Neuropsychol Rev
DOI 10.1007/s11065-015-9281-1

REVIEW

Genetics and Brain Morphology

Lachlan T. Strike · Baptiste Couvy-Duchesne · Narelle K. Hansell · Gabriel Cuellar-Partida · Sarah E. Medland · Margaret J. Wright

Received: 31 October 2014 / Accepted: 8 February 2015

Abstract A wealth of empirical evidence is accumulating on the genetic mediation of brain structure phenotypes. This comes from twin studies that assess heritability and genetic covariance between traits, candidate gene associations, and genome-wide association studies (GWAS) that can identify specific genetic variants. Here we review the major findings from each of these approaches and consider how they inform on the genetic architecture of brain structure. The findings from twin studies show there is a strong genetic influence (heritability) on brain structure, and overlap of genetic effects (pleiotropy) between structures, and between structure and cognition. However, there is also evidence for genetic specificity, with distinct genetic effects across some brain regions. Candidate gene associations show little convergence; most have been under powered to detect effect sizes of the magnitude now expected. GWAS have identified 19 genetic variants for brain structure, though no replicated associations account for more than 1 % of the variance. Together these studies are revealing new insights into the genetic architecture of brain morphology. As the scope of inquiry broadens, including measures that capture the complexity of the brain, along with larger samples and new analyses, such as genome-wide common trait analysis (GCTA) and polygenic scores, which combine variant effects for a phenotype, as well as whole-genome sequencing, more genetic variants for brain structure will be identified. Increasingly, large-scale multi-site studies will facilitate this next wave of studies, and promise to enhance our understanding of the etiology of variation in brain morphology, as well as brain disorders.

Keywords Magnetic resonance imaging · Brain · Genes · Twins

LS, BCD and NH contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s11065-015-9281-1) contains supplementary material, which is available to authorized users.

L. T. Strike (✉) · B. Couvy-Duchesne · N. K. Hansell · M. J. Wright
Neuroimaging Genetics, QIMR Berghofer Medical Research Institute, Brisbane, QLD 4006, Australia
e-mail: lachlan.strike@qimrberghofer.edu.au

L. T. Strike · B. Couvy-Duchesne · M. J. Wright
School of Psychology, University of Queensland, Brisbane, QLD 4072, Australia

L. T. Strike · B. Couvy-Duchesne
Centre for Advanced Imaging, University of Queensland, Brisbane, QLD 4072, Australia

G. Cuellar-Partida
Statistical Genetics, QIMR Berghofer Medical Research Institute, Brisbane, QLD 4006, Australia

S. E. Medland
Quantitative Genetics, QIMR Berghofer Medical Research Institute, Brisbane, QLD 4006, Australia

Introduction

Imaging genetics is a rapidly emerging field integrating imaging and genetic approaches to better understand the aetiology of normal brain variation, and, especially mental illness. As magnetic resonance imaging (MRI) measures of brain morphology are thought to be closer to the biology of genetic function than a behavioural phenotype or a disease syndrome itself, they are often used as endophenotypes to provide an objective tool to search for genetic variability in the brain (Gottesman and Gould 2003; Glahn et al. 2007). To be the

Published online: 14 March 2015

most useful, endophenotypes must meet certain criteria (Gottesman and Shields 1967; Gottesman and Gould 2003), including that the endophenotype is heritable. Endophenotypes have been hypothesized to have a simpler genetic architecture (Gottesman and Gould 2003), or reduced genetic complexity (de Geus 2010) than other complex traits and have higher penetrance for genetic variants because they are biologically closer to the genes. However, there is a shift in current perspective such that the importance of endophenotypes is not in gene discovery itself, but rather, in the critical insights they will likely provide into the neurobiology underlying neurodegenerative, developmental, and neuropsychiatric conditions (Iacono et al. 2014).

Here we review imaging genetic studies focussing on brain morphology, including i) twin studies that assess not only the heritability of brain structure, but, importantly, the genetic covariance between different brain structure phenotypes (e.g. brain regions), ii) candidate gene association studies, which are hypothesis driven, and iii) genome-wide association studies (GWAS) that can identify specific genetic variants and provide insight into the biological mechanisms influencing the structure of the brain. We review the major findings from each of these approaches, and consider how they inform on the genetic architecture of brain structure. We also provide a brief overview of the methodology, highlighting important design considerations and central themes that emerge from the work, which are relevant to future research.

Twin and Family Imaging Studies

Twin and family studies provide a powerful method to establish the heritability of brain imaging phenotypes (see below for a brief overview of how twin studies estimate heritability). The very first study examined total brain volume and included ten identical (monozygotic (MZ)) and nine non-identical (dizygotic (DZ)) twin pairs (Bartley et al. 1997). Today there are several large imaging cohorts for which structural scans have been acquired in over 100 pairs of twins (200 individuals) or multiple members from the one family. These studies include imaging cohorts from North America (CLDRC, GOBS/SAFHS, FHS, NIMH/PTS, NHLBI, QNTS, UNC, VETSA), Europe and the UK (FNTR, TEDS, NTR/UMCU) and Australia (OATS, QTIM, TWIN-E) (Table 1). Together, the cohorts cover a wide age range. Some cohorts themselves cover much of the lifespan (FHS, GOBS/SAFHS, NTR/UMCU, TWIN-E), whereas others are more restricted with respect to age, focussing on either neonates and children (NIMH/PTS, QNTS, TEDS, UNC), adolescents and young adults (CLDRC, QTIM), middle-aged adults (FNTR, VETSA) or the elderly (NHLBI, OATS). Encouragingly, for a number of the cohorts (NIMH/PTS, NHLBI, NTR/UMCU, OATS, QTIM) imaging data are being acquired at more than one time point using the same imaging protocol. Longitudinal studies in genetically informative samples are vital to characterise the extent that changes in brain structure during the life course are due to genes and/or the environment. As can be seen in Table 1, for almost all cohorts global and regional brain volumes have been extracted, with more recent efforts examining cortical topography and white matter integrity. For many cohorts there is also extensive phenotyping on a wide range of measures, including cognition, as well as resting state and/or task based functional magnetic resonance imaging (fMRI) (Glahn et al. 2010; Koten et al. 2009; Blokland et al. 2011).

Estimating Heritability and Multivariate Genetic Analysis

The pattern of MZ and DZ twin correlations provides a first indication of whether there is a genetic contribution to brain morphology. Increased similarity between MZ twins compared to DZ suggests that variation in brain structure is influenced by additive genetic factors, while increased similarity between DZ twins (DZ twin correlation is more than half the MZ twin correlation) suggests an effect of common (i.e. shared) environment. Non-additive genetic effects will further increase the degree of similarity between MZ twins compared to DZ, though large samples are required to detect non-additive genetic effects (Martin et al. 1978). As well as examining twin correlations most studies use standard ACE model fitting analysis in Mx (Neale and Cardon 1992) to partition the variance in a structure into additive genetic (A), common environmental (C) and unique or non-shared environmental (E) sources. Variation due to dominant genetic (D) effects can also be partitioned; though not simultaneously with C in samples where twins are reared together (C and D are confounded in this instance). See Evans et al. (2002); Neale and Cardon (1992) and Verweij et al. (2012) for excellent reviews of the twin method and genetic analyses.

The twin design can be extended to include additional family members such as siblings, estimate the effects of covariates such as age to improve the fit of the model, allow multiple phenotypes to be studied simultaneously or to analyse more complex genetic and environmental influences (e.g. interaction terms such as gene-environment correlation). By including multiple phenotypes, as shown in Fig. 1, it is possible to decompose the variance both within and between variables into genetic and environmental components. From this, genetic, environmental and phenotypic correlations can be calculated, with the genetic correlation indicating the extent to which two different structural measures share genetic influence. For example, a genetic correlation of 1 indicates two traits are influenced completely by the same latent genetic factor(s). Several studies have examined the genetic overlap among different brain measures, and a few have investigated whether there is any genetic overlap across brain structure and cognition. We review these studies in the sections below.

Table 1 Twin and family cohorts with imaging data

Cohort	Reference	*n* pairs MZ/DZ[a]	Age range (years)	Structural imaging Phenotypes[b]	Cognitive phenotypes
Twin					
Colorado Learning Disabilities Research Centre (CLDRC)	Betjemann et al. (2010)	41/30	12–24	Global, regional	General cognitive ability, reading
Finnish National Twin Registry (FNTR)	Thompson et al. (2001)	10/10	Mean = 48 ± 3	Voxel-wise morphometry	General cognitive ability
National Heart, Lung, and Blood Institute (NHLBI)	Pfefferbaum et al. (2000)	72/70	69–80	Global, regional, subcortical, white matter hyperintensities	General cognitive ability
National Institute of Mental Health (NIMH)/Pediatric Twin Study (PTS)	Schmitt et al. (2007)	127/36	6–19	Global, regional, subcortical, cortical thickness	General cognitive ability
Netherlands Twin Registry (NTR)/University Medical Centre Utrecht (UMCU)	Brans et al. (2010)	102/131	9–69	Global, regional, subcortical, cortical thickness & surface area, white matter integrity, voxel-wise morphometry.	General cognitive ability, working memory
Older Adult Twins Study (OATS)	Sachdev et al. (2009)	77/41	65–88	Global, regional, subcortical	General cognitive ability, memory, language
Quebec Newborn Twin Study (QNTS)	Yoon et al. (2010)	57/35	Mean = 8	Global, regional, subcortical, cortical thickness, voxel-wise morphometry	General cognitive ability
Queensland Twin IMaging Study (QTIM)	Blokland et al. (2014)	148/202	12–30	Global, regional, subcortical, cortical thickness & surface area, white matter integrity, voxel-wise morphometry	General cognitive ability, reading
Twins Early Development Study (TEDS)	Rijsdijk et al. (2010)	31/35	Mean = 9	Voxel-wise morphometry	General cognitive ability, reading, mathematics
The Twin study in Wellbeing using Integrative Neuroscience of Emotion (TWIN-E)[b]	Gatt et al. (2012)	122 pairs	18–65	Global, regional, cortical thickness & surface area, white matter integrity	General cognitive ability
University of North Carolina (UNC)	Gilmore et al. (2010)	41/50	0–1 week	Global, regional, subcortical, white matter integrity	
Vietnam Era Twin Study of Aging (VETSA)	Kremen et al. (2013)	110/92	51–59	Global, regional, cortical thickness & surface area, voxel-wise morphometry	General cognitive ability, working memory
Family/pedigree					
Framingham Heart Study (FHS)	DeStefano et al. (2009)	1538 subjects	34–97	Global, regional, subcortical	Verbal memory, visuospatial memory
Genetics of Brain Structure and Function (GOBS)/San Antonio Family Heart Study (SAFHS)	Kochunov et al. (2011)	1129 subjects	19–85	Global, regional, subcortical, cortical thickness & surface area, white matter integrity, voxel-wise morphometry	General cognitive ability, working memory, declarative memory, language processing, emotional processing

[a] Largest published complete pairs

[b] sample size and age range are projected figures. Several cohorts function as combinations (NIMH/PTS, NTR/UMCU) or subsamples (GOBS/SAFHS). Cohorts with small sample sizes (<20 pairs) are not included.

DZ dizygotic, *MZ* monozygotic

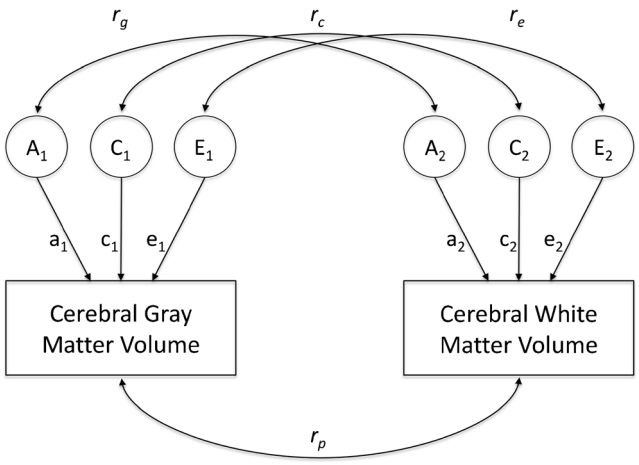


Fig. 1 Correlated factors bivariate twin model. *a*, *c* and *e* are the latent additive genetic, common environment and unique environment factors respectively, which influence cerebral gray and white matter volume. r_g, r_c and r_e are the genetic, common environmental and unique environmental correlations respectively, with r_p the phenotypic correlation between the two observed traits. *a*, *c* and *e* represent parameter estimates, conceptualised as the strength of the latent factor on the phenotype

Besides twin studies, heritability can also be estimated from family pedigrees (Kochunov et al. 2009). More recently it has become possible to estimate heritability directly from DNA (Yang et al. 2011). This method, often called GCTA (Genome-wide Complex Trait Analysis), estimates the genetic influence (heritability) using genome-wide genotypes from very large samples (i.e. thousands) of unrelated individuals, rather than comparing MZ and DZ twins. GCTA detects only the genetic effects that are tagged by common single nucleotide polymorphisms (SNPs) (allele frequencies greater than 1 %) available on a DNA array used in GWAS. Thus GCTA heritability (or SNP heritability) is not expected to be as large as the heritability estimated from twin studies that can detect genetic influences due to DNA variants of any kind (Vinkhuyzen et al. 2013). GCTA can also be extended to multivariate analysis to estimate the SNP heritability of each phenotype and the SNP-correlation between phenotypes (Wray et al. 2014). To date only a few studies have estimated SNP heritability of brain imaging phenotypes (Bryant et al. 2013; Adib-Samii et al. 2015), mainly because very large samples are necessary, requiring genotypes and phenotypes to be shared among several research groups (which is not always possible). We discuss these results later.

Heritability of Brain Structure

The growing number of imaging studies in twins and families provides strong evidence that individual differences in brain structure are due to both genetic factors and the environment. Fig. 2 shows heritability estimates for a range of measures, including global and regional brain volumes, subcortical volumes, cortical thickness and white matter integrity. These results are from a recent meta- and mega-analysis (Blokland et al. 2012; Kochunov et al. 2014), or a new meta-analysis, including data that has accumulated since 2012. Heritability estimates range from moderate to high i.e. 40 % to >80 %. As noted previously (Blokland et al. 2012), lower heritability estimates are generally found for smaller brain structures relative to larger global or lobar structures. However, whether one structure has a higher heritability than the other needs to be interpreted with care. As seen in Fig. 2, confidence intervals for heritability estimates overlap for many structures.

Notably, few studies have examined whether the heritability of imaging phenotypes is a function of gender (Chiang et al. 2011). By dividing twin pairs into five groups (MZ male, MZ female, DZ male, DZ female, DZ opposite sex pairs), sex limitation modelling can estimate both qualitative and quantitative differences between males and females (i.e. is heritability larger in one sex than in the other (same genetic source, different magnitude), or whether a factor (genetic or environmental) effects one sex but not the other). As different neurodevelopmental trajectories for males and females have been reported (Dennison et al. 2013; Koolschijn and Crone 2013; Lenroot and Giedd 2010), and as the majority of the cohorts include both opposite and same sex pairs, these studies will be well placed to examine sex-dependent effects as their samples increase. Additionally, throughout development into old age, the structure of the brain changes over the life course. While cross sectional studies have demonstrated heritability changes associated with age (Lenroot et al. 2009; Wallace et al. 2006), longitudinal twin studies (Schmitt et al. 2014; van Soelen et al. 2012) that are well powered (large samples) are needed to detect changes in heritability for different brain structures as a function of age. Recently, some progress was made when several groups with longitudinal imaging data on twins worked together as part of the ENIGMA Consortium to estimate the heritability of subcortical brain plasticity (Brouwer et al. 2015).

Though heritability tells us nothing about the underlying genes or the number of genes involved or their effect size, the magnitude of the heritability estimate does provide some indication of the statistical power for discovering the causal genes of a trait (Bochud et al. 2012). For example, if the heritability estimates of several brain structures are available, the structure with the highest heritability estimate can be chosen for genome-wide association studies (e.g. Stein et al. (2012) selected hippocampus and intracranial volume on this basis). A general benchmark is to consider heritability estimates below 20 % as low, those between 20 and 50 % as moderate, and estimates above 50 % as high. Of the 67 brain structure phenotypes in Fig. 2, two structures (bilateral lateral ventricles, corticospinal tract) are moderately heritable, with all others highly heritable. Nonetheless, a high heritability estimate does not imply that genetic variants associated with

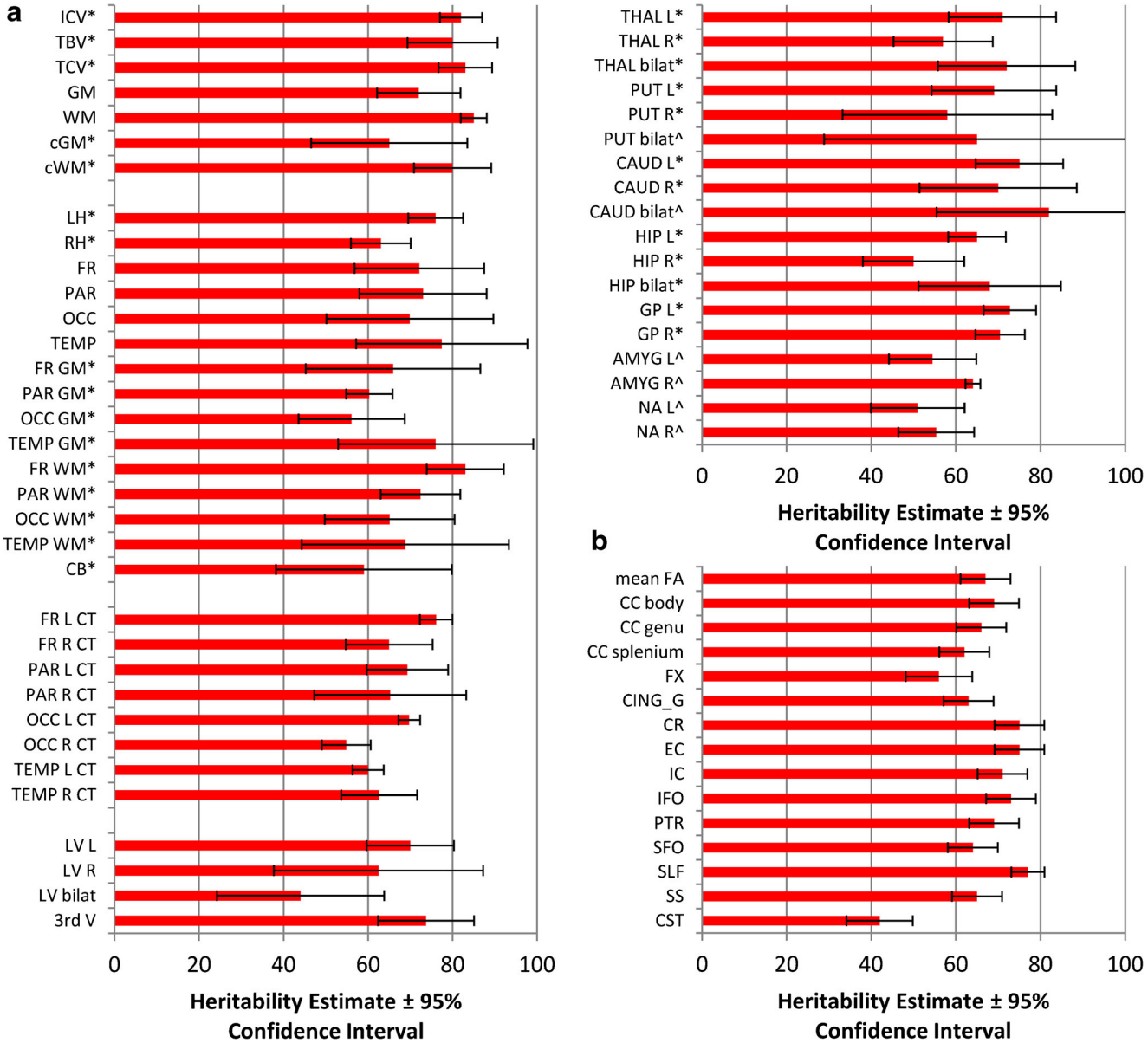


Fig. 2 Additive genetic estimates and 95 % confidence intervals for **a** volumetric and cortical thickness and **b** fractional anisotropy (FA) white matter integrity. **a** Meta-analysis combining estimates from Blokland et al. (2012) and 3 additional studies, some of which include new imaging phenotypes not included in Blokland et al. 2012; * denotes revised heritability estimate, and ^ denotes new phenotype. The meta-analyses including the additional studies (Batouli et al. 2014; den Braber et al. 2013; Rentería et al. 2014) used the same methodology as Blokland et al. (2012). Full ACE estimates available in Online Resource 1. **b** FA white matter mega-analysis estimates (Kochunov et al. 2014) publically available from http://enigma.ini.usc.edu/ongoing/dti-working-group/2014_nimg/. A harmonisation protocol combined DTI for 2248 participants from five cohorts followed by a mega-analysis to estimate heritability. *3rd* third ventricle*; AMYG* amygdala; *bilat* bilateral; *CB* cerebellum; *CAUD* caudate; *CC* corpus callosum; *CING_G* cingulum gyrus; *cGM* cerebral gray matter; *CST* corticospinal tract; *CR* corona radiate; *CT* cortical thickness; *cWM* cerebral white matter; *EC* external capsule; *FA* fractional anisotropy; *FR* frontal; *FX* fornix; *GM* gray matter; *GP* globus pallidus; *HIP* hippocampus; *IC* internal capsule; *ICV* intracranial volume; *IFO* inferior fronto-occipital fasciculus; *L* left; *LH* left hemisphere; *LV* lateral ventricle; *NA* nucleus accumbens; *OCC* occipital; *PAR* parietal; *PTR* posterior thalamic radiation; *PUT* putamen; *R* right; *RH* right hemisphere; *SFO* superior fronto-occipital fasciculus; *SLF* superior longitudinal fasciculus; *SS* sagittal stratum; *TBV* total brain volume; *TCV* total cerebral volume; *TEMP* temporal; *THAL* thalamus; *WM* white matter

the brain structure will have a large effect on the phenotype (Visscher et al. 2008). Limitations of heritability estimation include a lack of information on the **mode of inheritance** of the trait, and the possibility that heritability estimates may vary across populations or with time. This may be especially relevant to child and adolescent samples, with Schmitt et al. (2014) reporting increases in cortical thickness heritability estimates during the first two decades of life. The finding of a similar genetic architecture behind subcortical structures in young (Rentería et al. (2014) and middle-aged (Eyler et al.

2011) adults possibly suggests greater stability in heritability estimates following adolescence. In samples with a wide age range there may be cohort effects, which may confound the "equal environment" within the sample population (Bochud et al. 2012).

The pervasiveness of a genetic contribution to brain structural measures does not mean that environmental influences are not important. Both genetic and environmental (non-genetic) factors contribute to the variance in brain structure. Non-genetic factors include identifiable environmental factors and measurement error, though both are not well assessed. However, recently collaborative studies such as IMAG EMEND (Frangou 2014) have begun collecting information on environmental factors as well as genetic data. This will facilitate considerations of the interplay between genes and the environment in the development of the brain, brain ageing, as well as psychiatric disorders.

