitofrontal	26900
45.	Right lingual	26901
46.	Right medial orbitofrontal	26902
47.	Right middle temporal	26903
48.	Right parahippocampal	26904
49.	Right paracentral	26905
50.	Right pars opercularis	26906
51.	Right pars orbitalis	26907
52.	Right pars triangularis	26908
53.	Right pericalcarine	26909
54.	Right postcentral	26910
55.	Right posterior cingulate	26911
56.	Right precentral	26912
57.	Right precuneus	26913
58.	Right rostral anterior cingulate	26914
59.	Right rostral middle frontal	26915
60.	Right superior frontal	26916
61.	Right superior parietal	26917
62.	Right superior temporal	26918
63.	Right supramarginal	26919
64.	Right frontal pole	26920
65.	Right transverse temporal	26921
66.	Right insula	26922
67.	Left thalamus proper	26558
68.	Left caudate	26559
69.	Left putamen	26560
70.	Left pallidum	26561
71.	Left hippocampus	26562
72.	Left amygdala	26563
73.	Left accumbens area	26564
74.	Left ventral diencephalon (DC)	26565
75.	Right thalamus proper	26589
76.	Right caudate	26590

77.	Right putamen	26591
78.	Right pallidum	26592
79.	Right hippocampus	26593
80.	Right amygdala	26594
81.	Right accumbens area	26595
82.	Right ventral diencephalon (DC)	26596
83.	Brain stem	26526

Note. This list excludes the right and left temporal pole. Both have been removed from category 192 in UK Biobank due to overwhelming amounts of missing data based on difficulties segmenting this area. https://www.jiscmail.ac.uk/cgi-bin/webadmin?A2=ind2005&L=UKB-NEUROIMAGING&O=D&P=4857

Supplementary Table 2. Network characterisation

	left inferior parietal (InfP)
	left rostral middle frontal (RosMF)
	left superior frontal (SupF)
Central Executive (8 nodes)	left superior parietal (SupP)
	right inferior parietal (InfP)
	right rostral middle frontal (RosMF)
	right superior frontal (SupF)
	right superior parietal (SupP)
	left thalamus proper (Tha)
	right thalamus proper (Tha)
	left caudal anterior cingulate (caACg)
	left rostral middle frontal (RosMF)
Cingulo-opercular (10 nodes)	left insula (Ins)
	left frontal pole (FPo)
	right caudal anterior cingulate (caACg)
	right rostral middle frontal (RosMF)
	right insula (Ins)
	right frontal pole (FPo)
	left inferior parietal (InfP)
	left isthmus cingulate (IsCg)
	left medial orbitofrontal (MedO)
	left parahippocampal (PaHip)
	left precuneus (PrCun)
	left rostral anterior cingulate (RosACg)
	left superior frontal (SupF)
Default mode (16 nodes)	left frontal pole (FPo)
	right inferior parietal (InfP)
	right isthmus cingulate (IsCg)
	right medial orbitofrontal (MedO)
	right parahippocampal (PaHip)
	right precuneus (PrCun)
	right rostral anterior cingulate (RosACg)
	right superior frontal (SupF)
	right frontal pole (FPo)
	left hippocampus (Hip)
	left ventral DC (VDC)
Hippocampal-Diencephalic (12 nodes)	right hippocampus (Hip)
	right ventral DC (VDC)
	left entorhinal (Ent)

	left fusiform (Fus)
	left isthmus cingulate (IsCg)
	left parahippocampal (PaHip)
	right entorhinal (Ent)
	right fusiform (Fus)
	right isthmus cingulate (IsCg)
	right parahippocampal (PaHip)
Multiple Demand (12 nodes)	left caudal anterior cingulate (CaACg)
	left caudal middle frontal (CaMF)
	left paracentral (PaC)
	left rostral middle frontal (RosMF)
	left superior parietal (SupP)
	left frontal pole (FPo)
	right caudal anterior cingulate (CaACg)
	right caudal middle frontal (CaMF)
	right paracentral (PaC)
	right rostral middle frontal (RosMF)
	right superior parietal (SupP)
	right frontal pole (FPo)
Parieto-Frontal Integration Theory (36 nodes)	left caudal anterior cingulate
	left superior temporal sulcus
	left caudal middle frontal
	left fusiform
	left inferior parietal
	left lateral occipital
	left middle temporal
	left pars opercularis
	left pars orbitalis
	left pars triangularis
	left precuneus
	left rostral anterior cingulate
	left rostral middle frontal
	left superior frontal
	left superior parietal
	left superior temporal
	left frontal pole
	left transverse temporal
	right superior temporal sulcus
	right caudal anterior cingulate
	right caudal middle frontal

	right fusiform
	right inferior parietal
	right lateral occipital
	right middle temporal
	right pars opercularis
	right pars orbitalis
	right pars triangularis
	right precuneus
	right rostral anterior cingulate
	right rostral middle frontal
	right superior frontal
	right superior parietal
	right superior temporal
	right frontal pole
	right transverse temporal
Salience (10 nodes)	left thalamus proper (Tha)
	left amygdala (Amg)
	left ventral DC (VDC)
	right thalamus proper (Tha)
	right amygdala (Amg)
	right ventral DC (VDC)
	left caudal anterior cingulate (caACg)
	left insula (Ins)
	right caudal anterior cingulate (caACg)
	right insula (Ins)
Sensorimotor (12 nodes)	left thalamus proper (Tha)
	left putamen (Put)
	right thalamus proper proper (Tha)
	right putamen (Put)
	right caudal middle frontal (CaMF)
	left paracentral (PaC)
	left postcentral (PosC)
	left precentral (PrC)
	left caudal middle frontal (CaMF)
	right paracentral (PaC)
	right postcentral (PosC)
	right precentral (PrC)
Temporo-Amygdala-Orbitofrontal (30 nodes)	left amygdala (Amg)
	right amygdala (Amg)
	left superior temporal sulcus (SupT)

	right superior temporal sulcus (SupT)
	left caudal anterior cingulate (caACg)
	left entorhinal (Ent)
	left fusiform (Fus)
	left inferior temporal (InfT)
	left isthmus cingulate (IsCg)
	left lateral orbitofrontal (LatO)
	left medial orbitofrontal (MedO)
	left middle temporal (MidT)
	left parahippocampal (PaHip)
	left posterior cingulate (PosCg)
	left rostral anterior cingulate (RosACg)
	left superior temporal (SupT)
	left transverse temporal (TraT)
	right caudal anterior cingulate (caACg)
	right entorhinal (Ent)
	right fusiform (Fus)
	right inferior temporal (InfT)
	right isthmus cingulate (IsCg)
	right lateral orbitofrontal (LatO)
	right medial orbitofrontal (MedO)
	right middle temporal (MidT)
	right parahippocampal (PaHip)
	right posterior cingulate (PosCg)
	right rostral anterior cingulate (RosACg)
	right superior temporal (SupT)
	right transverse temporal (TraT)

Note. The region abbreviations have been added with reference to Figure 1. There are no abbreviations for the regions included in the P-FIT, because Figure 1 does not display them for aesthetical reasons.

Supplementary Table 3a. Descriptive statistics of PC loadings within genetic networks

Network	Included volumes	% variance explained by 1st PC	PC loadings			
			Mean	SD	Median	Range
Whole brain	83	39.67	0.62	0.13	0.65	0.30-0.81
Central executive	8	65.23	0.81	0.06	0.81	0.74-0.88
Cingulo-opercular	10	47.5	0.68	0.12	0.67	0.52-0.90
Default mode	16	51.78	0.71	0.13	0.75	0.42-0.85
Hippocampal-diencephalic	12	47.63	0.69	0.04	0.68	0.64-0.74
Multiple demand	12	55.23	0.73	0.12	0.77	0.48-0.88
P-FIT	36	47.56	0.68	0.09	0.69	0.43-0.83
Salience	10	49.05	0.69	0.13	0.73	0.47-0.86
Sensorimotor	12	55.42	0.73	0.16	0.79	0.43-0.89
Temporo-amygdala-orbitofrontal	30	46.92	0.67	0.13	0.72	0.41-0.85

PC loadings were extracted using PCA of genetic correlations indicating genetic overlap between brain volumes. *Included volumes* indicates the number of regional volumes included in a network. *Explained variance* shows the genetic variance explained by the first genetic PC. Mean, standard deviation (SD), median, and range relate to the distribution of PC loadings within a network (Figure 4).

Supplementary Table 3b. Descriptive statistics of PC loadings within phenotypic canonical networks

Network	Included volumes	Explained variance (%)	Mean	SD	Median	Mode	Minimum	Maximum
Whole brain	83	30.93	0.55	0.11	0.56	0.59	0.29	0.73
Central executive	8	53.06	0.73	0.06	0.74	0.77	0.64	0.78
Cingulo-opercular	10	38.69	0.6	0.17	0.7	0.72	0.33	0.78
Default mode	16	36.49	0.59	0.13	0.61	0.67	0.4	0.76
Hippocampal-diencephalic	12	38.09	0.61	0.1	0.58	0.54	0.5	0.76
Multiple demand	12	41.17	0.63	0.15	0.68	0.68	0.34	0.77
P-FIT	36	34.12	0.57	0.12	0.57	0.56	0.3	0.74
Salience	10	44.43	0.63	0.22	0.67	0.78	0.24	0.84
Sensorimotor	12	45.62	0.67	0.07	0.67	0.67	0.56	0.79
Temporo-amygdala-orbitofrontal	30	32.2	0.55	0.12	0.56	0.64	0.34	0.74

PC loadings were extracted using PCA of phenotypic correlations between brain volumes. *Included volumes* indicates the number of regional volumes included in a network. *Explained variance* shows the phenotypic variance explained by the first genetic PC. Mean, standard deviation (SD), median, and range relate to the distribution of PC loadings within a network (Figure 4).

Supplementary Table 4. Model fit for genetic correlations between genetic general cognitive ability and each canonical network

Network	χ^2	df	p-value	AIC	CFI	SRMR
Whole brain	157.75	20	2.04×10^{-23}	189.748	0.962	0.081
Central executive	142.04	20	2.07×10^{-20}	174.042	0.967	0.08
Cingulo-opercular	156.5	20	3.56×10^{-23}	188.497	0.964	0.08
Default mode	150.83	20	4.36×10^{-22}	182.83	0.965	0.082
Hippocampal-diencephalic	143.69	20	1.01×10^{-20}	175.686	0.965	0.08
Multiple demand	143.06	20	1.33×10^{-20}	175.057	0.966	0.079
P-FIT	161.08	20	4.65×10^{-24}	193.077	0.962	0.082
Salience	144.92	20	5.87×10^{-21}	176.919	0.965	0.08
Sensorimotor	146.2	20	3.34×10^{-21}	178.204	0.965	0.078
Temporo-amygdala-orbitofrontal	165.98	20	5.23×10^{-25}	197.98	0.96	0.084

AIC = Akaike Information Criterion, CFI = Comparative Fit Index, SRMR = Standardised Root Mean Square Residuals

Supplementary Table 5. Fit indices for the comparison between freely-varying or constrained correlations with general cognitive ability between central executive and other networks

Comparison with central executive	χ² test			constrained model			freely-varying model		
	Δ χ²	Δ df	p	AIC	CFI	SRMR	AIC	CFI	SRMR
Cingulo-opercular	0.83	1	0.361	209.981	0.988	0.073	211.148	0.988	0.074
Default Mode	-0.17	1	1	198.202	0.994	0.075	200.375	0.994	0.075
Hippocampal-diencephalic	1.1	1	0.294	201.567	0.974	0.074	202.466	0.974	0.074
Multiple Demand	-0.58	1	1	199.321	0.993	0.073	201.906	0.993	0.073
P-FIT	0.79	1	0.375	217.691	0.995	0.075	218.904	0.995	0.075
Salience	0.05	1	0.831	214.194	0.976	0.073	216.149	0.976	0.073
Sensorimotor	0.04	1	0.85	204.424	0.982	0.072	206.388	0.982	0.072
Temporo-amygdala-orbitofrontal	0.96	1	0.328	213.984	0.988	0.077	215.028	0.988	0.077
Whole brain	3.23	1	0.072	218.243	0.994	0.075	217.016	0.994	0.075

Akaike Information Criterion, CFI = Comparative Fit Index, SRMR = Standardised Root Mean Square Residuals. This table indicates fit indices for two GenomicSEM models: a model where genetic correlations between networks and general cognitive abilities are estimated freely-varying, and another model where correlation magnitudes are forced to be the same. Δ χ² is the difference in χ² values between the nested constrained and freely-varying model, Δ df indicates their difference in degrees of freedom. AIC, CFI, and SRMR are also reported for the constrained and freely-varying models. These fit indices indicate no evidence for differing correlation magnitudes with cognitive abilities between the central executive network and all other networks.

Supplementary Table 6. Fit indices for the *adjusted* comparison between freely-varying or constrained correlations with general cognitive ability between central executive and other networks

Comparison with central executive	χ² test			constrained model			freely-varying model		
	Δ χ²	**Δ df**	*p*	AIC	CFI	SRMR	AIC	CFI	SRMR
Cingulo-opercular	0.72	1	0.4	209.868	0.988	0.073	211.148	0.988	0.074
Default Mode	-1.3	1	1	197.076	0.994	0.075	200.375	0.994	0.075
Hippocampal-diencephalic	1.6	1	0.21	202.065	0.974	0.075	202.466	0.974	0.074
Multiple Demand	-1.01	1	1	198.893	0.993	0.073	201.906	0.993	0.073
P-FIT	-2.65	1	1	214.258	0.995	0.075	218.904	0.995	0.075
Salience	-0.04	1	1	214.107	0.976	0.074	216.149	0.976	0.073
Sensorimotor	0.64	1	0.42	205.032	0.982	0.073	206.388	0.982	0.072
Temporo-amygdala-orbitofrontal	-1.61	1	1	211.416	0.988	0.077	215.028	0.988	0.077
Whole brain	-0.44	1	1	214.577	0.994	0.075	217.016	0.994	0.07

Information Criterion, CFI = Comparative Fit Index, SRMR = Standardised Root Mean Square Residuals. This table indicates fit indices for two GenomicSEM models in which correlation estimates have been adjusted for network size: (1) a model where genetic correlations between networks and general cognitive abilities are estimated freely-varying, and (2) another model where correlation magnitudes are forced to be the same, after they have been divided by the total number of included brain volumes in a respective network. **Δ χ²** is the difference in χ² values between the nested constrained and freely-varying model, Δ df indicates their difference in degrees of freedom. AIC, CFI, and SRMR are also reported for the constrained and freely-varying models. These fit indices indicate no evidence for differing correlation magnitudes with cognitive abilities between the central executive network and all other networks, even when adjusting for the size of a network.

Supplementary Table 7. Genetic correlations between three cognitive abilities and brain networks

Brain network	Cognitive trait	r_g	SE	95% CI	p-value	FDR- adjusted p-value
Central executive	Matrix	0.2	0.09	0.03 - 0.37	0.0227	0.031
	Memory	0.12	0.04	0.04 - 0.19	0.0024	0.0193
	Symbol Digit	0.14	0.05	0.05 - 0.23	0.0024	0.0193
Cingulo-opercular	Matrix	0.19	0.1	0.01 - 0.38	0.0446	0.0558
	Memory	0.09	0.04	0.02 - 0.16	0.0131	0.022
	Symbol Digit	0.12	0.05	0.03 - 0.21	0.0083	0.019
Default Mode	Matrix	0.22	0.09	0.04 - 0.40	0.0151	0.0238
	Memory	0.11	0.04	0.03 - 0.19	0.0075	0.019
	Symbol Digit	0.14	0.04	0.05 - 0.23	0.0019	0.019
Hippocampal-diencephalic	Matrix	0.19	0.09	0.0008 - 0.37	0.0491	0.059
	Memory	0.03	0.04	-0.16	0.4866	0.487
	Symbol Digit	0.13	0.05	0.04 - 0.21	0.0052	0.0193
Multiple Demand	Matrix	0.17	0.09	-0.35	0.0593	0.0684
	Memory	0.11	0.04	0.03 - 0.19	0.0048	0.0193
	Symbol Digit	0.11	0.05	0.02 - 0.20	0.018	0.026
P-FIT	Matrix	0.23	0.09	0.06 - 0.40	0.0087	0.019
	Memory	0.11	0.04	0.04 - 0.19	0.0029	0.019
	Symbol Digit	0.12	0.05	0.03 - 0.21	0.0083	0.019
Salience	Matrix	0.13	0.09	-0.36	0.1619	0.17
	Memory	0.05	0.04	-0.14	0.16	0.17
	Symbol Digit	0.13	0.05	0.04 - 0.22	0.0063	0.019
Sensorimotor	Matrix	0.11	0.09	-0.34	0.2053	0.212
	Memory	0.09	0.04	0.01 - 0.17	0.0267	0.0348
	Symbol Digit	0.12	0.05	0.03 - 0.21	0.0104	0.0195
Temporo-amygdala-orbitofrontal	Matrix	0.24	0.09	0.07 - 0.41	0.0061	0.019
	Memory	0.1	0.04	0.02 - 0.17	0.0133	0.022
	Symbol Digit	0.11	0.05	0.02 - 0.20	0.0182	0.026
Whole brain	Matrix	0.23	0.09	0.06 - 0.40	0.009	0.019
	Memory	0.1	0.04	0.02 - 0.17	0.0099	0.02
	Symbol Digit	0.14	0.05	0.04 - 0.23	0.0036	0.019

This table shows genetic correlations (r_g) between brain networks and three specific cognitive traits. SE stands for the standard error of the genetic correlation, 95% CI = 95% confidence interval, FDR adjusted p-value = false discovery rate adjusted p-value

Supplementary Table 8. Canonical networks explain more variance than networks with randomly included volumes

Comparison with which network	Variance explained by the canonical network	No. of included volumes	Random network R^2	Confidence interval
Central Executive	65.23	8	46.37	46.05 - 46.68
Salience	49.05	10	45.01	44.72 - 45.31
Cingulo-opercular	47.5	10	45.01	44.72 - 45.31
Hippocampal-diencephalic	47.63	12	43.99	43.72 - 44.25
Multiple Demand	55.23	12	43.99	43.72 - 44.25
Sensorimotor	55.42	12	43.99	43.72 - 44.25
Default Mode	51.78	16	42.61	42.38 - 42.83
Temporo-amygdala-orbitrofrontal	46.92	30	40.86	40.72 - 41.01
P-FIT	47.56	36	40.59	40.45 - 40.72

Note. Random networks have been simulated 800 times per number of randomly included volumes. Comparisons with canonical networks must be made for matching number of included volumes. For example, the central executive network, which is constituted of 8 volumes, must be compared to a network including 8 randomly selected volumes.

Supplementary Table 9. Associations between brain volumes and potential covariates

Region	age	sex	time of year	x coordinate	y coordinate	z coordinate	site
Brain stem	-0.052	0.489	-0.002	0.016	-0.038	-0.064	0.032
Left accumbens area	-0.372	0.187	-0.002	-0.016	-0.133	0.045	0.117
Left amygdala	-0.296	0.351	0.001	-0.012	-0.064	0.022	0.063
Left bankssts	-0.123	0.257	0.003	0.008	-0.047	0.026	0.029
Left caudal anterior cingulate	-0.095	0.07	0.007	0.002	-0.034	0.053	0.01
Left caudal middle frontal	-0.149	0.257	-0.001	-0.002	-0.048	-0.066	0.031
Left caudate	-0.001	0.292	-0.008	-0.002	-0.05	-0.055	0.059
Left cuneus	-0.075	0.313	0.004	<0.001	-0.064	-0.03	0.047
Left DC	-0.202	0.461	-0.001	0.003	-0.096	-0.011	0.08
Left entorhinal	-0.036	0.23	-0.005	-0.011	-0.029	-0.037	0.022
Left frontal pole	-0.135	0.217	-0.004	0.003	-0.041	-0.005	0.003
Left fusiform	-0.161	0.374	-0.005	0.01	-0.074	0.004	0.046
Left hippocampus	-0.286	0.313	-0.001	0.007	-0.073	-0.01	0.055
Left inferior parietal	-0.161	0.284	-0.006	0.012	-0.068	0.015	0.046
Left inferior temporal	-0.127	0.41	-0.001	0.012	-0.067	0.015	0.045
Left insula	-0.023	0.456	-0.006	0.03	-0.053	-0.031	0.05
Left isthmus cingulate	-0.052	0.405	-0.004	0.026	-0.048	-0.034	0.044
Left lateral occipital	-0.137	0.419	0.006	0.006	-0.061	-0.022	0.046
Left lateral orbitofrontal	-0.164	0.405	0.002	0.016	-0.078	-0.031	0.055
Left lingual	-0.111	0.267	<0.001	0.005	-0.056	-0.056	0.03
Left medial orbitofrontal	-0.134	0.394	-0.003	<0.001	-0.078	0.012	0.049
Left middle temporal	-0.178	0.412	0.002	0.004	-0.071	0.03	0.055
Left pallidum	-0.097	0.399	0.003	0.006	-0.05	-0.032	0.033
Left paracentral	-0.168	0.224	0.001	0.009	-0.084	0.029	0.064
Left parahippocampal	-0.154	0.073	-0.002	-0.003	-0.051	0.026	0.023
Left pars opercularis	-0.176	0.239	-0.009	0.01	-0.054	-0.024	0.023
Left pars orbitalis	-0.182	0.325	-0.003	0.007	-0.065	-0.006	0.048
Left pars triangularis	-0.182	0.287	-0.007	0.011	-0.06	-0.01	0.039

Left pericalcarine	-0.011	0.247	0.003	-0.004	-0.063	-0.058	0.039
Left postcentral	-0.143	0.321	-0.005	0.001	-0.067	-0.031	0.049
Left posterior cingulate	-0.1	0.306	-0.003	0.002	-0.046	-0.018	0.045
Left precentral	-0.188	0.353	<0.001	-0.006	-0.069	-0.033	0.064
Left precuneus	-0.195	0.357	-0.004	0.012	-0.088	0.001	0.051
Left putamen	-0.177	0.4	-0.006	0.011	-0.064	-0.037	0.073
Left rostral anterior cingulate	-0.084	0.323	0.004	0.003	-0.057	0.014	0.045
Left rostral middle frontal	-0.193	0.422	-0.003	0.002	-0.056	-0.028	0.039
Left superior frontal	-0.203	0.408	-0.008	<0.001	-0.06	-0.07	0.046
Left superior parietal	-0.173	0.278	0.002	0.001	-0.07	-0.011	0.041
Left superior temporal	-0.176	0.396	-0.001	0.001	-0.064	0.01	0.05
Left supramarginal	-0.128	0.398	-0.002	0.019	-0.058	-0.007	0.035
Left thalamus proper	-0.276	0.367	-0.001	0.018	-0.09	-0.026	0.063
Left transverse temporal	-0.048	0.227	0.008	0.009	-0.055	-0.023	0.032
Right accumbens area	-0.303	0.246	-0.004	0.005	-0.104	0.021	0.085
Right amygdala	-0.21	0.414	-0.006	0.017	-0.061	-0.049	0.071
Right bankssts	-0.135	0.228	-0.002	-0.011	-0.065	0.021	0.042
Right caudal anterior cingulate	-0.099	0.11	-0.002	-0.005	-0.041	0.032	0.032
Right caudal middle frontal	-0.135	0.244	-0.001	0.006	-0.049	-0.063	0.029
Right caudate	0.029	0.318	-0.007	0.014	-0.027	-0.071	0.046
Right cuneus	-0.039	0.344	0.006	0.003	-0.046	-0.047	0.035
Right DC	-0.214	0.46	-0.003	0.006	-0.087	-0.029	0.068
Right entorhinal	-0.022	0.22	-0.01	0.01	-0.038	-0.024	0.03
Right frontal pole	-0.102	0.235	-0.006	-0.009	-0.052	-0.004	0.023
Right fusiform	-0.165	0.41	-0.004	0.001	-0.063	-0.026	0.043
Right hippocampus	-0.287	0.293	-0.007	-0.003	-0.072	-0.025	0.054
Right inferior parietal	-0.177	0.375	-0.008	0.007	-0.067	0.008	0.047
Right inferior temporal	-0.13	0.419	-0.001	0.016	-0.063	-0.016	0.051
Right insula	-0.037	0.485	-0.007	0.014	-0.054	-0.017	0.058
Right isthmus cingulate	-0.053	0.348	-0.005	-0.011	-0.041	-0.041	0.032

Region	age	sex	time of year	x coordinate	y coordinate	z coordinate	site
Right lateral occipital	-0.12	0.442	-0.003	0.009	-0.053	-0.047	0.046
Right lateral orbitofrontal	-0.153	0.392	0.002	0.004	-0.085	-0.024	0.065
Right lingual	-0.08	0.252	-0.002	-0.003	-0.042	-0.071	0.015
Right medial orbitofrontal	-0.15	0.407	0.001	0.012	-0.099	-0.011	0.082
Right middle temporal	-0.182	0.415	0.001	0.007	-0.067	0.002	0.054
Right pallidum	-0.077	0.399	0.003	0.026	-0.06	-0.038	0.045
Right paracentral	-0.149	0.266	0.003	-0.006	-0.073	0.029	0.059
Right parahippocampal	-0.138	0.108	-0.006	-0.019	-0.048	0.015	0.018
Right pars opercularis	-0.165	0.258	-0.007	-0.014	-0.053	-0.031	0.038
Right pars orbitalis	-0.196	0.315	-0.003	0.009	-0.06	-0.019	0.045
Right pars triangularis	-0.176	0.307	0.001	-0.007	-0.055	-0.019	0.041
Right pericalcarine	0.014	0.256	0.005	0.012	-0.055	-0.06	0.017
Right postcentral	-0.14	0.295	0.001	0.002	-0.071	-0.022	0.059
Right posterior cingulate	-0.129	0.281	-0.002	-0.003	-0.059	-0.011	0.057
Right precentral	-0.182	0.344	<0.001	0.003	-0.067	-0.036	0.071
Right precuneus	-0.168	0.398	-0.003	0.01	-0.073	-0.014	0.04
Right putamen	-0.176	0.395	-0.007	0.007	-0.053	-0.052	0.067
Right rostral anterior cingulate	-0.071	0.274	0.003	<0.001	-0.069	0.031	0.064
Right rostral middle frontal	-0.179	0.428	-0.001	0.009	-0.06	-0.024	0.054
Right superior frontal	-0.193	0.406	-0.006	0.002	-0.067	-0.061	0.05
Right superior parietal	-0.186	0.295	-0.002	0.01	-0.078	-0.018	0.051
Right superior temporal	-0.185	0.348	-0.003	0.001	-0.076	-0.005	0.06
Right supramarginal	-0.135	0.342	0.003	0.001	-0.069	-0.015	0.047
Right thalamus proper	-0.24	0.422	-0.002	-0.014	-0.071	-0.054	0.062
Right transverse temporal	-0.047	0.224	0.009	-0.004	-0.059	-0.01	0.045
Region	**age**	**sex**	**time of year**	**x coordinate**	**y coordinate**	**z coordinate**	**site**
Average r (*SD*)	**-0.14** (0.07)	**0.33** (0.09)	**-0.002** (0.004)	**0.004** (0.01)	**-0.06** (0.02)	**-0.02** (0.03)	**0.05** (0.02)

We pre-registered an arbitrary cut-off of *r* = .10 to include covariates in the GWAS. Therefore, only age and sex (of the covariates displayed in this table) were included as covariates in the GWAS.

Supplementary Table 10. Genetic quality control exclusion criteria resulting in a total GWAS sample of 36,778 out of 39,947 participants

Exclusion criterion	Excluded	Remaining
No genetic data	990	38,957
Non-European	919	38,038
Failed quality assurance by UK Biobank	72	37,966
High missingness rates	204	37,762
Relatedness within the sample	956	36,806
Mismatch between biological and self-reported sex	28	**36,778 (final)**

Supplementary Material for Chapter 5

Power calculation

Power calculations for the likelihood ratio test are not trivial and have only been derived for the LRT to test for significance of a vertex-wise morphometricity estimate, and not for ROI-based estimates (Couvy-Duchesne et al., 2020). To estimate power, we calculate the probability of rejecting the null hypothesis when the alternative hypothesis is true, as the probability that the likelihood ratio is larger than the quantile value of the central χ^2 distribution (null model) with degrees of freedom equal to the amount of variance components tested (with $a = 0.05/588$, $\chi^2_{(1-a)}(1) = 15.34$). If the alternative hypothesis is true, the LHR statistic is non-central χ^2 distributed with a non-centrality parameter obtained through the hypothesised variance accounted for (in this case the difference of explained variance between the null and the alternative model), and the standard error which can be approximated with Haseman-Elston regression (Sham & Purcell, 2001; Visscher et al., 2014) based on the variance of the off-diagonal BRM (var(B_{ij}); see Couvy-Duchesne et al. (2020)). We use realistic var(B_{ij}) values from a simulation presented in Couvy-Duchesne et al. (2020) which differ between surface area (var(B_{ij}) = 0.00047) and cortical thickness (var(B_{ij}) = 0.0017). We consider N = 38,000 and calculate non-centrality parameters of 33.93 and 30.69, respectively, for surface area (hypothesised difference in R^2 = 0.01%, SE = 0.0017) and cortical thickness (hypothesised difference in R^2 = 0.005%; SE = 0.0009). With these assumptions, power can be calculated using the R formula 1-pchisq(qchisq(1-(0.05/588), 1), 1, ncp), resulting in 97.5% power for surface area at a R^2 difference between null and alternative model of 0.01%, and a power of 95.4% for

cortical thickness at an even smaller R^2 difference of 0.005% (*STable 1*). Power could also be calculated online (https://shiny.cnsgenomics.com/gctaPower/)

Parameters for power calculations

	Surface area	**Cortical thickness**
N	38,000	38,000
effect size	0.01	0.005
Type 1 error rate *a*	0.05/ 588 = 8.5 x10^{-5}	8.5 x10^{-5}
$\chi^2_{(1-a)}(1)$	15.34	15.34
var(B_{ij})	0.00047	0.0017
standard error	0.0017	0.0009
ncp	33.93	30.69
Power (%)	97.10	94.63

ncp = non-centrality parameter

Linear mixed model (LMM) definitions

To estimate morphometricity, we fitted linear mixed models (LMMs, Fig.2). This approach recognises the correlation structure between the many independent variables, and was presented and validated elsewhere, using vertex-wise data (Couvy-Duchesne et al., 2020). The LMM models covariates as fixed effects and fits all cortical measurements as a vector of random effect. The vector of random effects (**b**) is constrained to a normal distribution with a mean of zero, and a structure of variance-covariance derived from the brain relatedness matrix (BRM, **B**), which quantifies the resemblance between each pair of individuals' brain. $\mathbf{b} \sim \mathcal{N}(0, \mathbf{B}\sigma_b^2)$ In practice, the BRM is an *NxN* variance-covariance matrix calculated for the standardised and centred vertices across pairs of participants. The BRM diagonal holds the variance s^2 of a participant's vertex-wise (or ROI-wise) measurements.

$$var(indiv\_k) = \frac{1}{n-1}\sum_i (x_{ik} - \bar{x_i})^2$$

With x_{ik} the value of vertex i for individual k. $\bar{x_i}$ the average vertex value of vertex i. n denotes the number of vertices. Large variances indicate that an individual has a bigger proportion of extremely small or large measurements, relative to the sample mean.

The BRM off-diagonal holds the covariance of participant's measurements, indicating how strongly participants resemble each other based on their grey-matter structure.

$$cov(indiv_k, indiv_j) = \frac{1}{n-1}\sum_i \left(x_{i_k} - \bar{x_i}\right)(x_{ij} - \bar{x_i})$$

We removed pairs of participants with outlying covariance (+-8SD from mean) from the analyses as they indicate oddly similar or dissimilar brains and could bias the

LMM results. The model will also include a standard vector of error terms, normally distributed with a mean of zero and a variance-covariance structure derived from the identity matrix (I): $e \sim \mathcal{N}(0, \mathbf{I}\sigma_e^2)$, which assumes that the errors are independent and identically distributed.

Supplementary Figures
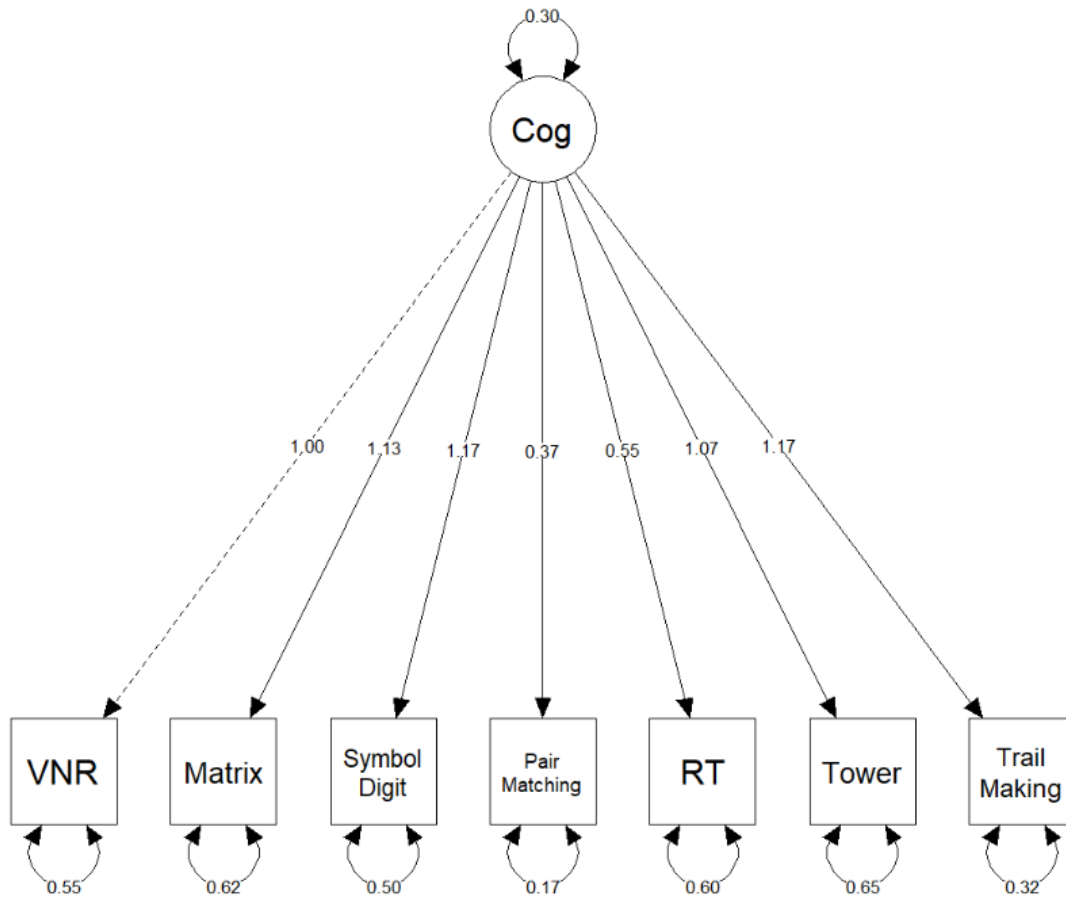
SFig.1. General cognitive ability factor model used to define cognitive phenotype considered for analysis. The total variance explained by this model is 34% of the variance contained in individual cognitive measures. Comparative Fit Index = 0.968, Tucker-Lewis Index = 0.951, SRMR = 0.024, RMSEA = 0.05. According to commonly-used rules of thumb (Hu & Bentler, 1998), this model has excellent model fit.
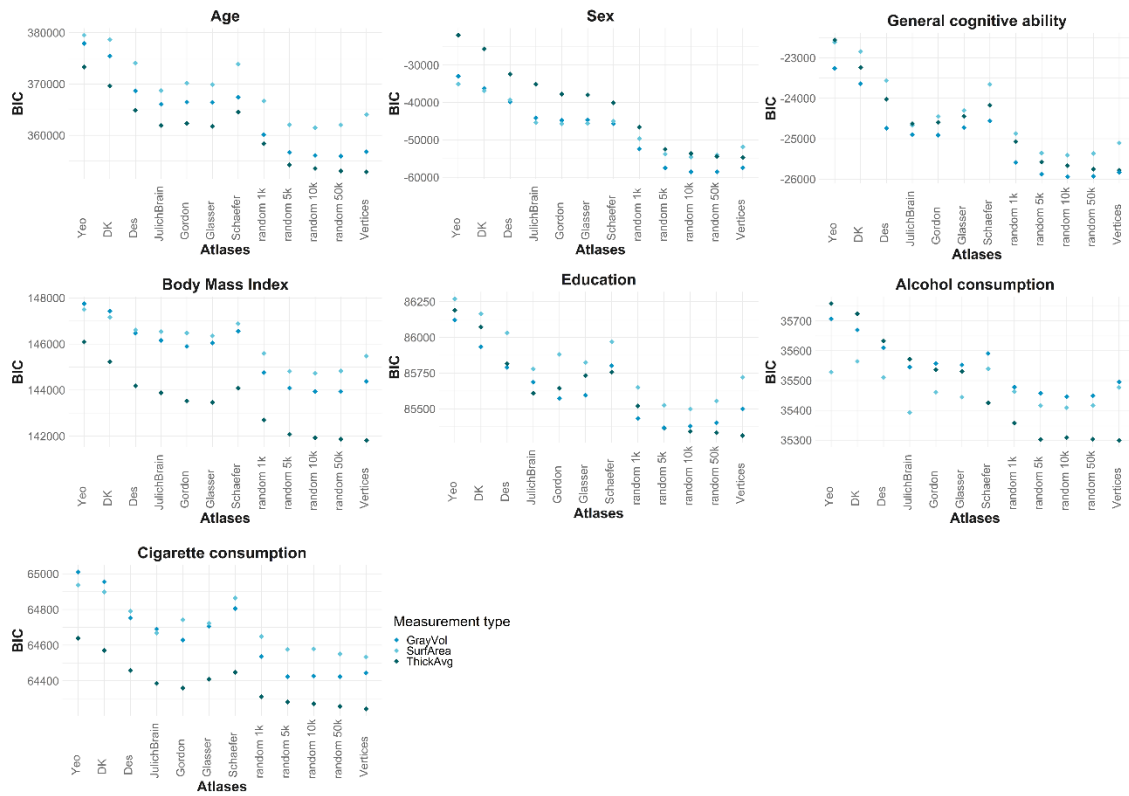
SFig.2. Bayesian Information Criterion (BIC). The BIC is defined as the difference between model complexity log(n)*p and model performance 2*LogL, meaning that smaller values indicate better model fit. In our case, p is always 1 because we only fit one variance component for each atlas, and n is always the same within one phenotype.
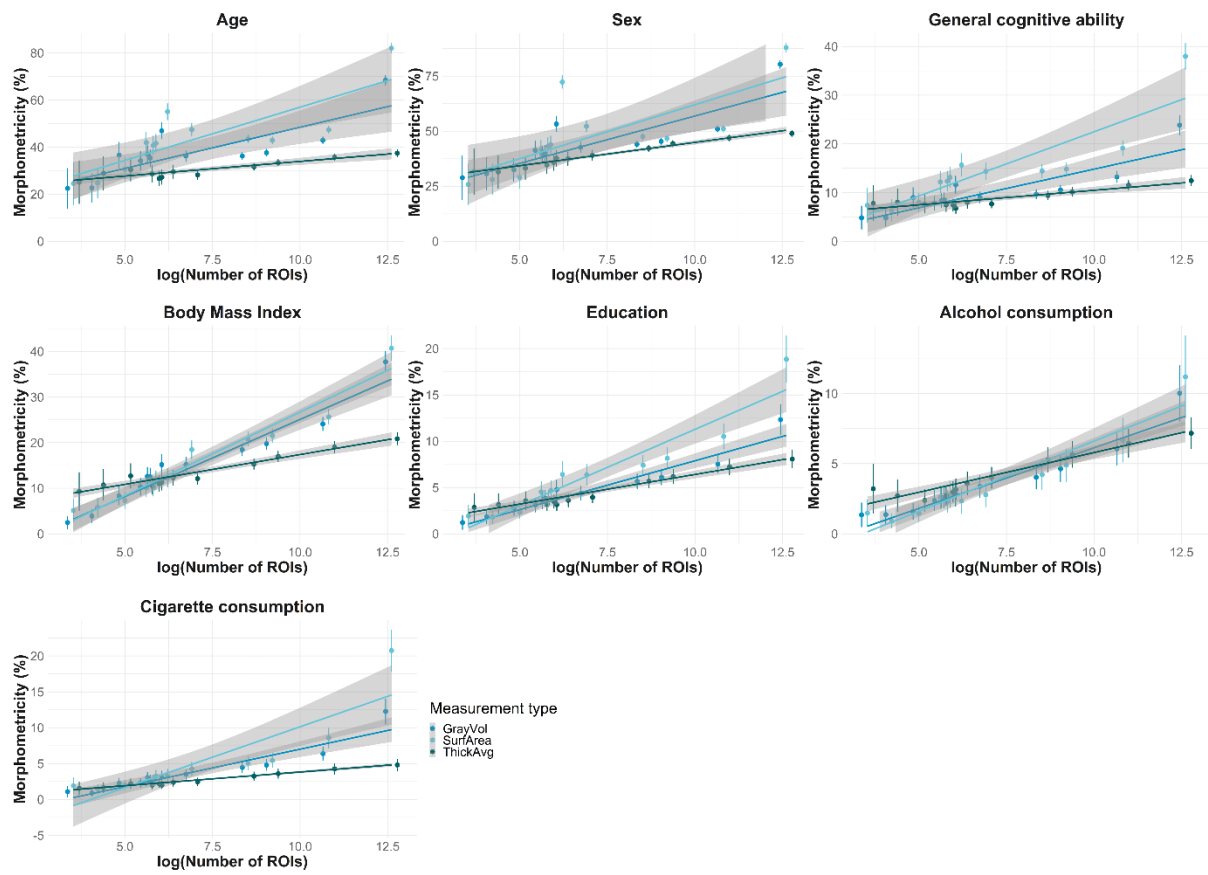
SFig.3. Non-linear relationships between atlas dimensionality and morphometricity. Linear log models were used to model these relationships, model parameters are reported in *Fig.1*.
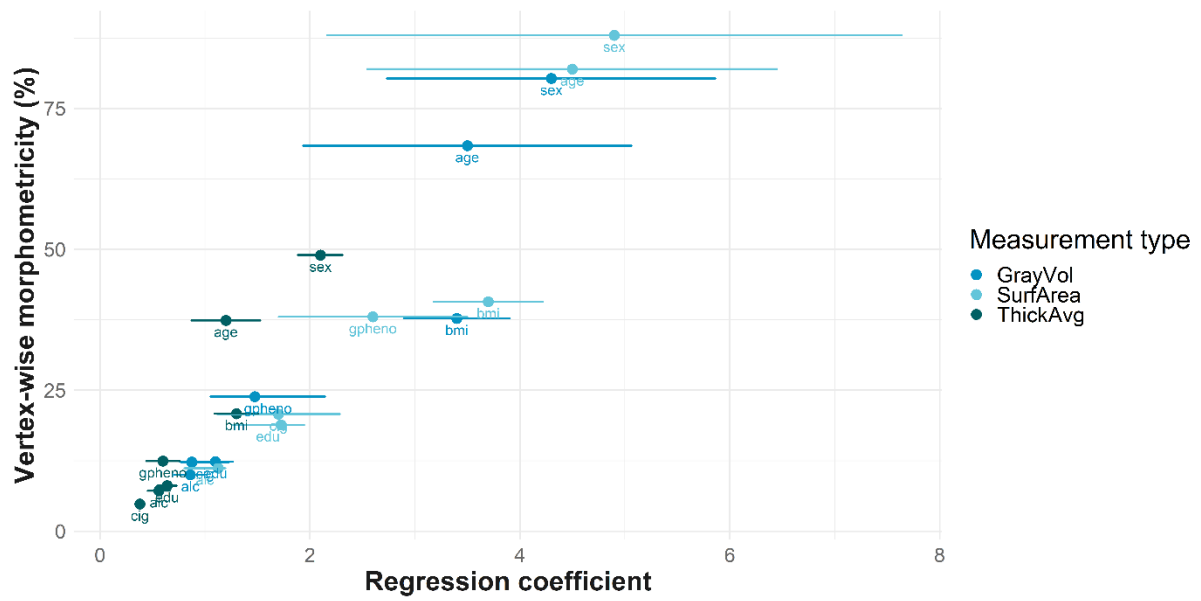
SFig.4. Relationship between trait morphometricity and regression coefficients obtained from linear log-models reported in *Table 1* and *SFig.3*. The more morphometric a trait, the stronger its morphometricity estimation is impacted by atlases dimensionality.
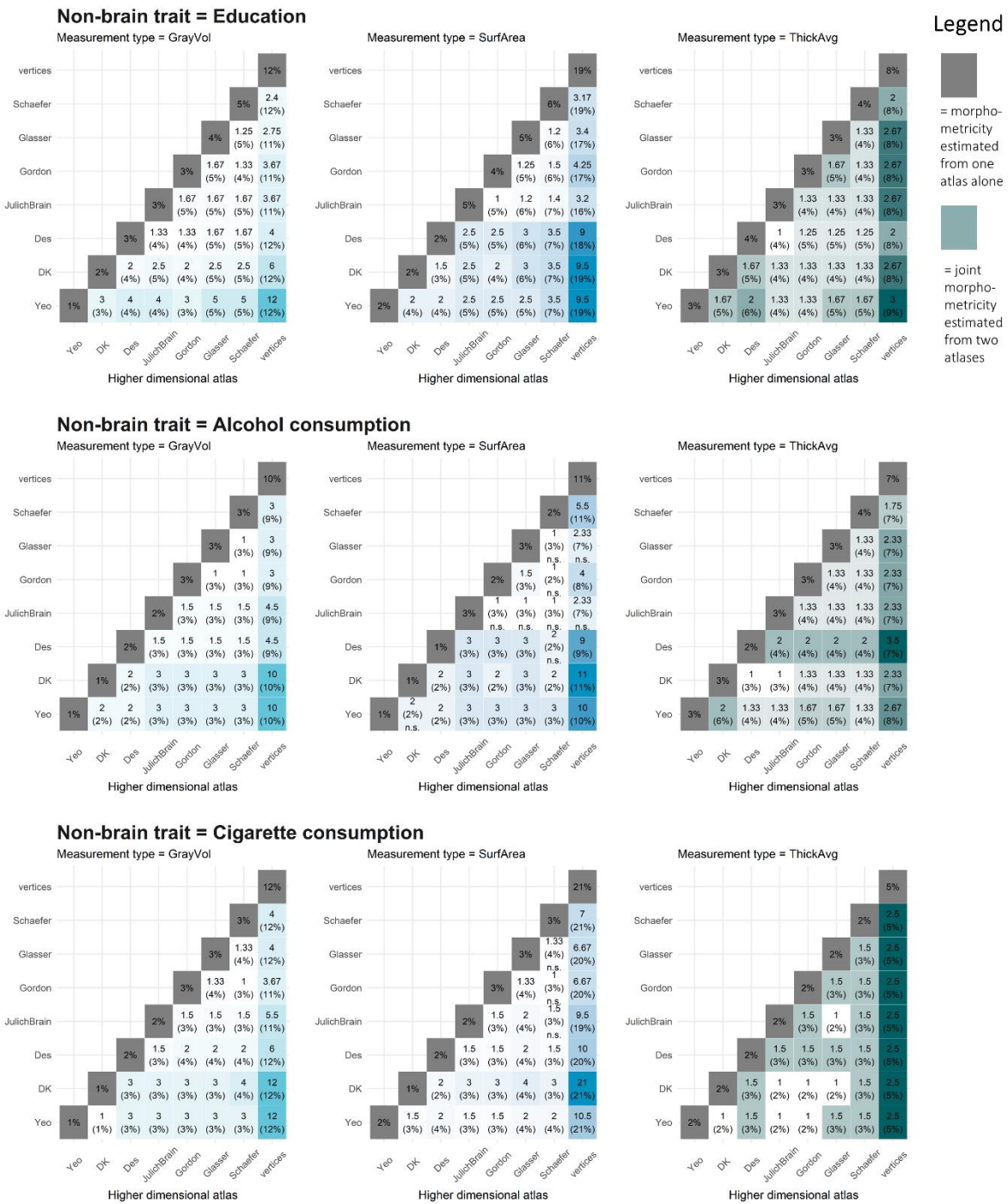
SFig5. Atlas comparisons in calculated from likelihood ratio tests for education, alcohol, and cigarette consumption. Find the equivalent visualisation for age, sex, body mass index, and cognitive ability in Fig.6. Percentages displayed on the diagonal are morphometricity estimates for atlases alone, as indicated in Fig.4. Indices on the off-diagonal show the relative improvement made to the model by adding the higher dimensional atlas, which we calculated as the sum of the variance explained by two atlases together, divided by the morphometricity estimate of the lower dimensional atlas alone ($\text{joint}_{\text{morphometricity}}$/ $\text{individual}_{\text{morphometricity}}$). Squares are coloured according to this ratio, i.e., larger ratio, darker colour. The raw sum of variance explained is printed in brackets below the respective index. Non-significant results are marked with *n.s.*
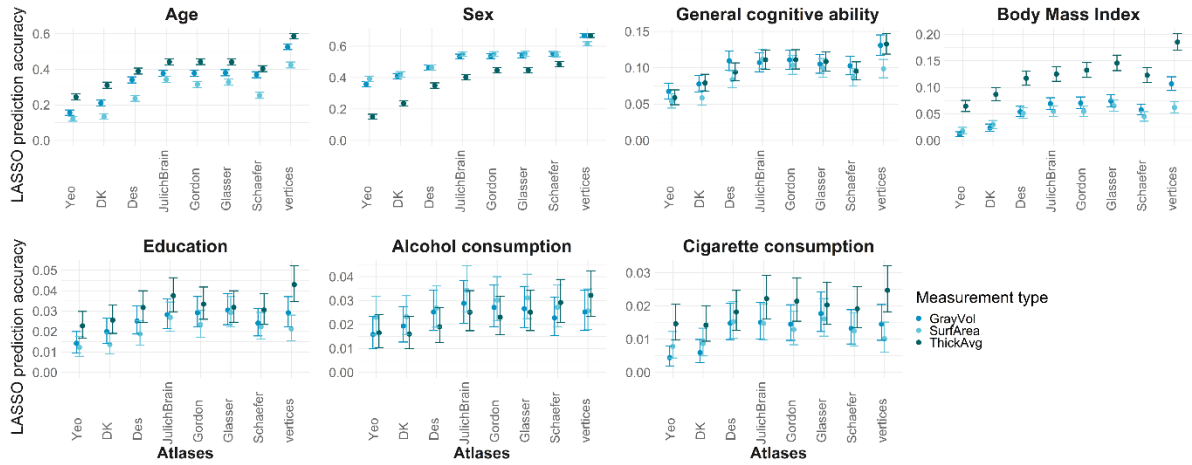
SFig.6. Prediction accuracy from LASSO models (i.e., R^2 between observed and LASSO predicted values). We split participants with available behavioural data 80-by-20% to perform out-of-sample prediction. Age, sex, cognitive ability, body mass index: $N_{discovery}$ = 30,802, $N_{evaluation}$ = 7,701; education: $N_{discovery}$ = 30,787, $N_{evaluation}$ = 7,697; alcohol consumption: $N_{discovery}$ = 20,811, $N_{evaluation}$ = 5,203; cigarette consumption: $N_{discovery}$ = 29,866, $N_{evaluation}$ = 7,467