Issues in Estimating Heritability For Imaging Phenotypes

Sample size, namely whether it is large enough to provide sufficient power to detect genetic variance (heritability) is an important consideration. Generally, the global brain measures in Fig. 2 are available on large samples; the smaller confidence intervals for the heritability estimates reflect this increased power. Many of the early imaging studies in twins were restricted by small sample size (Bartley et al. 1997; Pennington et al. 2000; Wright et al. 2002) so the degree to which the results can be generalized to larger populations is somewhat limited. The high cost of brain imaging and twin recruitment are ever present factors impacting sample size for imaging genetics studies. The mega-analysis approach undertaken by Kochunov et al. (2014) is an engaging future research direction for addressing such concerns. In this method, imaging data from multiple sites are pooled together and analysed as one data set, ensuring homogeneity of processing. Though the sharing of data across research groups is not always possible, this technique allows for large sample sizes to be amassed and removes one level in which error variance may be introduced (measurement variability).

Another issue is that some brain structures are more easily measured than others. In assessing or ranking heritability estimates, any differences in measurement error need to be considered. For example, the lower heritability estimates for smaller brain structures (e.g. nucleus accumbens) may be due to increased measurement error, as reflected in lower test-retest reliability (Morey et al. 2010; Nugent et al. 2013; Rentería et al. 2014). Measurement error artificially reduces the similarity between MZ twins and thereby reduces heritability estimates (i.e. reliability sets the upper limit for heritability by constraining the maximum possible correlation between MZ twins). How well a brain region or white matter tract can be reliably measured is often overlooked, but in the absence of test-retest reliability statistics it is difficult to determine if low heritability estimates are valid or simply reflect measurement error (Kuntsi et al. 2006). If test-retest statistics were provided alongside heritability estimates (e.g. as per Rentería et al. (2014)), any differences in the genetic contribution across measures could be better evaluated. For example, as the intra-class correlation (ICC) is conceptually analogous to the r^2 obtained in regression analysis (Rousson et al. 2002), it is possible to make a direct comparison between the portion of variability not explained between the repeats ($1 - r^2$ test-retest or 1 – ICC test-retest) and e^2 (variance due to non-shared environment and measurement error). Where e^2 exceeds unreliability ($1 - r^2$ test-retest), non-shared environmental influences are greater than measurement error. Conversely, measurement error is greater than non-shared environment when unreliability ($1 - r^2$ test-retest) is greater than e^2. See Online Resource 2 for an example. Further, the accuracy of the heritability estimate can be increased by modelling unreliability (Kuntsi et al. 2006; Luciano et al. 2001), though this can result in an overestimation of the genetic effects (Luciano et al. 2001). Thus it is better to improve measurement of a phenotype rather than statistically adjust for measurement error (Kuntsi et al. 2006). For genetic association studies to build upon the findings of heritability estimates, imaging phenotypes, and the resulting heritability estimates, should be as reliable as possible.

A further consideration, especially when comparing or meta-analyzing heritability estimates, is the use of different software to segment and measure the brain (Gronenschild et al. 2012; Morey et al. 2009), as well as the MRI platform or field strength used for scanning (Jovicich et al. 2009; Han et al. 2006). Variation in measurement can arise through differences in software methodologies (Freesurfer, FSL, SPM), but also differences within the same analysis package (i.e. different versions of the same software). For instance, Rentería et al. (2014) reported test-retest Pearson's correlations of 0.51 and 0.52 for mean nucleus accumbens and amygdala volume respectively, extracted through Freesurfer 5.1. Correlations from the same sample extracted through Freesurfer 5.3, and with improved intensity inhomogeneity correction, increased to 0.68 and 0.77 respectively. Recently, standardised imaging and analysis protocols have been adopted for multi-site analyses (e.g. as implemented by the ENIGMA consortium) to address this issue and enhance phenotypic heritability.

Are there Overlapping or Distinct Genetic Factors Contributing to Brain Structure?

While the brain can be segmented into many regions and quantified in different ways, several brain structure phenotypes are highly correlated. For instance, individuals with a large gray matter volume also have a large white matter

volume (Baare et al. 2001). Several twin studies have investigated whether different brain structure phenotypes are genetically correlated (Table 2). A number of these show high genetic correlations, such as between intracranial and thalamus volume (Rentería et al. 2014), cortical gray and white matter volume (Baare et al. 2001), gray matter volume and surface area (Winkler et al. 2010) and longitudinal change in total cerebral and cerebellum volume (van Soelen et al. 2013). The average phenotypic correlation for these studies is .60. The genetic correlations are slightly higher, ranging from 0.68 to 0.88, indicating an overlapping set of genes are responsible for the heritability of these brain structure phenotypes. Genetic correlations estimate the extent that genetic effects on one phenotype are correlated with the genetic effects on another phenotype, independent of their heritability. Overlapping genetic effects provide evidence for **pleiotropy** (each gene affects many traits) across the brain, and this has mostly been demonstrated between regional brain volumes and/or global measures.

Examining cortical thickness in 54 cortical sub-regions, Schmitt et al. (2008) identified a single factor accounting for over 60 % of the genetic variability in cortical thickness measurements. Using the same sample they also reported a strong genetic factor influencing variation between lobar volumes (Schmitt et al. 2010). Interestingly, in both studies, after correcting for mean cortical thickness or total brain volume, regionally specific patterns of genetic influence emerged, showing that there is genetic specificity, with distinct genetic effects across brain regions. Examining cortical thickness, Chen et al. (2013) reported distinct clusters of genetic influence throughout the cortex, i.e. different sets of genes influenced cortical thickness in different areas of the cortex. The structuring of these clusters appeared to reflect differences in maturation timing, and also the division between primary and association cortex. However, in an older sample, genetic effects on lobar volumes tended to be general rather than specific (Batouli et al. 2014). Differences in sample age and methodology across studies may account for these findings.

The low phenotypic and genetic correlations found for some of the studies shown in Table 2 also suggest that a common genetic influence does not extend to all areas of the brain, or two measures of the same structure. For example, Panizzon et al. (2009) and Winkler et al. (2010) found that two standard measures of gray matter (cortical thickness and surface area), shared little genetic overlap. Further, Rentería et al. (2014) identified region specific genetic factors explaining over 50 % of the heritability in hippocampus, caudate nucleus, amygdala and nucleus accumbens volumetric measures. These findings confirmed those of a prior study (Eyler et al. 2011), and demonstrate a similar genetic framework for subcortical structures across young and older adults.

Clearly, further work is needed to determine the pattern of genetic effects on brain structure phenotypes. A limited range of brain structure phenotypes have been examined and we have little understanding of whether the genetic covariance across brain structure phenotypes changes throughout the lifespan, or how such genetic organization reflects brain function.

Shared Genetic Influences Across Brain Structure and Cognitive Ability

Both brain structure (Blokland et al. 2012; Kremen et al. 2010) and cognitive ability (Deary et al. 2010; van Soelen et al. 2011) show substantial individual variability, much of which appears to be influenced by genetic factors. Multivariate studies have revealed that, to some degree, these genetic factors overlap. Table 3 displays phenotypic, genetic, and environmental correlations between measures of brain morphology and intelligence. Notably, the phenotypic correlations peak at approximately ±.30, thus brain structure is at best moderately associated with cognitive ability. However, genetic and environmental influences can sometimes have opposing effects that cancel each other out (Brans et al. 2010). Alternatively, both genetic and environmental factors can have the same direction of influence and thus enhance the phenotypic correlation (Brouwer et al. 2014b), and in some cases, the association is largely due to a common genetic source (Brans et al. 2010). The majority of studies to-date have focused on total measures of gray and white matter tissue (Posthuma et al. 2002, 2003; Betjemann et al. 2010; van Leeuwen et al. 2009). Interesting, Vuoksimaa et al. (2014) recently demonstrated genetic associations between cortical volume and cognitive ability to be the result of variation in cortical surface area, but not cortical thickness. White matter tracts have also shown phenotypic and genetic overlap with cognition (Karlsgodt et al. 2010; Chiang et al. 2009), possibly illustrating biological networks underpinning intelligence. Overall, the low phenotypic correlations reported by studies signify that while there does appear to be pleiotropy across brain structure and function, overlapping genes have only a small association with intelligence. Future studies focusing on the neural circuitry associated with cognitive ability (as opposed to global or anatomical brain divisions) may be able to tap into a greater phenotypic relationship between brain morphology and function, thereby identifying genetic factors with stronger genetic influences on cognition. Further, the relative importance of genetic and environmental factors may vary across the lifespan, with periods such as adolescence being of particular interest (Lenroot and Giedd 2008) and a potential target for future work.

Examining associations between brain development trajectories and whether this predicts cognitive performance, and whether shared genetic factors underlie such an association (and if this represents causality) are important future research questions. Additionally as studies have shown environmental

Table 2 Shared genetic influence across brain structures

Reference	Cohort	*n* pairs MZ/DZ (age in years)	Phenotypes[a]	r_p	r_g	r_e
Posthuma et al. (2000)	NTR/UMCU	53/58 (19–69)	ICV & CB	–	.57	.44
Pfefferbaum et al. (2000)	NHLBI	45/40 (68–78)	CC msa height & LV bilat	.66	.68	.58
Baare et al. (2001)	NTR/UMCU	54/58 (19–69)	cGM & cWM	.59	.68	.03
			ICV & cWM	.81	.83	.66
			ICV & cGM	.84	.90	.49
			ICV & TBV	.93	.95	.79
Schmitt et al. (2007)	NIMH	127/36 (5–18)	(Lowest r_g) LV bilat & THAL bilat	–	0	−.22
			(Highest r_g) TCV & THAL bilat	–	.97	.35
DeStefano et al. (2009)	FHS	1538 subjects (34–97)	(Lowest r_g) OCC lobe & LV bilat	−.01	.01	.01
			(Highest r_g) TEMP lobe & HIP bilat	.38	.76	−.11
Panizzon et al. (2009)	VETSA	110/92 (51–59)	cGM CT & SA	.01	.08	−.13
			(Lowest r_g) PCC R CT & SA	−.06	.01	−.03
			(Highest r_g) lat OCC R CT & SA	−.03	.88	−.11
Kochunov et al. (2009)	SAFHS	459 subjects (19/85)	Subcortical WMHI & ependymal WMHI	–	.46	.07
Kochunov et al. (2010)	SAFHS	467 subjects (19–85)	FA mean & $L_\perp$ mean	–	−.68	–
			FA mean & $L_\parallel$ mean	–	−.70	–
			$L_\perp$ mean & $L_\parallel$ mean	–	.95	–
Rogers et al. (2010)	GOBS	242 subjects (19–85)	SA mean & GI mean	–	−.60	–
			C mean & GI mean	–	-.73	–
Schmitt et al. (2010)	NIMH	127/36 (5–18)	(Lowest r_g) PAR GM & TEMP WM	–	.56	−.10
			(Highest r_g) FR GM & TEMP GM	–	.94	.86
Winkler et al. (2010)	GOBS	486 subjects (26–85)	(Lowest r_g) PCUN GM & PCUN CT	.10	.03	.16
			cGM & SA	.85	.85	.85
			(Highest r_g) TEMP pole GM & TEMP pole SA	.81	1	.76
Eyler et al. (2011)	VETSA	110/92 (51–59)	(Lowest r_g) THAL L & LV L	–	0	–
			THAL L & HIP R	–	.51	–
			(Highest r_g) CAUD L & CAUD R	–	1	–
Panizzon et al. (2012)	VETSA	130/97 (51–60)	(Lowest r_g) PCUN L WM/GM contrast & PCUN L CT	.15 −.54	−.01 −.89	.33 −.34
			(Highest r_g) PCAL R WM/GM contrast & PCAL R CT			
van Soelen et al. (2012)	NTR/UMCU	23/28 (T1 Mean = 9) (T2 Mean = 12)	(Lowest r_g) $^{\Delta}$ in FR R CT & $^{\Delta}$ in sup TEMP L CT over 3 years	0 .30	−.01 1	– –
			(Highest r_g) $^{\Delta}$ in inf PAR L CT & $^{\Delta}$ in inf FR L CT$^{\Delta}$ over 3 years			
van Soelen et al. (2013)	NTR/UMCU	23/28 (T1 Mean = 9) (T2 Mean = 12)	$^{\Delta}$ in TCV & $^{\Delta}$ in height over 3 years	.09	.39	−.13
			$^{\Delta}$ in CB & $^{\Delta}$ in height over 3 years	.24	.48	−.05
			$^{\Delta}$ in TCV & $^{\Delta}$ in CB over 3 years	.49	.88	.34
Batouli et al. (2014)	OATS	77/41 (65–88)	(Lowest r_g) cGM & cWM	–	.17	.56
			(Highest r_g) TEMP & TCV	–	.95	.9
Rentería et al. (2014)	QTIMS	148/202 (16–29)	(Lowest r_g) CAUD mean & AMYG mean	.10 .76	.13 .86	– –
			(Highest r_g) ICV & THAL mean			

[a] These phenotypes are volumes unless otherwise stated. Where results for a large number of structures are reported, the regions with lowest and highest genetic correlations are shown. Multivariate studies were identified through a PubMed search of the following terms: *twin - brain - imaging - MRI - genetic*. With a cut-off date of 1 July 2014, 27 studies were identified, with an additional 18 studies identified manually through references. Those that could be quantified are summarised in Tables 2 and 3, with those that could not be quantified (Chen et al. 2011, 2012, 2013; Chiang et al. 2009, 2012; Giedd et al. 2007; Schmitt et al. 2008, 2009; Wallace et al. 2010; Wright et al. 2002; Rimol et al. 2010; Eyler et al. 2011; Jahanshad et al. 2013a) consistent with tabled findings

AMYG amygdala, *bilat* bilateral, *C* cerebral; *CAUD* caudate, *CB* cerebellum, *CC* corpus callosum, *cGM* cerebral gray matter, *cWM* cerebral white matter, *CT* cortical thickness, *FA* fractional anisotropy, *FR* frontal, *GI* gyrification index, *GM* gray matter, *HIP* hippocampus, *inf* inferior, *ICV* intracranial volume, *L* left, *lat* lateral, $L_{||}$ axial diffusivity, $L_\perp$ radial diffusivity, *LV* lateral ventricle, *msa* midsagittal area, *OCC* occipital, *PAR* parietal, *PCAL* pericalcarine cortex, *PCC* posterior cingulate cortex, *PCUN* precuneus, *R* right, r_e environmental correlation, r_g genetic correlation, r_p phenotypic correlation; *SA* surface area, *sup* superior, *T1/T2* time one/time two, *TBV* total brain volume, *TCV* total cerebral volume, *TEMP* temporal, *THAL* thalamus, W*M* white matter, *WMHI* white matter hyperintensities, $^{\Delta}$ change

Table 3 Shared genetic influence between brain structure and cognitive ability

Reference	Cohort	*n* pairs MZ/DZ (age in years)	Phenotypes[a]	r_p	r_g	r_e
Pennington et al. (2000)	CLDRC	25/23 (>12)	TCV & FSIQ	–	.48	–
Carmelli et al. (2002a)	NHLBI	72/70 (69–80)	WMHI & EF	−.20	−.24	−.22
			WMHI & MMSE	−.20	−.36	0
Carmelli et al. (2002b)	NHLBI	72/67 (69–80)	LV L & EF	−.25	−.57	–
			LV R & EF	−.26	−.25	–
Posthuma et al. (2002)	NTR/UMCU	24/31 (19–69)	cGM & *g*	.25	.29	–
			cGM & WMem	.29	.38	–
			cWM & *g*	.24	.24	–
			cWM & Wmem	.29	.35	–
Posthuma et al. (2003)	NTR/UMCU	102/131 (19–69)	(Lowest r_g) CB & VC	–	.03	−.23
			(Highest r_g) cGM & WMem	–	.40	−.13
Hulshoff Pol et al. (2006)	NTR/UMCU	54/58 (19–69)	(Lowest r_g) CC WM & VIQ	.14	.15	–
			(Highest r_g) PH GM & PIQ	.23	.40	–
van Leeuwen et al. (2009)	NTR	48/64 (Mean = 9)	(Lowest r_g) cGM & PS	.06	.09	.05
			(Highest r_g) cGM & Raven	.22	.36	−.16
Betjemann et al. (2010)	CLDRC	41/30 (12–24)	(Lowest r_g) cGM & reading ability	.15	.14	–
			(Highest r_g) cWM & PS	.28	.89	–
Brans et al. (2010)	NTR/UMCU	77/84 (T1 Mean = 30)	(Lowest r_g) $^{\Delta}$ in MF L CT over 5 years & FSIQ	.08	.56	−.66
			(Highest r_g) $^{\Delta}$ in OP R CT over 5 years & FSIQ	.34	1	.22
Karlsgodt et al. (2010)	GOBS	467 subjects (19–85)	(Lowest r_g) CING FA & DB	–	.02	.01
			(Highest r_g) SLF FA & SDRT	–	.59	.25
Glahn et al. (2013)	GOBS	1129 subjects (18–83)	(Lowest r_g) EC FA & FMD	.06	0	.11
			(Highest r_g) CING_G FA & SDRT	−.08	−.47	.11
Brouwer et al. (2014a)	NTR/UMCU	11/21 (T1 19–56)	(Lowest r_g) $^{\Delta}$ in whole brain L over 5 years & FSIQ	.27	.50	−.17
			(Highest r_g) $^{\Delta}$ in cWM L over 5 years & VIQ	.29	1	−.41
Brouwer et al. (2014b)	NTR/UMCU	23/28 (Mean = 12)	(Lowest r_g) PCL L CT & FSIQ	−.29	−.32	−.24
			(Highest r_g) CUN L CT & VIQ	−.28	−1	.39
Bohlken et al. (2014)	NTR/UMCU	50/56 (19–55)	(Lowest r_g) PUT bilat & FSIQ	.01	0	.08
			(Highest r_g) THAL bilat & FSIQ	.26	.29	.08
Vuoksimaa et al. (2014)	VETSA	131/96 (51–60)	cGM CT & general cognitive ability	.08	.09	.10
			cGM SA & general cognitive ability	.21	.24	.21

[a] These phenotypes are volumes unless otherwise stated. Where results for a large number of structures are reported, the regions with lowest and highest genetic correlations are shown. Refer to Table 2 for details regarding how studies were identified.

bilat bilateral, *CB* cerebellum, *CC* corpus callosum; *cGM* cerebral gray matter; *CING* cingulum; *CING_G* cingulum gyrus; *CT* cortical thickness, *CUN* cuneus, *cWM* cerebral white matter; *DB* digits backwards, *EC* external capsule, *EF* executive function; *FA* fractional anisotropy, *FMD* facial memory display; *FSIQ* full scale intelligence, *g* general intelligence, *GM* gray matter; *L* left, *LV* lateral ventricle, *MF* medial frontal, *MMSE* Mini-mental state examination, *OP* occipital pole, *PCL* paracentral lobule; *PIQ* performance intelligence, *PH* parahippocampal; *PS* processing speed; *PUT* putamen, *R* right, *RAVEN* Raven standard progressive matrices, r_e environmental correlation, r_g genetic correlation, r_p phenotypic correlation, *SA* surface area, *SDRT* spatial delayed response task, *SLF* superior longitudinal fasciculus, *TCV* total cerebral volume, *THAL* thalamus, *VC* verbal comprehension, *VIQ* verbal intelligence, *WM* white matter, *WMem* working memory, *WMHI* white matter hyperintensities; $^{\Delta}$ change

factors can influence intelligence (Bates et al. 2013; Tucker-Drob et al. 2011; Turkheimer et al. 2003), determining if gene by environment interactions modify genetic associations between structure and cognition could prove crucial to our understanding of brain development. Also worth exploration is the gene-environment correlation (Plomin et al. 1977), which denotes the possibility that differential environmental effects can occur as a function of genetics (active gene-environment correlation). For example, individuals may choose to undertake mentally challenging activities (due to genetic influences) which may then influence cognitive ability or brain morphology. Furthermore examining if the magnitude of genetic overlap between brain morphology and cognitive ability differs by brain regions and structural phenotype (cortical thickness, volume, white matter integrity) will aid in identifying "higher order" brain structures especially sensitive to genetic or environmental influences. Multivariate twin studies provide unparalleled insights into the genetic relationships between brain structures and behavioral traits, and offer basic insights into the genetic mechanisms behind the human brain, which genetic association studies can build upon.

Candidate Genes for Brain Morphology

Candidate gene studies of brain phenotypes, especially prior to the GWAS era, offered the possibility of revealing genetic impact on neurobiological mechanisms that underlie human behaviour (Bigos and Hariri 2007; Bigos and Weinberger 2010; Addington and Rapoport 2012). Consequently, there is a large and ever growing number of publications, particularly for genetic polymorphisms known to be relevant to psychiatric illness and behavioural problems, as reviewed (e.g. Iofrida et al. 2014; Erhardt and Spoormaker 2013; Duff et al. 2013; Flint and Munafo 2013; Tost et al. 2012; Hasler and Northoff 2011; Meyer-Lindenberg 2010; Durston 2010). This work builds on the broader literature of gene characterisation, including studies of gene structure and the identification of critical polymorphisms, protein expression in the brain, as well as physiological function and implication in brain-related illnesses (e.g. Adachi et al. 2014).

The first candidate gene study of an MRI-based phenotype was published in 2000 (Bookheimer et al. 2000; Bigos and Weinberger 2010) and was for the *APOE* (apolipoprotein E) gene, an Alzheimer's disease risk gene involved in lipid metabolism and injury repair in the brain (Liu et al. 2013). The study showed that carriers of the ε4 risk allele had greater memory-related activation in brain regions affected by Alzheimer's disease than non-carriers. Many other studies followed (see reviews Liu et al. 2015; Cherbuin et al. 2007), exploring the effect of the risk allele on indicators of brain vulnerability (such as smaller volumes (Knickmeyer et al. 2014) and greater atrophy over time (Hostage et al. 2014) in ε4 carriers compared to non-carriers), particularly in regions associated with Alzheimer's disease (e.g. Knickmeyer et al. 2014). Other candidate genes that have received extensive attention are *BDNF* (brain-derived neurotrophic factor) and *COMT* (catechol-O-methyltransferase), both of which have been implicated in cognitive and emotional processing. BDNF is the most studied neurotrophin in the central nervous system due to its importance in the development and maintenance of normal brain functions (Adachi et al. 2014) and decreased levels have been associated with neurodegenerative, developmental, and neuropsychiatric disease (Allen et al. 2013; Ignacio et al. 2014; Malter Cohen et al. 2013). It is highly expressed in the prefrontal cortex and hippocampus (Pezawas et al. 2004), making these regions of interest in imaging studies. *COMT* is a dopaminergic system-related gene that dominates the regulation of dopamine metabolism in the prefrontal cortex (Scheggia et al. 2012). It has been widely studied in relation to neurocognitive processes and conditions such as depression, with considerable attention paid to brain systems linked to emotional processing (Antypa et al. 2013). Neuroimaging candidate gene studies have targeted specific structures and networks (e.g. hippocampal volume (Shi et al. 2013; Harrisberger et al. 2014; Wang et al. 2013b), Alzheimer's disease-related network (Hostage et al. 2014)), or alternatively, taken a whole brain approach (e.g. Brooks et al. 2014; Zhang et al. 2013; Dean et al. 2014).