Non-brain trait = Age
Vertex-wise estimates
LMM = 37.4%
LASSO = 58.65%
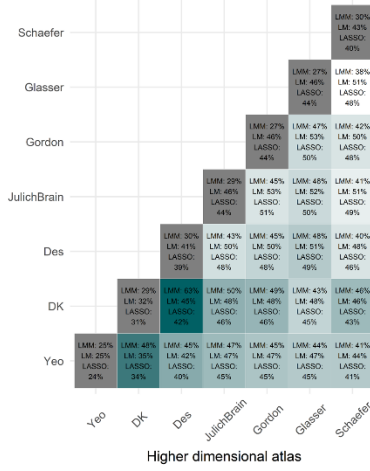
Non-brain trait = Sex
Vertex-wise estimates
LMM = 48.97%
LASSO = 66.74%
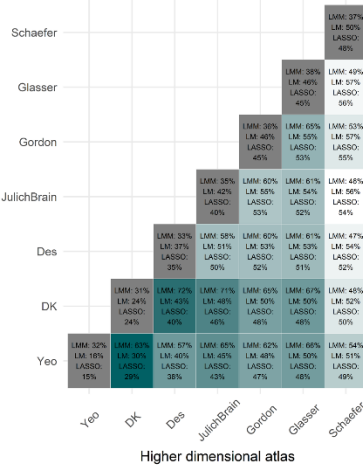
Non-brain trait = General cognitive ability
Vertex-wise estimates
LMM = 12.45%
LASSO = 13.25%
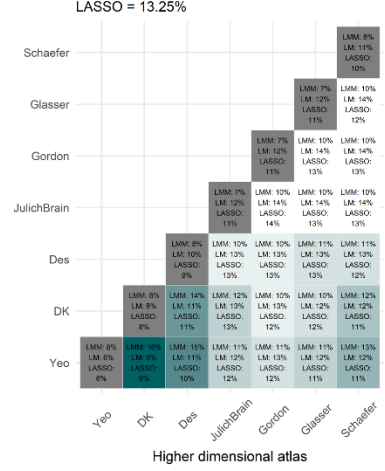
Non-brain trait = Body Mass Index
Vertex-wise estimates
LMM = 20.84%
LASSO = 18.55%
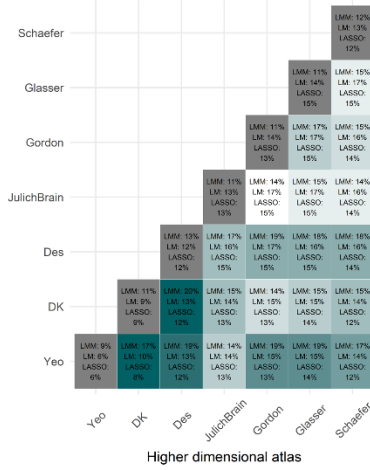
Non-brain trait = Education
Vertex-wise estimates
LMM = 8.08%
LASSO = 4.3%
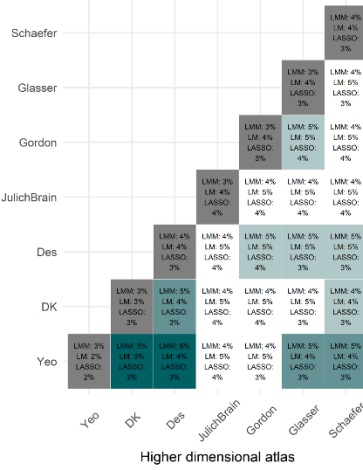
Non-brain trait = Alcohol consumption
Vertex-wise estimates
LMM = 7.17%
LASSO = 3.22%
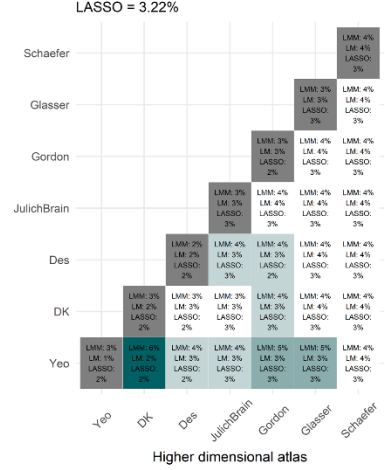
Non-brain trait = Cigarette consumption
Vertex-wise estimates
LMM = 4.81%
LASSO = 2.47%
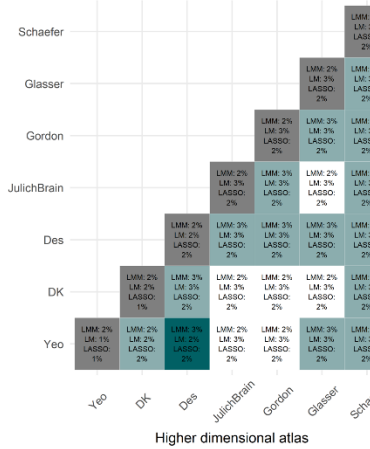
Figure legend

= morphometricity estimated by LMM, LM, and LASSO

= joint atlas variance explained estimated by LMM, LM, and LASSO

266

SFig.7. Morphometricity and joint-atlas variance explained estimated using LMM, LM, and LASSO. Grey squares indicate morphometricity explained by an individual atlas, and green squares indicate the joint-atlas variance explained. The darker the green colour, the more divergent LMM estimates manifest from LM estimates.

Supplementary Tables

Estimates from empirical atlases, large random atlases and vertices (as displayed in Figure 4)

Atlas	Measurement type	Pheno-type	Morpho-metricity	Standard error	LRT	BIC	N	95% CI (lower bound)	95% CI (upper bound)
Des	GrayVol	age	36.5803	0.027408	15577.415	368673.4185	38503	31.208332	41.952268
DK	GrayVol	age	22.6325	0.030564	8777.434	375473.3985	38503	16.641956	28.623044
Glasser	GrayVol	age	35.3488	0.01784	17834.44	366416.3925	38503	31.85216	38.84544
Gordon	GrayVol	age	36.9436	0.018793	17762.98	366487.8525	38503	33.260172	40.627028
JulichBrain	GrayVol	age	34.1846	0.020196	18187.976	366062.8565	38503	30.226184	38.143016
Schaefer	GrayVol	age	46.9021	0.017139	16838.813	367412.0205	38503	43.542856	50.261344
Yeo	GrayVol	age	22.455	0.042824	6329.41	377921.4225	38503	14.061496	30.848504
vertices	GrayVol	age	68.3814	0.010032	27459.212	356791.6205	38503	66.415128	70.347672
10000randomROIs	GrayVol	age	37.6078	0.007033	28165.322	356085.5105	38503	36.229332	38.986268
1000randomROIs	GrayVol	age	36.1911	0.011737	24126.624	360124.2085	38503	33.890648	38.491552
50000randomROIs	GrayVol	age	42.8436	0.007112	28339.453	355911.3785	38503	41.449648	44.237552
5000randomROIs	GrayVol	age	36.17	0.007494	27571.21	356679.6225	38503	34.701176	37.638824
vertices	SurfArea	age	81.9581	0.010746	20202.352	364048.4805	38503	79.851884	84.064316
Des	SurfArea	age	30.4885	0.025191	10168.479	374082.3545	38503	25.551064	35.425936
DK	SurfArea	age	24.6918	0.032497	5605.172	378645.6605	38503	18.322388	31.061212
Glasser	SurfArea	age	41.5744	0.019203	14385.586	369865.2465	38503	37.810612	45.338188
Gordon	SurfArea	age	40.6923	0.019725	14051.706	370199.1265	38503	36.8262	44.5584
JulichBrain	SurfArea	age	41.9453	0.021891	15511.899	368738.9325	38503	37.654664	46.235936
Schaefer	SurfArea	age	55.005	0.01761	10386.267	373864.5645	38503	51.55344	58.45656
Yeo	SurfArea	age	24.5316	0.045579	4718.96	379531.8725	38503	15.598116	33.465084

10000randomROIs	SurfArea	age	42.8583	0.00717	22766.912	361483.9205	38503	41.45298	44.26362
1000randomROIs	SurfArea	age	47.457	0.013078	17565.564	366685.2685	38503	44.893712	50.020288
50000randomROIs	SurfArea	age	47.3281	0.007034	22241.394	362009.4385	38503	45.949436	48.706764
5000randomROIs	SurfArea	age	43.4295	0.008015	22210.062	362040.7705	38503	41.85856	45.00044
Des	ThickAvg	age	30.4347	0.025097	19377.057	364873.7745	38503	25.515688	35.353712
DK	ThickAvg	age	28.8271	0.035641	14637.241	369613.5905	38503	21.841464	35.812736
Glasser	ThickAvg	age	27.2135	0.015435	22480.168	361770.6645	38503	24.18824	30.23876
Gordon	ThickAvg	age	26.6176	0.015767	21934.801	362316.0325	38503	23.527268	29.707932
JulichBrain	ThickAvg	age	28.6363	0.01813	22316.182	361934.6505	38503	25.08282	32.18978
Schaefer	ThickAvg	age	29.5182	0.01418	19719.441	364531.3925	38503	26.73892	32.29748
Yeo	ThickAvg	age	25.0744	0.046074	10912.119	373338.7125	38503	16.043896	34.104904
vertices	ThickAvg	age	37.396	0.00704	31386.593	352864.2405	38503	36.01616	38.77584
10000randomROIs	ThickAvg	age	33.3778	0.006943	30705.477	353545.3545	38503	32.016972	34.738628
1000randomROIs	ThickAvg	age	28.1622	0.010327	25889.527	358361.3045	38503	26.138108	30.186292
50000randomROIs	ThickAvg	age	35.6854	0.006863	31228.901	353021.9305	38503	34.340252	37.030548
5000randomROIs	ThickAvg	age	31.6107	0.007183	30023.431	354227.4025	38503	30.202832	33.018568
Des	GrayVol	alc	1.5987	0.002926	436.228	35610.57039	26014	1.025204	2.172196
DK	GrayVol	alc	1.3743	0.003229	376.858	35669.94039	26014	0.741416	2.007184
Glasser	GrayVol	alc	2.5695	0.003672	494.271	35552.52839	26014	1.849788	3.289212
Gordon	GrayVol	alc	2.5813	0.003738	489.288	35557.51039	26014	1.848652	3.313948
JulichBrain	GrayVol	alc	2.3703	0.003556	501.675	35545.12239	26014	1.673324	3.067276
Schaefer	GrayVol	alc	2.8286	0.004379	455.973	35590.82639	26014	1.970316	3.686884
Yeo	GrayVol	alc	1.3492	0.0043	339.284	35707.51439	26014	0.5064	2.192
vertices	GrayVol	alc	10.0093	0.01008	551.53	35495.26839	26014	8.03362	11.98498
10000randomROIs	GrayVol	alc	4.627	0.004654	600.724	35446.07439	26014	3.714816	5.539184
1000randomROIs	GrayVol	alc	3.3884	0.003939	568.624	35478.17439	26014	2.616356	4.160444
50000randomROIs	GrayVol	alc	6.0485	0.005907	597.816	35448.98239	26014	4.890728	7.206272

5000randomROIs	GrayVol	alc	4.0255	0.004215	589.268	35457.53039	26014	3.19936	4.85164
vertices	SurfArea	alc	11.1893	0.014866	569.879	35476.91839	26014	8.275564	14.103036
Des	SurfArea	alc	1.4612	0.002851	535.817	35510.98239	26014	0.902404	2.019996
DK	SurfArea	alc	0.908	0.002461	482.393	35564.40639	26014	0.425644	1.390356
Glasser	SurfArea	alc	2.8049	0.004481	602.446	35444.35239	26014	1.926624	3.683176
Gordon	SurfArea	alc	2.3068	0.003838	585.594	35461.20439	26014	1.554552	3.059048
JulichBrain	SurfArea	alc	2.5949	0.004174	654.716	35392.08239	26014	1.776796	3.413004
Schaefer	SurfArea	alc	2.3422	0.004553	507.735	35539.06239	26014	1.449812	3.234588
Yeo	SurfArea	alc	1.4788	0.004782	517.939	35528.86039	26014	0.541528	2.416072
10000randomROIs	SurfArea	alc	4.886	0.00589	637.877	35408.92039	26014	3.73156	6.04044
1000randomROIs	SurfArea	alc	2.781	0.004288	584.291	35462.50839	26014	1.940552	3.621448
50000randomROIs	SurfArea	alc	6.6733	0.007953	630.411	35416.38639	26014	5.114512	8.232088
5000randomROIs	SurfArea	alc	4.2168	0.005347	631.136	35415.66239	26014	3.168788	5.264812
Des	ThickAvg	alc	2.3874	0.003873	414.339	35632.46039	26014	1.628292	3.146508
DK	ThickAvg	alc	2.7092	0.005725	322.37	35724.42839	26014	1.5871	3.8313
Glasser	ThickAvg	alc	3.1646	0.003784	515.833	35530.96639	26014	2.422936	3.906264
Gordon	ThickAvg	alc	2.9878	0.003687	510.787	35536.01039	26014	2.265148	3.710452
JulichBrain	ThickAvg	alc	2.7212	0.003597	474.59	35572.20839	26014	2.016188	3.426212
Schaefer	ThickAvg	alc	3.5868	0.004155	621.634	35425.16439	26014	2.77242	4.40118
Yeo	ThickAvg	alc	3.2113	0.008809	288.452	35758.34639	26014	1.484736	4.937864
vertices	ThickAvg	alc	7.1666	0.005646	747.67	35299.12839	26014	6.059984	8.273216
10000randomROIs	ThickAvg	alc	5.6488	0.004625	738.076	35308.72239	26014	4.7423	6.5553
1000randomROIs	ThickAvg	alc	3.9997	0.003829	689.854	35356.94439	26014	3.249216	4.750184
50000randomROIs	ThickAvg	alc	6.4538	0.005146	743.795	35303.00439	26014	5.445184	7.462416
5000randomROIs	ThickAvg	alc	5.2768	0.004387	744.537	35302.26239	26014	4.416948	6.136652
Des	GrayVol	bmi	8.2979	0.009735	1650.336	146467.5557	38394	6.38984	10.20596
DK	GrayVol	bmi	3.9021	0.007364	689.94	147427.9517	38394	2.458756	5.345444
Glasser	GrayVol	bmi	12.4799	0.00961	2069.865	146048.0277	38394	10.59634	14.36346

Gordon	GrayVol	bmi	12.5993	0.009946	2226.739	145891.1537	38394	10.649884	14.548716
JulichBrain	GrayVol	bmi	10.3028	0.009308	1963.047	146154.8457	38394	8.478432	12.127168
Schaefer	GrayVol	bmi	15.1659	0.010923	1562.128	146555.7637	38394	13.024992	17.306808
Yeo	GrayVol	bmi	2.4865	0.006773	367.409	147750.4837	38394	1.158992	3.814008
vertices	GrayVol	bmi	37.7222	0.011286	3743.048	144374.8437	38394	35.510144	39.934256
10000randomROIs	GrayVol	bmi	19.7824	0.006545	4178.249	143939.6437	38394	18.49958	21.06522
1000randomROIs	GrayVol	bmi	15.1967	0.007852	3359.237	144758.6557	38394	13.657708	16.735692
50000randomROIs	GrayVol	bmi	24.0953	0.007364	4181.224	143936.6677	38394	22.651956	25.538644
5000randomROIs	GrayVol	bmi	18.4372	0.006542	4030.3	144087.5917	38394	17.154968	19.719432
vertices	SurfArea	bmi	40.6913	0.013847	2643.224	145474.6677	38394	37.977288	43.405312
Des	SurfArea	bmi	7.3278	0.00886	1508.53	146609.3617	38394	5.59124	9.06436
DK	SurfArea	bmi	5.819	0.010345	960.355	147157.5377	38394	3.79138	7.84662
Glasser	SurfArea	bmi	11.7969	0.010086	1760.988	146356.9037	38394	9.820044	13.773756
Gordon	SurfArea	bmi	10.0965	0.009122	1636.682	146481.2097	38394	8.308588	11.884412
JulichBrain	SurfArea	bmi	10.0585	0.009852	1576.721	146541.1717	38394	8.127508	11.989492
Schaefer	SurfArea	bmi	12.2649	0.010434	1238.405	146879.4877	38394	10.219836	14.309964
Yeo	SurfArea	bmi	5.1223	0.012816	621.353	147496.5397	38394	2.610364	7.634236
10000randomROIs	SurfArea	bmi	21.5056	0.007555	3396.28	144721.6117	38394	20.02482	22.98638
1000randomROIs	SurfArea	bmi	18.4467	0.010318	2528.566	145589.3257	38394	16.424372	20.469028
50000randomROIs	SurfArea	bmi	25.6494	0.008244	3290.6	144827.2917	38394	24.033576	27.265224
5000randomROIs	SurfArea	bmi	20.8025	0.007917	3305.028	144812.8637	38394	19.250768	22.354232
Des	ThickAvg	bmi	12.6981	0.013735	3934.032	144183.8597	38394	10.00604	15.39016
DK	ThickAvg	bmi	10.7267	0.017151	2883.993	145233.8997	38394	7.365104	14.088296
Glasser	ThickAvg	bmi	11.2713	0.008449	4651.009	143466.8837	38394	9.615296	12.927304
Gordon	ThickAvg	bmi	11.0039	0.008557	4585.202	143532.6897	38394	9.326728	12.681072
JulichBrain	ThickAvg	bmi	10.5138	0.00898	4237.877	143880.0157	38394	8.75372	12.27388
Schaefer	ThickAvg	bmi	12.2845	0.008217	4035.523	144082.3697	38394	10.673968	13.895032
Yeo	ThickAvg	bmi	9.2688	0.021127	2024.23	146093.6617	38394	5.127908	13.409692

vertices	ThickAvg	bmi	20.8368	0.006509	6300.044	141817.8477	38394	19.561036	22.112564
10000randomROIs	ThickAvg	bmi	16.9517	0.005867	6191.57	141926.3217	38394	15.801768	18.101632
1000randomROIs	ThickAvg	bmi	12.0968	0.006343	5415.254	142702.6377	38394	10.853572	13.340028
50000randomROIs	ThickAvg	bmi	19.0368	0.006144	6250.103	141867.7897	38394	17.832576	20.241024
5000randomROIs	ThickAvg	bmi	15.3007	0.005682	6037.764	142080.1277	38394	14.187028	16.414372
Des	GrayVol	cig	2.2598	0.00345	333.874	64753.03563	37333	1.5836	2.936
DK	GrayVol	cig	0.9	0.002178	132.089	64954.81963	37333	0.473112	1.326888
Glasser	GrayVol	cig	2.6875	0.003316	381.534	64705.37563	37333	2.037564	3.337436
Gordon	GrayVol	cig	3.061	0.003627	458.813	64628.09563	37333	2.350108	3.771892
JulichBrain	GrayVol	cig	2.3245	0.003088	396.54	64690.36963	37333	1.719252	2.929748
Schaefer	GrayVol	cig	3.118	0.004164	281.812	64805.09763	37333	2.301856	3.934144
Yeo	GrayVol	cig	1.0802	0.003523	77.382	65009.52763	37333	0.389692	1.770708
vertices	GrayVol	cig	12.258	0.008859	642.038	64444.87163	37333	10.521636	13.994364
10000randomROIs	GrayVol	cig	4.7965	0.003788	660.061	64426.84763	37333	4.054052	5.538948
1000randomROIs	GrayVol	cig	3.4926	0.003344	549.76	64537.14963	37333	2.837176	4.148024
50000randomROIs	GrayVol	cig	6.3847	0.004808	662.883	64424.02563	37333	5.442332	7.327068
5000randomROIs	GrayVol	cig	4.4567	0.003587	662.797	64424.11163	37333	3.753648	5.159752
vertices	SurfArea	cig	20.7596	0.014552	552.997	64533.91363	37333	17.907408	23.611792
Des	SurfArea	cig	2.0252	0.003254	296.989	64789.91963	37333	1.387416	2.662984
DK	SurfArea	cig	1.4143	0.003191	188.837	64898.07163	37333	0.788864	2.039736
Glasser	SurfArea	cig	3.2144	0.004286	363.721	64723.18763	37333	2.374344	4.054456
Gordon	SurfArea	cig	2.6525	0.003641	343.877	64743.03163	37333	1.938864	3.366136
JulichBrain	SurfArea	cig	2.4581	0.003563	417.952	64668.95763	37333	1.759752	3.156448
Schaefer	SurfArea	cig	3.2597	0.004764	222.276	64864.63363	37333	2.325956	4.193444
Yeo	SurfArea	cig	1.9126	0.005805	150.884	64936.02563	37333	0.77482	3.05038
10000randomROIs	SurfArea	cig	5.4338	0.004937	508.454	64578.45563	37333	4.466148	6.401452
1000randomROIs	SurfArea	cig	4.2653	0.004675	438.204	64648.70563	37333	3.349	5.1816

50000randomROIs	SurfArea	cig	8.5699	0.006974	535.696	64551.21363	37333	7.202996	9.936804
5000randomROIs	SurfArea	cig	5.0672	0.004726	509.92	64576.98963	37333	4.140904	5.993496
Des	ThickAvg	cig	2.215	0.003475	627.478	64459.43163	37333	1.5339	2.8961
DK	ThickAvg	cig	1.6324	0.003541	515.817	64571.09163	37333	0.938364	2.326436
Glasser	ThickAvg	cig	1.9949	0.002479	676.743	64410.16563	37333	1.509016	2.480784
Gordon	ThickAvg	cig	2.189	0.002686	727.747	64359.16163	37333	1.662544	2.715456
JulichBrain	ThickAvg	cig	1.9901	0.002616	700.322	64386.58763	37333	1.477364	2.502836
Schaefer	ThickAvg	cig	2.3803	0.002858	638.862	64448.04763	37333	1.820132	2.940468
Yeo	ThickAvg	cig	1.5256	0.004409	448.256	64638.65363	37333	0.661436	2.389764
vertices	ThickAvg	cig	4.8146	0.003988	843.529	64243.37963	37333	4.032952	5.596248
10000randomROIs	ThickAvg	cig	3.5798	0.003156	815.32	64271.58963	37333	2.961224	4.198376
1000randomROIs	ThickAvg	cig	2.4518	0.002537	775.034	64311.87563	37333	1.954548	2.949052
50000randomROIs	ThickAvg	cig	4.2619	0.003604	830.007	64256.90163	37333	3.555516	4.968284
5000randomROIs	ThickAvg	cig	3.2324	0.002936	805.076	64281.83363	37333	2.656944	3.807856
Des	GrayVol	edu	2.7618	0.004003	1075.12	85789.7	38484	1.977212	3.546388
DK	GrayVol	edu	1.8352	0.003791	929.977	85934.844	38484	1.092164	2.578236
Glasser	GrayVol	edu	3.8352	0.00413	1270.185	85594.636	38484	3.02572	4.64468
Gordon	GrayVol	edu	3.4974	0.00389	1291.662	85573.16	38484	2.73496	4.25984
JulichBrain	GrayVol	edu	3.1271	0.003799	1178.088	85686.732	38484	2.382496	3.871704
Schaefer	GrayVol	edu	4.763	0.005184	1062.749	85802.072	38484	3.746936	5.779064
Yeo	GrayVol	edu	1.2342	0.003809	743.534	86121.288	38484	0.487636	1.980764
vertices	GrayVol	edu	12.3335	0.008259	1365.868	85498.954	38484	10.714736	13.952264
10000randomROIs	GrayVol	edu	6.1104	0.004126	1486.988	85377.832	38484	5.301704	6.919096
1000randomROIs	GrayVol	edu	4.394	0.003732	1432.626	85432.196	38484	3.662528	5.125472
50000randomROIs	GrayVol	edu	7.531	0.004953	1463.376	85401.444	38484	6.560212	8.501788
5000randomROIs	GrayVol	edu	5.6653	0.00394	1498.529	85366.292	38484	4.89306	6.43754
vertices	SurfArea	edu	18.8645	0.012829	1144.854	85719.966	38484	16.350016	21.378984