However, many candidate gene studies of brain morphology have not been replicated and there are considerable inconsistencies across studies (e.g. see reviews for *APOE* (Chetelat and Fouquet 2013), *BDNF* (Harrisberger et al. 2014), and *COMT* (Barnes et al. 2012; Ira et al. 2013)). In an effort to gain some clarity several meta-analyses have been conducted. For example, a meta-analysis of *APOE* genotype effects on hippocampal volume (Liu et al. 2015) showed that ε4 carrier status was associated with smaller volumes ($N = 760$; 6 studies including Alzheimer's disease, mild cognitive impairment, and cognitively normal individuals). However, a study of *BDNF* Val^{66}Met effects on hippocampal volume in healthy individuals (Harrisberger et al. 2014) concluded that there was no effect ($N = 5298$; data from 32 independent studies), consistent with GWAS; see http://enigma.ini.usc.edu/enigma-vis/ (Stein et al. 2012)). Similarly, meta-analyses assessing genetic effects on white matter hyperintensities (Paternoster et al. 2009) did not find convincing evidence of association for *APOE* ($N = 8546$; 24 studies), *MTHFR* ($N = 2796$; 3 studies), *AGT* ($N = 2702$; 6 studies), or *ACE* ($N = 2316$; 9 studies).

This lack of replication is not specific to imaging phenotypes. Candidate gene studies of complex traits in general (Hewitt 2012), including intermediate phenotypes such as subjective and physiological responses (Hart et al. 2013), are often not replicated. The reasons for this inconsistency and the problem of false positives, have been addressed in several

reviews (Hewitt 2012; Ioannidis et al. 2008; Little et al. 2009; Moonesinghe et al. 2008; Munafo and Gage 2013; Sullivan 2007). The main issues are (i) *sample size*, (ii) the level of statistical evidence required, as for example, based on *prior probability* and *multiple testing*, (iii) confounding factors such as *population stratification* and *genotyping errors*, and (iv) the need for *replication* and/or *meta-analyses* to provide an additional level of validation, which we have summarised in the form of a checklist (Table 4). It consists of seven checkpoints for reducing false positives (first column), the reasons why these points should be considered (middle column), and some recommendations (last column). A recommendation with respect to sample size is to use effect sizes from well-powered GWAS as a guide. Of course, insights derived from GWAS have only been possible since 2007, when the first neuroimaging GWAS was published (Seshadri et al. 2007), and thus far these studies have largely focussed on brain volumes (see Table 5). Prior to GWAS, it was believed that effect sizes for candidate genes would be much larger for associated biological traits than for target behavioural and psychiatric traits (e.g. Inoue and Lupski 2003).

To assess the extent that candidate gene studies of brain morphology are now taking steps to reduce false positives, we examined publications from January 2013 to July 2014 for the frequently studied *APOE*, *BDNF*, and *COMT* genes. In total, we identified 58 publications (27 *APOE*, 22 *BDN*F, 9 *COMT*; Online Resources 3–5). While this is a narrow band, the studies are generally representative of the larger body of work in terms of the range of imaging phenotypes examined (from single measures through to whole brain approaches), sample size (from small studies of 30 individuals to meta-analysis of >5000), age range (neonates to the elderly), cohort (both healthy and patient groups), study design (including longitudinal, gene-gene interactions, gene-environment interactions), and differing ancestral populations. While the considerable between-study diversity makes comparison (and thereby consistency of findings) difficult to assess, we can nonetheless assess overall study power as well as the issue of replication.

Excluding meta-analyses, *sample sizes* for these studies ranged from 36 to 1147 for *APOE*, from 30 to 645 for *BDNF*, and from 30 to 274 for *COMT*. Figure 3 shows the power to reject the null hypothesis and declare a significant association, by sample size and amount of variance explained, for three different significance thresholds ($\alpha = 0.05$ (Fig. 3a), $\alpha = 5.0 \times 10^{-5}$ (Fig. 3b), and $\alpha = 5.0 \times 10^{-8}$ (Fig. 3c)). Broer et al. (2013) recommended a threshold of 0.05 be used when the prior probability of true association is high (i.e. $\alpha < 1.0 \times 10^{-7}$ in GWAS), but that a threshold of 1.0×10^{-5} be adopted for candidate genes not having prior probability established through GWAS, while 5.0×10^{-8} is the standard threshold required for GWAS significance. At the 0.05 threshold, a sample of 1566 would provide 80 % power to detect an additive 0.5 % effect, while a sample of 4785 would be required for the 5.0×10^{-5} threshold and 7900 for the 5.0×10^{-8} threshold. Superimposed upon the curves in Fig. 3a, b, and c, are the maximal sample sizes of the candidate gene studies of *APOE* (green line), *BDNF* (blue line), and *COMT* (red line). As is evident from these figures, even at the most lenient threshold for significance no studies other than meta-analysis (Harrisberger et al. 2014) achieved the canonical 80 % power threshold (at which we expect true effect will be missed 20 % of the time) to detect an effect accounting for 0.5 % of the variance. Note that large-scale GWAS studies thus far indicate that common genetic variants are likely to account for less than 0.5 % of phenotypic variance in complex traits, including structural brain imaging phenotypes (Flint and Munafo 2013; Hibar et al. 2015), and further, indicating that effect sizes reported in candidate gene studies have in some cases been overestimated (the accuracy of estimates of effect size is influenced by study power). Last, for a continuous distributed trait, which is the case for most measures of brain morphology, the amount of variance accounted for is a function of the **minor allele frequency (MAF)** of the candidate gene (Fig. 3d).

While it is possible to increase power by only genotyping individuals with extreme phenotypes (i.e. ±2 standard deviations from the mean) this strategy is not commonly used in imaging studies as it requires a large pool of phenotyped individuals to select from. However, the related practice of over sampling cases (or collecting phenotypes in a case control sample) is often used in imaging studies. Conceptually this method involves over sampling from one end of the distribution; however, the gain in power is dependent on the correlation between the trait studied and the disease on which the participants are ascertained. An example of a case–control imaging genetics cohort is the Alzheimer's Disease Neuroimaging Initiative (ADNI, http://www.adni-info.org/), for which the target sample was 800 individuals, comprising 200 healthy elderly, 400 with mild cognitive impairment, and 200 with Alzheimer's disease (Mueller et al. 2005).

Candidate gene studies of brain morphology often examine multiple phenotypes and include correction for *multiple testing*, as seen in recent studies (e.g. Aas et al. 2013; Brown et al. 2014; Hostage et al. 2014; Knickmeyer et al. 2014; Rabl et al. 2014; Zhang et al. 2013). On the related issue of multiple testing from the same dataset, it is becoming more common for studies to report where analyses and prior studies have used overlapping samples (e.g. Blokland et al. 2014). Similarly, transparent reporting can make clear the potential impact of confounding factors such as *genotyping error* and *sample ancestry*. To infer ancestry, ancestry informative marker (AIM) panels comprising a small number of highly informative genetic variants have been developed (Kidd et al. 2014; Phillips et al. 2014) and may become the new gold standard for assessing population stratification in candidate gene studies.

Table 4 Checklist for reducing false positives (and negatives) in candidate gene studies

Issues	Points to consider	Recommendations
Sample size	• Large samples are required when effect sizes are small (i.e. when variants account for a small percentage of variance, Fig. 3). • GWAS studies indicate small allelic effect sizes for complex traits, typically less than 1 % (Munafo and Gage 2013), and less than ~0.5 % thus far for brain structure (Flint and Munafo 2013). • Much larger samples are required for rare genetic variants. • Gene by environment interaction studies require considerably larger sample sizes than that needed for the main effect (Thomas 2010).	• Expect small effect sizes and plan sample size accordingly (i.e. do power calculations* based on effect sizes indicated in the largest available GWAS studies).
Prior probability	• Broer et al. (2013) suggest a conventional p-value threshold (0.05) for genetic variants with high prior probability of association (previously reported in GWAS at 1×10^{-7}), but a more stringent threshold (1×10^{-5}) for novel candidate gene studies.	• Consider prior probability and the level of statistical evidence required.
Multiple testing	• Testing multiple comparisons increases risk of type 1 error and must be appropriately corrected for (e.g. Little et al. 2009; Sullivan 2007). • Large samples allow the preservation of statistical power after correction for multiple testing (Sullivan 2007). • Transparency in reporting of hypotheses tested (number of genetic variants/genes, number of brain structures, number of sub-groups (e.g. males vs. females, young vs. old), number of statistical approaches etc.) allows for more informed interpretation of reported results. • A threshold of 5.0×10^{-6} is recommended for testing a single genetic variant at the whole brain voxel level (adopting a conservative correction for 10,000 independent traits (Medland et al. 2014)).	• Be transparent in reporting of hypotheses tested and include appropriate correction for multiple testing.
Population stratification	• Minor allele frequencies can vary considerably between populations. • Chance ancestral differences in genotype are confounding factors that can lead to false positives (or false negatives). • Self-identified race and ethnicity can be inaccurate ways to assess genomic ancestry (Mersha and Abebe 2015). • Ancestry informative marker (AIM) panels can characterise an individual's ancestry (Mersha and Abebe 2015; Kidd et al. 2014; Phillips et al. 2014). • Report methods used to assess and address population stratification (Little et al. 2009).	• Assess ancestry and aim for sample homogeneity.
Genotyping errors	• Systematic genotyping errors can lead to biased results (Ioannidis et al. 2008; Little et al. 2009). • Genotyping details (e.g. DNA source and storage, methods and platforms, error rates, call rates) and quality control procedures undertaken (e.g. blind replicate genotyping, assessment of missing data, tests of Hardy-Weinberg equilibrium) can provide insight into the potential for genotyping errors (Little et al. 2009).	• Detailed description of laboratory methods and quality control protocols should be available. • Seek replication in samples genotyped on different platforms (Ioannidis et al. 2008).
Replication	• Replication in an independent sample (same variant, same direction) enhances credibility. • Successful replication should be the result of primary analysis, rather than sub-group analyses (such as males vs. females) (Munafo and Gage 2013). • Between-study heterogeneity (both genetic and phenotypic) can reduce, or prevent, the possibility that true associations will replicate (Moonesinghe et al. 2008).	• Independent replication is crucial. • Minimize between-study heterogeneity.
Meta-analyses	• Meta-analysis of summary data from multiple studies (or mega-analysis of raw data from multiple cohorts) can increase power to identify influential genetic variants. • Factors influencing between-study heterogeneity (including genotyping errors, variation in phenotype classification, and population stratification) negatively impact the power of meta- and mega-analyses to identify true associations (Moonesinghe et al. 2008). • Large consortia using methods to reduce between-study heterogeneity maximise their ability to identify true association.	• Collaborate with independent groups to achieve larger samples with minimal between-study inconsistencies.

* http://pngu.mgh.harvard.edu/~purcell/gpc/qtlassoc.html

Table 5 Published GWAS for brain morphology

Authors	Phenotype	Discovery sample size	Replication sample	Number of sites (if meta-analysis)	Samples Ethnicity Discovery/ Replication	Genome –Wide significant association[a]	Replication[a]
Seshadri et al. (2007)	TBV, lobe volume (frontal, parietal, occipital, temporal), hippocampal volume, lateral ventricle volume, temporal horn volume, WMHI	Framingham: 705 stroke and dementia free controls	NA	NA	European	No	NA
Baranzini et al. (2009)	T2 lesion load, Total brain volume	1677 (794 MS, 883 HC)	NA	NA	European	No	NA
Potkin et al. (2009a)	Hippocampal volume	ADNI: 381 (172 AD, 209 HC)	NA	NA	European	No	NA
Debette et al. (2010)	Subclinical brain infarcts	CHARGE: 9,401 participants	CHARGE: 1,822 Caucasian and 644 African-American	7 and 2 (replication)	European/European and African	No	No
Shen et al. (2010)	142 measures of GM density, volume and cortical thickness	ADNI: 733 (175 AD, 354 MCI, 204 HC)	NA	NA	European	No	NA
Stein et al. (2010)	Temporal lobe volume, Hippocampal volume	ADNI: 742 (173 AD, 361 MCI, 208 HC)	NA	NA	European	No	NA
Bakken et al. (2011)	Mean cortical thickness	TOP: 94 SCZ	TOP: 327 (181 HC, 97 BD, 49 P&AD)[1]	NA	European/European	2 SNP in perfect LD	No
Fornage et al. (2011)	WMHI	CHARGE: 9,361	CHARGE:3,024 ARIC: 807 African-American	7 and 2 (replication)	European/European and African	6 SNP in high LD ($r^2 > 0.8$)	2 SNP in CHARGE, No replication in ARIC
Furney et al. (2011)	Hippocampal volume, entorhinal cortical volume, whole brain volume, ventricular volume, entorhinal cortical thickness	AddNeuroMed and ADNI: 939 (236 AD, 424 MCI, 279 HC)	NA	NA	European	No	NA
Hibar et al. (2011b)	Voxel-wise regional brain volume differences (vs. template)	ADNI: 731 (413 MCI, 237 HC)	NA	NA	European	No	No
Stein et al. (2011)	Left and right caudate volume	ADNI: 734 (172 AD, 205 HC, 357 MCI) QTIM: 464 twins and siblings	NA	2	European	No	NA
Bakken et al. (2012)	Visual cortical surface	TOP: 421 (Bakken et al. 2011) ADNI : 482 (180 HC, 302 MCI)[2] PING: 278	NA	3	All European	2 SNP in LD ($r^2 = 0.88$)	NA
Bis et al. (2012)	Hippocampal volume	CHARGE: 9,232	CHARGE: 2,318	12 and 2 (replication)	European/European	3 SNP	All replicated
Ikram et al. (2012)	Brain volume, intracranial volume	CHARGE: 8,175	CHARGE: 1,752	7 and 2 (replication)	European/European	2 SNP for IV	All replicated
Melville et al. (2012)	Hippocampal volume, total brain volume, WMHI	MIRAGE: 981 (454 AD and 537 HC) ADNI: 692 (168 AD, 336 MCI, 188 HC)	MIRAGE: 319 (188 AD, 231 HC)	2 and 1 (replication)	European/African-American	1 SNP and APOE ε4 for HV 1 SNP for HV [c] APOE ε4 for TBV	Only APOE ε4 replicated NA No
Paus et al. (2012)[b]	Total brain volume interaction with cigarette exposure during pregnancy	SYS: 599 (290 females, 309 males)	ALSPAC: 2,601 females	NA	European/European	8 SNP in high LD ($r^2 > 0.7$) in females	No
Stein et al. (2012)	Total brain volume	ENIGMA1: 6,500	NA	20	European & Hispanic	No	NA
	Intracranial volume	ENIGMA1: 7,607	CHARGE: 8,175	22 and 1 (replication)	European & mixed/ European	1 SNP [d]	Yes

Table 5 (continued)

Authors	Phenotype	Discovery sample size	Replication sample	Number of sites (if meta-analysis)	Samples Ethnicity Discovery/ Replication	Genome –Wide significant association[a]	Replication[a]
	Hippocampal volume	ENIGMA1: 10,372 CHARGE: 10,779	NA	23	European & mixed	1 SNP [d]	NA
Taal et al. (2012)	Head circumference	EGG EAGLE CHAR GE:19,089	NA	7 and 6 (replication)	European	2 SNP	NA
Hibar et al. (2013b)	Lentiform Nucleus volume (bilateral)	ADNI: 706 (162 AD, 246 MCI, 198 HC) QTIM: 639 (364 families)	NA	2	European	1 SNP	NA
Hibar et al. (2013a)	Regional hippocampal volume	ADNI:511 (323 MCI) QTIM: 571 (335 families) TOP: 172	NA	3	European	1 SNP	NA
Jahanshad et al. (2013b)	59 measures of brain fibre connectivity	QTIM:169 twins and siblings	QTIM:162 twins and siblings	NA	European	1 SNP	Yes
Hass et al. (2013)	Hippocampal volume	MCIC: 328	ENIGMA: 7,795 (5,775 HC, 2,020 cases) IMAGEN: 1,663 HC	NA	All European	No	No
Sprooten et al. (2013)	Fractional anisotropy	BFS: 150 (70 at risk for BD, 80 HC)	NA	NA	European	No	NA
Wang et al. (2013a)	Gray matter volume reduction in 4 regions associated with SCZ	125: (74 SCZ, 54 HC)	NA	NA	Chinese	No	NA
Hohman et al. (2014)	Left Inferior lateral ventricle dilation	ADNI: 700 (197 HC, 388 MCI, 115 AD)	NA	NA	European	No	NA
Hibar et al. (2015)	Intracranial volume, subcortical volume (hippocampus putamen, pallidium, amygdala, thalamus, nucleus accumbens, caudate nucleus)	ENIGMA2: 13,163	ENIGMA2: 1,878 ENIGMA2: 3,046\ ENIGMA2: 13,113	28 and 22 (replication)	European/European & Mixed	1 SNP for ICV 2 SNPs for HV [d] 3 SNPs for PV [d]	All but the SNP associated with ICV
		ENIGMA2: 15,031				1 SNPs for CNV [c] 1 for PV [c]	NA

[a] At the discovery stage the Genome-Wide (GW) significance threshold for one phenotype = 5.10^{-8}. When several traits are tested a correction for multiple testing is required. When this step had been ignored we calculated and used a threshold of 5.10^{-8} /Ntrait (Bonferroni correction). A similar correction for multiple testing has been applied at the replication stage to control for the number of tests performed. Thus in Melville et al. (2012) the SNP rs2298948 associated with hippocampus volume was reported to be genome-wide significant (p-value = 4.98×10^{-8}) but after correction for multiple phenotype testing the association did not survive. Likewise, the SNP rs1970546 association with total brain volume (Seshadri et al. 2007) is considered non-significant after correction for testing of multiple phenotypes.

[b] Paus et al. (2012) study differs from the other GWAS on head size in that it aimed to identify SNP effects increased by cigarette exposure during pregnancy, thus measuring a GxE (genetic-environmental) interaction effect (Gauderman et al. 2013).

[c] GW significance using discovery and replication samples

[d] SNP significant overall, but heterogeneous effect across ethnicity groups suggests the SNP might have an effect only in individuals of European ethnicity

AD Alzheimer's disease, *BD* bipolar disorder, *CSF* cerebral spinal fluid, *GM* gray matter, *HC* healthy controls, *ICV* intracranial volume, *MD* major depressive disorder, *MCI* mild cognitive impairment, *MS* multiple sclerosis cases, *P&AD* psychotic and affective disorder cases, *SCZ* schizophrenic cases, *WMHI* white matter hyperintensities

AddNeuroMed a European collaboration for the discovery of novel biomarkers for Alzheimer's disease (Lovestone et al. 2009), *ADNI* Alzheimer's Disease Neuroimaging Initiative, *ALSPAC* Avon Longitudinal Study of Parents And Children, *ARIC* The Atherosclerosis Risk in Communities Study, *BFS* Bipolar Family Study, *CHARGE* Cohorts for Heart and Aging Research in Genomic Epidemiology, *DNS* Duke Neurogenetics Study, *EGG* Early Growth Genetics consortium, *ENIGMA* Enhancing NeuroImaging Genetics through Meta-Analysis (1 and 2 refer to study phases; Thompson et al. (2014)), *FBIRN* Functional Biomedical Informatics Research Network consortium, *MCIC* Mind Clinical Imaging Consortium, *QTIM* Queensland Twin Imaging Study (de Zubicaray et al. 2008), *SYS* Saquenay Youth Study, *TOP* Thematically Organized Psychosis study

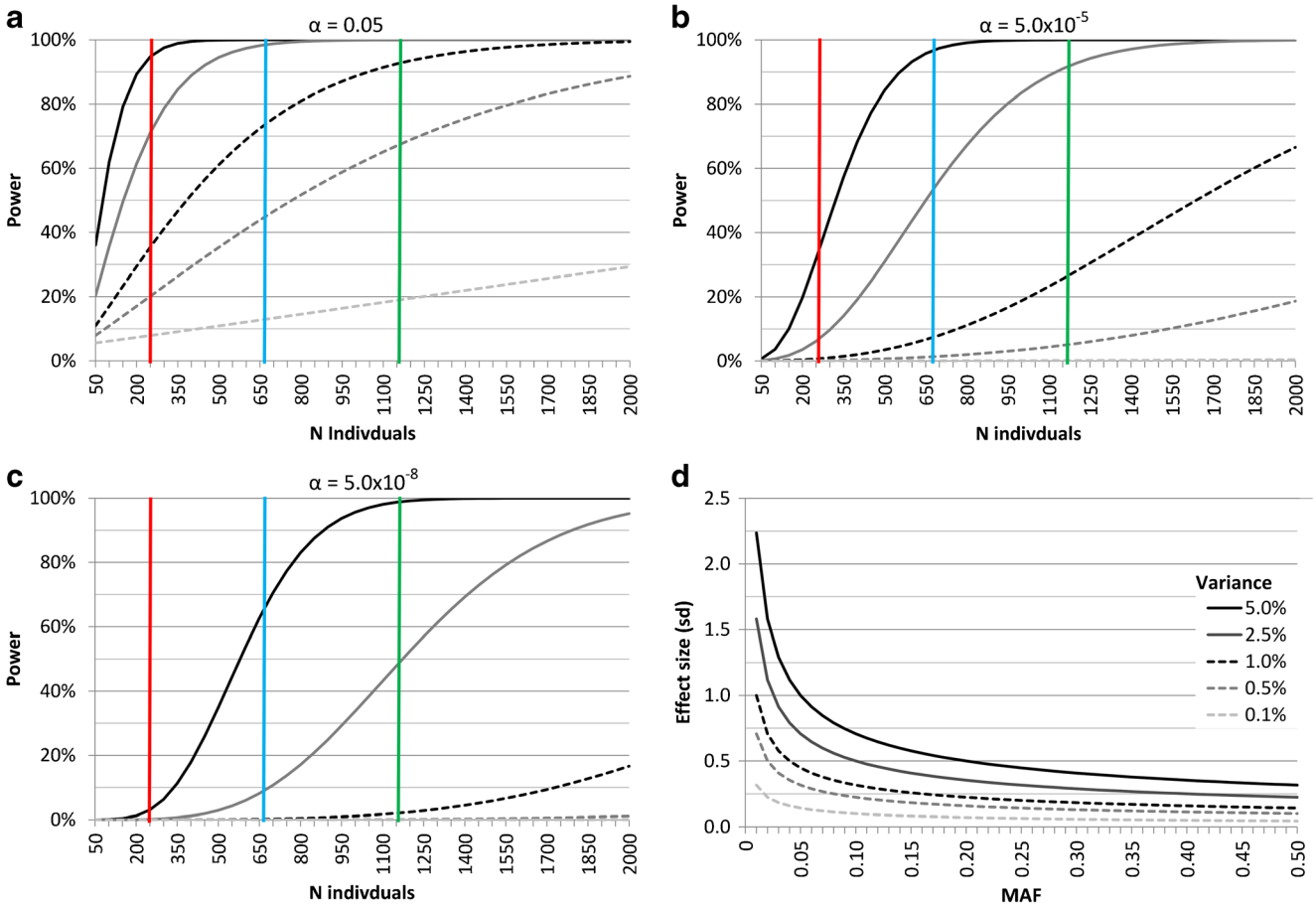


Fig. 3 The power of studies examining continuous measures (such as brain morphology) is typically discussed in terms of variance explained. Here, variance explained is plotted as a function of sample size (X axis) and power to reject the null hypothesis and declare a significant association (Y axis) given thresholds of **a** 0.05, **b** 5.0×10^{-5}, and **c** 5.0×10^{-8}. Superimposed on the power curves for **a**, **b**, and **c** are lines indicating the largest sample found for recent candidate gene studies for brain morphology (January 2013 to July 2014) for *APOE* (*green line*), *BDNF* (*blue line*), and *COMT* (*red line*). Section **d** shows that for a continuous normally distributed phenotype, the amount of variance explained by an individual variant is a function of the minor allele frequency (MAF) and the size of the effect in standard deviations. For a common variant with a MAF of 0.5, an increase in the mean of 0.14 standard deviations per risk allele will account for 1 % of the variance, whereas for a less common variant with a MAF of 0.1, the increase in the mean per risk allele required to generate a 1 % effect is around twice the size, at 0.32 standard deviations

The credibility of novel findings can be greatly enhanced by including replication from an independent sample in the same publication (as is required by some journals (Hewitt 2012)). However, *replication* is not typically included in candidate gene studies of brain morphology. Of the *APOE*, *BDNF*, and *COMT* studies (Online Resources 3–5), the majority report significant effects or suggestive trends (many of which are for novel findings), but none include independent replication within the same paper. The challenge for imaging studies is finding an available replication sample, which for the more complex phenotypes becomes difficult from a practical perspective. Recent approaches have sought validation across age groups (e.g. *OPRD1* effects on brain volume in independent young adult and elderly cohorts (Roussotte et al. 2014)) and through splitting a single cohort into discovery and replication subgroups (e.g. *ATXN7* effects on cerebellar volume (van der Heijden et al. 2013)). Further, with the growing use of common software, and facilitated by the sharing of processing pipelines and scripts, opportunities for imagers to work together and seek independent validation in comparable samples will increase. In addition, groups with imaging phenotypes and genotyping are coming together to form large consortia (e.g. ENIGMA (Thompson et al. 2014), CHARGE (Psaty and Sitlani 2013), EGG (Taal et al. 2012), IMAGEN (Schumann et al. 2010)), thereby furthering opportunities for collaboration and replication, and maximising samples for *meta-analysis* and/or pooling data (mega-analysis).