Des	SurfArea	edu	2.351	0.003549	832.816	86032.004	38484	1.655396	3.046604
DK	SurfArea	edu	1.8279	0.003855	699.431	86165.39	38484	1.07232	2.58348
Glasser	SurfArea	edu	4.6733	0.005317	1040.717	85824.104	38484	3.631168	5.715432
Gordon	SurfArea	edu	3.8112	0.004545	983.405	85881.416	38484	2.92038	4.70202
JulichBrain	SurfArea	edu	4.5268	0.005481	1086.383	85778.438	38484	3.452524	5.601076
Schaefer	SurfArea	edu	6.4127	0.007004	895.3	85969.52	38484	5.039916	7.785484
Yeo	SurfArea	edu	1.9382	0.005648	594.289	86270.532	38484	0.831192	3.045208
10000randomROIs	SurfArea	edu	8.1597	0.005665	1366.181	85498.64	38484	7.04936	9.27004
1000randomROIs	SurfArea	edu	6.3654	0.005787	1214.71	85650.112	38484	5.231148	7.499652
50000randomROIs	SurfArea	edu	10.5087	0.006974	1310.383	85554.438	38484	9.141796	11.875604
5000randomROIs	SurfArea	edu	7.41	0.005445	1340.454	85524.366	38484	6.34278	8.47722
Des	ThickAvg	edu	3.5751	0.004885	1049.46	85815.36	38484	2.61764	4.53256
DK	ThickAvg	edu	3.149	0.005982	792.686	86072.134	38484	1.976528	4.321472
Glasser	ThickAvg	edu	3.1444	0.003284	1132.107	85732.714	38484	2.500736	3.788064
Gordon	ThickAvg	edu	3.4664	0.003572	1221.373	85643.448	38484	2.766288	4.166512
JulichBrain	ThickAvg	edu	3.2149	0.003567	1255.833	85608.988	38484	2.515768	3.914032
Schaefer	ThickAvg	edu	3.6159	0.003518	1108.14	85756.68	38484	2.926372	4.305428
Yeo	ThickAvg	edu	2.8678	0.00749	673.968	86190.852	38484	1.39976	4.33584
vertices	ThickAvg	edu	8.0762	0.004743	1552.282	85312.538	38484	7.146572	9.005828
10000randomROIs	ThickAvg	edu	6.1863	0.003885	1523.818	85341.004	38484	5.42484	6.94776
1000randomROIs	ThickAvg	edu	3.9632	0.003187	1345.22	85519.6	38484	3.338548	4.587852
50000randomROIs	ThickAvg	edu	7.2198	0.004338	1531.755	85333.066	38484	6.369552	8.070048
5000randomROIs	ThickAvg	edu	5.687	0.003686	1500.969	85363.852	38484	4.964544	6.409456
Des	GrayVol	gpheno	9.0163	0.010352	4429.702	-24743.25574	38494	6.987308	11.045292
DK	GrayVol	gpheno	4.7708	0.008509	3326.048	-23639.60174	38494	3.103036	6.438564
Glasser	GrayVol	gpheno	8.4806	0.007183	4411.057	-24724.60974	38494	7.072732	9.888468
Gordon	GrayVol	gpheno	8.3738	0.007202	4598.841	-24912.39374	38494	6.962208	9.785392
JulichBrain	GrayVol	gpheno	7.8453	0.007406	4583.89	-24897.44374	38494	6.393724	9.296876

Schaefer	GrayVol	gpheno	11.5861	0.009067	4249.802	-24563.35574	38494	9.808968	13.363232
Yeo	GrayVol	gpheno	4.8282	0.012046	2945.553	-23259.10574	38494	2.467184	7.189216
vertices	GrayVol	gpheno	23.8461	0.009943	5521.144	-25834.69774	38494	21.897272	25.794928
10000randomROIs	GrayVol	gpheno	10.5224	0.004952	5628.222	-25941.77574	38494	9.551808	11.492992
1000randomROIs	GrayVol	gpheno	9.0041	0.005555	5277.679	-25591.23174	38494	7.91532	10.09288
50000randomROIs	GrayVol	gpheno	13.2281	0.005876	5618.84	-25932.39374	38494	12.076404	14.379796
5000randomROIs	GrayVol	gpheno	9.6652	0.004774	5568.103	-25881.65574	38494	8.729496	10.600904
vertices	SurfArea	gpheno	38.0167	0.013545	4793.002	-25106.55574	38494	35.36188	40.67152
Des	SurfArea	gpheno	7.985	0.009422	3249.071	-23562.62374	38494	6.138288	9.831712
DK	SurfArea	gpheno	6.3434	0.011015	2529.653	-22843.20574	38494	4.18446	8.50234
Glasser	SurfArea	gpheno	13.1181	0.01054	3989.046	-24302.59974	38494	11.05226	15.18394
Gordon	SurfArea	gpheno	12.3541	0.010192	4136.145	-24449.69774	38494	10.356468	14.351732
JulichBrain	SurfArea	gpheno	12.1949	0.010934	4355.53	-24669.08374	38494	10.051836	14.337964
Schaefer	SurfArea	gpheno	15.6662	0.011776	3343.229	-23656.78174	38494	13.358104	17.974296
Yeo	SurfArea	gpheno	7.4033	0.017649	2301.338	-22614.89174	38494	3.944096	10.862504
10000randomROIs	SurfArea	gpheno	14.8144	0.00655	5095.369	-25408.92174	38494	13.5306	16.0982
1000randomROIs	SurfArea	gpheno	14.3532	0.008736	4562.491	-24876.04374	38494	12.640944	16.065456
50000randomROIs	SurfArea	gpheno	19.1314	0.007631	5052.247	-25365.79974	38494	17.635724	20.627076
5000randomROIs	SurfArea	gpheno	14.4226	0.006747	5040.741	-25354.29374	38494	13.100188	15.745012
Des	ThickAvg	gpheno	7.6382	0.008962	3707.942	-24021.49374	38494	5.881648	9.394752
DK	ThickAvg	gpheno	8.0358	0.013374	2923.854	-23237.40774	38494	5.414496	10.657104
Glasser	ThickAvg	gpheno	6.7863	0.005663	4130.47	-24444.02374	38494	5.676352	7.896248
Gordon	ThickAvg	gpheno	7.4162	0.006219	4283.403	-24596.95574	38494	6.197276	8.635124
JulichBrain	ThickAvg	gpheno	7.4544	0.006782	4317.352	-24630.90574	38494	6.125128	8.783672
Schaefer	ThickAvg	gpheno	7.9811	0.005977	3860.522	-24174.07574	38494	6.809608	9.152592
Yeo	ThickAvg	gpheno	7.7953	0.018191	2243.191	-22556.74374	38494	4.229864	11.360736
vertices	ThickAvg	gpheno	12.4454	0.005328	5465.607	-25779.15974	38494	11.401112	13.489688

10000randomROIs	ThickAvg	gpheno	10.1429	0.004636	5352.953	-25666.50574	38494	9.234244	11.051556
1000randomROIs	ThickAvg	gpheno	7.6923	0.004668	4760.328	-25073.88174	38494	6.777372	8.607228
50000randomROIs	ThickAvg	gpheno	11.5186	0.005005	5440.074	-25753.62774	38494	10.53762	12.49958
5000randomROIs	ThickAvg	gpheno	9.4038	0.004461	5266.532	-25580.08574	38494	8.529444	10.278156
Des	GrayVol	sex	32.4564	0.025979	24355.883	-39843.94751	38503	27.364516	37.548284
DK	GrayVol	sex	30.7012	0.036965	20753.089	-36241.15351	38503	23.45606	37.94634
Glasser	GrayVol	sex	39.8419	0.018609	29214.95	-44703.01551	38503	36.194536	43.489264
Gordon	GrayVol	sex	39.2417	0.01916	29284.021	-44772.08551	38503	35.48634	42.99706
JulichBrain	GrayVol	sex	41.2438	0.02165	28628.744	-44116.80951	38503	37.0004	45.4872
Schaefer	GrayVol	sex	53.2406	0.016855	30183.676	-45671.74151	38503	49.93702	56.54418
Yeo	GrayVol	sex	28.8669	0.05027	17498.987	-32987.05151	38503	19.01398	38.71982
vertices	GrayVol	sex	80.3006	0.00916	41952.281	-57440.34551	38503	78.50524	82.09596
10000randomROIs	GrayVol	sex	45.3718	0.006831	43080.15	-58568.21551	38503	44.032924	46.710676
1000randomROIs	GrayVol	sex	42.6399	0.012202	36919.45	-52407.51551	38503	40.248308	45.031492
50000randomROIs	GrayVol	sex	50.9883	0.006716	43061.294	-58549.35951	38503	49.671964	52.304636
5000randomROIs	GrayVol	sex	44.0609	0.007447	42063.716	-57551.78151	38503	42.601288	45.520512
vertices	SurfArea	sex	87.9437	0.010031	36376.903	-51864.96751	38503	85.977624	89.909776
Des	SurfArea	sex	28.9921	0.024522	23833.076	-39321.13951	38503	24.185788	33.798412
DK	SurfArea	sex	28.0655	0.035182	21431.093	-36919.15751	38503	21.169828	34.961172
Glasser	SurfArea	sex	43.9439	0.019438	30068.518	-45556.58151	38503	40.134052	47.753748
Gordon	SurfArea	sex	43.0843	0.019975	30175.589	-45663.65351	38503	39.1692	46.9994
JulichBrain	SurfArea	sex	41.4479	0.021787	29924.808	-45412.87151	38503	37.177648	45.718152
Schaefer	SurfArea	sex	72.2982	0.014179	29523.331	-45011.39551	38503	69.519116	75.077284
Yeo	SurfArea	sex	25.7536	0.047016	19560.624	-35048.68951	38503	16.538464	34.968736
10000randomROIs	SurfArea	sex	46.6497	0.007001	39145.185	-54633.24951	38503	45.277504	48.021896
1000randomROIs	SurfArea	sex	52.2062	0.012916	34160.029	-49648.09351	38503	49.674664	54.737736
50000randomROIs	SurfArea	sex	51.0285	0.006817	38582.746	-54070.81151	38503	49.692368	52.364632

5000randomROIs	SurfArea	sex	47.5001	0.007872	38353.915	-53841.97951	38503	45.957188	49.043012
Des	ThickAvg	sex	33.2084	0.026242	16923.183	-32411.24751	38503	28.064968	38.351832
DK	ThickAvg	sex	31.4338	0.037399	10209.324	-25697.38951	38503	24.103596	38.764004
Glasser	ThickAvg	sex	37.7057	0.018061	22503.219	-37991.28351	38503	34.165744	41.245656
Gordon	ThickAvg	sex	35.7916	0.018329	22271.648	-37759.71151	38503	32.199116	39.384084
JulichBrain	ThickAvg	sex	34.647	0.019982	19652.852	-35140.91751	38503	30.730528	38.563472
Schaefer	ThickAvg	sex	37.4534	0.015681	24608.916	-40096.98151	38503	34.379924	40.526876
Yeo	ThickAvg	sex	31.5469	0.052826	6509.414	-21997.47751	38503	21.193004	41.900796
vertices	ThickAvg	sex	48.9738	0.006766	39244.686	-54732.75151	38503	47.647664	50.299936
10000randomROIs	ThickAvg	sex	44.3096	0.006851	38165.898	-53653.96351	38503	42.966804	45.652396
1000randomROIs	ThickAvg	sex	39.0063	0.011697	31094.832	-46582.89751	38503	36.713688	41.298912
50000randomROIs	ThickAvg	sex	46.8694	0.006648	38951.451	-54439.51551	38503	45.566392	48.172408
5000randomROIs	ThickAvg	sex	42.1844	0.007316	37012.651	-52500.71551	38503	40.750464	43.618336

Here we describe the relative improvement in morphometricty between Yeo and vertex-wise measures. The last column 'improvement' indicates the vertex-wise morphometricity (300.000 ROIs) divided by the morphometricity yielded by the Yeo atlas (34 ROIs)

Measurement type	Phenotype	Yeo atlas estimate	Vertices estimates	improvement
GrayVol	bmi	2.4865	37.7222	15.17
GrayVol	cig	1.0802	12.258	11.35
SurfArea	cig	1.9126	20.7596	10.85
GrayVol	edu	1.2342	12.3335	9.99
SurfArea	edu	1.9382	18.8645	9.73
SurfArea	bmi	5.1223	40.6913	7.94
SurfArea	alc	1.4788	11.1893	7.57
GrayVol	alc	1.3492	10.0093	7.42
SurfArea	gpheno	7.4033	38.0167	5.14
GrayVol	gpheno	4.8282	23.8461	4.94
SurfArea	sex	25.7536	87.9437	3.41
SurfArea	age	24.5316	81.9581	3.34
ThickAvg	cig	1.5256	4.8146	3.16
GrayVol	age	22.455	68.3814	3.05
ThickAvg	edu	2.8678	8.0762	2.82
GrayVol	sex	28.8669	80.3006	2.78
ThickAvg	bmi	9.2688	20.8368	2.25
ThickAvg	alc	3.2113	7.1666	2.23
ThickAvg	gpheno	7.7953	12.4454	1.6
ThickAvg	sex	31.5469	48.9738	1.55
ThickAvg	age	25.0744	37.396	1.49

We noticed that the raw Schaefer atlas data as given by FreeSurfer (orig) yielded unusually large morphometricity estimates with surface area and volume measures. We suspected that Schaefer may violate normality assumptions, transformed Schaefer data atlas with rank-based inverse normal transformation (rint). Estimates displayed here are from both the raw data (orig) and the normalised data (rint)

Atlas	Measure-ment	Pheno-type	Original morpho-metricity	Original BIC	Original N	Original 95% CI (lower bound)	Original 95% CI (upper bound)	Rint morpho-metricity	Rint BIC	Rint 95% CI (lower bound)	Rint 95% CI (upper bound)
Schaefer	GrayVol	age	46.9021	367412.0205	38503	43.542856	50.261344	36.4559	366626.1804	33.379288	39.532512
Schaefer	GrayVol	alc	2.8286	35590.82639	26014	1.970316	3.686884	2.8322	35609.79835	2.033304	3.631096
Schaefer	GrayVol	bmi	15.1659	146555.7637	38394	13.024992	17.306808	13.0069	146491.3876	11.220556	14.793244
Schaefer	GrayVol	cig	3.118	64805.09763	37333	2.301856	3.934144	2.8475	64866.18151	2.146996	3.548004
Schaefer	GrayVol	edu	4.763	85802.072	38484	3.746936	5.779064	4.4964	85857.75192	3.599896	5.392904
Schaefer	GrayVol	gpheno	11.5861	-24563.35574	38494	9.808968	13.363232	9.4439	-24922.55782	8.028976	10.858824
Schaefer	GrayVol	sex	53.2406	-45671.74151	38503	49.93702	56.54418	44.4569	-47498.85759	41.232112	47.681688
Schaefer	SurfArea	age	55.005	373864.5645	38503	51.55344	58.45656	38.9862	373481.2704	35.763764	42.208636
Schaefer	SurfArea	alc	2.3422	35539.06239	26014	1.449812	3.234588	2.3184	35569.34635	1.510684	3.126116
Schaefer	SurfArea	bmi	12.2649	146879.4877	38394	10.219836	14.309964	10.841	146947.2296	9.131684	12.550316
Schaefer	SurfArea	cig	3.2597	64864.63363	37333	2.325956	4.193444	2.8292	64944.14351	2.061664	3.596736
Schaefer	SurfArea	edu	6.4127	85969.52	38484	5.039916	7.785484	5.5305	86039.16192	4.396248	6.664752
Schaefer	SurfArea	gpheno	15.6662	-23656.78174	38494	13.358104	17.974296	12.7207	-23956.57982	10.86948	14.57192
Schaefer	SurfArea	sex	72.2982	-45011.39551	38503	69.519116	75.077284	48.3203	-46515.03759	45.019856	51.620744
Schaefer	ThickAvg	age	29.5182	364531.3925	38503	26.73892	32.29748	27.0261	365130.4724	24.407932	29.644268
Schaefer	ThickAvg	alc	3.5868	35425.16439	26014	2.77242	4.40118	3.4511	35495.39235	2.673372	4.228828
Schaefer	ThickAvg	bmi	12.2845	144082.3697	38394	10.673968	13.895032	11.4105	144228.7316	9.919724	12.901276
Schaefer	ThickAvg	cig	2.3803	64448.04763	37333	1.820132	2.940468	2.2148	64571.38351	1.689716	2.739884
Schaefer	ThickAvg	edu	3.6159	85756.68	38484	2.926372	4.305428	3.3981	85930.05192	2.7513	4.0449
Schaefer	ThickAvg	gpheno	7.9811	-24174.07574	38494	6.809608	9.152592	7.5137	-24234.96982	6.4161	8.6113
Schaefer	ThickAvg	sex	37.4534	-40096.98151	38503	34.379924	40.526876	35.2774	-40766.64959	32.304864	38.249936

We created parcellations with random regions boundaries (100 for each level of dimensionality), mapped these random parcellations onto our participant data, and used this to estimate morphometricity. The results displayed here are summary statistics of 100 estimates at each atlas dimensionality, non-brain trait, and measurement type.

Number of ROIs	Measurement	Pheno-type	Mean	Standard deviation	Mini-mum	Maxi-mum	Median	Range
34	SurfArea	age	20.24	3.32	13.0591	29.4638	20.3538	16.4047
34	SurfArea	alc	0.8	0.15	0.5278	1.3238	0.77425	0.796
34	SurfArea	bmi	6.51	0.82	4.6024	9.0372	6.5747	4.4348
34	SurfArea	cig	0.89	0.27	0.4093	1.4829	0.8749	1.0736
34	SurfArea	edu	1.85	0.41	1.0747	3.0017	1.79515	1.927
34	SurfArea	gpheno	5.48	1.15	3.2464	8.3907	5.3372	5.1443
34	SurfArea	sex	24.31	3.26	17.6613	33.6637	24.22455	16.0024
34	GrayVol	age	22.96	4.32	13.7503	35.4775	23.1704	21.7272
34	GrayVol	alc	1.14	0.2	0.7243	1.8084	1.1164	1.0841
34	GrayVol	bmi	3.36	0.79	1.5518	5.3022	3.2771	3.7504
34	GrayVol	cig	0.8	0.23	0.3293	1.7284	0.7727	1.3991
34	GrayVol	edu	1.52	0.35	0.9413	2.5215	1.4559	1.5802
34	GrayVol	gpheno	4.49	0.86	2.1894	6.4619	4.5318	4.2725
34	GrayVol	sex	26.18	3.45	19.322	35.757	26.28915	16.435
34	ThickAvg	age	27.93	2.6	21.1393	34.686	27.93805	13.5467
34	ThickAvg	alc	2.52	0.38	1.4985	3.5505	2.51575	2.052
34	ThickAvg	bmi	11.96	1.74	7.1794	16.3333	11.84725	9.1539
34	ThickAvg	cig	2.11	0.4	1.1461	2.9955	2.10325	1.8494
34	ThickAvg	edu	3.38	0.39	2.4513	4.2118	3.34105	1.7605
34	ThickAvg	gpheno	8.5	1.07	5.6611	11.1978	8.3959	5.5367
34	ThickAvg	sex	31.88	3.32	21.3362	38.1168	31.73435	16.7806
68	SurfArea	age	27.09	2.98	18.459	35.0895	27.2711	16.6305
68	SurfArea	alc	1.01	0.17	0.6523	1.4997	0.99465	0.8474
68	SurfArea	bmi	6.66	0.68	5.0985	8.5899	6.69625	3.4914
68	SurfArea	cig	1.49	0.32	0.707	2.2208	1.5129	1.5138
68	SurfArea	edu	2.51	0.43	1.52	3.5308	2.5397	2.0108
68	SurfArea	gpheno	7.73	1.16	4.7301	10.6638	7.7398	5.9337
68	SurfArea	sex	31.91	2.89	24.4323	38.9053	31.86715	14.473
68	GrayVol	age	29.83	2.85	23.4911	38.1919	29.75905	14.7008
68	GrayVol	alc	1.37	0.19	0.8896	1.7335	1.3723	0.8439
68	GrayVol	bmi	4.92	0.93	2.9377	7.5594	4.8073	4.6217
68	GrayVol	cig	1.35	0.25	0.6868	1.985	1.3357	1.2982
68	GrayVol	edu	2.29	0.39	1.4213	3.3213	2.2551	1.9
68	GrayVol	gpheno	6.37	0.83	4.3627	9.7028	6.26535	5.3401
68	GrayVol	sex	32.46	2.52	26.5826	38.3035	32.6795	11.7209
68	ThickAvg	age	27.77	1.91	23.1032	32.7326	27.698	9.6294
68	ThickAvg	alc	2.68	0.29	2.1812	3.4582	2.6382	1.277
68	ThickAvg	bmi	11.21	0.98	8.7599	13.7748	11.05635	5.0149
68	ThickAvg	cig	2.05	0.25	1.3074	2.579	2.0542	1.2716
68	ThickAvg	edu	3.27	0.3	2.5942	4.2038	3.2612	1.6096

68	ThickAvg	gpheno	8.26	0.71	6.5893	9.9867	8.3652	3.3974
68	ThickAvg	sex	33.74	2.54	26.7485	39.8847	33.6942	13.1362
148	SurfArea	age	36.59	2.4	30.5963	42.7875	36.451	12.1912
148	SurfArea	alc	1.47	0.2	1.0475	2.0494	1.4446	1.0019
148	SurfArea	bmi	7.91	0.74	6.241	10.5281	7.9305	4.2871
148	SurfArea	cig	2.34	0.34	1.7005	2.9838	2.32935	1.2833
148	SurfArea	edu	3.67	0.4	2.6512	4.8664	3.6518	2.2152
148	SurfArea	gpheno	11.39	1	8.6436	13.7535	11.5283	5.1099
148	SurfArea	sex	40.12	2.3	35.1418	45.2621	39.8313	10.1203
148	GrayVol	age	35.64	2.36	30.33	45.3096	35.80185	14.9796
148	GrayVol	alc	1.84	0.2	1.3805	2.2809	1.862	0.9004
148	GrayVol	bmi	8.21	1.05	5.456	11.7022	8.0495	6.2462
148	GrayVol	cig	2.1	0.29	1.3702	2.7795	2.1106	1.4093
148	GrayVol	edu	3.21	0.37	2.3665	4.2028	3.238	1.8363
148	GrayVol	gpheno	8.14	0.7	6.5458	10.6105	8.02265	4.0647
148	GrayVol	sex	38.31	1.92	33.1105	42.2053	38.58225	9.0948
148	ThickAvg	age	28.14	1.49	24.4292	31.6093	28.142	7.1801
148	ThickAvg	alc	2.89	0.22	2.4741	3.5173	2.8626	1.0432
148	ThickAvg	bmi	10.53	0.73	8.9862	12.9375	10.5404	3.9513
148	ThickAvg	cig	1.99	0.14	1.6132	2.3242	1.97635	0.711
148	ThickAvg	edu	3.32	0.22	2.8864	4.0668	3.3151	1.1804
148	ThickAvg	gpheno	7.98	0.44	7.0253	9.2351	7.941	2.2098
148	ThickAvg	sex	35.88	1.49	32.8439	39.102	35.9336	6.2581
274	SurfArea	age	44.59	1.91	39.373	49.0971	44.62265	9.7241
274	SurfArea	alc	2.01	0.24	1.5652	2.6508	1.9925	1.0856
274	SurfArea	bmi	10.39	0.89	8.7966	13.9481	10.2488	5.1515
274	SurfArea	cig	2.95	0.26	2.344	3.5561	2.949	1.2121
274	SurfArea	edu	4.71	0.38	3.783	5.6136	4.72645	1.8306
274	SurfArea	gpheno	14.15	0.96	12.1801	16.565	14.1856	4.3849
274	SurfArea	sex	45.8	1.82	40.4245	50.0019	45.7619	9.5774
274	GrayVol	age	37.4	1.77	33.6607	43.26	37.5457	9.5993
274	GrayVol	alc	2.36	0.21	1.8426	2.9693	2.34745	1.1267
274	GrayVol	bmi	11.04	0.86	8.7425	13.4061	11.01885	4.6636
274	GrayVol	cig	2.66	0.2	2.1251	3.2927	2.66315	1.1676
274	GrayVol	edu	3.77	0.29	3.1109	4.4677	3.75945	1.3568
274	GrayVol	gpheno	8.88	0.55	7.6751	10.3499	8.8518	2.6748
274	GrayVol	sex	41.3	1.38	38.0404	46.1708	41.2116	8.1304
274	ThickAvg	age	27.94	0.94	25.2968	30.4152	27.88385	5.1184
274	ThickAvg	alc	3.12	0.18	2.6958	3.5709	3.10695	0.8751
274	ThickAvg	bmi	10.79	0.52	9.3664	12.0759	10.7727	2.7095
274	ThickAvg	cig	2.05	0.12	1.7809	2.3151	2.0591	0.5342
274	ThickAvg	edu	3.44	0.18	3.0109	3.9594	3.4305	0.9485
274	ThickAvg	gpheno	7.53	0.33	6.6215	8.4037	7.51035	1.7822
274	ThickAvg	sex	37.34	1.18	33.7488	41.4591	37.2547	7.7103
334	SurfArea	age	46.28	1.75	42.0046	50.5587	46.2346	8.5541
334	SurfArea	alc	2.17	0.22	1.6259	2.6435	2.1809	1.0176
334	SurfArea	bmi	11.47	0.97	9.4608	14.0247	11.4246	4.5639