In summary, underpowered candidate gene studies have led to considerable inconsistency in the literature. However, in looking only at larger studies, meta-analyses, and large-scale GWAS results, there is some evidence that the *APOE* genotype may influence hippocampal volume in patient groups ($N = 760$, meta-analysis, smaller volume in ε4 carriers

versus non-carriers (Y. Liu et al. 2015)), but not in healthy adolescents (N = 1412 (Khan et al. 2014)). *BDNF* Val^{66}Met does not appear to influence hippocampal volume in healthy individuals or patient groups (N = 5298, meta-analysis (Harrisberger et al. 2014); $N \geq 5775$, GWAS, http://enigma.ini.usc.edu/enigma-vis/ (Stein et al. 2012)), nor does *COMT* Val^{158}Met ($N \geq 5775$, GWAS, (Stein et al. 2012)). Larger samples and independent replication will provide more robust insights into the role of these and other candidate genes.

The way forward for candidate gene studies of brain morphology is changing as we gain more knowledge regarding the genetic architecture of brain phenotypes (e.g. GWAS show that small effect sizes are expected for common genetic variants). Moreover, the performance of well-studied genetic variants can be assessed in large GWAS to help resolve existing conflicts (although the number of brain phenotypes investigated in large GWAS is still limited). Note that lack of replication in GWAS may reflect selective reporting and publication bias among candidate gene studies (i.e. the bias of both researchers and journal editors to treat positive results differently to null results) and implies the publication of false-positive findings (Bosker et al. 2011; Hart et al. 2013; Ioannidis et al. 2008). To address these biases, both researchers and journals should routinely seek to publish high-quality null results (Shields 2000; Ioannidis et al. 2008). In addition, new avenues of research are made possible through GWAS. The top performing genetic variants provide new candidates for closer research, and further, rather than focussing on the effect of a single variant, multiple variants can be combined (summing "risk" alleles weighted by discovery sample effect size) to create individual polygenic scores (Wray et al. 2014). Finally, collaboration is a growing requirement for candidate gene studies in order to obtain large samples and to provide independent replication of novel findings. These and other steps to reduce false positives (Table 4) will enhance the validity of future candidate gene studies.

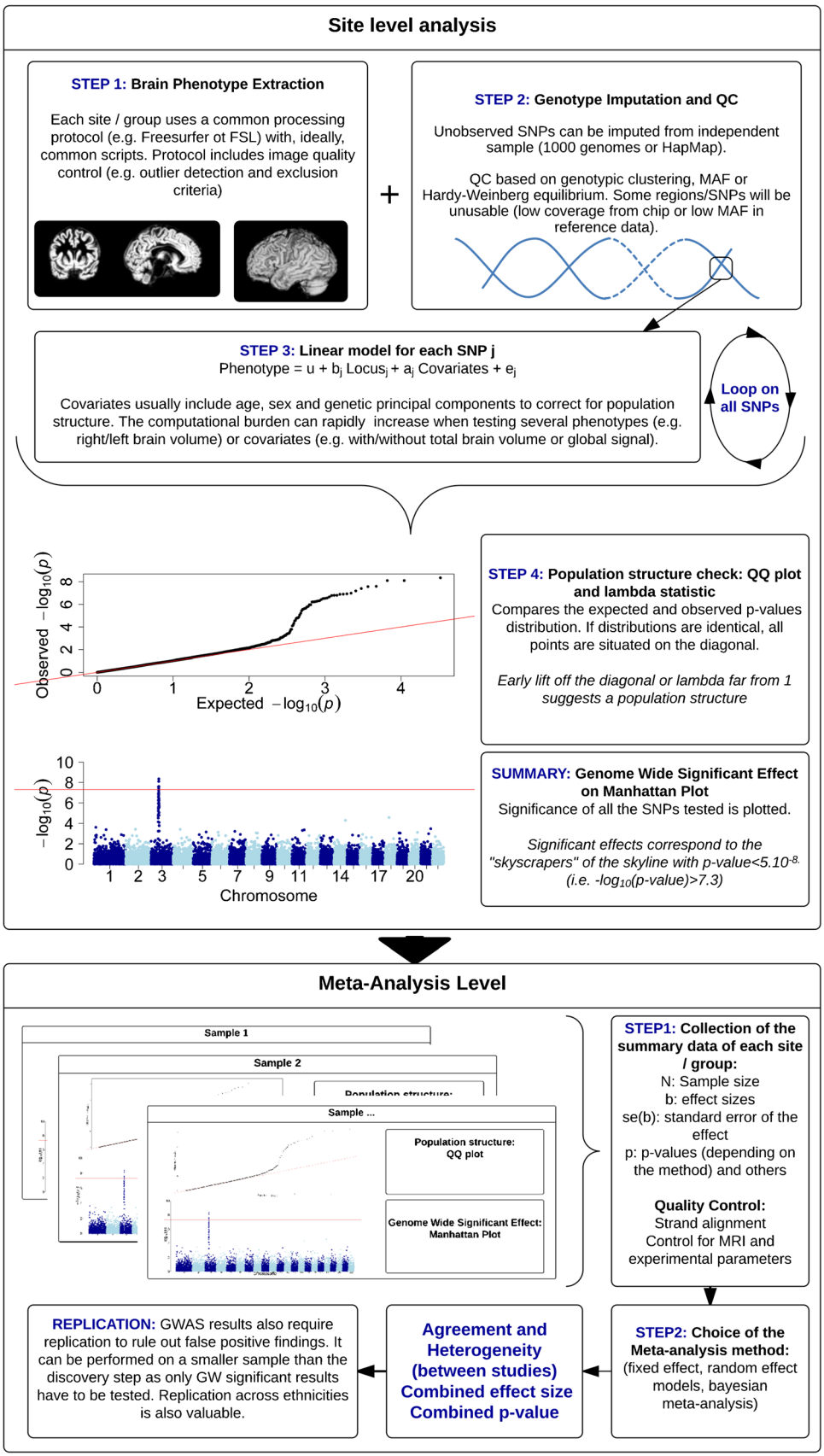
GWAS: a Decade of Publications

The past decade has seen a number of milestones for GWAS, beginning with the first report of genome-wide association (found for age-related macular degeneration in a modest sample of 96 cases and 50 controls (Klein et al. 2005)). Just two years later, the first large-scale GWAS was conducted (for common familial diseases; ~2000 cases per disease and 3000 shared controls; total N = 16,179; (Wellcome Trust Case Control 2007)). Modest effect sizes were observed, with significant loci explaining only a small proportion of trait heritability. This study was followed by more than 1900 GWAS (Mailman et al. 2007; Welter et al. 2014), in which hundreds of SNPs, all with small effect sizes, were associated with Mendelian and complex disorders (Visscher et al. 2012a; Bentham and Vyse 2013; Stranger et al. 2011). More recently, imaging phenotypes have been the target of GWAS in order to identify specific genes responsible for the moderate to high heritability found for brain morphology. The largest study thus far is for subcortical volumes (N > 30,000, Hibar et al. (2015)).

A Well Described Methodology

Several excellent reviews have described the main steps in a GWAS (McCarthy et al. 2008; Hirschhorn and Daly 2005), including meta-analysis methods (Evangelou and Ioannidis 2013; Pereira et al. 2009) and software (Marchini and Howie 2010; Begum et al. 2012; Benyamin et al. 2009; Evangelou and Ioannidis 2013), as well as specific issues such as imputation (Marchini and Howie 2010) and genome-wide significance levels (Dudbridge and Gusnanto 2008; Pe'er et al. 2008; Rosenberg et al. 2010; Teo et al. 2010; Medland et al. 2014). An overview of GWAS and GWAS meta-analysis methods is given in Fig. 4 (summarising the GWAS processing and analyses undertaken at each site, with an emphasis on the further methods and requirements of a meta-analysis, as well as key references).

First, at each site the phenotype is extracted from the MRI image(s), using a rigorous processing pipeline that maximises signal to noise ratio and reliability of the measurement. In a meta-analysis it is essential to establish a common protocol that can be used across sites (participant inclusion and exclusion criteria, software for extracting the brain phenotype, software version and quality control (QC) checks, inclusion of covariates) to ensure harmonization of the phenotype (Evangelou and Ioannidis 2013; de Bakker et al. 2008). Similarly, at each site the same QC checks (e.g. call rate, minor allele frequency, Hardy–Weinberg outliers, technical artefacts, cryptic relatedness, and population outliers) of the genotypes are essential and can prevent false positives (de Bakker et al. 2008; Teo 2008). Each site then imputes their genotype data from the same reference panel (International HapMap, C 2005; Genomes Project et al. 2010). Imputation allows the same set of SNPs to be analysed across sites, despite the use of different SNP chips with distinct SNP content (de Bakker et al. 2008). Once both the phenotypic and genetic data have been extracted the sites use standard software (Benyamin et al. 2009; Galesloot et al. 2014; Anderson et al. 2010) to test the association of every SNP with the phenotype. For a quantitative trait such as a brain volume this is the difference in mean score associated with the probability of having the alternate SNP allele. Sites with



General methodology and introduction (McCarthy et al.. 2008; Hirschohorn and Daly, 2005)

Review on imputation methods (Marchini & Howie 2010)

SNP quality control (Gondro et al., 2013)

Family based GWAS can be performed, with specific limitations and advantages (Benyamin et al., 2009; Chen & Abecasis, 2007)

List of most popular software for GWAS (Begum et al., 2012)

Plots from R package qqman (Turner, 2014)

Genome-wide significance is derived from an estimation of the number of independent loci. This threshold may depend on the ethnicity (Dudbridge & Gusnanto, 2008; Pe'er et al., 2008)

GWAS specificities on African and non-European populations (Rosenberg et al., 2010; Teo et al.. 2010)

Genome-wide brain-wide significance (Medland et al., 2014)

Review of the meta-analyses methods and software (Evangelou & Ioannidis, 2013; Pereira et al., 2009)

Practicality of GWAS meta-analysis after imputation (de Bakker et al., 2008)

Review of meta-analysis GWAS (Panagiotou et al., 2013)

Relevance of the findings and future of GWAS (Stranger et al., 2011; Visscher et al., 2012)

Fig. 4 Flowchart of genome-wide association study meta-analysis

twin or family data control for individual relatedness in the regression (Benyamin et al. 2009; Chen and Abecasis 2007; Laird and Lange 2006).

In a meta-analysis, each site uploads the regression summary statistics (sample size, strength of the association and p-value for every SNP) to a central site. The data are then quality controlled to remove results from SNPs with low MAF (e.g. less than 1 %) and poor imputation, as well as checking for evidence of population substructure, which is a major confound when correlated with the trait/disease studied (Marchini et al. 2004; Rosenberg et al. 2010; Campbell et al. 2005; Teo et al. 2010). Because different ethnic groups can have different SNP allele frequencies, even if geographically close (Rosenberg et al. 2010; Teo et al. 2010), random sampling within a heterogeneous population can create false positive association or reduce the power of detection (Campbell et al. 2005; Marchini et al. 2004). Population heterogeneity is assessed by calculating the ratio of the median test statistic to that of the expected (lambda) and by plotting the expected and observed distribution of p-values (quantile-quantile or QQ plot), see Fig. 4. A greater frequency of "high" p-values (e.g. $>10^{-4}$) than expected by chance and a lambda different to 1 suggest an unaccounted population structure. These spurious effects can be reduced by excluding genetic outliers, regressing out principal genetic components (Stranger et al. 2011; Price et al. 2006), using Genomic Control (Devlin and Roeder 1999), mixed models (Price et al. 2010), population stratification (Price et al. 2010) or family data (Benyamin et al. 2009).

The threshold for declaring genome wide significance of an association test is a p-value $<5.0 \times 10^{-8}$. This corresponds to a Bonferroni correction from an estimated 1 million independent loci (0.05/1,000,000) in a European genome (Pe'er et al. 2008). The GWAS results can be summarised in a Manhattan plot in which the significance level is represented (Y-axis) for all SNPs tested (X-axis). Several meta-analysis approaches allow aggregating the results (Evangelou and Ioannidis 2013; Pereira et al. 2009) with more or less power and sensitivity to heterogeneity (Evangelou and Ioannidis 2013). Remaining differences between sites (e.g. MRI field strength, population differences) can be included as covariates. Ultimately, a meta effect size and p-value summarise the SNP association across sites, and heterogeneity metrics (Q and I^2) used to assess the remaining between-study variance and heterogeneity not due to chance (Evangelou and Ioannidis 2013).

Replication of the results in an independent sample is important in GWAS in order to validate findings and eliminate false positives (Stranger et al. 2011). Often only a set of significant SNPs from the discovery stage are carried on to the replication phase, which reduces the burden of multiple testing correction. To provide a more accurate estimate of the effect size, discovery and replication samples are often meta-analysed.

GWAS for Brain Morphology

We conducted an extensive search in PubMed for GWAS using brain morphology phenotypes. The search terms "GWAS AND brain AND (MRI or DTI)" returned 168 articles, among which 27 were GWAS of brain morphology, including one GWAS of head circumference (Taal et al. 2012). An additional 6 articles used fMRI or magnetic resonance spectroscopy (MRS) (Baranzini et al. 2010; Liu et al. 2010; Liu et al. 2011; Brown et al. 2012; Ousdal et al. 2012; Potkin et al. 2009b) (these are described in Online Resource 4). We confirmed that our PubMed search was exhaustive by cross-checking with the GWAS catalogues (Mailman et al. 2007; Welter et al. 2014; Hindorff et al. 2013), and identified no additional studies. We omitted two studies that selected SNPs based on a round of penalised regression (Kohannim et al. 2012a; Kohannim et al. 2012b), which limits their comparison with standard GWAS. Table 5 lists the 25 GWAS for brain morphology, as well as the most recent GWAS on brain and subcortical volumes (Hibar et al. 2015) – a total of 26 GWAS. The number of publications over the last decade is presented in Fig. 5.

Volumetric phenotypes are the most widely studied (19 out of 26 publications). A main reason being that brain volumes are more straightforward to extract from the images. The most common is hippocampal volume with 10 studies (Seshadri et al. 2007; Potkin et al. 2009a; Stein et al. 2010; Furney et al. 2011; Bis et al. 2012; Melville et al. 2012; Stein et al. 2012; Hass et al. 2013; Hibar et al. 2013a, 2015) followed by global head size (i.e. total brain volume, intracranial volume, head circumference) with 8 publications (Seshadri et al. 2007; Baranzini et al. 2009; Furney et al. 2011; Ikram et al. 2012; Melville et al. 2012; Paus et al. 2012; Stein et al. 2012; Hibar et al. 2015). Hippocampal volume has been a main focus as it is commonly associated with early Alzheimer's disease (Jack et al. 2011; Simic et al. 1997), schizophrenia (Wright et al. 2000), depression (Videbech and Ravnkilde 2004) and memory ability (Van Petten, 2004). Other volumetric phenotypes include lobe volumes (Seshadri et al. 2007; Stein et al. 2010), ventricle (Seshadri et al. 2007; Furney et al. 2011), caudate (Stein et al. 2011; Hibar et al. 2015), lentiform nucleus (Hibar et al. 2013b) and temporal horn volume (Seshadri et al. 2007); gray matter (GM) volumes (Shen et al. 2010; Wang et al. 2013a), tensor based morphometry (Hibar et al. 2011b), and most recently nucleus accumbens, putamen, palladium, thalamus and amygdala volumes (Hibar et al. 2015). There are also three GWAS for white matter hyperintensity (Seshadri et al. 2007; Fornage et al. 2011; Melville et al. 2012) and one for

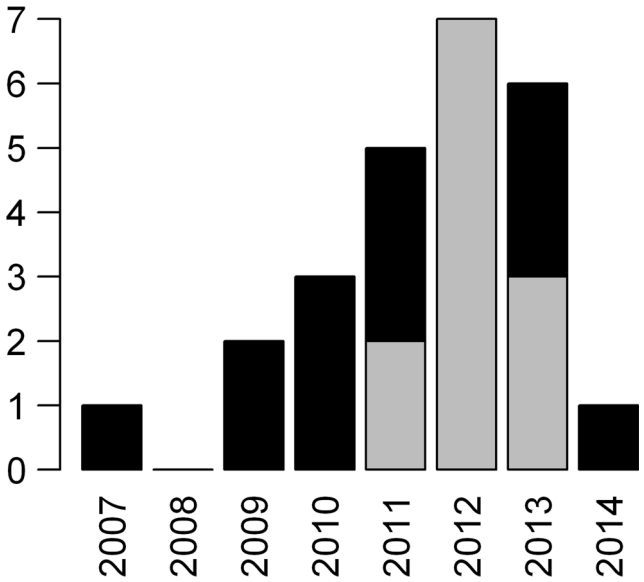


Fig. 5 Number of published GWAS for brain morphology per year, studies with replicated results appear in gray. The GWAS that focused on brain phenotypes were identified in PubMed using search terms "GWAS AND brain AND (MRI OR DTI)" and manually filtered to select GWAS. We ensured that our list was exhaustive by running PheGenI (Phenotype-Genotype Integrator (Welter et al. 2014)) and genome.gov (Mailman et al. 2007) searches for each phenotype previously identified. The second ENIGMA paper (Hibar et al. 2015), was not included for 2015

subclinical brain infarcts (Debette et al. 2010). White matter lesions are frequently associated with ageing (Launer 2004), mood disorders (Steffens et al. 1999) stroke (Vermeer et al. 2003; Launer 2004), and with decreases in cognitive functions (de Groot et al. 2000). More recently, with DTI becoming acquired in larger samples, two studies focused on fractional anisotropy (Sprooten et al. 2013) and brain fibre connectivity (Jahanshad et al. 2013b).

Of the 26 GWAS, 12 are meta-analyses, of which 11 (92 %) report a genome-wide significant result. In contrast, for single site studies only three (out of 14, or 21 %) found an association that was genome-wide significant (Table 5). The higher rate of genome-wide associations in meta-analyses is due to the larger sample sizes, hence power. For example, the mean sample size in the meta-analyses discovery stage was 6891, range [1181–19,089], while in the single site or single centre studies the mean was 577, range [94–1677]. The largest (and most recent) meta-analysis included >13000 individuals from 28 sites in the discovery sample, and >17,000 scans from 22 sites for the replications (Hibar et al. 2015). This illustrates the trend towards larger samples through the formation of consortia (e.g. CHARGE (Psaty et al. 2009; Psaty and Sitlani 2013) or ENIGMA (Thompson et al. 2014)). ENIGMA has seen its discovery sample size almost double in three years (Stein et al. 2012; Hibar et al. 2015) with an additional six groups (sites) joining, and the availability of larger samples at many of the participating sites. This increase in **statistical power** resulted in more genome-wide significant associations (e.g. for hippocampal volume from one SNP to two) and allowed more phenotypes to be studied (e.g. from three to eight traits). Ultimately, increased power leads to greater precision in estimating SNP effect sizes, as well as enhancing the probability of detecting rare variants with small effects (Panagiotou et al. 2013; Sham and Purcell 2014; Anderson et al. 2010; Pereira et al. 2009).

Eleven of the 26 GWAS included a replication sample (Debette et al. 2010; Bakken et al. 2011; Fornage et al. 2011; Bis et al. 2012; Ikram et al. 2012; Melville et al. 2012; Paus et al. 2012; Stein et al. 2012; Jahanshad et al. 2013b; Hass et al. 2013; Hibar et al. 2015), though two (Debette et al. 2010; Hass et al. 2013) used a more lenient significance threshold, which increases the probability of reporting false positives in the discovery phase as well as limiting the confidence in the replication because it is based on a weak prior (Broer et al. 2013; Nuzzo 2014). Debette et al. (2010) did not replicate their association and the results from Hass et al. (2013) were suggestive. Among the other nine publications that identified genome wide significant results, seven replicated their findings, at least partially (Fornage et al. 2011; Bis et al. 2012; Ikram et al. 2012; Melville et al. 2012; Stein et al. 2012; Jahanshad et al. 2013b; Hibar et al. 2015). Replication samples were smaller than the discovery sample in about half of the studies (Fornage et al. 2011; Bis et al. 2012; Ikram et al. 2012; Melville et al. 2012; Jahanshad et al. 2013b). However, due to the smaller number of tests at replication compared to discovery, most replication samples were well-powered. Failure to replicate can arise from a false positive result at the discovery stage (Lazzeroni et al. 2014) or insufficient statistical power in the replication phase (Lazzeroni et al. 2014).

At a SNP level, the 26 GWAS identified 29 genome-wide significant associations **(**excluding loci in strong **linkage disequilibrium**) with one of the brain morphological phenotypes (Table 6). Of these, 21 findings (73 %) replicated, five (17 %) were only suggestive (no replication) and three (10 %) did not replicate. One study used head circumference as a replication proxy for brain volume and found suggestive support (Paus et al. 2012). In addition, some results were replicated in European populations only (Fornage et al. 2011; Hibar et al. 2015; Stein et al. 2012) suggesting a different SNP effect across ethnicities, while others replicated in ethnically diverse populations (Melville et al. 2012).

As can be seen in Table 6 and Fig. 6, the associated SNPs are scattered over 13 different chromosomes, with Chromosome 12 having the largest number of **variants** (5) associated with hippocampal volume (Hibar et al. 2015; Bis et al. 2012; Stein et al. 2012), intracranial volume (Stein et al. 2012) or head circumference (Taal et al. 2012). On several chromosomes, including the sex chromosomes (X and Y), which are often excluded from an analysis, no genome-wide

Table 6 Genome-Wide significant associations reported in the GWAS literature of brain morphology

Phenotype	SNP	Chr	Minor allele/Other	MAF	Intronic (gene)/ intergenic	*p*-value[a]	% Change in structure[b] (% variance explained[b])	Replicated	Combined Sample size[a]	Study
Hippocampal volume (dorsal region only)	rs6703865	1	A/G	0.08 (European)/ 0.17	Intronic (F5/SELP)	1.14×10^{-9}	NA (1.67)	Yes	2,102	Melville et al. (2012)
	rs145212527	3	G/T	0.04	Intronic (FBLN2)	1.25×10^{-8}	NA (2.26)	NA	1,345	Hibar et al. (2015)
	rs17178006	12	G/T	0.10	Intronic (MSRB3)	5.3×10^{-11}	−2.4 (0.66)	Yes	6,252	Bis et al. (2012)
	rs6581612	12	C/A	0.27	Intergenic	7.1×10^{-11}	−1.2 (0.36)	Yes	11,501	Bis et al. (2012)
	rs61921502	12	G/T	0.16	Intronic (MSRB3)	6.87×10^{-11}	−1.01 (0.26)	Yes	16,209	Hibar et al. (2015)
	rs7294919	12	C/T	0.09	Intergenic	2.9×10^{-11}	+2.1 (0.32)	Yes	9,662	Bis et al. (2012)
	rs77956314	12	C/T	0.09	Intergenic	2.82×10^{-15}	−1.40 (0.36)	Yes	17,190	Hibar et al. (2015)
	rs7294919	12	C/T	0.099	Intergenic	6.70×10^{-16}	+2.1 (0.27)	Yes	21,151	Stein et al. (2012)
	rs9315702	13	A/C	0.44/0.22	Intronic (LHFP)	1.52×10^{-08}	NA (1.44)	Yes	2,102	Melville et al. (2012)
	–	19	ε4	0.19	APOE	1.58×10^{-33}	NA (6.43)	Yes	2,102	Melville et al. (2012)
Total brain volume	rs716890[c]	4	G/C	0.24	Intronic (KCTD8)	5.4×10^{-09}	−5.2 (5.57)	(different phenotype)	557	Paus et al. (2012)
	–	19	ε4	0.19	APOE	4.25×10^{-10}	NA (2.19)	Yes (Caucasian)	1,683	Melville et al. (2012)
Intracranial volume	rs4273712	6	G/A	0.27	Intergenic	1.8×10^{-13}	+1.0 (0.64)	Yes	8,175	Ikram et al. (2012)
	rs10784502	12	C/T	0.49	Intronic (HMGA2)	1.12×10^{-12}	+0.60 (0.17)	Yes	15,782	Stein et al. (2012)
	rs9303525	17	G/A	0.22	Intronic (KANSL1)	7.6×10^{-15}	−1.1 (0.72)	Yes	8,175	Ikram et al. (2012)
	rs17689882	17	A/G	0.22	Intronic (CRHR1)	7.72×10^{-9}	−0.96 (0.26)	No	12,822	Hibar et al. (2015)
Head circumference	rs1042725	12	C/T	0.48	Intronic (HMGA2)	2.8×10^{-10}	−0.14 (0.20)	NA	19,089	Taal et al. (2012)
	rs7980687	12	A/G	0.20	Intronic (SBNO1)	8.1×10^{-9}	+0.16 (0.24)	NA	13,134	Taal et al. (2012)
Putamen Volume	rs945270	14	G/C	0.42	Intergenic	1.08×10^{-33}	−0.94 (0.52)	Yes	28,275	Hibar et al. (2015)
	rs62097986	18	A/C	0.44	Intronic (DCC)	1.01×10^{-13}	+0.58 (0.20)	Yes	28,036	Hibar et al. (2015)
	rs6087771	20	C/T	0.29	Intronic (BCL2L1)	1.28×10^{-12}	−0.64 (0.20)	Yes	25,540	Hibar et al. (2015)
	rs683250	11	G/A	0.37	Intonic (DLG2)	3.94×10^{-11}	+0.51 (0.17)	Yes	26,258	Hibar et al. (2015)
Caudate Nucleus Volume	rs1318862	11	C/T	0.42	Intergenic	6.17×10^{-9}	−0.74 (0.22)	Yes	15,031	Hibar et al. (2015)
White matter hyperintensity	rs3744028[c]	17	C/A	0.18	Intronic (TRIM65)	4.0×10^{-15}	+3.4 (0.46)	Yes (in Europeans)	13,084	Fornage et al. (2011)
	rs1055129	17	G/C	0.30	Intronic (TRIM47)	2.3×10^{-11}	+2.0 (0.33)	Yes (in Europeans)	13,084	Fornage et al. (2011)
Average Cortical Thickness	rs4906844[c]	15	A/G	0.47	Intronic (LOC100128714)	1.08×10^{-8}	−3.0 (2.90)	No	1,054	Bakken et al. (2011)
Visual Cortical Surface	rs238295[c]	2	T/G	0.36	Intronic (GPCPD1)	6.5×10^{-9}	+0.6 (2.4)	Yes	1181	Bakken et al. (2012)
Left PCC - left SPC brain fibre connectivity	rs2618516	11	T/C	0.36	Intronic (SPON1)	3.23×10^{-9}	+0.2 (10.4)	No	331	Jahanshad et al. (2013b)

Table 6 (continued)

Phenotype	SNP	Chr	Minor allele/Other	MAF	Intronic (gene)/ intergenic	*p*-value[a]	% Change in structure[b] (% variance explained[b])	Replicated	Combined Sample size[a]	Study
Lentiform Nucleus volume (bilateral)	rs1795240	1	A/G	0.43	Intergenic	4.79×10^{-08}	−1.7 (2.08)	NA	1,345	Hibar et al. (2013b)

[a] Reported for combined discovery and replication samples

[b] Percentage change in volume per effect allele was based on the absolute value of the final combined effect (discovery + replication) divided by a weighted average of the brain structure across all sites in the discovery sample and then multiplied by 100. Percentage variance explained was based on the combined discovery and replication samples using the following equation ($t^2/((n\text{-}k\text{-}1)+t^2)$ where the t-statistic is the beta-coefficient from the regression model (controlling for covariates) divided by the standard error of the beta estimate, where n is the total number of subjects, and where k is the total number of covariates (given as 2 for all studies) (Hibar et al. 2015)

[c] Several SNPs in LD also passed the Genome-Wide significance threshold

SNPs are ordered by phenotype and location on the genome. Multiple significant hits within a haplotype are reported together, effect allele, MAF, effect size, and *p*-values are reported for the most significant SNP that appears in bold. The reference SNP cluster ID, starting with "rs", corresponds to a unique SNP in the dbSNP database (Sherry et al. 2001). The genomic region contains the chromosome number followed by letters "p" (short arm of the chromosome) or "q" (long arm) and the SNP position from the centromere (e.g. 12q24 for chromosome 12, position 22 off the long arm)

Chr chromosome, *NA* not available, *NR+* not reported, but positive result, *NR-* not reported, but negative result, *PCC* posterior cingulate cortex, *SPC* superior parietal cortex

associations have been identified. Multiple **variants** were reported across independent studies at locations: 12q14.3 (rs17178006, rs6581612, rs61921502) and 12q24.22 (rs7294919, rs77956314) for hippocampal volume (Hibar et al. 2015; Bis et al. 2012; Stein et al. 2012) and 17q21.31 (rs9303525, rs17689882) for intracranial volume (Ikram et al. 2012; Hibar et al. 2015). For some of the regions there were several SNPs in linkage disequalibrium that passed the genome-wide threshold (Paus et al. 2012; Fornage et al. 2011; Bakken et al. 2012; Ousdal et al. 2012; Ikram et al. 2012). Though it is not possible to identify the **causal variant(s)** from a set of variants in linkage disequilibrium using p-values (Visscher et al. 2012a), together with **expression quantitative trait loci *(*eQTL),** and regulatory information they can provide suggestive evidence (Pickrell 2014; Veyrieras et al. 2008; Kindt et al. 2013; Gagliano et al. 2014).

In Table 6, we also report the minor and major alleles, the **MAF,** and the nature (**intronic, intergenic**) of the published SNPs. Most (83 %) are common SNPs with a MAF > 10 %, ranging from 4 % (Hibar et al. 2013a) to 49 % (Stein et al. 2012). In addition, 19 of the reported SNPs are **intronic** (within gene sequence), with the other 10 being intergenic (between genes). This overrepresentation of intronic SNPs aligns with the global trend observed across all published GWAS (Schaub et al. 2012), though, as already noted above, the **causal variants** could be different from the reported ones or located in different genetic regions (Visscher et al. 2012a; Schaub et al. 2012). Even so, the finding of a large number of causal variants in intergenic regions suggests the need for a better understanding of gene expression regulation in non-coding regions (Edwards et al. 2013). Several techniques, such as **eQTL** or chromosome conformation capture (3C) are now being used to charactrerise non-coding regulatory variants (Edwards et al. 2013). In addition, epigenome-wide association studies (EWAS), where traits are associated with **methylated sites** (Michels et al. 2013; Rakyan et al. 2011), can help characterise the effect of non-coding **epigenetic** variants. However, as recent studies suggest **methylation** of some CpG sites could have a genetic origin (McRae et al. 2014), a better integration of EWAS and GWAS (van der Sijde et al. 2014) may be necessary to disentangle the genetic and **epigenetic** components of a trait.

Also shown in Table 6 are p-values and effect sizes. We show effect sizes as both percentage change in structure size (e.g. from mL or mm^3), and as percentage of variance explained, to allow comparisons across studies. Percentage change in structure was not possible where means and variances were not reported. Most of the SNPs had small effects on structure size, that is, they accounted for less than a 1 % change in size per effect allele. The largest percentage change in size (−5.2 %) was identified in a relatively small sample ($N = 599$) and did not replicate (Paus et al. 2012), and therefore is likely to be overestimated (Ioannidis et al. 2001).

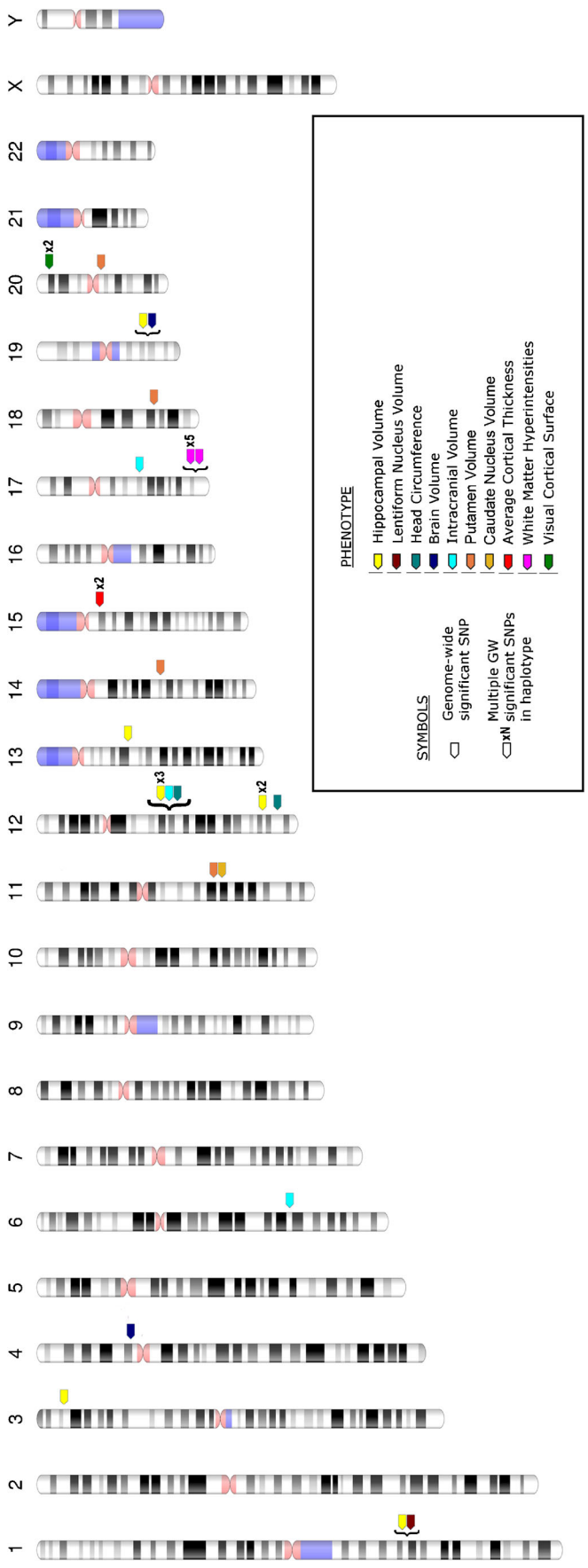


Fig. 6 Karyotype of the genom- wide significant hits of the GWAS on brain morphology. *DTI* diffusion tensor imaging; *PCG* posterior cingulate gyrus; *SPL* superior parietal lobe; *VBM* voxel-based morphometry. For an interactive karyotype of all GWAS findings, see the GWAS diagram browser of the NHGRI GWAS Catalog (Welter et al. 2014; Hindorff et al. 2013)

Percentage of variance explained was also small, being less than 1 % in the larger studies. Much larger estimates of variance explained were found in smaller studies. This phenomenon is well known in genome-wide association studies and reflects a combination of imprecision due to poor power and "winner's curse", where the effect within discovery samples is affected by bias away from the null. The combination of evidence from multiple studies via meta-analysis has been shown to produce more accurate estimates.

It is common for **variants to** have small effects on complex traits, leaving the total variance explained by all the GWAS hits (genome-wide significant heritability or h^2_{GWS}) often far below the heritability measured from twin studies (Craddock et al., 2011; So et al. 2011). For example, genome-wide significant variants currently explain only 7 % of variance for schizophrenia (from 108 genome-wide significant SNPs) (Schizophrenia Working Group of the Psychiatric Genomics Consortium 2014). This gap, often referred to as **missing heritability** (Maher 2008), remains substantial for complex traits, even though the percentage of variance explained by genome-wide associations has risen, as the power of studies has increased (Visscher et al. 2012a).

As we indicated earlier, a greater proportion of variance can be explained by considering all SNPs in aggregate (SNP heritability) using tools such as GCTA (Yang et al. 2010; Yang et al. 2011). Assuming brain structure phenotypes are similar to other complex traits, SNP heritability estimates will likely be about half that estimated from twin studies (Sullivan et al. 2012; Silventoinen et al. 2003). Of the first SNP heritability estimates reported for brain structure (Adib-Samii et al. 2015; Bryant et al. 2013), the best powered study is consistent with this assumption. This study found a SNP heritability of 23 % for white matter hyperintensity, which was considered conservative due to limited SNP coverage (N = 2,336; <500,000 SNPs (Adib-Samii et al. 2015), compared to estimates of 55–80 % from twin and family studies (Atwood et al. 2004; Carmelli et al. 1998; Turner et al. 2004). As GCTA requires thousands of unrelated individuals, the earlier study of SNP heritability for multiple GM and WM volumes was considerably underpowered (N = 747), with confidence intervals covering almost the entire range of possibilities (Bryant et al. 2013). Even so, as larger samples are amassed, it is likely that common SNPs from current genotyping will account for more of the missing heritability in brain morphology.

Accumulating evidence shows that most of the missing heritability is "hidden" (Visscher et al. 2012b): due to low power, and in causal SNPs neither tagged nor imputed (Schaub et al. 2012; Visscher et al. 2012a; 2012b). These include unknown causal SNPs in LD with identified hits, unobserved structural variations (insertions, deletions, inversions and translocations) as well as less common (MAF<1 %)) SNPs (Maher 2008), which could be identified in the future using Whole-Genome Sequencing, (i.e. the entire genome is sequenced) (Bansal et al. 2010; Panoutsopoulou et al. 2013). In addition, non-additive genetic variance could also explain some of the missing heritability (Hemani et al. 2013).

Challenges for Future GWAS in Neuroimaging and Emerging Approaches

A major challenge for genome-wide studies in neuroimaging is to work out how best to handle the high dimensional imaging and genetic data without generating too many false positives or false negatives. Voxel-wise brain maps are an example of high dimensional phenotypes that, when combined with genotype data can cause the number of tests to skyrocket in GWAS. Several multivariate approaches have been tried, for example, parallel independent component analysis (ParaICA) (Liu et al. 2009; Hibar et al. 2011a), sparse-reduced rank regression (sRRR) (Vounou et al. 2010), or reduction of the brain map based on its genetic structure (Chiang et al. 2012; Hibar et al. 2013a). Though sample sizes were limited (less than 500 individuals), these exploratory analyses reported a handful of SNP associations with parietal lobe activation (Liu et al. 2009) or a gene network associated with brain wiring (Chiang et al. 2012). These methods are not easily adapted to meta-analysis and ideally require the sharing of scans and genotypes across studies (Vounou et al. 2010; Liu et al. 2009).

Gene-based analyses can also be used to reduce dimensionality by aggregating SNP effects across genes and excluding intergenic loci (e.g. Kohannim et al. 2012b). While the number of tests performed per phenotype is reduced (there are only around 18,000 genes in the human genome), the number of phenotypes across a voxel-wise brain map is considerable and thus the burden of correcting for these multiple tests remains. As GWAS have become better powered gene-based analyses are often used to complement GWAS, as illustrated in (Hibar et al. 2015). In addition, pathway association analyses, which groups genes in terms of their biological interactions can reduce the multiple testing (Ramanan et al. 2012).

Ultimately, the reduction of dimension, as well as the prioritisation of SNPs or voxels, leads to a (over) simplification of the relations between phenotype and genotype and does not make complete use of the data collected. Thus, the optimal solution is not to reduce the number of tests performed, but rather to correct for multiple testing while taking into account the correlated structure of genotype and brain maps. For example, imaging specialists can evaluate the significance of a signal peak, or that of a cluster of voxels within a brain map (Ge et al. 2012), using results from random field theory (RFT) (Nichols and Hayasaka 2003). A concurrent approach derived from genome-wide significance threshold calculation, has proposed a genome-wide brain-wide significance threshold of 5.0×10^{-12}, based on the empirical evaluation of correlational structure

of voxel level data from two independent cohorts of only 10,000 independent tests brain wide (Medland et al. 2014). Further development of these and new methods should arise from conjoint methodological efforts and the cross-fertilization of ideas between the fields of neuroimaging and genetics.

Lastly, with GWAS findings for brain volumes starting to emerge, it is now possible to construct a **polygenic score** from an individual's SNP data (Wray et al. 2014). These scores are created by summing an individual's effect (risk) alleles for a given phenotype, each weighted by the effect size identified in an independent discovery sample. The set of loci is limited to those with a *p*-value below a defined threshold in the discovery sample, though in principal, all loci can be included. Polygenic scores provide an estimate of the genetic liability for a trait (or disease) and can be used to test associations or to predict a complex trait, even if larger GWAS samples are required to achieve useful levels of prediction (Dudbridge 2013). Ultimately, the quality of a polygenic score depends on the precision of the SNP effect estimates, which requires large(r) GWAS samples. As yet, no polygenic scores have been computed for measures of brain morphology, mainly due to the small number of genome-wide significant hits and small discovery sample sizes, which would contribute to noise in the score. However, a polygenic risk score based on 26 SNPs that have shown replicated associations with Alzheimer's disease (AD) identified an association between cortical thickness and genetic liability for AD (Sabuncu et al. 2012).

Conclusions

The findings emerging from the work reviewed here show that the genetic architecture of the brain is indeed complex. However, sample sizes are increasing in parallel with deep phenotyping and measures that capture the complexity of the brain. Large scale multi-site studies of brain structure phenotypes will enable well powered GWAS to discover more loci, which may be tied in with other information to further our understanding of the function of a particular genetic variant, and link to the neurobiology. The next stage of multi-site studies will also enable us to gain a more complete picture by examining the effect of rare variants, environmental risk factors, and gene x environment interactions, incorporating whole genome sequencing data, as well as the impact of epigenetic factors on brain structure.

Acknowledgments We acknowledge funding support from National Institute of Child Health and Human Development (R01 HD050735), and the National Health and Medical Research Council (NHMRC 486682, 1009064), *Australia*. Baptiste Couvy-Duchesne is supported by a UQ International Scholarship (UQI). Special thanks to Gabriëlla Blokland for personal communications and Seyed Amir Hossein Batouli for providing unpublished results.

Glossary

Bonferroni correction	simplest method of correcting for multiple testing (α/N_{tests}). In GWAS, the number of independent tests has been estimated to 10^6, leading to a genome wide significance threshold of 5×10^{-8}.
Causal variant	a variant that has a direct or indirect functional effect on disease risk. Because of the LD structure of the genome, identifying the causal variant among highly correlated signals is not straightforward. A lower p-value, expression data or a biological understanding of the causal mechanism are only suggestive evidence.
Endophenotypes	endophenotypes are heritable phenotypes; genetically correlated with the disease/complex trait (John and Lewis 1966; Gottesman and Gould 2003).
Epigenetic	molecular process that causes gene expression to change in time through environmental changes in the DNA methylation sites or RNA sequence.
eQTL	(expression Quantitative Trait Loci): Loci that regulate the expression of a gene by regulating the number of micro RNA copies.
(Genetic) Variant	used to refer to observed/tagged SNPs but also to SNPs or any structural variants imputed from the reference panel.
Haplotype	set of SNPs, within one chromosome, that show non random association (linkage disequilibrium). Within one haplotype, SNPs can be highly correlated, which reduces the number of independent testing but creates collinearity issues when simultaneously testing the effect of several SNPs (e.g. gene based tests).
Imputation	the statistical method consisting in inferring a missing value. In genetics, it uses the LD structure of a fully sequenced reference panel to predict the unobserved SNPs (up to 3 millions) based on the set of tagged SNPs (usually 500,000 from a SNP chip).
Linkage Disequilibrium (LD)	non-random association of alleles across a population genome. The LD structure of a genome, defines blocks of strongly correlated alleles called haplotypes.