334	SurfArea	cig	3.14	0.23	2.6972	3.7533	3.1714	1.0561
334	SurfArea	edu	5.01	0.37	4.2266	6.3526	5.0468	2.126
334	SurfArea	gpheno	14.67	0.82	12.7568	16.4227	14.6914	3.6659
334	SurfArea	sex	47.33	1.63	43.4617	51.8989	47.3039	8.4372
334	GrayVol	age	37.56	1.69	33.7734	41.5952	37.65785	7.8218
334	GrayVol	alc	2.5	0.2	1.8835	3.0648	2.49805	1.1813
334	GrayVol	bmi	11.86	0.82	9.8429	13.5468	11.85955	3.7039
334	GrayVol	cig	2.83	0.21	2.433	3.7539	2.8002	1.3209
334	GrayVol	edu	3.92	0.24	3.3247	4.659	3.9275	1.3343
334	GrayVol	gpheno	9.07	0.49	7.8203	10.2586	9.03875	2.4383
334	GrayVol	sex	41.73	1.18	38.4316	45.0578	41.64365	6.6262
334	ThickAvg	age	27.76	0.89	25.9754	30.0638	27.7598	4.0884
334	ThickAvg	alc	3.19	0.16	2.7553	3.7434	3.184	0.9881
334	ThickAvg	bmi	10.92	0.45	9.7099	12.0409	10.9482	2.331
334	ThickAvg	cig	2.09	0.11	1.8065	2.3439	2.0951	0.5374
334	ThickAvg	edu	3.49	0.17	3.156	3.9765	3.4805	0.8205
334	ThickAvg	gpheno	7.42	0.29	6.7231	8.0549	7.4352	1.3318
334	ThickAvg	sex	37.68	0.98	33.7894	40.4839	37.6542	6.6945
360	SurfArea	age	46.94	1.58	43.4063	51.0843	46.8599	7.678
360	SurfArea	alc	2.24	0.21	1.7594	2.77	2.23505	1.0106
360	SurfArea	bmi	11.88	0.94	9.6643	14.1492	11.97065	4.4849
360	SurfArea	cig	3.23	0.23	2.7003	3.8642	3.2564	1.1639
360	SurfArea	edu	5.13	0.35	4.3431	6.1686	5.13585	1.8255
360	SurfArea	gpheno	14.83	0.76	13.1391	16.6705	14.80045	3.5314
360	SurfArea	sex	47.73	1.5	43.3953	51.4321	47.7894	8.0368
360	GrayVol	age	37.53	1.65	33.9354	41.8426	37.4878	7.9072
360	GrayVol	alc	2.57	0.22	2.0792	3.1563	2.5438	1.0771
360	GrayVol	bmi	12.14	0.8	9.7835	13.9045	12.14485	4.121
360	GrayVol	cig	2.87	0.19	2.4727	3.533	2.8704	1.0603
360	GrayVol	edu	3.96	0.24	3.4241	4.6501	3.96045	1.226
360	GrayVol	gpheno	9.07	0.5	7.7737	10.7572	9.02585	2.9835
360	GrayVol	sex	41.77	1.09	38.8508	44.0326	41.6848	5.1818
360	ThickAvg	age	27.83	0.86	25.9245	30.3051	27.7722	4.3806
360	ThickAvg	alc	3.24	0.17	2.8945	3.8969	3.2353	1.0024
360	ThickAvg	bmi	10.94	0.42	9.8153	12.0364	10.9302	2.2211
360	ThickAvg	cig	2.1	0.1	1.8776	2.3422	2.099	0.4646
360	ThickAvg	edu	3.52	0.15	3.2011	3.9627	3.517	0.7616
360	ThickAvg	gpheno	7.39	0.27	6.7105	7.9647	7.43915	1.2542
360	ThickAvg	sex	37.77	1	34.1975	40.2861	37.74855	6.0886
500	SurfArea	age	48.47	1.35	45.0789	51.613	48.39545	6.5341
500	SurfArea	alc	2.47	0.21	2.0326	3.0099	2.44535	0.9773
500	SurfArea	bmi	13.82	0.98	11.5838	16.6298	13.71425	5.046
500	SurfArea	cig	3.49	0.21	3.0024	4.13	3.4545	1.1276
500	SurfArea	edu	5.6	0.36	4.6927	6.4495	5.57115	1.7568
500	SurfArea	gpheno	15.14	0.67	13.7105	16.6643	15.18275	2.9538
500	SurfArea	sex	49.55	1.22	46.0471	52.2456	49.5363	6.1985
500	GrayVol	age	37.5	1.28	34.6478	40.7905	37.50145	6.1427

500	GrayVol	alc	2.78	0.21	2.3561	3.2321	2.7417	0.876
500	GrayVol	bmi	13.32	0.66	11.8516	14.97	13.30075	3.1184
500	GrayVol	cig	3.09	0.18	2.6124	3.5693	3.0813	0.9569
500	GrayVol	edu	4.13	0.24	3.6202	4.6688	4.1155	1.0486
500	GrayVol	gpheno	9.03	0.37	8.2046	9.8486	8.9967	1.644
500	GrayVol	sex	42.2	0.97	40.0876	44.5617	42.29315	4.4741
500	ThickAvg	age	27.78	0.72	26.1928	29.6335	27.7175	3.4407
500	ThickAvg	alc	3.41	0.15	3.0676	3.8165	3.4174	0.7489
500	ThickAvg	bmi	11.08	0.42	9.9145	11.9686	11.06265	2.0541
500	ThickAvg	cig	2.17	0.1	1.9453	2.5227	2.1728	0.5774
500	ThickAvg	edu	3.64	0.14	3.2957	3.9451	3.6374	0.6494
500	ThickAvg	gpheno	7.36	0.25	6.817	8.3213	7.35155	1.5043
500	ThickAvg	sex	38.09	0.84	35.5351	40.6454	38.0153	5.1103

Here we aim to compare how well empirical estimates did compared with estimates obtained from random parcellations. We calculate the percentage of random estimates that were smaller than the empirical estimate. A percentage of 99%, for example, means that the empirical estimate was larger than 99% of the random estimates. Keep in mind this is comparing point estimates only, and Fig.5 displays that the confidence intervals from the empirical estimates map well onto the spread in estimates from the random parcellations.

Atlas	Pheno-type	Measurement	Empirical morphometricity estimate	Percentage larger
Yeo	alc	SurfArea	1.48	1
Yeo	cig	SurfArea	1.91	1
Glasser	alc	SurfArea	2.8	1
Schaefer	bmi	ThickAvg	12.28	1
Des	bmi	ThickAvg	12.7	0.99
Schaefer	age	ThickAvg	29.52	0.99
Schaefer	cig	ThickAvg	2.38	0.99
Schaefer	gpheno	ThickAvg	7.98	0.99
Yeo	alc	ThickAvg	3.21	0.98
JulichBrain	alc	SurfArea	2.59	0.98
Schaefer	sex	GrayVol	44.46	0.98
Yeo	gpheno	SurfArea	7.4	0.94
Des	age	ThickAvg	30.43	0.93
Des	cig	ThickAvg	2.21	0.93
Yeo	age	SurfArea	24.53	0.92
Yeo	cig	GrayVol	1.08	0.92
Schaefer	edu	GrayVol	4.5	0.91
Des	gpheno	GrayVol	9.02	0.9
Des	edu	ThickAvg	3.58	0.89
Schaefer	alc	ThickAvg	3.59	0.87
Yeo	alc	GrayVol	1.35	0.86
Gordon	cig	GrayVol	3.06	0.86
Schaefer	gpheno	GrayVol	9.44	0.85
Gordon	cig	ThickAvg	2.19	0.8
Gordon	bmi	GrayVol	12.6	0.79
Glasser	bmi	ThickAvg	11.27	0.79
Yeo	sex	GrayVol	28.87	0.76
Yeo	sex	SurfArea	25.75	0.74
JulichBrain	age	ThickAvg	28.64	0.74
Gordon	alc	SurfArea	2.31	0.74
DK	age	ThickAvg	28.83	0.68
Des	cig	GrayVol	2.26	0.68
Gordon	alc	GrayVol	2.58	0.68
Yeo	gpheno	GrayVol	4.83	0.67
Des	age	GrayVol	36.58	0.67
Glasser	bmi	GrayVol	12.48	0.67
Yeo	edu	SurfArea	1.94	0.61
Schaefer	alc	GrayVol	2.83	0.61

Des	bmi	GrayVol	8.3	0.59
Gordon	bmi	ThickAvg	11	0.59
DK	alc	ThickAvg	2.71	0.56
JulichBrain	alc	GrayVol	2.37	0.54
Des	alc	SurfArea	1.46	0.53
Glasser	alc	GrayVol	2.57	0.53
JulichBrain	sex	GrayVol	41.24	0.51
DK	alc	GrayVol	1.37	0.5
Yeo	sex	ThickAvg	31.55	0.48
Gordon	gpheno	ThickAvg	7.42	0.48
JulichBrain	gpheno	ThickAvg	7.45	0.46
Gordon	edu	ThickAvg	3.47	0.45
Glasser	sex	ThickAvg	37.71	0.45
Glasser	bmi	SurfArea	11.8	0.43
Glasser	cig	SurfArea	3.21	0.43
Yeo	age	GrayVol	22.45	0.42
DK	cig	SurfArea	1.41	0.42
Schaefer	edu	SurfArea	5.53	0.42
JulichBrain	edu	SurfArea	4.53	0.38
Schaefer	edu	ThickAvg	3.62	0.38
JulichBrain	bmi	SurfArea	10.06	0.37
Gordon	age	GrayVol	36.94	0.36
DK	edu	ThickAvg	3.15	0.35
DK	bmi	ThickAvg	10.73	0.34
DK	gpheno	ThickAvg	8.04	0.34
Glasser	alc	ThickAvg	3.16	0.34
JulichBrain	bmi	ThickAvg	10.51	0.32
Glasser	edu	GrayVol	3.84	0.32
JulichBrain	cig	ThickAvg	1.99	0.31
Schaefer	bmi	GrayVol	13.01	0.3
Yeo	gpheno	ThickAvg	7.8	0.26
Schaefer	alc	SurfArea	2.32	0.26
Yeo	edu	GrayVol	1.23	0.25
DK	alc	SurfArea	0.91	0.25
Glasser	age	ThickAvg	27.21	0.25
DK	sex	GrayVol	30.7	0.23
Des	cig	SurfArea	2.03	0.23
Des	gpheno	ThickAvg	7.64	0.23
Schaefer	sex	ThickAvg	37.45	0.22
DK	age	SurfArea	24.69	0.21
Schaefer	age	GrayVol	36.46	0.2
Des	bmi	SurfArea	7.33	0.19
JulichBrain	bmi	GrayVol	10.3	0.18
Glasser	cig	GrayVol	2.69	0.17
Yeo	bmi	GrayVol	2.49	0.16
Glasser	cig	ThickAvg	1.99	0.16
DK	sex	ThickAvg	31.43	0.15

DK	gpheno	SurfArea	6.34	0.14
DK	bmi	GrayVol	3.9	0.14
Des	alc	GrayVol	1.6	0.13
Glasser	edu	SurfArea	4.67	0.13
Yeo	age	ThickAvg	25.07	0.12
Schaefer	sex	SurfArea	48.32	0.12
DK	bmi	SurfArea	5.82	0.11
Des	edu	GrayVol	2.76	0.11
JulichBrain	edu	ThickAvg	3.21	0.11
Gordon	alc	ThickAvg	2.99	0.1
DK	sex	SurfArea	28.07	0.09
DK	edu	GrayVol	1.84	0.09
JulichBrain	age	SurfArea	41.95	0.09
Glasser	age	GrayVol	35.35	0.09
Glasser	gpheno	GrayVol	8.48	0.09
Yeo	cig	ThickAvg	1.53	0.08
Yeo	edu	ThickAvg	2.87	0.08
Gordon	bmi	SurfArea	10.1	0.08
Gordon	age	ThickAvg	26.62	0.08
Gordon	gpheno	GrayVol	8.37	0.07
Yeo	bmi	SurfArea	5.12	0.06
DK	edu	SurfArea	1.83	0.05
DK	cig	ThickAvg	1.63	0.05
JulichBrain	cig	GrayVol	2.32	0.05
Schaefer	cig	GrayVol	2.85	0.05
Yeo	bmi	ThickAvg	9.27	0.04
Gordon	edu	GrayVol	3.5	0.04
DK	cig	GrayVol	0.9	0.03
Des	sex	ThickAvg	33.21	0.03
JulichBrain	cig	SurfArea	2.46	0.02
JulichBrain	sex	SurfArea	41.45	0.02
JulichBrain	sex	ThickAvg	34.65	0.02
JulichBrain	age	GrayVol	34.18	0.02
JulichBrain	gpheno	GrayVol	7.85	0.02
Gordon	sex	ThickAvg	35.79	0.02
Glasser	sex	GrayVol	39.84	0.02
DK	gpheno	GrayVol	4.77	0.01
JulichBrain	gpheno	SurfArea	12.19	0.01
JulichBrain	alc	ThickAvg	2.72	0.01
JulichBrain	edu	GrayVol	3.13	0.01
Gordon	sex	GrayVol	39.24	0.01
Glasser	sex	SurfArea	43.94	0.01
Glasser	gpheno	ThickAvg	6.79	0.01
DK	age	GrayVol	22.63	0
Des	age	SurfArea	30.49	0
Des	edu	SurfArea	2.35	0
Des	gpheno	SurfArea	7.98	0

Des	sex	SurfArea	28.99	0
Des	alc	ThickAvg	2.39	0
Des	sex	GrayVol	32.46	0
Gordon	age	SurfArea	40.69	0
Gordon	cig	SurfArea	2.65	0
Gordon	edu	SurfArea	3.81	0
Gordon	gpheno	SurfArea	12.35	0
Gordon	sex	SurfArea	43.08	0
Glasser	age	SurfArea	41.57	0
Glasser	gpheno	SurfArea	13.12	0
Glasser	edu	ThickAvg	3.14	0
Schaefer	age	SurfArea	38.99	0
Schaefer	bmi	SurfArea	10.84	0
Schaefer	cig	SurfArea	2.83	0
Schaefer	gpheno	SurfArea	12.72	0

LRTs were calculated for each atlas pair to quantify whether adding a second atlas to the equation would significantly improve morphometricity. Here we display summary statistics for the ratio of and difference between sum_G (joint variance explained by two atlases) and individual morphometricity (explained by the lower dimensional only). We then report summary statistics of this ratio and difference per measurement type and non-brain trait.

Measure-ment	Pheno-type	Ratio minimum	Ratio maximum	Ratio median	Ratio mean	Ratio SD	Difference minimum	Difference maximum	Difference median	Difference mean	Difference SD
SurfArea	age	1.27	3.36	1.94	2.04	0.57	11.41	57.8	31.425	30.01	12.35
GrayVol	age	1.21	3.18	1.94	1.95	0.55	7.5	48.86	27.4	25.33	10.7
ThickAvg	age	1.31	2.2	1.58	1.62	0.19	9.14	34.49	16.205	16.92	4.91
SurfArea	alc	0.93	11.73	2.02	2.75	2.3	-0.16	9.75	1.45	2.55	2.93
GrayVol	alc	1.19	7.29	1.99	2.5	1.67	0.5	8.48	1.36	2.6	2.74
ThickAvg	alc	1.22	3	1.46	1.7	0.5	0.6	4.77	1.215	1.99	1.42
SurfArea	bmi	1.34	8.14	2.61	2.96	1.56	3.85	36.57	10.025	13.93	10.48
GrayVol	bmi	1.34	15.63	3.42	4.14	3.07	4.27	36.37	11.22	14.57	9.05
ThickAvg	bmi	1.3	2.52	1.54	1.7	0.35	3.39	14.08	6.05	7.46	3.44
SurfArea	cig	1.15	14.56	1.87	3.57	3.51	0.48	19.18	1.655	5.46	7.22
GrayVol	cig	1.11	13.6	2.87	3.34	2.85	0.33	11.34	1.8	3.53	3.66
ThickAvg	cig	1.15	3.24	1.45	1.66	0.55	0.34	3.42	0.765	1.2	0.98
SurfArea	edu	1.21	10.24	2.72	3.22	2.33	0.96	17.4	3.345	5.65	5.28
GrayVol	edu	1.22	10.05	2.6	2.83	1.82	0.78	11.17	2.32	3.76	2.97
ThickAvg	edu	1.18	3.04	1.43	1.71	0.53	0.62	5.85	1.39	2.25	1.68
SurfArea	gpheno	1.25	6.03	2.02	2.38	1.17	3.01	31.92	7.79	11.63	9.05
GrayVol	gpheno	1.21	4.99	2.35	2.29	0.98	1.87	19.19	6.87	7.71	5.11
ThickAvg	gpheno	1.25	2.03	1.46	1.54	0.23	1.99	8.06	3.575	4.15	1.78
SurfArea	sex	1.22	3.49	2.02	2.18	0.62	15.64	64.08	37.485	35.62	14.4
GrayVol	sex	1.37	2.8	1.79	1.82	0.36	14.64	52.09	24.65	27.28	10.55
ThickAvg	sex	1.3	2.28	1.79	1.78	0.28	11.06	40.2	26.05	25.65	8.29

Note. The ratio described here is an index of the relative improvement made to the model by adding the higher dimensional atlas, which is calculated as the morphometricity estimate of the lower dimensional atlas alone, divided by the sum of the variance explained by both atlases together. A ratio of 1 indicates no increase in morphometricity when adding the higher dimensional atlas to the equation, and a ratio larger than 1 indicates the proportional increase in

morphometricity grained by adding the higher dimensional atlas. The summary statistics here describe the ratios we obtained within measurement types and non-brain traits.

STable 6: LASSO prediction accuracy

LASSO prediction accuracy calculated as the correlation (R2) between predicted and observed trait values.

Atlas	Measurement type	Pheno-type	R2	95% CI (lower bound)	95% CI (upper bound)	p-value
Des	GrayVol	age	0.340379488	0.323183559	0.357561648	0
Des	SurfArea	age	0.236488228	0.220016432	0.253181535	0
Des	ThickAvg	age	0.389839799	0.372771206	0.406805307	0
DK	GrayVol	age	0.209874899	0.19385628	0.22618565	0
DK	SurfArea	age	0.133909923	0.120016154	0.148320796	1.1065E-242
DK	ThickAvg	age	0.310441724	0.293306276	0.327624574	0
Glasser	GrayVol	age	0.380387639	0.363275455	0.397412687	0
Glasser	SurfArea	age	0.327605264	0.310422447	0.344799758	0
Glasser	ThickAvg	age	0.441228946	0.424559808	0.457717729	0
Gordon	GrayVol	age	0.378305303	0.361184523	0.395342529	0
Gordon	SurfArea	age	0.31468071	0.297530143	0.331869691	0
Gordon	ThickAvg	age	0.442902594	0.426250641	0.459371941	0
JulichBrain	GrayVol	age	0.376515106	0.359387301	0.393562456	0
JulichBrain	SurfArea	age	0.343399487	0.326204647	0.360574687	0
JulichBrain	ThickAvg	age	0.441952166	0.425290423	0.45843258	0
Schaefer	GrayVol	age	0.369694014	0.352542569	0.386776994	0
Schaefer	SurfArea	age	0.25310955	0.236418179	0.269980427	0
Schaefer	ThickAvg	age	0.404045045	0.38706191	0.420902164	0
vertices	SurfArea	age	0.424209385	0.407380207	0.440882053	0
vertices	ThickAvg	age	0.586488605	0.572184708	0.600479182	0
vertices	GrayVol	age	0.525945003	0.510450647	0.541166106	0
Yeo	GrayVol	age	0.155943635	0.141283668	0.1710519	8.4882E-286
Yeo	SurfArea	age	0.121327196	0.107939491	0.1352726	1.5103E-218
Yeo	ThickAvg	age	0.24411233	0.227534799	0.260891884	0
Des	GrayVol	sex	0.462980348	0.446553162	0.4791992	0
Des	SurfArea	sex	0.46324725	0.446823308	0.479462534	0
Des	ThickAvg	sex	0.349168248	0.331975162	0.366330624	0
DK	GrayVol	sex	0.409487587	0.392542108	0.426298579	0
DK	SurfArea	sex	0.420813663	0.403955811	0.43751997	0
DK	ThickAvg	sex	0.235324817	0.218871235	0.252002882	0
Glasser	GrayVol	sex	0.540434506	0.52519986	0.555384315	0
Glasser	SurfArea	sex	0.551854416	0.536835827	0.56657992	0
Glasser	ThickAvg	sex	0.447090849	0.430483019	0.463510512	0
Gordon	GrayVol	sex	0.536278931	0.520968137	0.551308082	0
Gordon	SurfArea	sex	0.550319819	0.535271621	0.565075995	0
Gordon	ThickAvg	sex	0.445576199	0.428952229	0.462014002	0
JulichBrain	GrayVol	sex	0.534410495	0.519065897	0.549474919	0
JulichBrain	SurfArea	sex	0.549770753	0.534712005	0.564537864	0
JulichBrain	ThickAvg	sex	0.402800513	0.385809179	0.419667793	0
Schaefer	GrayVol	sex	0.550943308	0.535907119	0.565687042	0
Schaefer	SurfArea	sex	0.547206748	0.532099032	0.562024641	0

Schaefer	ThickAvg	sex	0.484959277	0.468820889	0.500863992	0
vertices	SurfArea	sex	0.614744657	0.60108772	0.628077085	0
vertices	ThickAvg	sex	0.66740929	0.655103631	0.679382344	0
vertices	GrayVol	sex	0.667704347	0.655406774	0.679669317	0
Yeo	GrayVol	sex	0.358126087	0.340946061	0.375258619	0
Yeo	SurfArea	sex	0.392062198	0.375006565	0.40901109	0
Yeo	ThickAvg	sex	0.15136831	0.136855918	0.166343048	9.3996E-277
Des	GrayVol	bmi	0.054154499	0.044721221	0.064378314	3.33349E-95
Des	SurfArea	bmi	0.051261016	0.042070004	0.061253199	4.37151E-90
Des	ThickAvg	bmi	0.117184327	0.103974656	0.130965343	1.1958E-210
DK	GrayVol	bmi	0.023727614	0.017465561	0.030894401	4.19158E-42
DK	SurfArea	bmi	0.029549421	0.022542275	0.037439004	3.76526E-52
DK	ThickAvg	bmi	0.086785165	0.075109666	0.099134672	5.182E-154
Glasser	GrayVol	bmi	0.074282255	0.063375609	0.085906767	3.0331E-131
Glasser	SurfArea	bmi	0.065457387	0.055156118	0.076508041	2.3507E-115
Glasser	ThickAvg	bmi	0.145948448	0.131618436	0.160757611	4.2099E-266
Gordon	GrayVol	bmi	0.070594146	0.059933572	0.081985701	1.3813E-124
Gordon	SurfArea	bmi	0.05456434	0.045098049	0.064819564	6.08955E-96
Gordon	ThickAvg	bmi	0.1325689	0.118726529	0.146932655	4.2925E-240
JulichBrain	GrayVol	bmi	0.06920981	0.058644107	0.080511439	4.294E-122
JulichBrain	SurfArea	bmi	0.055019498	0.045516094	0.065310168	9.50036E-97
JulichBrain	ThickAvg	bmi	0.125114783	0.111569008	0.139205938	8.9194E-226
Schaefer	GrayVol	bmi	0.057795309	0.048070017	0.06829773	1.1195E-101
Schaefer	SurfArea	bmi	0.044780473	0.036167589	0.054218417	1.09448E-78
Schaefer	ThickAvg	bmi	0.122946124	0.109490229	0.136954442	1.239E-221
vertices	SurfArea	bmi	0.062248544	0.052182315	0.073075734	1.2959E-109
vertices	ThickAvg	bmi	0.185510239	0.170024791	0.201356216	0
vertices	GrayVol	bmi	0.106684469	0.093958811	0.12001626	6.8784E-191
Yeo	GrayVol	bmi	0.011277233	0.007065855	0.01644171	9.27295E-21
Yeo	SurfArea	bmi	0.018261689	0.012801172	0.024648061	1.02933E-32
Yeo	ThickAvg	bmi	0.064546824	0.05431139	0.075534921	1.0054E-113
Des	GrayVol	gpheno	0.109359955	0.096506122	0.122811051	6.9613E-196
Des	SurfArea	gpheno	0.083706324	0.072210726	0.095886733	2.3349E-148
Des	ThickAvg	gpheno	0.094009197	0.08192932	0.106738252	2.761E-167
DK	GrayVol	gpheno	0.077695754	0.066569583	0.089527724	1.9775E-137
DK	SurfArea	gpheno	0.058646203	0.048854375	0.069212062	3.4308E-103
DK	ThickAvg	gpheno	0.079093305	0.067879428	0.091008061	5.7155E-140
Glasser	GrayVol	gpheno	0.105071523	0.092424111	0.118330481	7.1888E-188
Glasser	SurfArea	gpheno	0.099617162	0.087242846	0.112621446	1.0631E-177
Glasser	ThickAvg	gpheno	0.10844289	0.095633189	0.121852839	3.4663E-194
Gordon	GrayVol	gpheno	0.110914303	0.097988658	0.124431958	7.8387E-199
Gordon	SurfArea	gpheno	0.103551645	0.090978997	0.116740953	4.9738E-185
Gordon	ThickAvg	gpheno	0.111208245	0.098268976	0.124738546	2.1921E-199
JulichBrain	GrayVol	gpheno	0.107224417	0.094472813	0.120580344	6.6926E-192
JulichBrain	SurfArea	gpheno	0.111557746	0.098602323	0.125103039	4.8155E-200
JulichBrain	ThickAvg	gpheno	0.110973481	0.09804509	0.124493685	6.0652E-199
Schaefer	GrayVol	gpheno	0.102778689	0.090244455	0.115932192	1.3778E-183

291

Schaefer	SurfArea	gpheno	0.086556076	0.074893833	0.098893122	1.3627E-153
Schaefer	ThickAvg	gpheno	0.095265109	0.083118559	0.108056452	1.252E-169
vertices	SurfArea	gpheno	0.098554263	0.086234744	0.11150736	1.003E-175
vertices	ThickAvg	gpheno	0.132542702	0.118701341	0.146905532	4.8222E-240
vertices	GrayVol	gpheno	0.130877311	0.117100612	0.145180811	7.8082E-237
Yeo	GrayVol	gpheno	0.067579529	0.057127381	0.078773441	3.6614E-119
Yeo	SurfArea	gpheno	0.053829159	0.044423262	0.064026683	1.2219E-94
Yeo	ThickAvg	gpheno	0.058956502	0.049140577	0.069545328	9.6179E-104
Des	GrayVol	cig	0.014843162	0.009885318	0.020769592	4.31208E-26
Des	SurfArea	cig	0.015282243	0.010246069	0.021285239	8.05056E-27
Des	ThickAvg	cig	0.018189408	0.012663005	0.02467097	1.19085E-31
DK	GrayVol	cig	0.005932563	0.002963267	0.009906536	2.65576E-11
DK	SurfArea	cig	0.008630317	0.004951295	0.013303038	8.71882E-16
DK	ThickAvg	cig	0.014231434	0.009384838	0.020049079	4.46585E-25
Glasser	GrayVol	cig	0.017755781	0.012299627	0.024168831	6.26318E-31
Glasser	SurfArea	cig	0.016093532	0.010915775	0.022234851	3.62006E-28
Glasser	ThickAvg	cig	0.020310442	0.014453088	0.027114481	3.51996E-35
Gordon	GrayVol	cig	0.014510816	0.009612803	0.020378883	1.54696E-25
Gordon	SurfArea	cig	0.012901312	0.00830557	0.018473594	7.23512E-23
Gordon	ThickAvg	cig	0.021415196	0.015392697	0.028380102	5.13974E-37
JulichBrain	GrayVol	cig	0.015062752	0.010065579	0.021027627	1.86291E-26
JulichBrain	SurfArea	cig	0.014749181	0.009808265	0.020659061	6.17551E-26
JulichBrain	ThickAvg	cig	0.022203857	0.016067	0.029279857	2.46335E-38
Schaefer	GrayVol	cig	0.013219495	0.008562736	0.018851381	2.1321E-23
Schaefer	SurfArea	cig	0.012499657	0.007982726	0.017994618	3.33263E-22
Schaefer	ThickAvg	cig	0.019151822	0.013472738	0.025782217	2.98577E-33
vertices	SurfArea	cig	0.010136311	0.006111399	0.015148813	2.76771E-18
vertices	ThickAvg	cig	0.024740471	0.018249489	0.032160526	1.43747E-42
vertices	GrayVol	cig	0.014550949	0.009645931	0.020425729	1.31726E-25
Yeo	GrayVol	cig	0.004421129	0.001925051	0.007928054	8.86858E-09
Yeo	SurfArea	cig	0.007822016	0.004341757	0.012299258	1.91555E-14
Yeo	ThickAvg	cig	0.014634098	0.00971399	0.020523633	9.58685E-26
Des	GrayVol	alc	0.025213706	0.017472999	0.034285125	9.97286E-31
Des	SurfArea	alc	0.0268478	0.018848856	0.036168184	1.23179E-32
Des	ThickAvg	alc	0.019042098	0.012373692	0.027076445	1.53676E-23
DK	GrayVol	alc	0.019385407	0.012652636	0.02748215	6.13025E-24
DK	SurfArea	alc	0.023279772	0.015857376	0.032043884	1.79602E-28
DK	ThickAvg	alc	0.016026822	0.009953488	0.02348343	4.89147E-20
Glasser	GrayVol	alc	0.026612281	0.018649988	0.035897349	2.32124E-32
Glasser	SurfArea	alc	0.031096138	0.022465832	0.041023891	1.31186E-37
Glasser	ThickAvg	alc	0.025144628	0.017415047	0.034205315	1.20071E-30
Gordon	GrayVol	alc	0.027136192	0.019092623	0.03649957	5.66901E-33
Gordon	SurfArea	alc	0.0301722	0.021674684	0.039972369	1.58786E-36
Gordon	ThickAvg	alc	0.023058694	0.015673627	0.031786735	3.25053E-28
JulichBrain	GrayVol	alc	0.028824606	0.020525128	0.038434342	6.00921E-35
JulichBrain	SurfArea	alc	0.034262757	0.025194345	0.044610838	2.51143E-41
JulichBrain	ThickAvg	alc	0.025092241	0.017371108	0.034144776	1.38221E-30

Schaefer	GrayVol	alc	0.02276916	0.01543329	0.031449656	7.06832E-28
Schaefer	SurfArea	alc	0.027149285	0.019103696	0.036514607	5.47274E-33
Schaefer	ThickAvg	alc	0.029137067	0.020791195	0.038791431	2.58873E-35
vertices	SurfArea	alc	0.025705176	0.017885816	0.034852457	2.66152E-31
vertices	ThickAvg	alc	0.032203094	0.023416736	0.042280675	6.59694E-39
vertices	GrayVol	alc	0.025216883	0.017475666	0.034288795	9.88807E-31
Yeo	GrayVol	alc	0.015888973	0.00984425	0.023317764	7.071E-20
Yeo	SurfArea	alc	0.022990551	0.015617031	0.031707433	3.90265E-28
Yeo	ThickAvg	alc	0.016543868	0.010364394	0.024103638	1.22769E-20
Des	GrayVol	edu	0.025193668	0.018733296	0.032553608	1.31625E-44
Des	SurfArea	edu	0.018821829	0.013272174	0.025295645	1.171E-33
Des	ThickAvg	edu	0.031788651	0.024516216	0.039935454	5.32364E-56
DK	GrayVol	edu	0.019939241	0.014218582	0.026579731	1.41957E-35
DK	SurfArea	edu	0.0136576	0.008981793	0.019277679	8.10469E-25
DK	ThickAvg	edu	0.025569605	0.019059827	0.032977507	2.96226E-45
Glasser	GrayVol	edu	0.030472137	0.023353287	0.038470365	1.01351E-53
Glasser	SurfArea	edu	0.029297732	0.022319237	0.037160083	1.08921E-51
Glasser	ThickAvg	edu	0.031807323	0.024532737	0.039956206	4.94149E-56
Gordon	GrayVol	edu	0.029270253	0.022295081	0.037129385	1.21511E-51
Gordon	SurfArea	edu	0.023394656	0.017176753	0.030519046	1.64499E-41
Gordon	ThickAvg	edu	0.033419544	0.025962021	0.041745243	7.91491E-59
JulichBrain	GrayVol	edu	0.028245279	0.021395401	0.035983031	7.17259E-50
JulichBrain	SurfArea	edu	0.026876603	0.020198244	0.034448073	1.65259E-47
JulichBrain	ThickAvg	edu	0.03756314	0.029658681	0.046320115	4.95142E-66
Schaefer	GrayVol	edu	0.024147216	0.017826632	0.031371389	8.34166E-43
Schaefer	SurfArea	edu	0.022322977	0.016254584	0.029301987	1.1451E-39
Schaefer	ThickAvg	edu	0.030561002	0.023431662	0.038569381	7.11274E-54
vertices	SurfArea	edu	0.021216926	0.015307151	0.028041589	9.09985E-38
vertices	ThickAvg	edu	0.042997915	0.034551239	0.052276724	1.59651E-75
vertices	GrayVol	edu	0.029209671	0.022241833	0.037061701	1.54646E-51
Yeo	GrayVol	edu	0.014252292	0.009467607	0.019978926	7.79621E-26
Yeo	SurfArea	edu	0.012255832	0.007847178	0.017614244	2.01758E-22
Yeo	ThickAvg	edu	0.02277921	0.016646683	0.029820596	1.88178E-40

Wilcoxon tests based on LASSO prediction errors (which tests whether there is a sig. difference between the difference in observed vs. predicted values from two atlases)

atlas1	atlas2	Measurement	Pheno-type	N	Statistic	p-value	Effect size
Des	Glasser	GrayVol	age	7701	15021636	0.312	0.011521208
Des	Glasser	GrayVol	alc	5203	6656389	0.298	0.014421095
Des	Glasser	GrayVol	bmi	7701	14912064	0.653	0.005119933
Des	Glasser	GrayVol	cig	7467	14134502	0.299	0.012027884
Des	Glasser	GrayVol	edu	7697	14898177	0.662	0.004987201
Des	Glasser	GrayVol	gpheno	7742	15427595	0.877	0.001740268
Des	Glasser	GrayVol	sex	7701	15077551	0.194	0.014787802
Des	Glasser	SurfArea	age	7701	15243592	0.0316	0.024488036
Des	Glasser	SurfArea	alc	5203	6712057	0.599	0.007298701
Des	Glasser	SurfArea	bmi	7701	14691033	0.494	0.007792856
Des	Glasser	SurfArea	cig	7467	14045434	0.575	0.006494683
Des	Glasser	SurfArea	edu	7697	14413268	0.0404	0.023363616
Des	Glasser	SurfArea	gpheno	7742	15179915	0.166	0.015616246
Des	Glasser	SurfArea	sex	7701	14916861	0.636	0.005400177
Des	Glasser	ThickAvg	age	7701	14539162	0.144	0.016665269
Des	Glasser	ThickAvg	alc	5203	6655255	0.293	0.014566184
Des	Glasser	ThickAvg	bmi	7701	14709838	0.557	0.006694255
Des	Glasser	ThickAvg	cig	7467	13570966	0.0471	0.022980847
Des	Glasser	ThickAvg	edu	7697	14747911	0.739	0.00379829
Des	Glasser	ThickAvg	gpheno	7742	15940915	0.0166	0.027017876
Des	Glasser	ThickAvg	sex	7701	15665597	0.0000162	0.04914187
Des	Gordon	GrayVol	age	7701	15035805	0.279	0.012348971
Des	Gordon	GrayVol	alc	5203	6549693	0.0429	0.028072222
Des	Gordon	GrayVol	bmi	7701	14965367	0.47	0.008233933
Des	Gordon	GrayVol	cig	7467	13757756	0.335	0.011147879
Des	Gordon	GrayVol	edu	7697	15123728	0.111	0.018174325
Des	Gordon	GrayVol	gpheno	7742	15334206	0.536	0.006972276
Des	Gordon	GrayVol	sex	7701	14952220	0.512	0.007465875
Des	Gordon	SurfArea	age	7701	15453403	0.00126	0.036745345
Des	Gordon	SurfArea	alc	5203	6702630	0.54	0.00850483
Des	Gordon	SurfArea	bmi	7701	14765901	0.764	0.003419014
Des	Gordon	SurfArea	cig	7467	14023130	0.644	0.005342453
Des	Gordon	SurfArea	edu	7697	14630776	0.35	0.010646736
Des	Gordon	SurfArea	gpheno	7742	14738672	0.000348	0.040336361
Des	Gordon	SurfArea	sex	7701	15080848	0.189	0.014980415
Des	Gordon	ThickAvg	age	7701	14868434	0.822	0.002571037
Des	Gordon	ThickAvg	alc	5203	6629144	0.196	0.017906933
Des	Gordon	ThickAvg	bmi	7701	14915362	0.641	0.005312605
Des	Gordon	ThickAvg	cig	7467	13408588	0.00454	0.032845168
Des	Gordon	ThickAvg	edu	7697	14686241	0.516	0.007403904
Des	Gordon	ThickAvg	gpheno	7742	15489920	0.877	0.001751416
Des	Gordon	ThickAvg	sex	7701	15451266	0.00131	0.0366205

Des	JulichBrain	GrayVol	age	7701	14825388	0.996	5.62591E-05
Des	JulichBrain	GrayVol	alc	5203	6672713	0.374	0.012332535
Des	JulichBrain	GrayVol	bmi	7701	15192151	0.0594	0.021482816
Des	JulichBrain	GrayVol	cig	7467	14007355	0.721	0.004129089
Des	JulichBrain	GrayVol	edu	7697	14819990	0.971	0.0004159
Des	JulichBrain	GrayVol	gpheno	7742	15486585	0.89	0.001564577
Des	JulichBrain	GrayVol	sex	7701	14924424	0.608	0.005842013
Des	JulichBrain	SurfArea	age	7701	14942151	0.546	0.006877637
Des	JulichBrain	SurfArea	alc	5203	6671435	0.367	0.012496048
Des	JulichBrain	SurfArea	bmi	7701	14710643	0.56	0.006647226
Des	JulichBrain	SurfArea	cig	7467	14138774	0.288	0.012293274
Des	JulichBrain	SurfArea	edu	7697	14321580	0.0117	0.02872427
Des	JulichBrain	SurfArea	gpheno	7742	15440475	0.928	0.001018681
Des	JulichBrain	SurfArea	sex	7701	14846780	0.909	0.001305995
Des	JulichBrain	ThickAvg	age	7701	14509962	0.107	0.018371154
Des	JulichBrain	ThickAvg	alc	5203	6774897	0.957	0.000741308
Des	JulichBrain	ThickAvg	bmi	7701	14685912	0.478	0.008092029
Des	JulichBrain	ThickAvg	cig	7467	13468290	0.0112	0.029359422
Des	JulichBrain	ThickAvg	edu	7697	14844126	0.873	0.001827041
Des	JulichBrain	ThickAvg	gpheno	7742	15881475	0.0357	0.023687821
Des	JulichBrain	ThickAvg	sex	7701	15370463	0.00512	0.031899931
Des	Schaefer	GrayVol	age	7701	15258261	0.0261	0.02534501
Des	Schaefer	GrayVol	alc	5203	6712560	0.602	0.007234345
Des	Schaefer	GrayVol	bmi	7701	15000047	0.368	0.010259963
Des	Schaefer	GrayVol	cig	7467	13785875	0.405	0.009629985
Des	Schaefer	GrayVol	edu	7697	14898230	0.662	0.0049903
Des	Schaefer	GrayVol	gpheno	7742	15387418	0.723	0.003991136
Des	Schaefer	GrayVol	sex	7701	15101938	0.155	0.016212508
Des	Schaefer	SurfArea	age	7701	15136675	0.109	0.018241868
Des	Schaefer	SurfArea	alc	5203	6738161	0.775	0.003958847
Des	Schaefer	SurfArea	bmi	7701	14906387	0.674	0.004788279
Des	Schaefer	SurfArea	cig	7467	14237716	0.111	0.018439881
Des	Schaefer	SurfArea	edu	7697	14451125	0.0635	0.021150258
Des	Schaefer	SurfArea	gpheno	7742	15219812	0.235	0.013381064
Des	Schaefer	SurfArea	sex	7701	14841767	0.929	0.001013132
Des	Schaefer	ThickAvg	age	7701	14784256	0.837	0.002346702
Des	Schaefer	ThickAvg	alc	5203	6781957	0.906	0.001644594
Des	Schaefer	ThickAvg	bmi	7701	14702850	0.533	0.007102499
Des	Schaefer	ThickAvg	cig	7467	13468883	0.0113	0.029322583
Des	Schaefer	ThickAvg	edu	7697	14992720	0.356	0.010514777
Des	Schaefer	ThickAvg	gpheno	7742	15898736	0.0288	0.024654848
Des	Schaefer	ThickAvg	sex	7701	15794952	0.000000651	0.056698882
Des	vertices	GrayVol	age	7701	14973193	0.446	0.008691133
Des	vertices	GrayVol	alc	5203	6589329	0.0971	0.023001028
Des	vertices	GrayVol	bmi	7701	14892184	0.728	0.003958529
Des	vertices	GrayVol	cig	7467	13891215	0.79	0.003085914
Des	vertices	GrayVol	edu	7697	15035769	0.253	0.013031691

Des	vertices	GrayVol	gpheno	7742	15523062	0.749	0.003608158
Des	vertices	GrayVol	sex	7701	15133860	0.113	0.018077414
Des	vertices	SurfArea	age	7701	15083082	0.185	0.015110927
Des	vertices	SurfArea	alc	5203	6736257	0.762	0.004202453
Des	vertices	SurfArea	bmi	7701	14742407	0.674	0.00479155
Des	vertices	SurfArea	cig	7467	14054092	0.543	0.007032547
Des	vertices	SurfArea	edu	7697	14979828	0.392	0.00976103
Des	vertices	SurfArea	gpheno	7742	15389638	0.732	0.003866764
Des	vertices	SurfArea	sex	7701	14891583	0.731	0.003923418
Des	vertices	ThickAvg	age	7701	14839925	0.937	0.000905521
Des	vertices	ThickAvg	alc	5203	6387774	0.000433	0.048788808
Des	vertices	ThickAvg	bmi	7701	14720247	0.593	0.006086153
Des	vertices	ThickAvg	cig	7467	13658337	0.129	0.01755307
Des	vertices	ThickAvg	edu	7697	14375268	0.0248	0.025585334
Des	vertices	ThickAvg	gpheno	7742	16006003	0.00655	0.030664354
Des	vertices	ThickAvg	sex	7701	15330580	0.00947	0.029569937
Des	Yeo	GrayVol	age	7701	14627055	0.312	0.011530497
Des	Yeo	GrayVol	alc	5203	6782123	0.904	0.001665833
Des	Yeo	GrayVol	bmi	7701	14954679	0.504	0.007609532
Des	Yeo	GrayVol	cig	7467	13813229	0.493	0.007930664
Des	Yeo	GrayVol	edu	7697	14959439	0.452	0.008568962
Des	Yeo	GrayVol	gpheno	7742	15453225	0.978	0.000304377
Des	Yeo	GrayVol	sex	7701	14971119	0.452	0.008569968
Des	Yeo	SurfArea	age	7701	14848705	0.901	0.001418455
Des	Yeo	SurfArea	alc	5203	6863378	0.384	0.012061933
Des	Yeo	SurfArea	bmi	7701	14888106	0.744	0.00372029
Des	Yeo	SurfArea	cig	7467	14016905	0.683	0.004722367
Des	Yeo	SurfArea	edu	7697	14671740	0.469	0.008251724
Des	Yeo	SurfArea	gpheno	7742	15335339	0.54	0.006908801
Des	Yeo	SurfArea	sex	7701	15075435	0.198	0.014664184
Des	Yeo	ThickAvg	age	7701	14956483	0.498	0.007714923
Des	Yeo	ThickAvg	alc	5203	6833547	0.552	0.008245232
Des	Yeo	ThickAvg	bmi	7701	14593757	0.237	0.01347579
Des	Yeo	ThickAvg	cig	7467	13840491	0.59	0.006237058
Des	Yeo	ThickAvg	edu	7697	14634048	0.359	0.010455434
Des	Yeo	ThickAvg	gpheno	7742	15808733	0.082	0.019612536
Des	Yeo	ThickAvg	sex	7701	15383701	0.00414	0.032673304
DK	Des	GrayVol	age	7701	14951389	0.515	0.007417328
DK	Des	GrayVol	alc	5203	7140436	0.00061	0.047509879
DK	Des	GrayVol	bmi	7701	14839686	0.938	0.000891559
DK	Des	GrayVol	cig	7467	13992705	0.781	0.003218982
DK	Des	GrayVol	edu	7697	14454522	0.066	0.020951648
DK	Des	GrayVol	gpheno	7742	15412208	0.818	0.002602306
DK	Des	GrayVol	sex	7701	14777933	0.812	0.002716096
DK	Des	SurfArea	age	7701	14582055	0.214	0.014159429
DK	Des	SurfArea	alc	5203	6889016	0.268	0.015342165
DK	Des	SurfArea	bmi	7701	15094110	0.167	0.015755191

DK	Des	SurfArea	cig	7467	13867658	0.694	0.004549353
DK	Des	SurfArea	edu	7697	14864013	0.793	0.00298976
DK	Des	SurfArea	gpheno	7742	15613687	0.441	0.008685316
DK	Des	SurfArea	sex	7701	14675666	0.446	0.008690607
DK	Des	ThickAvg	age	7701	15063021	0.221	0.013938949
DK	Des	ThickAvg	alc	5203	6733715	0.744	0.004527687
DK	Des	ThickAvg	bmi	7701	14960185	0.486	0.007931196
DK	Des	ThickAvg	cig	7467	14170506	0.218	0.014264572
DK	Des	ThickAvg	edu	7697	14704597	0.579	0.006330698
DK	Des	ThickAvg	gpheno	7742	15044202	0.0395	0.023219405
DK	Des	ThickAvg	sex	7701	14469580	0.0689	0.020730299
DK	Glasser	GrayVol	age	7701	15128346	0.124	0.017525781
DK	Glasser	GrayVol	alc	5203	6955296	0.0857	0.023822302
DK	Glasser	GrayVol	bmi	7701	14951684	0.527	0.007207741
DK	Glasser	GrayVol	cig	7467	14123264	0.328	0.011329742
DK	Glasser	GrayVol	edu	7697	14547189	0.173	0.015533755
DK	Glasser	GrayVol	gpheno	7742	15358489	0.605	0.005830652
DK	Glasser	GrayVol	sex	7701	14983382	0.427	0.00905908
DK	Glasser	SurfArea	age	7701	15074672	0.207	0.014390922
DK	Glasser	SurfArea	alc	5203	6840115	0.512	0.009085569
DK	Glasser	SurfArea	bmi	7701	14889139	0.755	0.003554766
DK	Glasser	SurfArea	cig	7467	14067988	0.495	0.007895813
DK	Glasser	SurfArea	edu	7697	14372991	0.024	0.025718461
DK	Glasser	SurfArea	gpheno	7742	15121927	0.0906	0.019080387
DK	Glasser	SurfArea	sex	7701	14785613	0.827	0.002491727
DK	Glasser	ThickAvg	age	7701	14723919	0.593	0.006094998
DK	Glasser	ThickAvg	alc	5203	6611039	0.145	0.020223361
DK	Glasser	ThickAvg	bmi	7701	14828242	1	1.95659E-06
DK	Glasser	ThickAvg	cig	7467	13825464	0.536	0.007170585
DK	Glasser	ThickAvg	edu	7697	14644120	0.387	0.009866562
DK	Glasser	ThickAvg	gpheno	7742	15581187	0.556	0.006642567
DK	Glasser	ThickAvg	sex	7701	15275914	0.0218	0.026144571
DK	Gordon	GrayVol	age	7701	15044670	0.267	0.012638639
DK	Gordon	GrayVol	alc	5203	6827412	0.59	0.007460294
DK	Gordon	GrayVol	bmi	7701	15023632	0.317	0.011409903
DK	Gordon	GrayVol	cig	7467	13805945	0.481	0.008153416
DK	Gordon	GrayVol	edu	7697	14882248	0.722	0.004055892
DK	Gordon	GrayVol	gpheno	7742	15421578	0.839	0.002297064
DK	Gordon	GrayVol	sex	7701	14880661	0.788	0.003059604
DK	Gordon	SurfArea	age	7701	15253325	0.0294	0.024825248
DK	Gordon	SurfArea	alc	5203	6850886	0.45	0.010463655
DK	Gordon	SurfArea	bmi	7701	14968569	0.472	0.008193918
DK	Gordon	SurfArea	cig	7467	14085458	0.426	0.009215513
DK	Gordon	SurfArea	edu	7697	14545681	0.171	0.015621922
DK	Gordon	SurfArea	gpheno	7742	14784236	0.000754	0.037994313
DK	Gordon	SurfArea	sex	7701	14837339	0.963	0.000529359
DK	Gordon	ThickAvg	age	7701	15074358	0.207	0.014372583