Intergenic region	DNA region located between genes. Intergenic regions are non-coding and tend to regulate nearby gene expression.
Intronic region	DNA region located inside a gene.
Minor Allele Frequency (MAF)	the frequency of the least frequent nucleotide version at one SNP. MAF can be different across ethnic groups.
Meta-analysis	statistical method that allows merging results from different centres without having to share the (raw) individual data. Only summary statistics of the associations (size effect, number of individuals, p-values) are shared and combined. Studies where the raw data is shared are usually called mega-analyses.
Mode of inheritance	The manner in which a particular genetic trait or disorder is passed from one generation to the next.
(DNA) Methylation	epigenetic process that sees a methyl group binding to the DNA molecule with consequences on gene expression. In the human adult, methylation only happens on sites where *a* cytosine is directly followed by a guanine in the DNA sequence (CpG site). However, highly conserved non-CpG methylation has been shown to accumulate in neurons from fetal to early adult age, to compose the main form of methylation in the neuronal genome (Lister et al. 2013).
Missing Heritability	discrepancy between heritability estimated from twin or family studies and heritability calculated from genome-wide significant SNPs identified through GWAS.
Pleiotropy	One gene influences multiple traits
Polygenic Score	Also known as a genomic profile score. Individual scores are calculated from participant's genotype data by summing the number of effect (risk) alleles weighted by the variant effect size as determined in an independent Discovery sample. The SNP list is limited to variants with a *p*-value less than a defined threshold (or several thresholds may be considered, e.g. <.00001, .001, .001, .01, .1, etc. (Wray et al. 2014)).
SNP	Single Nucleotide Polymorphism, a variant in the genome where the nucleotide can differ between individuals.
Statistical Power	Probability of detecting a significant association, in presence of true association i.e. 1- probability of false negative. In GWAS, the power is specific to a SNP and depends on the level of statistical significance (α), the SNP effect size and MAF, the sample size.

References

Aas M., Haukvik U. K., Djurovic S., Bergmann O., Athanasiu L., Tesli M. S., et al. (2013). BDNF val66met modulates the association between childhood trauma, cognitive and brain abnormalities in psychoses. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, *46*, 181–188. doi:10.1016/j.pnpbp.2013.07.008.

Adachi N., Numakawa T., Richards M., Nakajima S., & Kunugi H. (2014). New insight in expression, transport, and secretion of brain-derived neurotrophic factor: Implications in brain-related diseases. *World Journal of Biological Chemistry*, *5*(4), 409–428. doi:10.4331/wjbc.v5.i4.409.

Addington A. M., & Rapoport J. L. (2012). Annual research review: impact of advances in genetics in understanding developmental psychopathology. *Journal of Child Psychology and Psychiatry*, *53*(5), 510–518. doi:10.1111/j.1469-7610.2011.02478.x.

Adib-Samii P., Devan W., Traylor M., Lanfranconi S., Zhang C. R., Cloonan L., et al. (2015). Genetic architecture of white matter hyperintensities differs in hypertensive and nonhypertensive ischemic stroke. *Stroke*, *46*(2), 348–353. doi:10.1161/strokeaha.114.006849.

Allen S. J., Watson J. J., Shoemark D. K., Barua N. U., & Patel N. K. (2013). GDNF, NGF and BDNF as therapeutic options for neurodegeneration. *Pharmacology & Therapeutics*, *138*(2), 155–175. doi:10.1016/j.pharmthera.2013.01.004.

Anderson C. A., Pettersson F. H., Clarke G. M., Cardon L. R., Morris A. P., & Zondervan K. T. (2010). Data quality control in genetic case–control association studies. *Nature Protocols*, *5*(9), 1564–1573. doi:10.1038/nprot.2010.116.

Antypa N., Drago A., & Serretti A. (2013). The role of COMT gene variants in depression: bridging neuropsychological, behavioral and clinical phenotypes. *Neuroscience and Biobehavioral Reviews*, *37*(8), 1597–1610. doi:10.1016/j.neubiorev.2013.06.006.

Atwood L. D., Wolf P. A., Heard-Costa N. L., Massaro J. M., Beiser A., D'Agostino R. B., et al. (2004). Genetic variation in white matter hyperintensity volume in the Framingham Study. *Stroke*, *35*(7), 1609–1613. doi:10.1161/01.str.0000129643.77045.10.

Baare W. F., Hulshoff Pol H. E., Boomsma D. I., Posthuma D., de Geus E. J., Schnack H. G., et al. (2001). Quantitative genetic modeling of variation in human brain morphology. *Cerebral Cortex*, *11*(9), 816–824.

Bakken T. E., Bloss C. S., Roddey J. C., Joyner A. H., Rimol L. M., Djurovic S., et al. (2011). Association of genetic variants on 15q12 with cortical thickness and cognition in schizophrenia. *Archives of General Psychiatry*, *68*(8), 781–790. doi:10.1001/archgenpsychiatry.2011.81.

Bakken T. E., Roddey J. C., Djurovic S., Akshoomoff N., Amaral D. G., Bloss C. S., et al. (2012). Association of common genetic variants in GPCPD1 with scaling of visual cortical surface area in humans. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(10), 3985–3990. doi:10.1073/pnas.1105829109.

Bansal, V., Libiger, O., Torkamani, A., & Schork, N. J. (2010). Statistical analysis strategies for association studies involving rare variants. Nature Reviews. Genetics, 11(11), 773–785, doi:10.1038/nrg2867.

Baranzini S. E., Wang J., Gibson R. A., Galwey N., Naegelin Y., Barkhof F., et al. (2009). Genome-wide association analysis of susceptibility and clinical phenotype in multiple sclerosis. *Human Molecular Genetics*, *18*(4), 767–778. doi:10.1093/hmg/ddn388.

Baranzini S. E., Srinivasan R., Khankhanian P., Okuda D. T., Nelson S. J., Matthews P. M., et al. (2010). Genetic variation influences glutamate concentrations in brains of patients with multiple sclerosis. *Brain*, *133*(9), 2603–2611. doi:10.1093/brain/awq192.

Barnes A., Isohanni M., Barnett J. H., Pietilainen O., Veijola J., Miettunen J., et al. (2012). No association of COMT (Val158Met) genotype with brain structure differences between men and women. *PloS One*, *7*(3), e33964. doi:10.1371/journal.pone.0033964.

Bartley A. J., Jones D. W., & Weinberger D. R. (1997). Genetic variability of human brain size and cortical gyral patterns. *Brain*, *120*(Pt 2), 257–269.

Bates T. C., Lewis G. J., & Weiss A. (2013). Childhood socioeconomic status amplifies genetic effects on adult intelligence. *Psychological Science*, *24*(10), 2111–2116. doi:10.1177/0956797613488394.

Batouli, S. A., Sachdev, P. S., Wen, W., Wright, M. J., Ames, D., & Trollor, J. N. (2014). Heritability of brain volumes in older adults: the Older Australian Twins Study. *Neurobiol Aging, 35*(4), 937.e935-918, doi:10.1016/j.neurobiolaging.2013.10.079.

Begum F., Ghosh D., Tseng G. C., & Feingold E. (2012). Comprehensive literature review and statistical considerations for GWAS meta-analysis. *Nucleic Acids Research*, *40*(9), 3777–3784. doi:10.1093/nar/gkr1255.

Bentham J., & Vyse T. J. (2013). The development of genome-wide association studies and their application to complex diseases, including lupus. *Lupus*, *22*(12), 1205–1213. doi:10.1177/0961203313492870.

Benyamin B., Visscher P. M., & McRae A. F. (2009). Family-based genome-wide association studies. *Pharmacogenomics*, *10*(2), 181–190. doi:10.2217/14622416.10.2.181.

Betjemann R. S., Johnson E. P., Barnard H., Boada R., Filley C. M., Filipek P. A., et al. (2010). Genetic covariation between brain volumes and IQ, reading performance, and processing speed. *Behavior Genetics*, *40*(2), 135–145. doi:10.1007/s10519-009-9328-2.

Bigos K. L., & Hariri A. R. (2007). Neuroimaging: technologies at the interface of genes, brain, and behavior. *Neuroimaging Clinics of North America*, *17*(4), 459–viii467. doi:10.1016/j.nic.2007.09.005.

Bigos K. L., & Weinberger D. R. (2010). Imaging genetics-days of future past. *NeuroImage*, *53*(3), 804–809. doi:10.1016/j.neuroimage.2010.01.035.

Bis J. C., DeCarli C., Smith A. V., van der Lijn F., Crivello F., Fornage M., et al. (2012). Common variants at 12q14 and 12q24 are associated with hippocampal volume. *Nature Genetics*, *44*(5), 545–551. doi:10.1038/ng.2237.

Blokland G. A., McMahon K. L., Thompson P. M., Martin N. G., de Zubicaray G. I., & Wright M. J. (2011). Heritability of working memory brain activation. *The Journal of Neuroscience*, *31*(30), 10882–10890. doi:10.1523/jneurosci.5334-10.2011.

Blokland G. A., 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.

Blokland G. A., McMahon K. L., Thompson P. M., Hickie I. B., Martin N. G., de Zubicaray G. I., et al. (2014). Genetic effects on the cerebellar role in working memory: same brain, different genes? *NeuroImage*, *86*, 392–403. doi:10.1016/j.neuroimage.2013.10.006.

Bochud, M. (2012). Estimating Heritability from Nuclear Family and Pedigree Data. In R. C. Elston, J. M. Satagopan, & S. Sun (Eds.), *Statistical Human Genetics* (Vol. 850, pp. 171–186, Methods in Molecular Biology). Humana Press.

Bohlken M. M., Brouwer R. M., Mandl R. C., van Haren N. E., Brans R. G., van Baal G. C., et al. (2014). Genes contributing to subcortical volumes and intellectual ability implicate the thalamus. *Human Brain Mapping*, *35*(6), 2632–2642. doi:10.1002/hbm.22356.

Bookheimer S. Y., Strojwas M. H., Cohen M. S., Saunders A. M., Pericak-Vance M. A., Mazziotta J. C., et al. (2000). Patterns of brain activation in people at risk for Alzheimer's disease. *The New England Journal of Medicine*, *343*(7), 450–456. doi:10.1056/NEJM200008173430701.

Bosker F. J., Hartman C. A., Nolte I. M., Prins B. P., Terpstra P., Posthuma D., et al. (2011). Poor replication of candidate genes for major depressive disorder using genome-wide association data. *Molecular Psychiatry*, *16*(5), 516–532. doi:10.1038/mp.2010.38.

Brans R. G., Kahn R. S., Schnack H. G., van Baal G. C., Posthuma D., van Haren N. E., et al. (2010). Brain plasticity and intellectual ability are influenced by shared genes. *The Journal of Neuroscience*, *30*(16), 5519–5524. doi:10.1523/jneurosci.5841-09.2010.

Broer L., Lill C. M., Schuur M., Amin N., Roehr J. T., Bertram L., et al. (2013). Distinguishing true from false positives in genomic studies: p values. *European Journal of Epidemiology*, *28*(2), 131–138. doi:10.1007/s10654-012-9755-x.

Brooks S. J., Nilsson E. K., Jacobsson J. A., Stein D. J., Fredriksson R., Lind L., et al. (2014). BDNF polymorphisms are linked to poorer working memory performance, reduced cerebellar and hippocampal volumes and differences in prefrontal cortex in a Swedish elderly population. *PloS One*, *9*(1), e82707. doi:10.1371/journal.pone.0082707.

Brouwer R. M., Hedman A. M., van Haren N. E., Schnack H. G., Brans R. G., Smit D. J., et al. (2014a). Heritability of brain volume change and its relation to intelligence. *NeuroImage*. doi:10.1016/j.neuroimage.2014.04.072.

Brouwer R. M., van Soelen I. L., Swagerman S. C., Schnack H. G., Ehli E. A., Kahn R. S., et al. (2014b). Genetic associations between intelligence and cortical thickness emerge at the start of puberty. *Human Brain Mapping*, *35*(8), 3760–3773. doi:10.1002/hbm.22435.

Brouwer, R. M., Glahn, D. C., Hibar, D. P., Hua, X., Jahanshad, N., Franz, C. E., et al. (2015). Genetic influences on longitudinal changes in subcortical volumes: Results of the ENIGMA Plasticity Working Group. Paper presented at the Organization for Human Brain Mapping 2015,

Brown A. A., Jensen J., Nikolova Y. S., Djurovic S., Agartz I., Server A., et al. (2012). Genetic variants affecting the neural processing of human facial expressions: evidence using a genome-wide functional imaging approach. *Translational Psychiatry*, *2*, e143. doi:10.1038/tp.2012.67.

Brown B. M., Bourgeat P., Peiffer J. J., Burnham S., Laws S. M., Rainey-Smith S. R., et al. (2014). Influence of BDNF Val66Met on the relationship between physical activity and brain volume. *Neurology*, *83*(15), 1345–1352. doi:10.1212/WNL.0000000000000867.

Bryant, C., Giovanello, K. S., Ibrahim, J. G., Chang, J., Shen, D., Peterson, B. S., et al. (2013). Mapping the Genetic Variation of Regional Brain Volumes as Explained by All Common SNPs from the ADNI Study. PloS One, 8(8), e71723, doi:10.1371/journal.pone.0071723.

Campbell C. D., Ogburn E. L., Lunetta K. L., Lyon H. N., Freedman M. L., Groop L. C., et al. (2005). Demonstrating stratification in a European American population. *Nature Genetics*, *37*(8), 868–872. doi:10.1038/ng1607.

Carmelli D., DeCarli C., Swan G. E., Jack L. M., Reed T., Wolf P. A., et al. (1998). Evidence for genetic variance in white matter hyperintensity volume in normal elderly male twins. *Stroke*, *29*(6), 1177–1181.

Carmelli D., Reed T., & DeCarli C. (2002a). A bivariate genetic analysis of cerebral white matter hyperintensities and cognitive performance in elderly male twins. *Neurobiology of Aging*, *23*(3), 413–420.

Carmelli D., Swan G. E., DeCarli C., & Reed T. (2002b). Quantitative genetic modeling of regional brain volumes and cognitive performance in older male twins. *Biological Psychology*, *61*(1–2), 139–155.

Chen W. M., & Abecasis G. R. (2007). Family-based association tests for genomewide association scans. *American Journal of Human Genetics*, *81*(5), 913–926. doi:10.1086/521580.

Chen C. H., Panizzon M. S., Eyler L. T., Jernigan T. L., Thompson W., Fennema-Notestine C., et al. (2011). Genetic influences on cortical regionalization in the human brain. *Neuron*, *72*(4), 537–544. doi:10.1016/j.neuron.2011.08.021.

Chen C. H., Gutierrez E. D., Thompson W., Panizzon M. S., Jernigan T. L., Eyler L. T., et al. (2012). Hierarchical genetic organization of human cortical surface area. *Science*, *335*(6076), 1634–1636. doi:10.1126/science.1215330.

Chen C. H., Fiecas M., Gutierrez E. D., Panizzon M. S., Eyler L. T., Vuoksimaa E., et al. (2013). Genetic topography of brain morphology. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(42), 17089–17094. doi:10.1073/pnas.1308091110.

Cherbuin N., Leach L. S., Christensen H., & Anstey K. J. (2007). Neuroimaging and APOE genotype: a systematic qualitative review. *Dementia and Geriatric Cognitive Disorders*, *24*(5), 348–362. doi:10.1159/000109150.

Chetelat G., & Fouquet M. (2013). Neuroimaging biomarkers for Alzheimer's disease in asymptomatic APOE4 carriers. *Revue Neurologique (Paris)*, *169*(10), 729–736. doi:10.1016/j.neurol.2013.07.025.

Chiang M. C., Barysheva M., Shattuck D. W., Lee A. D., Madsen S. K., Avedissian C., et al. (2009). Genetics of brain fiber architecture and intellectual performance. *The Journal of Neuroscience*, *29*(7), 2212–2224. doi:10.1523/jneurosci.4184-08.2009.

Chiang M. C., McMahon K. L., de Zubicaray G. I., Martin N. G., Hickie I., Toga A. W., et al. (2011). Genetics of white matter development: a DTI study of 705 twins and their siblings aged 12 to 29. *NeuroImage*, *54*(3), 2308–2317. doi:10.1016/j.neuroimage.2010.10.015.

Chiang M. C., Barysheva M., McMahon K. L., de Zubicaray G. I., Johnson K., Montgomery G. W., et al. (2012). Gene network effects on brain microstructure and intellectual performance identified in 472 twins. *The Journal of Neuroscience*, *32*(25), 8732–8745. doi:10.1523/jneurosci.5993-11.2012.

Craddock, R. C., LaConte, S., Castellanos, F. X., Zuo, X. N., Thompson, P. M., de Zubicaray, G. I., et al. (2011) Genetics influence inter-subject brain state prediction. In *Organization for human brain mapping,* Quebec City, June 2011.

de Bakker P. I., Ferreira M. A., Jia X., Neale B. M., Raychaudhuri S., & Voight B. F. (2008). Practical aspects of imputation-driven meta-analysis of genome-wide association studies. *Human Molecular Genetics*, *17*(R2), R122–R128. doi:10.1093/hmg/ddn288.

de Geus E. J. (2010). From genotype to EEG endophenotype: a route for post-genomic understanding of complex psychiatric disease? *Genome Medicine*, *2*(9), 63. doi:10.1186/gm184.

de Groot J. C., de Leeuw F. E., Oudkerk M., van Gijn J., Hofman A., Jolles J., et al. (2000). Cerebral white matter lesions and cognitive function: the Rotterdam Scan Study. *Annals of Neurology*, *47*(2), 145–151.

de Zubicaray G. I., Chiang M. C., McMahon K. L., Shattuck D. W., Toga A. W., Martin N. G., et al. (2008). Meeting the challenges of neuroimaging genetics. *Brain Imaging and Behavior*, *2*(4), 258–263.

Dean 3rd D. C., Jerskey B. A., Chen K., Protas H., Thiyyagura P., Roontiva A., et al. (2014). Brain differences in infants at differential genetic risk for late-onset Alzheimer disease: a cross-sectional imaging study. *JAMA Neurology*, *71*(1), 11–22. doi:10.1001/jamaneurol.2013.4544.

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.

Debette S., Bis J. C., Fornage M., Schmidt H., Ikram M. A., Sigurdsson S., et al. (2010). Genome-wide association studies of MRI-defined brain infarcts: meta-analysis from the CHARGE consortium. *Stroke*, *41*(2), 210–217. doi:10.1161/STROKEAHA.109.569194.

den Braber, A., Bohlken, M. M., Brouwer, R. M., van't Ent D., Kanai, R., Kahn, R. S., et al. (2013). Heritability of subcortical brain measures: a perspective for future genome-wide association studies. NeuroImage, 83, 98–102, doi:10.1016/j.neuroimage.2013.06.027.

Dennison M., Whittle S., Yucel M., Vijayakumar N., Kline A., Simmons J., et al. (2013). Mapping subcortical brain maturation during adolescence: evidence of hemisphere- and sex-specific longitudinal changes. *Developmental Science*, *16*(5), 772–791. doi:10.1111/desc.12057.

DeStefano A. L., Seshadri S., Beiser A., Atwood L. D., Massaro J. M., Au R., et al. (2009). Bivariate heritability of total and regional brain volumes: the Framingham Study. *Alzheimer disease and Associated Disorders*, *23*(3), 218–223. doi:10.1097/WAD.0b013e31819cadd8.

Devlin B., & Roeder K. (1999). Genomic control for association studies. *Biometrics*, *55*(4), 997–1004.

Dudbridge F. (2013). Power and predictive accuracy of polygenic risk scores. *PLoS Genetics*, *9*(3), e1003348. doi:10.1371/journal.pgen.1003348.

Dudbridge F., & Gusnanto A. (2008). Estimation of significance thresholds for genomewide association scans. *Genetic Epidemiology*, *32*(3), 227–234. doi:10.1002/gepi.20297.

Duff, B. J., Macritchie, K. A., Moorhead, T. W., Lawrie, S. M., & Blackwood, D. H. (2013). Human brain imaging studies of DISC1 in schizophrenia, bipolar disorder and depression: a systematic review. Schizophrenia Research, 147(1), 1–13, doi:10.1016/j.schres.2013.03.015.

Durston S. (2010). Imaging genetics in ADHD. *NeuroImage*, *53*(3), 832–838. doi:10.1016/j.neuroimage.2010.02.071.

Edwards S. L., Beesley J., French J. D., & Dunning A. M. (2013). Beyond GWASs: illuminating the dark road from association to function. *American Journal of Human Genetics*, *93*(5), 779–797. doi:10.1016/j.ajhg.2013.10.012.

Erhardt A., & Spoormaker V. I. (2013). Translational approaches to anxiety: focus on genetics, fear extinction and brain imaging. *Current Psychiatry Reports*, *15*(12), 417. doi:10.1007/s11920-013-0417-9.

Evangelou E., & Ioannidis J. P. (2013). Meta-analysis methods for genome-wide association studies and beyond. *Nature Reviews. Genetics*, *14*(6), 379–389. doi:10.1038/nrg3472.

Evans D. M., Gillespie N. A., & Martin N. G. (2002). Biometrical genetics. *Biological Psychology*, *61*(1–2), 33–51.

Eyler L. T., Prom-Wormley E., Fennema-Notestine C., Panizzon M. S., Neale M. C., Jernigan T. L., et al. (2011). Genetic patterns of correlation among subcortical volumes in humans: results from a magnetic resonance imaging twin study. *Human Brain Mapping*, *32*(4), 641–653. doi:10.1002/hbm.21054.

Flint J., & Munafo M. R. (2013). Candidate and non-candidate genes in behavior genetics. *Current Opinion in Neurobiology*, *23*(1), 57–61. doi:10.1016/j.conb.2012.07.005.

Fornage M., Debette S., Bis J. C., Schmidt H., Ikram M. A., Dufouil C., et al. (2011). Genome-wide association studies of cerebral white matter lesion burden: the CHARGE consortium. *Annals of Neurology*, *69*(6), 928–939. doi:10.1002/ana.22403.

Frangou, S. (2014). Developing individualized diagnostic and therapeutic tools with images: The IMAGEMEND network. In *Biol Psychiatry, 2014* (Vol. 75, pp. 142S-142S, Vol. 9). ELSEVIER SCIENCE INC 360 PARK AVE SOUTH, NEW YORK, NY 10010–1710 USA.

Furney S. J., Simmons A., Breen G., Pedroso I., Lunnon K., Proitsi P., et al. (2011). Genome-wide association with MRI atrophy measures

as a quantitative trait locus for Alzheimer's disease. *Molecular Psychiatry*, *16*(11), 1130–1138. doi:10.1038/mp.2010.123.

Gagliano S. A., Barnes M. R., Weale M. E., & Knight J. (2014). A Bayesian method to incorporate hundreds of functional characteristics with association evidence to improve variant prioritization. *PloS One*, *9*(5), e98122. doi:10.1371/journal.pone.0098122.

Galesloot T. E., van Steen K., Kiemeney L. A., Janss L. L., & Vermeulen S. H. (2014). A comparison of multivariate genome-wide association methods. *PloS One*, *9*(4), e95923. doi:10.1371/journal.pone.0095923.

Gatt J. M., Korgaonkar M. S., Schofield P. R., Harris A., Clark C. R., Oakley K. L., et al. (2012). The TWIN-E project in emotional wellbeing: study protocol and preliminary heritability results across four MRI and DTI measures. *Twin Research and Human Genetics*, *15*(3), 419–441. doi:10.1017/thg.2012.12.

Gauderman W. J., Zhang P., Morrison J. L., & Lewinger J. P. (2013). Finding novel genes by testing G x E interactions in a genome-wide association study. *Genetic Epidemiology*, *37*(6), 603–613. doi:10.1002/gepi.21748.

Ge T., Feng J., Hibar D. P., Thompson P. M., & Nichols T. E. (2012). Increasing power for voxel-wise genome-wide association studies: the random field theory, least square kernel machines and fast permutation procedures. *NeuroImage*, *63*(2), 858–873. doi:10.1016/j.neuroimage.2012.07.012.

Genomes Project, C., Abecasis, G. R., Altshuler, D., Auton, A., Brooks, L. D., Durbin, R. M., et al. (2010). A map of human genome variation from population-scale sequencing. Nature, 467(7319), 1061–1073, doi:10.1038/nature09534.

Giedd J. N., Schmitt J. E., & Neale M. C. (2007). Structural brain magnetic resonance imaging of pediatric twins. *Human Brain Mapping*, *28*(6), 474–481. doi:10.1002/hbm.20403.

Gilmore J. H., Schmitt J. E., Knickmeyer R. C., Smith J. K., Lin W., Styner M., et al. (2010). Genetic and environmental contributions to neonatal brain structure: a twin study. *Human Brain Mapping*, *31*(8), 1174–1182. doi:10.1002/hbm.20926.

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

Glahn D. C., Kent Jr. J. W., Sprooten E., Diego V. P., Winkler A. M., Curran J. E., et al. (2013). Genetic basis of neurocognitive decline and reduced white-matter integrity in normal human brain aging. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(47), 19006–19011. doi:10.1073/pnas.1313735110.

Glahn, D. C., Winkler, A. M., Kochunov, P., Almasy, L., Duggirala, R., Carless, M. A., et al. (2010). Genetic control over the resting brain. *Proc Natl Acad Sci U S A, 107*(3), 1223–1228. doi:10.1073/pnas.0909969107.

Gondro C., Lee S. H., Lee H. K., & Porto-Neto L. R. (2013). Quality control for genome-wide association studies. *Methods in Molecular Biology*, *1019*, 129–147. doi:10.1007/978-1-62703-447-0_5.

Gottesman I. I., & Gould T. D. (2003). The endophenotype concept in psychiatry: etymology and strategic intentions. *The American Journal of Psychiatry*, *160*(4), 636–645.

Gottesman I. I., & Shields J. (1967). A polygenic theory of schizophrenia. *Proceedings of the National Academy of Sciences of the United States of America*, *58*(1), 199–205.

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 VERS ION on anatomical volume and cortical thickness measurements. PloS One, 7(6), e38234.

Han X., Jovicich J., Salat D., van der Kouwe A., Quinn B., Czanner S., et al. (2006). Reliability of MRI-derived measurements of human cerebral cortical thickness: the effects of field strength, scanner upgrade and manufacturer. *NeuroImage*, *32*(1), 180–194. doi:10.1016/j.neuroimage.2006.02.051.

Harrisberger F., Spalek K., Smieskova R., Schmidt A., Coynel D., Milnik A., et al. (2014). The association of the BDNF Val66Met polymorphism and the hippocampal volumes in healthy humans: a joint meta-analysis of published and new data. *Neuroscience and Biobehavioral Reviews*, *42*, 267–278. doi:10.1016/j.neubiorev.2014.03.011.

Hart A. B., de Wit H., & Palmer A. A. (2013). Candidate gene studies of a promising intermediate phenotype: failure to replicate. *Neuropsychopharmacology*, *38*(5), 802–816. doi:10.1038/npp.2012.245.

Hasler G., & Northoff G. (2011). Discovering imaging endophenotypes for major depression. *Molecular Psychiatry*, *16*(6), 604–619. doi:10.1038/mp.2011.23.

Hass J., Walton E., Kirsten H., Liu J., Priebe L., Wolf C., et al. (2013). A genome-wide association study suggests novel loci associated with a schizophrenia-related brain-based phenotype. *PloS One*, *8*(6), e64872. doi:10.1371/journal.pone.0064872.

Hemani G., Knott S., & Haley C. (2013). An evolutionary perspective on epistasis and the missing heritability. *PLoS Genetics*, *9*(2), e1003295. doi:10.1371/journal.pgen.1003295.

Hewitt, J. K. (2012). Editorial policy on candidate gene association and candidate gene-by-environment interaction studies of complex traits. Behavior Genetics, 42(1), 1–2, doi:10.1007/s10519-011-9504-z.

Hibar D. P., Kohannim O., Stein J. L., Chiang M. C., & Thompson P. M. (2011a). Multilocus genetic analysis of brain images. *Frontiers in Genetics*, *2*, 73. doi:10.3389/fgene.2011.00073.

Hibar D. P., Stein J. L., Kohannim O., Jahanshad N., Saykin A. J., Shen L., et al. (2011b). Voxelwise gene-wide association study (vGeneWAS): multivariate gene-based association testing in 731 elderly subjects. *NeuroImage*, *56*(4), 1875–1891. doi:10.1016/j.neuroimage.2011.03.077.

Hibar D. P., Medland S. E., Stein J. L., Kim S., Shen L., Saykin A. J., et al. (2013a). Genetic clustering on the hippocampal surface for genome-wide association studies. *Medical Image Computing and Computer-Assisted Intervention*, *16*(Pt 2), 690–697.

Hibar D. P., Stein J. L., Ryles A. B., Kohannim O., Jahanshad N., Medland S. E., et al. (2013b). Genome-wide association identifies genetic variants associated with lentiform nucleus volume in $N = 1345$ young and elderly subjects. *Brain Imaging and Behavior*, *7*(2), 102–115. doi:10.1007/s11682-012-9199-7.

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

Hindorff, L. A., MacArthur, J., Morales, J., Junkins, H. A., Hall, P. N., Klemm, A. K., et al. (2013). A Catalog of published genome-wide association studies. http://www.genome.gov/gwastudies. Accessed 20 Jan 2015.

Hirschhorn J. N., & Daly M. J. (2005). Genome-wide association studies for common diseases and complex traits. *Nature Reviews. Genetics*, *6*(2), 95–108. doi:10.1038/nrg1521.

Hohman T. J., Koran M. E., Thornton-Wells T. A., & Alzheimer's Disease Neuroimaging I. (2014). Genetic modification of the relationship between phosphorylated tau and neurodegeneration. *Alzheimers Dement*. doi:10.1016/j.jalz.2013.12.022.

Hostage C. A., Choudhury K. R., Murali Doraiswamy P., Petrella J. R., & Alzheimer's Disease Neuroimaging I. (2014). Mapping the effect of the apolipoprotein E genotype on 4-year atrophy rates in an Alzheimer disease-related brain network. *Radiology*, *271*(1), 211–219. doi:10.1148/radiol.13131041.

Hulshoff Pol H. E., Schnack H. G., Posthuma D., Mandl R. C., Baare W. F., van Oel C., et al. (2006). Genetic contributions to human brain

morphology and intelligence. *The Journal of Neuroscience*, *26*(40), 10235–10242. doi:10.1523/jneurosci.1312-06.2006.

Iacono W. G., Vaidyanathan U., Vrieze S. I., & Malone S. M. (2014). Knowns and unknowns for psychophysiological endophenotypes: integration and response to commentaries. *Psychophysiology*, *51*(12), 1339–1347. doi:10.1111/psyp.12358.

Ignacio Z. M., Reus G. Z., Abelaira H. M., & Quevedo J. (2014). Epigenetic and epistatic interactions between serotonin transporter and brain-derived neurotrophic factor genetic polymorphism: Insights in depression. *Neuroscience*, *275C*, 455–468. doi:10.1016/j.neuroscience.2014.06.036.

Ikram M. A., Fornage M., Smith A. V., Seshadri S., Schmidt R., Debette S., et al. (2012). Common variants at 6q22 and 17q21 are associated with intracranial volume. *Nature Genetics*, *44*(5), 539–544. doi:10.1038/ng.2245.

Inoue K., & Lupski J. R. (2003). Genetics and genomics of behavioral and psychiatric disorders. *Current Opinion in Genetics & Development*, *13*(3), 303–309.

International HapMap, C (2005). A haplotype map of the human genome. *Nature*, *437*(7063), 1299–1320. doi:10.1038/nature04226.

Ioannidis J. P., Ntzani E. E., Trikalinos T. A., & Contopoulos-Ioannidis D. G. (2001). Replication validity of genetic association studies. *Nature Genetics*, *29*(3), 306–309. doi:10.1038/ng749.

Ioannidis J. P., Boffetta P., Little J., O'Brien T. R., Uitterlinden A. G., Vineis P., et al. (2008). Assessment of cumulative evidence on genetic associations: interim guidelines. *International Journal of Epidemiology*, *37*(1), 120–132. doi:10.1093/ije/dym159.

Iofrida C., Palumbo S., & Pellegrini S. (2014). Molecular genetics and antisocial behavior: where do we stand? *Experimental Biology and Medicine (Maywood, N.J.)*, *239*(11), 1514–1523. doi:10.1177/1535370214529508.

Ira E., Zanoni M., Ruggeri M., Dazzan P., & Tosato S. (2013). COMT, neuropsychological function and brain structure in schizophrenia: a systematic review and neurobiological interpretation. *Journal of Psychiatry & Neuroscience*, *38*(6), 366–380. doi:10.1503/jpn.120178.

Jack Jr. C. R., Barkhof F., Bernstein M. A., Cantillon M., Cole P. E., Decarli C., et al. (2011). Steps to standardization and validation of hippocampal volumetry as a biomarker in clinical trials and diagnostic criterion for Alzheimer's disease. *Alzheimers Dement*, *7*(4), 474–e474485. doi:10.1016/j.jalz.2011.04.007.

Jahanshad, N., Bhatt, P., Hibar, D., Villalon, J., Nir, T., Toga, A., et al. (2013a). Bivariate genome-wide association study of genetically correlated neuroimaging phenotypes from DTI and MRI through a seemingly unrelated regression model. In L. Shen, T. Liu, P.-T. Yap, H. Huang, D. Shen, & C.-F. Westin (Eds.), *Multimodal Brain Image Analysis* (Vol. 8159, pp. 189–201, Lecture Notes in Computer Science). Springer International Publishing.

Jahanshad N., Rajagopalan P., Hua X., Hibar D. P., Nir T. M., Toga A. W., et al. (2013b). Genome-wide scan of healthy human connectome discovers SPON1 gene variant influencing dementia severity. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(12), 4768–4773. doi:10.1073/pnas.1216206110.

John B., & Lewis K. R. (1966). Chromosome variability and geographic distribution in insects. *Science*, *152*(3723), 711–721. doi:10.1126/science.152.3723.711.

Jovicich J., Czanner S., Han X., Salat D., van der Kouwe A., Quinn B., et al. (2009). MRI-derived measurements of human subcortical, ventricular and intracranial brain volumes: Reliability effects of scan sessions, acquisition sequences, data analyses, scanner upgrade, scanner vendors and field strengths. *NeuroImage*, *46*(1), 177–192. doi:10.1016/j.neuroimage.2009.02.010.

Karlsgodt K. H., Kochunov P., Winkler A. M., Laird A. R., Almasy L., Duggirala R., et al. (2010). A multimodal assessment of the genetic control over working memory. *The Journal of Neuroscience*, *30*(24), 8197–8202. doi:10.1523/jneurosci.0359-10.2010.

Khan W., Giampietro V., Ginestet C., Dell'Acqua F., Bouls D., Newhouse S., et al. (2014). No differences in hippocampal volume between carriers and non-carriers of the ApoE epsilon4 and epsilon2 alleles in young healthy adolescents. *Journal of Alzheimer's disease*, *40*(1), 37–43. doi:10.3233/JAD-131841.

Kidd K. K., Speed W. C., Pakstis A. J., Furtado M. R., Fang R., Madbouly A., et al. (2014). Progress toward an efficient panel of SNPs for ancestry inference. *Forensic Science International. Genetics*, *10*, 23–32. doi:10.1016/j.fsigen.2014.01.002.

Kindt A. S., Navarro P., Semple C. A., & Haley C. S. (2013). The genomic signature of trait-associated variants. *BMC Genomics*, *14*, 108. doi:10.1186/1471-2164-14-108.

Klein R. J., Zeiss C., Chew E. Y., Tsai J. Y., Sackler R. S., Haynes C., et al. (2005). Complement factor H polymorphism in age-related macular degeneration. *Science*, *308*(5720), 385–389. doi:10.1126/science.1109557.

Knickmeyer R. C., Wang J., Zhu H., Geng X., Woolson S., Hamer R. M., et al. (2014). Common variants in psychiatric risk genes predict brain structure at birth. *Cerebral Cortex*, *24*(5), 1230–1246. doi:10.1093/cercor/bhs401.

Kochunov P., Glahn D., Winkler A., Duggirala R., Olvera R. L., Cole S., et al. (2009). Analysis of genetic variability and whole genome linkage of whole-brain, subcortical, and ependymal hyperintense white matter volume. *Stroke*, *40*(12), 3685–3690. doi:10.1161/strokeaha.109.565390.

Kochunov P., Glahn D. C., Lancaster J. L., Winkler A. M., Smith S., Thompson P. M., et al. (2010). Genetics of microstructure of cerebral white matter using diffusion tensor imaging. *NeuroImage*, *53*(3), 1109–1116. doi:10.1016/j.neuroimage.2010.01.078.

Kochunov P., Glahn D. C., Nichols T. E., Winkler A. M., Hong E. L., Holcomb H. H., et al. (2011). Genetic analysis of cortical thickness and fractional anisotropy of water diffusion in the brain. *Frontiers in Neuroscience*, *5*, 120. doi:10.3389/fnins.2011.00120.

Kochunov P., Jahanshad N., Sprooten E., Nichols T. E., Mandl R. C., Almasy L., et al. (2014). Multi-site study of additive genetic effects on fractional anisotropy of cerebral white matter: Comparing meta and megaanalytical approaches for data pooling. *NeuroImage*, *95*, 136–150. doi:10.1016/j.neuroimage.2014.03.033.

Kohannim, O., Hibar, D. P., Jahanshad, N., Stein, J. L., Hua, X., Toga, A. W., et al. (2012a) Predicting temporal lobe volume on MRI from genotypes using L 1-L 2 regularized regression. In *Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on, 2012a* (pp. 1160–1163). IEEE.

Kohannim O., Hibar D. P., Stein J. L., Jahanshad N., Hua X., Rajagopalan P., et al. (2012b). Discovery and replication of gene influences on brain structure using LASSO regression. *Frontiers in Neuroscience*, *6*, 115. doi:10.3389/fnins.2012.00115.

Koolschijn P. C., & Crone E. A. (2013). Sex differences and structural brain maturation from childhood to early adulthood. *Developmental Cognitive Neuroscience*, *5*, 106–118. doi:10.1016/j.dcn.2013.02.003.

Koten Jr. J. W., Wood G., Hagoort P., Goebel R., Propping P., Willmes K., et al. (2009). Genetic contribution to variation in cognitive function: an FMRI study in twins. *Science*, *323*(5922), 1737–1740. doi:10.1126/science.1167371.

Kremen W. S., Prom-Wormley E., Panizzon M. S., Eyler L. T., Fischl B., Neale M. C., et al. (2010). Genetic and environmental influences on the size of specific brain regions in midlife: the VETSA MRI study. *NeuroImage*, *49*(2), 1213–1223. doi:10.1016/j.neuroimage.2009.09.043.

Kremen W. S., Franz C. E., & Lyons M. J. (2013). VETSA: the Vietnam Era twin study of aging. *Twin Research and Human Genetics*, *16*(1), 399–402. doi:10.1017/thg.2012.86.

Kuntsi J., Rogers H., Swinard G., Borger N., van der Meere J., Rijsdijk F., et al. (2006). Reaction time, inhibition, working memory and 'delay aversion' performance: genetic influences and their interpretation. *Psychological Medicine*, *36*(11), 1613–1624. doi:10.1017/s0033291706008580.

Laird N. M., & Lange C. (2006). Family-based designs in the age of large-scale gene-association studies. *Nature Reviews Genetics*, *7*(5), 385–394. doi:10.1038/nrg1839.

Launer L. J. (2004). Epidemiology of white matter lesions. *Topics in Magnetic Resonance Imaging*, *15*(6), 365–367. doi:10.1097/01.rmr.0000168216.98338.8d.

Lazzeroni L. C., Lu Y., & Belitskaya-Levy I. (2014). P-values in genomics: apparent precision masks high uncertainty. *Molecular Psychiatry*, *19*(12), 1336–1340. doi:10.1038/mp.2013.184.

Lenroot R. K., & Giedd J. N. (2008). The changing impact of genes and environment on brain development during childhood and adolescence: initial findings from a neuroimaging study of pediatric twins. *Development and Psychopathology*, *20*(4), 1161–1175. doi:10.1017/s0954579408000552.

Lenroot R. K., & Giedd J. N. (2010). Sex differences in the adolescent brain. *Brain and Cognition*, *72*(1), 46–55. doi:10.1016/j.bandc.2009.10.008.

Lenroot R. K., Schmitt J. E., Ordaz S. J., Wallace G. L., Neale M. C., Lerch J. P., et al. (2009). Differences in genetic and environmental influences on the human cerebral cortex associated with development during childhood and adolescence. *Human Brain Mapping*, *30*(1), 163–174. doi:10.1002/hbm.20494.

Lister R., Mukamel E. A., Nery J. R., Urich M., Puddifoot C. A., Johnson N. D., et al. (2013). Global epigenomic reconfiguration during mammalian brain development. *Science*, *341*(6146), 1237905. doi:10.1126/science.1237905.

Little J., Higgins J. P., Ioannidis J. P., Moher D., Gagnon F., von Elm E., et al. (2009). Strengthening the reporting of genetic association studies (STREGA): an extension of the STROBE Statement. *Human Genetics*, *125*(2), 131–151. doi:10.1007/s00439-008-0592-7.

Liu J., Pearlson G., Windemuth A., Ruano G., Perrone-Bizzozero N. I., & Calhoun V. (2009). Combining fMRI and SNP data to investigate connections between brain function and genetics using parallel ICA. *Human Brain Mapping*, *30*(1), 241–255. doi:10.1002/hbm.20508.

Liu X., Akula N., Skup M., Brotman M. A., Leibenluft E., & McMahon F. J. (2010). A genome-wide association study of amygdala activation in youths with and without bipolar disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, *49*(1), 33–41.

Liu X., Cannon D. M., Akula N., Moya P. R., Knudsen G. M., Arentzen T. E., et al. (2011). A non-synonymous polymorphism in galactose mutarotase (GALM) is associated with serotonin transporter binding potential in the human thalamus: results of a genome-wide association study. *Molecular Psychiatry*, *16*(6), 584–585. doi:10.1038/mp.2011.1.

Liu C. C., Kanekiyo T., Xu H., & Bu G. (2013). Apolipoprotein E and Alzheimer disease: risk, mechanisms and therapy. *Nature Reviews Neurology*, *9*(2), 106–118. doi:10.1038/nrneurol.2012.263.

Liu Y., Yu J. T., Wang H. F., Han P. R., Tan C. C., Wang C., et al. (2015). APOE genotype and neuroimaging markers of Alzheimer's disease: systematic review and meta-analysis. *Journal of Neurology, Neurosurgery, and Psychiatry*, *86*(2), 127–134. doi:10.1136/jnnp-2014-307719.

Lovestone S., Francis P., Kloszewska I., Mecocci P., Simmons A., Soininen H., et al. (2009). AddNeuroMed-the European collaboration for the discovery of novel biomarkers for Alzheimer's disease. *Annals of the New York Academy of Sciences*, *1180*, 36–46. doi:10.1111/j.1749-6632.2009.05064.x.

Luciano M., Wright M. J., Smith G. A., Geffen G. M., Geffen L. B., & Martin N. G. (2001). Genetic covariance among measures of information processing speed, working memory, and IQ. *Behavior Genetics*, *31*(6), 581–592. doi:10.1023/A:1013397428612.

Maher B. (2008). Personal genomes: the case of the missing heritability. *Nature*, *456*(7218), 18–21. doi:10.1038/456018a.

Mailman M. D., Feolo M., Jin Y., Kimura M., Tryka K., Bagoutdinov R., et al. (2007). The NCBI dbGaP database of genotypes and phenotypes. *Nature Genetics*, *39*(10), 1181–1186. doi:10.1038/ng1007-1181.

Malter Cohen M., Tottenham N., & Casey B. J. (2013). Translational developmental studies of stress on brain and behavior: implications for adolescent mental health and illness? *Neuroscience*, *249*, 53–62. doi:10.1016/j.neuroscience.2013.01.023.

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

Marchini J., Cardon L. R., Phillips M. S., & Donnelly P. (2004). The effects of human population structure on large genetic association studies. *Nature Genetics*, *36*(5), 512–517. doi:10.1038/ng1337.

Martin N. G., Eaves L. J., Kearsey M. J., & Davies P. (1978). The power of the classical twin study. *Heredity (Edinburgh)*, *40*(1), 97–116.

McCarthy M. I., Abecasis G. R., Cardon L. R., Goldstein D. B., Little J., Ioannidis J. P., et al. (2008). Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nature Reviews. Genetics*, *9*(5), 356–369. doi:10.1038/nrg2344.

McRae A. F., Powell J. E., Henders A. K., Bowdler L., Hemani G., Shah S., et al. (2014). Contribution of genetic variation to transgenerational inheritance of DNA methylation. *Genome Biology*, *15*(5), R73. doi:10.1186/gb-2014-15-5-r73.

Medland S. E., Jahanshad N., Neale B. M., & Thompson P. M. (2014). Whole-genome analyses of whole-brain data: working within an expanded search space. *Nature Neuroscience*, *17*(6), 791–800. doi:10.1038/nn.3718.

Melville S. A., Buros J., Parrado A. R., Vardarajan B., Logue M. W., Shen L., et al. (2012). Multiple loci influencing hippocampal degeneration identified by genome scan. *Annals of Neurology*, *72*(1), 65–75. doi:10.1002/ana.23644.

Mersha, T. B., & Abebe, T. (2015). Self-reported race/ethnicity in the age of genomic research: its potential impact on understanding health disparities. Human Genomics, 9(1), 1, doi:10.1186/PREACCEPT-2695828013752627.

Meyer-Lindenberg A. (2010). Imaging genetics of schizophrenia. *Dialogues in Clinical Neuroscience*, *12*(4), 449–456.

Michels, K. B., Binder, A. M., Dedeurwaerder, S., Epstein, C. B., Greally, J. M., Gut, I., et al. (2013). Recommendations for the design and analysis of epigenome-wide association studies. Nature Methods, 10(10), 949–955, doi:10.1038/nmeth.2632.