DK	Gordon	ThickAvg	alc	5203	6573239	0.0707	0.02505965
DK	Gordon	ThickAvg	bmi	7701	15116263	0.14	0.016820067
DK	Gordon	ThickAvg	cig	7467	13635132	0.105	0.018767731
DK	Gordon	ThickAvg	edu	7697	14596323	0.267	0.012661074
DK	Gordon	ThickAvg	gpheno	7742	15190139	0.176	0.015259862
DK	Gordon	ThickAvg	sex	7701	15132137	0.119	0.017747197
DK	JulichBrain	GrayVol	age	7701	14858334	0.878	0.001755583
DK	JulichBrain	GrayVol	alc	5203	6984596	0.0467	0.027571065
DK	JulichBrain	GrayVol	bmi	7701	15215194	0.0474	0.022598186
DK	JulichBrain	GrayVol	cig	7467	13985189	0.812	0.002752063
DK	JulichBrain	GrayVol	edu	7697	14535085	0.154	0.016241431
DK	JulichBrain	GrayVol	gpheno	7742	15533210	0.726	0.003955396
DK	JulichBrain	GrayVol	sex	7701	14885311	0.77	0.003331189
DK	JulichBrain	SurfArea	age	7701	14943633	0.554	0.006737518
DK	JulichBrain	SurfArea	alc	5203	6818690	0.647	0.006344366
DK	JulichBrain	SurfArea	bmi	7701	14871899	0.823	0.002547854
DK	JulichBrain	SurfArea	cig	7467	14296285	0.0564	0.022078382
DK	JulichBrain	SurfArea	edu	7697	14252486	0.00405	0.032763938
DK	JulichBrain	SurfArea	gpheno	7742	15338608	0.538	0.006944178
DK	JulichBrain	SurfArea	sex	7701	14706923	0.534	0.007087659
DK	JulichBrain	ThickAvg	age	7701	14635617	0.323	0.011252325
DK	JulichBrain	ThickAvg	alc	5203	6715427	0.62	0.006867529
DK	JulichBrain	ThickAvg	bmi	7701	14935799	0.582	0.006279969
DK	JulichBrain	ThickAvg	cig	7467	13715749	0.227	0.013986446
DK	JulichBrain	ThickAvg	edu	7697	14737244	0.698	0.00442195
DK	JulichBrain	ThickAvg	gpheno	7742	15598790	0.499	0.007628503
DK	JulichBrain	ThickAvg	sex	7701	15131416	0.12	0.017705086
DK	Schaefer	GrayVol	age	7701	15141820	0.108	0.018312737
DK	Schaefer	GrayVol	alc	5203	7108314	0.00174	0.043400057
DK	Schaefer	GrayVol	bmi	7701	14875597	0.808	0.002763838
DK	Schaefer	GrayVol	cig	7467	13763659	0.341	0.011010117
DK	Schaefer	GrayVol	edu	7697	14649869	0.403	0.009530439
DK	Schaefer	GrayVol	gpheno	7742	15440150	0.911	0.001256855
DK	Schaefer	GrayVol	sex	7701	15013670	0.342	0.010828067
DK	Schaefer	SurfArea	age	7701	14884042	0.775	0.003257073
DK	Schaefer	SurfArea	alc	5203	6871149	0.346	0.013056187
DK	Schaefer	SurfArea	bmi	7701	15177728	0.0733	0.020409964
DK	Schaefer	SurfArea	cig	7467	14399002	0.0139	0.028459503
DK	Schaefer	SurfArea	edu	7697	14523093	0.137	0.016942558
DK	Schaefer	SurfArea	gpheno	7742	15333722	0.522	0.007217841
DK	Schaefer	SurfArea	sex	7701	14619089	0.284	0.012217652
DK	Schaefer	ThickAvg	age	7701	14948393	0.538	0.007015528
DK	Schaefer	ThickAvg	alc	5203	6771590	0.982	0.000318197
DK	Schaefer	ThickAvg	bmi	7701	14763376	0.739	0.003790491
DK	Schaefer	ThickAvg	cig	7467	13646266	0.114	0.018302961
DK	Schaefer	ThickAvg	edu	7697	14773988	0.842	0.002273665
DK	Schaefer	ThickAvg	gpheno	7742	15482261	0.922	0.001101764

DK	Schaefer	ThickAvg	sex	7701	15331731	0.00987	0.029404593
DK	vertices	GrayVol	age	7701	15027334	0.308	0.01162612
DK	vertices	GrayVol	alc	5203	6903806	0.214	0.017234458
DK	vertices	GrayVol	bmi	7701	14940434	0.565	0.006550678
DK	vertices	GrayVol	cig	7467	13923876	0.927	0.001056904
DK	vertices	GrayVol	edu	7697	14764337	0.803	0.002837923
DK	vertices	GrayVol	gpheno	7742	15498790	0.857	0.002027546
DK	vertices	GrayVol	sex	7701	15037459	0.284	0.012217476
DK	vertices	SurfArea	age	7701	14974342	0.454	0.008531094
DK	vertices	SurfArea	alc	5203	6905430	0.208	0.01744224
DK	vertices	SurfArea	bmi	7701	14888466	0.758	0.003515459
DK	vertices	SurfArea	cig	7467	14019902	0.671	0.00490855
DK	vertices	SurfArea	edu	7697	14918317	0.589	0.006164712
DK	vertices	SurfArea	gpheno	7742	15414945	0.813	0.002668576
DK	vertices	SurfArea	sex	7701	14856252	0.886	0.001633983
DK	vertices	ThickAvg	age	7701	14881667	0.784	0.00311836
DK	vertices	ThickAvg	alc	5203	6402599	0.000719	0.046892036
DK	vertices	ThickAvg	bmi	7701	14784462	0.822	0.002558951
DK	vertices	ThickAvg	cig	7467	13765145	0.345	0.010917802
DK	vertices	ThickAvg	edu	7697	14369749	0.023	0.025908009
DK	vertices	ThickAvg	gpheno	7742	15694497	0.249	0.01298901
DK	vertices	ThickAvg	sex	7701	15088530	0.182	0.015200306
DK	Yeo	GrayVol	age	7701	14762193	0.735	0.003859584
DK	Yeo	GrayVol	alc	5203	7237705	0.0000153	0.059954876
DK	Yeo	GrayVol	bmi	7701	14886202	0.767	0.003383229
DK	Yeo	GrayVol	cig	7467	13732351	0.263	0.012955074
DK	Yeo	GrayVol	edu	7697	14478032	0.0859	0.019577107
DK	Yeo	GrayVol	gpheno	7742	15525251	0.756	0.003509615
DK	Yeo	GrayVol	sex	7701	14984257	0.424	0.009110185
DK	Yeo	SurfArea	age	7701	14725527	0.598	0.006001082
DK	Yeo	SurfArea	alc	5203	7027664	0.017	0.033081363
DK	Yeo	SurfArea	bmi	7701	15114533	0.142	0.016719025
DK	Yeo	SurfArea	cig	7467	13922538	0.922	0.001140025
DK	Yeo	SurfArea	edu	7697	14554871	0.186	0.015084617
DK	Yeo	SurfArea	gpheno	7742	15292926	0.399	0.009502807
DK	Yeo	SurfArea	sex	7701	14831395	0.987	0.000182196
DK	Yeo	ThickAvg	age	7701	15222506	0.0433	0.023025247
DK	Yeo	ThickAvg	alc	5203	6835969	0.537	0.008555112
DK	Yeo	ThickAvg	bmi	7701	14850421	0.91	0.00129342
DK	Yeo	ThickAvg	cig	7467	14074181	0.474	0.008280542
DK	Yeo	ThickAvg	edu	7697	14416253	0.0419	0.023189094
DK	Yeo	ThickAvg	gpheno	7742	15459485	0.988	0.00017391
DK	Yeo	ThickAvg	sex	7701	15250339	0.0305	0.024650849
Glasser	Gordon	GrayVol	age	7701	14771085	0.769	0.003340242
Glasser	Gordon	GrayVol	alc	5203	6600106	0.119	0.021622175
Glasser	Gordon	GrayVol	bmi	7701	14808055	0.917	0.001180989
Glasser	Gordon	GrayVol	cig	7467	13636131	0.106	0.018705653

Glasser	Gordon	GrayVol	edu	7697	15131073	0.103	0.01860376
Glasser	Gordon	GrayVol	gpheno	7742	15460176	0.99	0.000135207
Glasser	Gordon	GrayVol	sex	7701	14561362	0.171	0.015589228
Glasser	Gordon	SurfArea	age	7701	15078477	0.2	0.014613155
Glasser	Gordon	SurfArea	alc	5203	6729142	0.712	0.005112775
Glasser	Gordon	SurfArea	bmi	7701	15008484	0.356	0.010525176
Glasser	Gordon	SurfArea	cig	7467	13891880	0.808	0.002813418
Glasser	Gordon	SurfArea	edu	7697	15097965	0.144	0.016668059
Glasser	Gordon	SurfArea	gpheno	7742	15043424	0.0373	0.023477306
Glasser	Gordon	SurfArea	sex	7701	14891242	0.747	0.003677593
Glasser	Gordon	ThickAvg	age	7701	15136737	0.114	0.018015862
Glasser	Gordon	ThickAvg	alc	5203	6744640	0.821	0.003129897
Glasser	Gordon	ThickAvg	bmi	7701	15009495	0.353	0.010584224
Glasser	Gordon	ThickAvg	cig	7467	13615011	0.0837	0.020018049
Glasser	Gordon	ThickAvg	edu	7697	14785526	0.888	0.001599082
Glasser	Gordon	ThickAvg	gpheno	7742	15043054	0.0372	0.023498029
Glasser	Gordon	ThickAvg	sex	7701	14625636	0.299	0.01183527
Glasser	JulichBrain	GrayVol	age	7701	14617521	0.28	0.012309231
Glasser	JulichBrain	GrayVol	alc	5203	6785993	0.876	0.002160976
Glasser	JulichBrain	GrayVol	bmi	7701	15104860	0.156	0.016154069
Glasser	JulichBrain	GrayVol	cig	7467	13848542	0.62	0.005736903
Glasser	JulichBrain	GrayVol	edu	7697	14787407	0.896	0.001489106
Glasser	JulichBrain	GrayVol	gpheno	7742	15543417	0.688	0.004527085
Glasser	JulichBrain	GrayVol	sex	7701	14644098	0.345	0.010756987
Glasser	JulichBrain	SurfArea	age	7701	14469786	0.0662	0.020937775
Glasser	JulichBrain	SurfArea	alc	5203	6696591	0.503	0.009277485
Glasser	JulichBrain	SurfArea	bmi	7701	14862357	0.861	0.001990549
Glasser	JulichBrain	SurfArea	cig	7467	13981505	0.827	0.002523201
Glasser	JulichBrain	SurfArea	edu	7697	14537470	0.158	0.016101989
Glasser	JulichBrain	SurfArea	gpheno	7742	15808616	0.0857	0.019380766
Glasser	JulichBrain	SurfArea	sex	7701	14695681	0.497	0.007744254
Glasser	JulichBrain	ThickAvg	age	7701	14783353	0.818	0.002623723
Glasser	JulichBrain	ThickAvg	alc	5203	6848366	0.464	0.010141236
Glasser	JulichBrain	ThickAvg	bmi	7701	14745814	0.673	0.004816209
Glasser	JulichBrain	ThickAvg	cig	7467	13722148	0.24	0.013588919
Glasser	JulichBrain	ThickAvg	edu	7697	14826991	0.942	0.000825222
Glasser	JulichBrain	ThickAvg	gpheno	7742	15461512	0.996	6.03783E-05
Glasser	JulichBrain	ThickAvg	sex	7701	14656251	0.378	0.010047185
Glasser	Schaefer	GrayVol	age	7701	15119802	0.135	0.017026764
Glasser	Schaefer	GrayVol	alc	5203	6875656	0.325	0.013632831
Glasser	Schaefer	GrayVol	bmi	7701	14884649	0.773	0.003292525
Glasser	Schaefer	GrayVol	cig	7467	13775196	0.374	0.0102934
Glasser	Schaefer	GrayVol	edu	7697	14637005	0.367	0.010282549
Glasser	Schaefer	GrayVol	gpheno	7742	15196102	0.186	0.014925877
Glasser	Schaefer	GrayVol	sex	7701	15027737	0.307	0.011649658
Glasser	Schaefer	SurfArea	age	7701	14680233	0.448	0.008646503
Glasser	Schaefer	SurfArea	alc	5203	6753899	0.888	0.001945263

Glasser	Schaefer	SurfArea	bmi	7701	15068284	0.219	0.014017827
Glasser	Schaefer	SurfArea	cig	7467	14172045	0.215	0.01436018
Glasser	Schaefer	SurfArea	edu	7697	14795107	0.927	0.001038916
Glasser	Schaefer	SurfArea	gpheno	7742	15538147	0.707	0.004231915
Glasser	Schaefer	SurfArea	sex	7701	14581951	0.207	0.014386717
Glasser	Schaefer	ThickAvg	age	7701	15013877	0.341	0.010840157
Glasser	Schaefer	ThickAvg	alc	5203	6978450	0.0534	0.02678472
Glasser	Schaefer	ThickAvg	bmi	7701	14853926	0.895	0.001498131
Glasser	Schaefer	ThickAvg	cig	7467	13757079	0.324	0.011418889
Glasser	Schaefer	ThickAvg	edu	7697	15206181	0.0437	0.022995044
Glasser	Schaefer	ThickAvg	gpheno	7742	15552782	0.654	0.005051615
Glasser	Schaefer	ThickAvg	sex	7701	15061448	0.232	0.013618567
Glasser	vertices	GrayVol	age	7701	14808397	0.919	0.001161015
Glasser	vertices	GrayVol	alc	5203	6707668	0.571	0.007860248
Glasser	vertices	GrayVol	bmi	7701	14877522	0.801	0.002876269
Glasser	vertices	GrayVol	cig	7467	13834106	0.566	0.006633715
Glasser	vertices	GrayVol	edu	7697	14995426	0.349	0.010672987
Glasser	vertices	GrayVol	gpheno	7742	15566385	0.606	0.005813513
Glasser	vertices	GrayVol	sex	7701	14899063	0.717	0.004134383
Glasser	vertices	SurfArea	age	7701	14645825	0.35	0.010656121
Glasser	vertices	SurfArea	alc	5203	6798518	0.786	0.003763477
Glasser	vertices	SurfArea	bmi	7701	14934676	0.586	0.006214379
Glasser	vertices	SurfArea	cig	7467	13876000	0.728	0.004031121
Glasser	vertices	SurfArea	edu	7697	15362591	0.00481	0.032139752
Glasser	vertices	SurfArea	gpheno	7742	15771770	0.125	0.017317038
Glasser	vertices	SurfArea	sex	7701	14815171	0.946	0.000765375
Glasser	vertices	ThickAvg	age	7701	15163264	0.086	0.019565186
Glasser	vertices	ThickAvg	alc	5203	6474013	0.00646	0.037755034
Glasser	vertices	ThickAvg	bmi	7701	14776432	0.79	0.003027948
Glasser	vertices	ThickAvg	cig	7467	13834282	0.567	0.006622781
Glasser	vertices	ThickAvg	edu	7697	14525225	0.14	0.016817908
Glasser	vertices	ThickAvg	gpheno	7742	15624009	0.423	0.009041008
Glasser	vertices	ThickAvg	sex	7701	14641531	0.338	0.010906914
Glasser	Yeo	GrayVol	age	7701	14522074	0.117	0.017883866
Glasser	Yeo	GrayVol	alc	5203	6899220	0.23	0.016647707
Glasser	Yeo	GrayVol	bmi	7701	14760081	0.727	0.003982937
Glasser	Yeo	GrayVol	cig	7467	13659286	0.131	0.017494115
Glasser	Yeo	GrayVol	edu	7697	14815722	0.988	0.000166366
Glasser	Yeo	GrayVol	gpheno	7742	15593442	0.516	0.007328964
Glasser	Yeo	GrayVol	sex	7701	14765848	0.749	0.003646112
Glasser	Yeo	SurfArea	age	7701	14498470	0.091	0.01926247
Glasser	Yeo	SurfArea	alc	5203	6902993	0.217	0.01713044
Glasser	Yeo	SurfArea	bmi	7701	15002843	0.371	0.01019571
Glasser	Yeo	SurfArea	cig	7467	13903448	0.841	0.002325959
Glasser	Yeo	SurfArea	edu	7697	15191484	0.0521	0.022135766
Glasser	Yeo	SurfArea	gpheno	7742	15811993	0.0827	0.019569911
Glasser	Yeo	SurfArea	sex	7701	14942905	0.557	0.006694998

Glasser	Yeo	ThickAvg	age	7701	15120253	0.135	0.017053105
Glasser	Yeo	ThickAvg	alc	5203	6979512	0.0522	0.026920597
Glasser	Yeo	ThickAvg	bmi	7701	14715584	0.564	0.006581809
Glasser	Yeo	ThickAvg	cig	7467	14120523	0.335	0.011159462
Glasser	Yeo	ThickAvg	edu	7697	14724494	0.65	0.005167394
Glasser	Yeo	ThickAvg	gpheno	7742	15409807	0.793	0.002956353
Glasser	Yeo	ThickAvg	sex	7701	14685088	0.463	0.008362944
Gordon	JulichBrain	GrayVol	age	7701	14675276	0.433	0.008936019
Gordon	JulichBrain	GrayVol	alc	5203	6865136	0.375	0.012286859
Gordon	JulichBrain	GrayVol	bmi	7701	15001294	0.375	0.01010524
Gordon	JulichBrain	GrayVol	cig	7467	14094462	0.398	0.009775021
Gordon	JulichBrain	GrayVol	edu	7697	14543476	0.167	0.01575084
Gordon	JulichBrain	GrayVol	gpheno	7742	15566505	0.606	0.005820234
Gordon	JulichBrain	GrayVol	sex	7701	14816991	0.954	0.000659077
Gordon	JulichBrain	SurfArea	age	7701	14297703	0.00654	0.030988376
Gordon	JulichBrain	SurfArea	alc	5203	6711644	0.596	0.007351542
Gordon	JulichBrain	SurfArea	bmi	7701	14701214	0.515	0.007421096
Gordon	JulichBrain	SurfArea	cig	7467	13999283	0.739	0.003860601
Gordon	JulichBrain	SurfArea	edu	7697	14380483	0.0266	0.025280432
Gordon	JulichBrain	SurfArea	gpheno	7742	16219819	0.000169	0.042412068
Gordon	JulichBrain	SurfArea	sex	7701	14614935	0.274	0.012460268
Gordon	JulichBrain	ThickAvg	age	7701	14445600	0.0498	0.022350371
Gordon	JulichBrain	ThickAvg	alc	5203	6853990	0.433	0.010860794
Gordon	JulichBrain	ThickAvg	bmi	7701	14594078	0.23	0.013678433
Gordon	JulichBrain	ThickAvg	cig	7467	14012427	0.686	0.004677369
Gordon	JulichBrain	ThickAvg	edu	7697	15008910	0.315	0.011461346
Gordon	JulichBrain	ThickAvg	gpheno	7742	15820211	0.0757	0.020030197
Gordon	JulichBrain	ThickAvg	sex	7701	14737169	0.641	0.005321125
Gordon	Schaefer	GrayVol	age	7701	15116720	0.139	0.016846758
Gordon	Schaefer	GrayVol	alc	5203	6944333	0.106	0.02241965
Gordon	Schaefer	GrayVol	bmi	7701	14879435	0.793	0.002987998
Gordon	Schaefer	GrayVol	cig	7467	14064368	0.495	0.007904981
Gordon	Schaefer	GrayVol	edu	7697	14486565	0.0942	0.019078214
Gordon	Schaefer	GrayVol	gpheno	7742	15327748	0.503	0.007552442
Gordon	Schaefer	GrayVol	sex	7701	15414222	0.00267	0.034222525
Gordon	Schaefer	SurfArea	age	7701	14464187	0.062	0.021264787
Gordon	Schaefer	SurfArea	alc	5203	6718509	0.641	0.006473205
Gordon	Schaefer	SurfArea	bmi	7701	14912819	0.665	0.004937809
Gordon	Schaefer	SurfArea	cig	7467	14080211	0.442	0.008889464
Gordon	Schaefer	SurfArea	edu	7697	14473970	0.0821	0.019814596
Gordon	Schaefer	SurfArea	gpheno	7742	15942487	0.0171	0.026878823
Gordon	Schaefer	SurfArea	sex	7701	14610708	0.265	0.012707148
Gordon	Schaefer	ThickAvg	age	7701	14629950	0.309	0.011583309
Gordon	Schaefer	ThickAvg	alc	5203	7003053	0.0308	0.02993253
Gordon	Schaefer	ThickAvg	bmi	7701	14647807	0.355	0.010540361
Gordon	Schaefer	ThickAvg	cig	7467	13926164	0.953	0.000683011
Gordon	Schaefer	ThickAvg	edu	7697	15433050	0.00147	0.036259227

Gordon	Schaefer	ThickAvg	gpheno	7742	16038556	0.00423	0.032259606
Gordon	Schaefer	ThickAvg	sex	7701	15334420	0.00948	0.029561645
Gordon	vertices	GrayVol	age	7701	14865816	0.847	0.002192573
Gordon	vertices	GrayVol	alc	5203	6811638	0.695	0.005442104
Gordon	vertices	GrayVol	bmi	7701	14797009	0.873	0.001826137
Gordon	vertices	GrayVol	cig	7467	14000709	0.733	0.003949212
Gordon	vertices	GrayVol	edu	7697	14690798	0.531	0.007137474
Gordon	vertices	GrayVol	gpheno	7742	15622594	0.427	0.008961755
Gordon	vertices	GrayVol	sex	7701	15061991	0.231	0.013650281
Gordon	vertices	SurfArea	age	7701	14445531	0.0498	0.022354401
Gordon	vertices	SurfArea	alc	5203	6790131	0.846	0.002690409
Gordon	vertices	SurfArea	bmi	7701	14855963	0.887	0.001617104
Gordon	vertices	SurfArea	cig	7467	13907414	0.873	0.001848136
Gordon	vertices	SurfArea	edu	7697	15081319	0.169	0.01569483
Gordon	vertices	SurfArea	gpheno	7742	16122306	0.00105	0.036950407
Gordon	vertices	SurfArea	sex	7701	14841610	0.946	0.000778809
Gordon	vertices	ThickAvg	age	7701	14766976	0.753	0.003580231
Gordon	vertices	ThickAvg	alc	5203	6525336	0.0245	0.031188557
Gordon	vertices	ThickAvg	bmi	7701	14671117	0.421	0.009178928
Gordon	vertices	ThickAvg	cig	7467	14079936	0.443	0.008872376
Gordon	vertices	ThickAvg	edu	7697	14565354	0.204	0.014471716
Gordon	vertices	ThickAvg	gpheno	7742	15939793	0.0178	0.026727933
Gordon	vertices	ThickAvg	sex	7701	14817223	0.955	0.000645527
Gordon	Yeo	GrayVol	age	7701	14604680	0.252	0.013059217
Gordon	Yeo	GrayVol	alc	5203	6993438	0.0384	0.028702347
Gordon	Yeo	GrayVol	bmi	7701	14708225	0.538	0.007011615
Gordon	Yeo	GrayVol	cig	7467	13934336	0.988	0.000175204
Gordon	Yeo	GrayVol	edu	7697	14574428	0.221	0.013941193
Gordon	Yeo	GrayVol	gpheno	7742	15541196	0.696	0.004402688
Gordon	Yeo	GrayVol	sex	7701	14833985	0.977	0.000333466
Gordon	Yeo	SurfArea	age	7701	14384518	0.0229	0.025917899
Gordon	Yeo	SurfArea	alc	5203	6950115	0.0948	0.023159423
Gordon	Yeo	SurfArea	bmi	7701	14915341	0.655	0.005085108
Gordon	Yeo	SurfArea	cig	7467	13936946	0.999	1.30183E-05
Gordon	Yeo	SurfArea	edu	7697	14918713	0.587	0.006187865
Gordon	Yeo	SurfArea	gpheno	7742	15987392	0.00914	0.029393932
Gordon	Yeo	SurfArea	sex	7701	14804358	0.902	0.001396915
Gordon	Yeo	ThickAvg	age	7701	14908031	0.683	0.004658163
Gordon	Yeo	ThickAvg	alc	5203	6994890	0.0372	0.028888122
Gordon	Yeo	ThickAvg	bmi	7701	14526047	0.121	0.01765182
Gordon	Yeo	ThickAvg	cig	7467	14255130	0.0878	0.019758925
Gordon	Yeo	ThickAvg	edu	7697	14776088	0.85	0.002150886
Gordon	Yeo	ThickAvg	gpheno	7742	15715003	0.21	0.014137543
Gordon	Yeo	ThickAvg	sex	7701	14769146	0.762	0.003453491
JulichBrain	vertices	GrayVol	age	7701	14958227	0.505	0.007589888
JulichBrain	vertices	GrayVol	alc	5203	6725511	0.687	0.005577341
JulichBrain	vertices	GrayVol	bmi	7701	14640132	0.335	0.010988624

JulichBrain	vertices	GrayVol	cig	7467	13902580	0.837	0.002379883
JulichBrain	vertices	GrayVol	edu	7697	14919068	0.586	0.00620862
JulichBrain	vertices	GrayVol	gpheno	7742	15540046	0.7	0.004338277
JulichBrain	vertices	GrayVol	sex	7701	15011099	0.349	0.010677906
JulichBrain	vertices	SurfArea	age	7701	15001309	0.375	0.010106116
JulichBrain	vertices	SurfArea	alc	5203	6828990	0.58	0.00766219
JulichBrain	vertices	SurfArea	bmi	7701	14899216	0.716	0.004143319
JulichBrain	vertices	SurfArea	cig	7467	13744290	0.291	0.012213384
JulichBrain	vertices	SurfArea	edu	7697	15424494	0.00171	0.035758989
JulichBrain	vertices	SurfArea	gpheno	7742	15416347	0.818	0.00259005
JulichBrain	vertices	SurfArea	sex	7701	14983777	0.425	0.00908215
JulichBrain	vertices	ThickAvg	age	7701	15079924	0.197	0.014697668
JulichBrain	vertices	ThickAvg	alc	5203	6419344	0.00125	0.044749612
JulichBrain	vertices	ThickAvg	bmi	7701	14831784	0.986	0.000204916
JulichBrain	vertices	ThickAvg	cig	7467	14051315	0.553	0.00686003
JulichBrain	vertices	ThickAvg	edu	7697	14444944	0.0591	0.021511638
JulichBrain	vertices	ThickAvg	gpheno	7742	15548002	0.671	0.004783889
JulichBrain	vertices	ThickAvg	sex	7701	14887370	0.762	0.003451446
JulichBrain	Yeo	GrayVol	age	7701	14770409	0.767	0.003379724
JulichBrain	Yeo	GrayVol	alc	5203	6927458	0.144	0.020260593
JulichBrain	Yeo	GrayVol	bmi	7701	14509008	0.102	0.018646992
JulichBrain	Yeo	GrayVol	cig	7467	13753948	0.316	0.011613397
JulichBrain	Yeo	GrayVol	edu	7697	14806444	0.974	0.000376084
JulichBrain	Yeo	GrayVol	gpheno	7742	15374195	0.661	0.004950966
JulichBrain	Yeo	GrayVol	sex	7701	14923569	0.625	0.005565669
JulichBrain	Yeo	SurfArea	age	7701	14722803	0.589	0.006160179
JulichBrain	Yeo	SurfArea	alc	5203	6918834	0.167	0.019157203
JulichBrain	Yeo	SurfArea	bmi	7701	14976142	0.449	0.008636224
JulichBrain	Yeo	SurfArea	cig	7467	13676540	0.156	0.016422239
JulichBrain	Yeo	SurfArea	edu	7697	15209943	0.0417	0.023214994
JulichBrain	Yeo	SurfArea	gpheno	7742	15447533	0.94	0.000843336
JulichBrain	Yeo	SurfArea	sex	7701	14931973	0.595	0.006056509
JulichBrain	Yeo	ThickAvg	age	7701	15218605	0.0454	0.022797407
JulichBrain	Yeo	ThickAvg	alc	5203	6889032	0.268	0.015344212
JulichBrain	Yeo	ThickAvg	bmi	7701	14716541	0.567	0.006525914
JulichBrain	Yeo	ThickAvg	cig	7467	14229815	0.121	0.017949044
JulichBrain	Yeo	ThickAvg	edu	7697	14715401	0.617	0.005699028
JulichBrain	Yeo	ThickAvg	gpheno	7742	15508206	0.821	0.002554932
JulichBrain	Yeo	ThickAvg	sex	7701	14827167	0.995	6.47425E-05
Schaefer	JulichBrain	GrayVol	age	7701	14421652	0.0372	0.023749067
Schaefer	JulichBrain	GrayVol	alc	5203	6693536	0.486	0.009668354
Schaefer	JulichBrain	GrayVol	bmi	7701	15000093	0.379	0.010035095
Schaefer	JulichBrain	GrayVol	cig	7467	14118408	0.341	0.011028071
Schaefer	JulichBrain	GrayVol	edu	7697	14856864	0.821	0.002571785
Schaefer	JulichBrain	GrayVol	gpheno	7742	15629053	0.408	0.00932352
Schaefer	JulichBrain	GrayVol	sex	7701	14615278	0.275	0.012440235
Schaefer	JulichBrain	SurfArea	age	7701	14672508	0.425	0.009097686

Schaefer	JulichBrain	SurfArea	alc	5203	6736451	0.763	0.004177632
Schaefer	JulichBrain	SurfArea	bmi	7701	14608112	0.259	0.012858769
Schaefer	JulichBrain	SurfArea	cig	7467	13944823	0.983	0.000244393
Schaefer	JulichBrain	SurfArea	edu	7697	14637890	0.369	0.010230807
Schaefer	JulichBrain	SurfArea	gpheno	7742	15666481	0.311	0.011419846
Schaefer	JulichBrain	SurfArea	sex	7701	14939670	0.568	0.006506057
Schaefer	JulichBrain	ThickAvg	age	7701	14648964	0.358	0.010472786
Schaefer	JulichBrain	ThickAvg	alc	5203	6749124	0.854	0.002556196
Schaefer	JulichBrain	ThickAvg	bmi	7701	14847144	0.923	0.001102025
Schaefer	JulichBrain	ThickAvg	cig	7467	13974156	0.858	0.002066657
Schaefer	JulichBrain	ThickAvg	edu	7697	14745897	0.731	0.003916041
Schaefer	JulichBrain	ThickAvg	gpheno	7742	15426220	0.857	0.002037068
Schaefer	JulichBrain	ThickAvg	sex	7701	14601400	0.245	0.013250787
Schaefer	vertices	GrayVol	age	7701	14770197	0.766	0.003392106
Schaefer	vertices	GrayVol	alc	5203	6644827	0.251	0.015900385
Schaefer	vertices	GrayVol	bmi	7701	14830056	0.993	0.000103991
Schaefer	vertices	GrayVol	cig	7467	14099108	0.396	0.009829091
Schaefer	vertices	GrayVol	edu	7697	14955778	0.464	0.008354917
Schaefer	vertices	GrayVol	gpheno	7742	15651781	0.347	0.010596506
Schaefer	vertices	GrayVol	sex	7701	14781639	0.811	0.00272383
Schaefer	vertices	SurfArea	age	7701	14796911	0.872	0.001831861
Schaefer	vertices	SurfArea	alc	5203	6786809	0.87	0.002265379
Schaefer	vertices	SurfArea	bmi	7701	14723693	0.592	0.006108198
Schaefer	vertices	SurfArea	cig	7467	13697528	0.191	0.015118395
Schaefer	vertices	SurfArea	edu	7697	15276968	0.0173	0.027133696
Schaefer	vertices	SurfArea	gpheno	7742	15686180	0.267	0.012523179
Schaefer	vertices	SurfArea	sex	7701	14890178	0.751	0.003615449
Schaefer	vertices	ThickAvg	age	7701	14940379	0.566	0.006547466
Schaefer	vertices	ThickAvg	alc	5203	6408153	0.000865	0.046181435
Schaefer	vertices	ThickAvg	bmi	7701	14809908	0.925	0.001072764
Schaefer	vertices	ThickAvg	cig	7467	14063094	0.512	0.007591781
Schaefer	vertices	ThickAvg	edu	7697	14311894	0.0102	0.029290574
Schaefer	vertices	ThickAvg	gpheno	7742	15521338	0.77	0.00329045
Schaefer	vertices	ThickAvg	sex	7701	14939115	0.57	0.006473641
Schaefer	Yeo	GrayVol	age	7701	14533296	0.131	0.017228439
Schaefer	Yeo	GrayVol	alc	5203	6827116	0.592	0.007422423
Schaefer	Yeo	GrayVol	bmi	7701	14866128	0.846	0.002210796
Schaefer	Yeo	GrayVol	cig	7467	14064462	0.507	0.007676766
Schaefer	Yeo	GrayVol	edu	7697	14701585	0.568	0.006506798
Schaefer	Yeo	GrayVol	gpheno	7742	15405752	0.778	0.003183472
Schaefer	Yeo	GrayVol	sex	7701	14829751	0.994	8.61774E-05
Schaefer	Yeo	SurfArea	age	7701	14607040	0.257	0.01292138
Schaefer	Yeo	SurfArea	alc	5203	6883371	0.292	0.01461992
Schaefer	Yeo	SurfArea	bmi	7701	14761146	0.731	0.003920735
Schaefer	Yeo	SurfArea	cig	7467	13611100	0.0767	0.02048759
Schaefer	Yeo	SurfArea	edu	7697	14927750	0.556	0.006716224
Schaefer	Yeo	SurfArea	gpheno	7742	15487554	0.901	0.001398223

Schaefer	Yeo	SurfArea	sex	7701	15012090	0.346	0.010735786
Schaefer	Yeo	ThickAvg	age	7701	15047333	0.262	0.012794173
Schaefer	Yeo	ThickAvg	alc	5203	6764905	0.969	0.000537109
Schaefer	Yeo	ThickAvg	bmi	7701	14814341	0.943	0.000813852
Schaefer	Yeo	ThickAvg	cig	7467	14304089	0.0512	0.022563192
Schaefer	Yeo	ThickAvg	edu	7697	14664882	0.448	0.008652685
Schaefer	Yeo	ThickAvg	gpheno	7742	15700815	0.237	0.013342879
Schaefer	Yeo	ThickAvg	sex	7701	14693496	0.49	0.00787187
Yeo	vertices	GrayVol	age	7701	15046720	0.263	0.01275837
Yeo	vertices	GrayVol	alc	5203	6540294	0.0347	0.029274769
Yeo	vertices	GrayVol	bmi	7701	14864516	0.853	0.002116646
Yeo	vertices	GrayVol	cig	7467	14025766	0.649	0.005272842
Yeo	vertices	GrayVol	edu	7697	14994615	0.351	0.010625571
Yeo	vertices	GrayVol	gpheno	7742	15545580	0.68	0.004648234
Yeo	vertices	GrayVol	sex	7701	14916268	0.652	0.00513925
Yeo	vertices	SurfArea	age	7701	15040789	0.276	0.012411967
Yeo	vertices	SurfArea	alc	5203	6702711	0.54	0.008494467
Yeo	vertices	SurfArea	bmi	7701	14680504	0.449	0.008630675
Yeo	vertices	SurfArea	cig	7467	14057760	0.53	0.007260415
Yeo	vertices	SurfArea	edu	7697	15102123	0.138	0.016911162
Yeo	vertices	SurfArea	gpheno	7742	15456285	0.975	0.00035314
Yeo	vertices	SurfArea	sex	7701	14738124	0.644	0.005265348
Yeo	vertices	ThickAvg	age	7701	14694374	0.493	0.00782059
Yeo	vertices	ThickAvg	alc	5203	6351377	0.000116	0.053445591
Yeo	vertices	ThickAvg	bmi	7701	14878435	0.797	0.002929593
Yeo	vertices	ThickAvg	cig	7467	13714473	0.224	0.014065715
Yeo	vertices	ThickAvg	edu	7697	14588244	0.249	0.013133423
Yeo	vertices	ThickAvg	gpheno	7742	15543086	0.689	0.004508546
Yeo	vertices	ThickAvg	sex	7701	14942578	0.558	0.0066759

Lessons from Statistical Genetics may help us overcome the Neuroimaging replication crisis

Anna Elisabeth Fürtjes [1]

[1] Social, Genetic and Developmental Psychiatry Centre, King's College, London UK

The article is displayed as submitted to *Cortex*. It is currently under review.

Abstract

The research fields of *Statistical Genetics* and *Neuroimaging* face similar challenges in identifying reliable biological correlates of common traits and diseases. This Viewpoint focuses on five major lessons that allowed genetics research to overcome many of its issues of replicability, and that may be directly applicable to neuroimaging research. First, the failure of candidate gene studies inspires abandoning overly simplistic studies mapping individual brain regions onto traits and diseases. Second, developments in genetics research demonstrate that robust study results can be achieved by increasing sample sizes. Third and fourth, the success of genome-wide association studies motivates the use of mass-univariate testing and sharing summary-level association data to boost large-scale collaboration and meta-analysis. Finally, applying genetics methods dealing with complex data structures to vertex-wise neuroimaging data promises novel discoveries without the need to develop novel neuroimaging-specific methods. Those practices – that are firmly established in genetics research – should either be further endorsed, or newly adopted by the neuroimaging community, promising to accelerate the evolution of Neuroimaging through robust discovery.

Viewpoint

Fifteen years ago, the scientific community was united in thinking we'd found a gene responsible for depression. Roughly 450 peer-reviewed studies, published in reputable journals, had delivered apparently supporting evidence for the hypothesis that the *Serotonin Transporter Gene* formed the biological basis of depression (Border et al., 2019). As the name suggests, the Serotonin Transporter Gene regulates serotonin levels in the brain, making it a logical therapeutic target that conformed with popular theories of depression at the time. Many other so-called *candidate gene studies,* which tested similar hypotheses about single genes forming the basis of other human traits and diseases, also claimed to have uncovered underlying genetic mechanisms. Those studies were cited countless times. Unfortunately, almost all this research later transpired to be based on oversimplified notions of human biology and could not be reliably replicated.

Candidate gene studies had mostly accumulated false results, and are now considered obsolete (Border et al., 2019). The full extent to which decades worth of research was erroneous became clear when studies with better methods and superior statistical power systematically contradicted candidate gene findings. The field overcame many of its flaws through drastically reforming approaches to analysing big genetic data, which eventually allowed novel insights into human biology. Contemporary genetic discoveries promise exciting translations into applications of personalised healthcare, according to which we will be able to predict disease risk from an individual's genetic make-up to guide treatment, or even prevent disease altogether.

Neuroimaging research now faces similar challenges of replicability. Specifically, studies looking for brain regions that may be "responsible" for functions (or dysfunctions) are just as difficult to replicate as candidate gene studies. For example, the Parieto-Frontal-Integration theory (P-FIT) – suggesting that enlarged frontal and parietal brain regions underpin good cognitive ability – shaped plenty of neuro-cognitive studies. However, meta-analytic evidence for this theory is inconsistent (Basten et al., 2015). It is possible, if not likely, that our future selves will remember the P-FIT, as the Serotonin Transporter Gene theory, as an abandoned piece of the self-correcting scientific process.

It is the aim of this essay to outline striking parallels between the challenges faced in the fields of Statistical Genetics and Neuroimaging. It is at the core of my reasoning that traits and diseases have complex genome-wide and brain-wide biology that both demonstrate similar characteristics. Based on their parallels, I will outline how genetics research overcame issues of replicability, and how it motivates practices that are either already endorsed by parts of the neuroimaging community, or even inspire adopting novel practices, that are firmly established in genetics research and will encourage an acceleration of the evolution of Neuroimaging.

Lesson 1: Abandon traditional studies mapping one biological variable onto traits and diseases

To illustrate how lessons from one research field may inspire change in another, I will first focus on advances in Statistical Genetics that were key to moving past the replication crisis. Most importantly, it was a conceptual shift from candidate gene towards *genome-wide approaches* that allowed the field to reliably identify genetic risk factors. In essence, candidate gene and genome-wide approaches differ in that the

former describe the statistical relationship between a trait and one pre-specified gene-of-interest that the researcher hypothesised to form its biological basis. Genome-wide approaches are hypothesis-free, and consider thousands, or millions of genetic markers.

Genome-wide methods successfully enabled robust discoveries as they accommodate two main characteristics of the genetic architecture of traits and diseases. First, genome-wide methods consider that markers across the genome are correlated among one another, which reflects the fact that genes are passed through families in conjunction with other genes. Geneticists call this *linkage disequilibrium*. Second, genome-wide methods recognise that most human traits have many genetic correlates that are weakly associated and diffusely distributed across the whole genome (as opposed to being controlled by only one strongly associated gene). Geneticists refer to this as *polygenicity*. It is now widely accepted that many genetic correlates (sometimes hundreds) account for why traits and diseases are heritable. For example, whether a person develops depression is influenced by how many genetic risk markers this person inherits at birth. From this perspective, it is intuitive that polygenic traits can only be modelled appropriately by methods that consider the entire genome, as opposed to one individual gene.

Those genetic data structures (i.e., linkage disequilibrium and polygenicity) both have close analogies in neuroimaging data. The parallel is strongest when neuroimaging data is represented in its raw vertex-wise (or voxel-wise) form, including hundreds of thousands of brain-wide measures. Like intercorrelated genetic markers, a measure of cortical thickness at a certain vertex is correlated with other vertex-wise brain measures, particularly with those in physical proximity. These interdependencies are organised along cortical gradients (Huntenburg et al., 2018).

Furthermore, we know from functional neuroimaging, and other modalities, that traits have many correlates spread across the brain (Marek et al., 2022), suggesting approaches considering only one brain region oversimplify matters to a substantial degree. Like in genetics research, it is more appropriate to model the brain based on thousands of brain measures, instead of considering one crudely averaged region-of-interest (ROI).

Lesson 2: Increase Sample Sizes

Using genome-wide approaches, the genetics community soon realised that polygenic traits have many genetic correlates with effect sizes much smaller than previously expected. As small effects require large samples to achieve adequate statistical power, it is widely accepted that insufficient samples had hindered the reliable identification of genetic mechanisms. For example, Serotonin Transporter Gene studies had a median sample size of 435 (Border et al., 2019), and resulting false discoveries were amplified by publication bias. Many efforts have since been devoted to increasing significance thresholds to counteract chance findings, as well as collecting large-scale genotyped samples, in some cases including millions of participants (Yengo et al., 2022).

While some neuroimaging samples are continually growing – for example, the UK Biobank cohort is on a trajectory to scanning 100,000 brains (Littlejohns et al., 2020) – overall, they remain small with a median sample size of 25 (Marek et al., 2022). Where big samples collected through consortiums improved the reliability of genetics studies, larger samples are imperative to improving neuroimaging studies too. Consortia like ENIGMA are already pioneering data sharing of tens of thousands

of participants which will unlock the reliable identification of many correlates spread across the brain.

Lesson 3: Use Mass-Univariate Testing

Beyond increasing sample sizes, genetics research established statistical techniques handling complicated data structures, that can also model vertex-wise neuroimaging data and account for complex brain-trait relationships. A popular genome-wide technique is mass-univariate testing, which geneticists call *genome-wide association studies* (GWAS). GWAS take a hypothesis-free approach to scanning the genome for any association between a trait and millions of genetic markers. GWAS results have been reliably replicated across many phenotypes and samples (Visscher et al., 2017). Resulting summary statistics, which conceal sensitive participant-level information, can be publicly shared, enabling large collaborative efforts and powerful meta-analyses. It has become routine for researchers to inform their genetic studies with the newest GWAS association data, in order to predict individual-level disease based on polygenic scores. Those scores reflect an individual's propensity towards disease and their predictive value is improving as sample sizes grow (Visscher et al., 2017).

Mass-univariate testing, which is what a GWAS does, has also been employed by neuroimaging studies, in which associations between traits and hundreds of thousands of vertex-wise brain measures are quantified. Vertex-wise mass-univariate testing is used in many neuroimaging studies (Ashburner & Friston, 2000), however, it has not fully replaced limited ROI-based studies. A downside to mass-univariate testing is the considerable power losses due to many significance tests that need correction for multiple testing. However, increasing neuroimaging sample sizes and

larger computational resources promise small but accurate estimates of vertex-trait associations, which will help uncover meaningful brain-wide association patterns in the future.

Lesson 4: Use and Share Summary-Level Association Data

GWAS summary statistics are routinely used as input data to infer estimates of genetic overlap, which quantifies the level of overlapping genetic biology shared between two traits. Many studies focus on genetic overlap to better understand comorbidity or disease risk factors. Based on estimates of genetic overlap, more advanced statistical approaches model relationships between traits at the level of their underlying genetic architecture, allowing to test specific theories about the shared biology between traits (e.g., Genomic SEM (Grotzinger et al., 2019), Genomic ICA (Soheili-Nezhad et al., 2021), Genomic PCA (Fürtjes et al., 2021)). Many more methods build on GWAS summary data, uncovering biologically interpretable mechanisms, for example, by linking them with gene expression or cell type profiles (de Leeuw et al., 2015).

Adopting practices that encourage collaboration, and meta-analysis also greatly benefits neuroimaging research. Just like geneticists share GWAS summary statistics, neuroimagers should calculate trait associations for all vertices across the brain and share it publicly. Meaningful summary data will require great, consortium-level efforts to reduce noise and (scanner) bias. Inspired by practices surrounding GWAS, vertex-wise association data may be used to infer brain-based etiology shared between traits (parallel to genetic overlap), or to uncover underlying biological mechanisms by mapping association data onto brain-specific gene expression (Shen et al., 2012) and neurotransmitter systems (Hansen et al., 2022).

Lesson 5: Use Multivariate Approaches That Were Originally Developed for Genetics Research

Alternative multivariate techniques exist that simultaneously map thousands of biological markers onto a trait or disease. Multivariate techniques do not require extensive multiple testing correction, and they therefore have more statistical power than mass-univariate methods. For example, the genetics technique *genome-wide complex trait analysis* (GCTA) (Yang et al., 2011) is ubiquitously used to estimate heritability. Implemented in efficient software, the GCTA framework employs linear mixed models fitting millions of variables as a vector of random effects, to quantify trait variance accounted for by genome-wide markers (i.e., heritability), while recognising the correlation structure between them. Recent neuroimaging studies repurposed GCTA which enabled the estimation of *morphometricity*, which is the trait variance explained by brain-wide measures (Couvy-Duchesne et al., 2020).

All traits are heritable (Turkheimer, 2000), and given heritability and morphometricity have an analogous statistical definition, it is unsurprising that most traits are also considerably morphometric (Couvy-Duchesne et al., 2020). Our recent study applied the GCTA framework to neuroimaging data, and compared the variance accounted for by ~300,000 cortical measures, with variance accounted for by coarser brain atlases (Fürtjes et al., 2023). It demonstrated that atlas-based representations of the cortex explained a fraction of the morphometricity that was explained by vertex-wise measures, which highlights that considering brain-wide vertex-wise measures maximises the potential of uncovering neuronal underpinnings of traits and diseases. Like in candidate gene approaches, coarse representations of the cortex do not reliably enough account for trait variance.

Critics may argue that modelling vertex-wise data – using genetics frameworks – would disregard the decade's worth of brain sciences that derived brain atlases to help interpret brain-trait associations. To facilitate more biologically meaningful interpretation, Couvy-Duchesne et al. (Couvy-Duchesne et al., 2020) demonstrate that the GCTA framework permits integrating prior knowledge about brain organisation by grouping vertex-wise measures based on the researchers input, and fitting each set of vertices as random effects. This analysis has the advantage that it still models vertex-wise cortical structure, while it drastically reduces multiple testing burden compared with mass-univariate testing, as it only performs a single association test per set of vertices. I suggest this framework has the potential and flexibility to fully replace ROI-based studies with robust vertex-wise approaches.

Limitations

It must be noted, however, that the discussion above only applies to studies researching traits that commonly vary across the general population. Depression is a prominent example, as it affects about 15% of people at some point in their lives (Bromet et al., 2011). It is precisely this variance across the population that the methods discussed above leverage to draw inference. Those methods would be inappropriate to model rarer monogenic traits. Huntington's, for example, only affects 1 in 10,000 people, and was linked to one single gene coding for a protein called *huntingtin* (Coleman et al., 2021), which would not map onto models of polygenicity. Nonetheless, parallels between the fields hold as monogenic traits in genetics research are analogous to lesion-based neuroimaging studies. The latter link localised brain lesions (that often result from rare accidents) with very specific loss of cognitive function (e.g., Scoville & Milner, 1957). Both monogenic and lesion-based correlates

are rare, they both have large effect sizes, and small clinical samples are sufficient to detect them.

A meaningful application of genetic methods to neuroimaging data must consider differences in data structures, which dictate the interpretation of results in their genetic- or brain-specific context. Primary among these differences is that genetic markers are inherited at conception and remain unaltered across the lifespan, while the brain evolves with its environment. Thus, genetic propensity towards a trait can imply directionality of effects, which cannot be inferred from neuroimaging studies. Genetic and neuroimaging data both contain interdependent measures, but the architecture of this interdependence is different, which may affect techniques deriving genetic or brain-based overlap. It complicates interpretation of genetic studies that trait correlates often sit in parts of the genome with complicated regulatory, and no direct coding functions (Visscher et al., 2017). In comparison, the interpretation of vertex-wise brain associations is trivial, as the strongest associations are between vertices in physical proximity.

Conclusion

Population traits have both complex genetic and brain-based biology, and here I argue that practices firmly established in genetics research can be directly applied to improving neuroimaging. Based on striking parallels between the two fields, this essay transferred lessons drawn from the field of Statistical Genetics to Neuroimaging, which promises more robust discoveries if widely endorsed by the neuroimaging community. The failure to produce reliable findings, by both candidate gene and neuroimaging studies mapping individual brain regions onto population traits, illustrate that future efforts should keep increasing sample sizes, and counteract noisy

findings by correcting for multiple testing. Genetics research teaches that we can improve replicability by abandoning hypothesis-driven, overly simplistic approaches, and by adopting hypothesis-free methods exemplified by GWAS and GCTA. Sharing summary-level association data will boost large-scale collaboration and meta-analysis. Those genetic practices, applied to vertex-wise neuroimaging data, promise an acceleration of the evolution of the field, without requiring brain sciences to painstakingly innovate neuroimaging methods, that would overcome the same challenges that Statistical Genetics already solved.