Moonesinghe R., Khoury M. J., Liu T., & Ioannidis J. P. (2008). Required sample size and nonreplicability thresholds for heterogeneous genetic associations. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(2), 617–622. doi:10.1073/pnas.0705554105.

Morey R. A., Petty C. M., Xu Y., Hayes J. P., Wagner 2nd H. R., Lewis D. V., et al. (2009). A comparison of automated segmentation and manual tracing for quantifying hippocampal and amygdala volumes. *NeuroImage*, *45*(3), 855–866. doi:10.1016/j.neuroimage.2008.12.033.

Morey R. A., Selgrade E. S., Wagner 2nd H. R., Huettel S. A., Wang L., & McCarthy G. (2010). Scan-rescan reliability of subcortical brain volumes derived from automated segmentation. *Human Brain Mapping*, *31*(11), 1751–1762. doi:10.1002/hbm.20973.

Mueller, S. G., Weiner, M. W., Thal, L. J., Petersen, R. C., Jack, C. R., Jagust, W., et al. (2005). Ways toward an early diagnosis in Alzheimer's disease: the Alzheimer's Disease Neuroimaging Initiative (ADNI). Alzheimers Dement, 1(1), 55–66, doi:10.1016/j.jalz.2005.06.003.

Munafo M. R., & Gage S. H. (2013). Improving the reliability and reporting of genetic association studies. *Drug and Alcohol Dependence*, *132*(3), 411–413. doi:10.1016/j.drugalcdep.2013.03.023.

Neale, M. C., & Cardon, L. R. (1992). *Methodology for genetic studies of twins and families* (Vol. 67). Springer.

Nichols T., & Hayasaka S. (2003). Controlling the familywise error rate in functional neuroimaging: a comparative review. *Statistical Methods in Medical Research*, *12*(5), 419–446.

Nugent A. C., Luckenbaugh D. A., Wood S. E., Bogers W., Zarate Jr. C. A., & Drevets W. C. (2013). Automated subcortical segmentation using FIRST: test-retest reliability, interscanner reliability, and comparison to manual segmentation. *Human Brain Mapping*, *34*(9), 2313–2329. doi:10.1002/hbm.22068.

Nuzzo R. (2014). Scientific method: statistical errors. *Nature*, *506*(7487), 150–152. doi:10.1038/506150a.

Ousdal O. T., Anand Brown A., Jensen J., Nakstad P. H., Melle I., Agartz I., et al. (2012). Associations between variants near a monoaminergic pathways gene (PHOX2B) and amygdala reactivity: a genome-wide functional imaging study. *Twin Research and Human Genetics*, *15*(3), 273–285. doi:10.1017/thg.2012.5.

Panagiotou O. A., Willer C. J., Hirschhorn J. N., & Ioannidis J. P. (2013). The power of meta-analysis in genome-wide association studies. *Annual Review of Genomics and Human Genetics*, *14*, 441–465. doi:10.1146/annurev-genom-091212-153520.

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

Panizzon M. S., Fennema-Notestine C., Kubarych T. S., Chen C. H., Eyler L. T., Fischl B., et al. (2012). Genetic and environmental influences of white and gray matter signal contrast: a new phenotype for imaging genetics? *NeuroImage*, *60*(3), 1686–1695. doi:10.1016/j.neuroimage.2012.01.122.

Panoutsopoulou K., Tachmazidou I., & Zeggini E. (2013). In search of low-frequency and rare variants affecting complex traits. *Human Molecular Genetics*, *22*(R1), R16–R21. doi:10.1093/hmg/ddt376.

Paternoster L., Chen W., & Sudlow C. L. (2009). Genetic determinants of white matter hyperintensities on brain scans: a systematic assessment of 19 candidate gene polymorphisms in 46 studies in 19,000 subjects. *Stroke*, *40*(6), 2020–2026. doi:10.1161/STROKEAHA.108.542050.

Paus T., Bernard M., Chakravarty M. M., Davey Smith G., Gillis J., Lourdusamy A., et al. (2012). KCTD8 gene and brain growth in adverse intrauterine environment: a genome-wide association study. *Cerebral Cortex*, *22*(11), 2634–2642. doi:10.1093/cercor/bhr350.

Pe'er I., Yelensky R., Altshuler D., & Daly M. J. (2008). Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. *Genetic Epidemiology*, *32*(4), 381–385. doi:10.1002/gepi.20303.

Pennington B. F., Filipek P. A., Lefly D., Chhabildas N., Kennedy D. N., Simon J. H., et al. (2000). A twin MRI study of size variations in human brain. *Journal of Cognitive Neuroscience*, *12*(1), 223–232.

Pereira T. V., Patsopoulos N. A., Salanti G., & Ioannidis J. P. (2009). Discovery properties of genome-wide association signals from cumulatively combined data sets. *American Journal of Epidemiology*, *170*(10), 1197–1206. doi:10.1093/aje/kwp262.

Pezawas L., Verchinski B. A., Mattay V. S., Callicott J. H., Kolachana B. S., Straub R. E., et al. (2004). The brain-derived neurotrophic factor val66met polymorphism and variation in human cortical morphology. *The Journal of Neuroscience*, *24*(45), 10099–10102. doi:10.1523/JNEUROSCI.2680-04.2004.

Pfefferbaum A., Sullivan E. V., Swan G. E., & Carmelli D. (2000). Brain structure in men remains highly heritable in the seventh and eighth decades of life. *Neurobiology of Aging*, *21*(1), 63–74.

Phillips C., Parson W., Lundsberg B., Santos C., Freire-Aradas A., Torres M., et al. (2014). Building a forensic ancestry panel from the ground up: The EUROFORGEN Global AIM-SNP set. *Forensic Science International. Genetics*, *11*, 13–25. doi:10.1016/j.fsigen.2014.02.012.

Pickrell J. K. (2014). Joint analysis of functional genomic data and genome-wide association studies of 18 human traits. *American Journal of Human Genetics*, *94*(4), 559–573. doi:10.1016/j.ajhg.2014.03.004.

Plomin R., DeFries J. C., & Loehlin J. C. (1977). Genotype-environment interaction and correlation in the analysis of human behavior. *Psychological Bulletin*, *84*(2), 309–322.

Posthuma D., de Geus E. J., Neale M. C., Hulshoff Pol H. E., Baare W. E. C., Kahn R. S., et al. (2000). Multivariate genetic analysis of brain structure in an extended twin design. *Behavior Genetics*, *30*(4), 311–319.

Posthuma D., De Geus E. J., Baare W. F., Hulshoff Pol H. E., Kahn R. S., & Boomsma D. I. (2002). The association between brain volume and intelligence is of genetic origin. *Nature Neuroscience*, *5*(2), 83–84. doi:10.1038/nn0202-83.

Posthuma D., Baare W. F., Hulshoff Pol H. E., Kahn R. S., Boomsma D. I., & De Geus E. J. (2003). Genetic correlations between brain volumes and the WAIS-III dimensions of verbal comprehension, working memory, perceptual organization, and processing speed. *Twin Research*, *6*(2), 131–139. doi:10.1375/136905203321536254.

Potkin S. G., Guffanti G., Lakatos A., Turner J. A., Kruggel F., Fallon J. H., et al. (2009a). Hippocampal atrophy as a quantitative trait in a genome-wide association study identifying novel susceptibility genes for Alzheimer's disease. *PloS One*, *4*(8), e6501. doi:10.1371/journal.pone.0006501.

Potkin S. G., Turner J. A., Guffanti G., Lakatos A., Fallon J. H., Nguyen D. D., et al. (2009b). A genome-wide association study of schizophrenia using brain activation as a quantitative phenotype. *Schizophrenia Bulletin*, *35*(1), 96–108. doi:10.1093/schbul/sbn155.

Price A. L., Patterson N. J., Plenge R. M., Weinblatt M. E., Shadick N. A., & Reich D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics*, *38*(8), 904–909. doi:10.1038/ng1847.

Price A. L., Zaitlen N. A., Reich D., & Patterson N. (2010). New approaches to population stratification in genome-wide association studies. *Nature Reviews. Genetics*, *11*(7), 459–463. doi:10.1038/nrg2813.

Psaty B. M., & Sitlani C. (2013). The Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium as a model of collaborative science. *Epidemiology*, *24*(3), 346–348. doi:10.1097/EDE.0b013e31828b2cbb.

Psaty B. M., O'Donnell C. J., Gudnason V., Lunetta K. L., Folsom A. R., Rotter J. I., et al. (2009). Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium: design of prospective meta-analyses of genome-wide association studies from 5 cohorts. *Circulation. Cardiovascular Genetics*, *2*(1), 73–80. doi:10.1161/CIRCGENETICS.108.829747.

Rabl U., Meyer B. M., Diers K., Bartova L., Berger A., Mandorfer D., et al. (2014). Additive gene-environment effects on hippocampal structure in healthy humans. *The Journal of Neuroscience*, *34*(30), 9917–9926. doi:10.1523/JNEUROSCI.3113-13.2014.

Rakyan V. K., Down T. A., Balding D. J., & Beck S. (2011). Epigenome-wide association studies for common human diseases. *Nature Reviews. Genetics*, *12*(8), 529–541. doi:10.1038/nrg3000.

Rentería, M. E., Hansell, N. K., Strike, L. T., McMahon, K. L., de Zubicaray, G. I., Hickie, I. B., et al. (2014). Genetic architecture of subcortical brain regions: common and region-specific genetic contributions. Genes Brain and Behavior, n/a-n/a, doi:10.1111/gbb.12177.

Rijsdijk F. V., Viding E., De Brito S., Forgiarini M., Mechelli A., Jones A. P., et al. (2010). Heritable variations in gray matter concentration as

a potential endophenotype for psychopathic traits. *Archives of General Psychiatry*, *67*(4), 406–413. doi:10.1001/archgenpsychiatry.2010.20.

Rimol L. M., Panizzon M. S., Fennema-Notestine C., Eyler L. T., Fischl B., Franz C. E., et al. (2010). Cortical thickness is influenced by regionally specific genetic factors. *Biological Psychiatry*, *67*(5), 493–499. doi:10.1016/j.biopsych.2009.09.032.

Rogers J., Kochunov P., Zilles K., Shelledy W., Lancaster J., Thompson P., et al. (2010). On the genetic architecture of cortical folding and brain volume in primates. *NeuroImage*, *53*(3), 1103–1108. doi:10.1016/j.neuroimage.2010.02.020.

Rosenberg N. A., Huang L., Jewett E. M., Szpiech Z. A., Jankovic I., & Boehnke M. (2010). Genome-wide association studies in diverse populations. *Nature Reviews. Genetics*, *11*(5), 356–366. doi:10.1038/nrg2760.

Rousson V., Gasser T., & Seifert B. (2002). Assessing intrarater, interrater and test-retest reliability of continuous measurements. *Statistics in Medicine*, *21*(22), 3431–3446. doi:10.1002/sim.1253.

Roussotte F. F., Jahanshad N., Hibar D. P., Sowell E. R., Kohannim O., Barysheva M., et al. (2014). A commonly carried genetic variant in the delta opioid receptor gene, OPRD1, is associated with smaller regional brain volumes: replication in elderly and young populations. *Human Brain Mapping*, *35*(4), 1226–1236. doi:10.1002/hbm.22247.

Sabuncu M. R., Buckner R. L., Smoller J. W., Lee P. H., Fischl B., Sperling R. A., et al. (2012). The Association between a Polygenic Alzheimer Score and Cortical Thickness in Clinically Normal Subjects. *Cerebral Cortex (New York, NY)*, *22*(11), 2653–2661. doi:10.1093/cercor/bhr348.

Sachdev P. S., Lammel A., Trollor J. N., Lee T., Wright M. J., Ames D., et al. (2009). A comprehensive neuropsychiatric study of elderly twins: the Older Australian Twins Study. *Twin Research and Human Genetics*, *12*(6), 573–582. doi:10.1375/twin.12.6.573.

Schaub M. A., Boyle A. P., Kundaje A., Batzoglou S., & Snyder M. (2012). Linking disease associations with regulatory information in the human genome. *Genome Research*, *22*(9), 1748–1759. doi:10.1101/gr.136127.111.

Scheggia D., Sannino S., Scattoni M. L., & Papaleo F. (2012). COMT as a drug target for cognitive functions and dysfunctions. *CNS & Neurological Disorders Drug Targets*, *11*(3), 209–221.

Schizophrenia Working Group of the Psychiatric Genomics Consortium (2014). Biological insights from 108 schizophrenia-associated genetic loci. *Nature*, *511*(7510), 421–427. doi:10.1038/nature13595.

Schmitt J. E., Wallace G. L., Rosenthal M. A., Molloy E. A., Ordaz S., Lenroot R., et al. (2007). A multivariate analysis of neuroanatomic relationships in a genetically informative pediatric sample. *NeuroImage*, *35*(1), 70–82. doi:10.1016/j.neuroimage.2006.04.232.

Schmitt J. E., Lenroot R. K., Wallace G. L., Ordaz S., Taylor K. N., Kabani N., et al. (2008). Identification of genetically mediated cortical networks: a multivariate study of pediatric twins and siblings. *Cerebral Cortex*, *18*(8), 1737–1747. doi:10.1093/cercor/bhm211.

Schmitt J. E., Lenroot R. K., Ordaz S. E., Wallace G. L., Lerch J. P., Evans A. C., et al. (2009). Variance decomposition of MRI-based covariance maps using genetically informative samples and structural equation modeling. *NeuroImage*, *47*(1), 56–64. doi:10.1016/j.neuroimage.2008.06.039.

Schmitt J. E., Wallace G. L., Lenroot R. K., Ordaz S. E., Greenstein D., Clasen L., et al. (2010). A twin study of intracerebral volumetric relationships. *Behavior Genetics*, *40*(2), 114–124. doi:10.1007/s10519-010-9332-6.

Schmitt J. E., Neale M. C., Fassassi B., Perez J., Lenroot R. K., Wells E. M., et al. (2014). The dynamic role of genetics on cortical patterning during childhood and adolescence. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(18), 6774–6779. doi:10.1073/pnas.1311630111.

Schumann G., Loth E., Banaschewski T., Barbot A., Barker G., Buchel C., et al. (2010). The IMAGEN study: reinforcement-related behaviour in normal brain function and psychopathology. *Molecular Psychiatry*, *15*(12), 1128–1139. doi:10.1038/mp.2010.4.

Seshadri S., DeStefano A. L., Au R., Massaro J. M., Beiser A. S., Kelly-Hayes M., et al. (2007). Genetic correlates of brain aging on MRI and cognitive test measures: a genome-wide association and linkage analysis in the Framingham study. *BMC Medical Genetics*, *8*(*Suppl 1*), S15. doi:10.1186/1471-2350-8-S1-S15.

Sham P. C., & Purcell S. M. (2014). Statistical power and significance testing in large-scale genetic studies. *Nature Reviews. Genetics*, *15*(5), 335–346. doi:10.1038/nrg3706.

Shen L., Kim S., Risacher S. L., Nho K., Swaminathan S., West J. D., et al. (2010). Whole genome association study of brain-wide imaging phenotypes for identifying quantitative trait loci in MCI and AD: a study of the ADNI cohort. *NeuroImage*, *53*(3), 1051–1063. doi:10.1016/j.neuroimage.2010.01.042.

Sherry S. T., Ward M. H., Kholodov M., Baker J., Phan L., Smigielski E. M., et al. (2001). dbSNP: the NCBI database of genetic variation. *Nucleic Acids Research*, *29*(1), 308–311.

Shi J., Thompson P. M., Gutman B., Wang Y., & Alzheimer's Disease Neuroimaging I. (2013). Surface fluid registration of conformal representation: application to detect disease burden and genetic influence on hippocampus. *NeuroImage*, *78*, 111–134. doi:10.1016/j.neuroimage.2013.04.018.

Shields P. G. (2000). Publication bias is a scientific problem with adverse ethical outcomes: the case for a section for null results. *Cancer Epidemiology, Biomarkers & Prevention*, *9*(8), 771–772.

Silventoinen K., Sammalisto S., Perola M., Boomsma D. I., Cornes B. K., Davis C., et al. (2003). Heritability of adult body height: a comparative study of twin cohorts in eight countries. *Twin Research*, *6*(5), 399–408. doi:10.1375/136905203770326402.

Simic G., Kostovic I., Winblad B., & Bogdanovi N. (1997). Volume and number of neurons of the human hippocampal formation in normal aging and Alzheimer's disease. *Journal of Comparative Neurology (1911)*, *379*(4), 482–494.

So H. C., Li M., & Sham P. C. (2011). Uncovering the total heritability explained by all true susceptibility variants in a genome-wide association study. *Genetic Epidemiology*, *35*(6), 447–456. doi:10.1002/gepi.20593.

Sprooten, E., Fleming, K. M., Thomson, P. A., Bastin, M. E., Whalley, H. C., Hall, J., et al. (2013). White matter integrity as an intermediate phenotype: exploratory genome-wide association analysis in individuals at high risk of bipolar disorder. Psychiatry Research, 206(2–3), 223–231, doi:10.1016/j.psychres.2012.11.002.

Steffens D. C., Helms M. J., Krishnan K. R., & Burke G. L. (1999). Cerebrovascular disease and depression symptoms in the cardiovascular health study. *Stroke*, *30*(10), 2159–2166.

Stein J. L., Hua X., Morra J. H., Lee S., Hibar D. P., Ho A. J., et al. (2010). Genome-wide analysis reveals novel genes influencing temporal lobe structure with relevance to neurodegeneration in Alzheimer's disease. *NeuroImage*, *51*(2), 542–554. doi:10.1016/j.neuroimage.2010.02.068.

Stein J. L., Hibar D. P., Madsen S. K., Khamis M., McMahon K. L., de Zubicaray G. I., et al. (2011). Discovery and replication of dopamine-related gene effects on caudate volume in young and elderly populations (N = 1198) using genome-wide search. *Molecular Psychiatry*, *16*(9), 927–881937. doi:10.1038/mp.2011.32.

Stein J. L., Medland S. E., Vasquez A. A., Hibar D. P., Senstad R. E., Winkler A. M., et al. (2012). Identification of common variants associated with human hippocampal and intracranial volumes. *Nature Genetics*, *44*(5), 552–561. doi:10.1038/ng.2250.

Stranger B. E., Stahl E. A., & Raj T. (2011). Progress and promise of genome-wide association studies for human complex trait genetics. *Genetics*, *187*(2), 367–383. doi:10.1534/genetics.110.120907.

Sullivan P. F. (2007). Spurious genetic associations. *Biological Psychiatry*, *61*(10), 1121–1126. doi:10.1016/j.biopsych.2006.11.010.

Sullivan P. F., Daly M. J., & O'Donovan M. (2012). Genetic architectures of psychiatric disorders: the emerging picture and its implications. *Nature Reviews. Genetics*, *13*(8), 537–551. doi:10.1038/nrg3240.

Taal H. R., St Pourcain B., Thiering E., Das S., Mook-Kanamori D. O., Warrington N. M., et al. (2012). Common variants at 12q15 and 12q24 are associated with infant head circumference. *Nature Genetics*, *44*(5), 532–538. doi:10.1038/ng.2238.

Teo Y. Y. (2008). Common statistical issues in genome-wide association studies: a review on power, data quality control, genotype calling and population structure. *Current Opinion in Lipidology*, *19*(2), 133–143. doi:10.1097/MOL.0b013e3282f5dd77.

Teo Y. Y., Small K. S., & Kwiatkowski D. P. (2010). Methodological challenges of genome-wide association analysis in Africa. *Nature Reviews. Genetics*, *11*(2), 149–160. doi:10.1038/nrg2731.

Thomas D. (2010). Gene-environment-wide association studies: emerging approaches. *Nature Reviews. Genetics*, *11*(4), 259–272. doi:10.1038/nrg2764.

Thompson P. M., Cannon T. D., Narr K. L., van Erp T., Poutanen V. P., Huttunen M., et al. (2001). Genetic influences on brain structure. *Nature Neuroscience*, *4*(12), 1253–1258. doi:10.1038/nn758.

Thompson P. M., Stein J. L., Medland S. E., Hibar D. P., Vasquez A. A., Renteria M. E., et al. (2014). The ENIGMA consortium: large-scale collaborative analyses of neuroimaging and genetic data. *Brain Imaging and Behavior*, *8*(2), 153–182. doi:10.1007/s11682-013-9269-5.

Tost H., Bilek E., & Meyer-Lindenberg A. (2012). Brain connectivity in psychiatric imaging genetics. *NeuroImage*, *62*(4), 2250–2260. doi:10.1016/j.neuroimage.2011.11.007.

Tucker-Drob E. M., Rhemtulla M., Harden K. P., Turkheimer E., & Fask D. (2011). Emergence of a Gene x socioeconomic status interaction on infant mental ability between 10 months and 2 years. *Psychological Science*, *22*(1), 125–133. doi:10.1177/0956797610392926.

Turkheimer E., Haley A., Waldron M., D'Onofrio B., & Gottesman I. I. (2003). Socioeconomic status modifies heritability of IQ in young children. *Psychological Science*, *14*(6), 623–628.

Turner, S. D. (2014). *qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots*.

Turner S. T., Jack C. R., Fornage M., Mosley T. H., Boerwinkle E., & de Andrade M. (2004). Heritability of leukoaraiosis in hypertensive sibships. *Hypertension*, *43*(2), 483–487. doi:10.1161/01.hyp.0000112303.26158.92.

van der Heijden C. D., Rijpkema M., Arias-Vasquez A., Hakobjan M., Scheffer H., Fernandez G., et al. (2013). Genetic variation in ataxia gene ATXN7 influences cerebellar grey matter volume in healthy adults. *Cerebellum*, *12*(3), 390–395. doi:10.1007/s12311-012-0423-1.

van der Sijde M. R., Ng A., & Fu J. (2014). Systems genetics: From GWAS to disease pathways. *Biochimica et Biophysica Acta*. doi:10.1016/j.bbadis.2014.04.025.

van Leeuwen M., Peper J. S., van den Berg S. M., Brouwer R. M., Hulshoff Pol H. E., Kahn R. S., et al. (2009). A genetic analysis of brain volumes and IQ in children. *Intelligence*, *37*(2), 181–191.

Van Petten C. (2004). Relationship between hippocampal volume and memory ability in healthy individuals across the lifespan: review and meta-analysis. *Neuropsychologia*, *42*(10), 1394–1413. doi:10.1016/j.neuropsychologia.2004.04.006.

van Soelen I. L., Brouwer R. M., van Leeuwen M., Kahn R. S., Hulshoff Pol H. E., & Boomsma D. I. (2011). Heritability of verbal and performance intelligence in a pediatric longitudinal sample. *Twin Research and Human Genetics*, *14*(2), 119–128. doi:10.1375/twin.14.2.119.

van Soelen I. L., Brouwer R. M., van Baal G. C., Schnack H. G., Peper J. S., Collins D. L., et al. (2012). Genetic influences on thinning of the cerebral cortex during development. *NeuroImage*, *59*(4), 3871–3880. doi:10.1016/j.neuroimage.2011.11.044.

van Soelen I. L., Brouwer R. M., van Baal G. C., Schnack H. G., Peper J. S., Chen L., et al. (2013). Heritability of volumetric brain changes and height in children entering puberty. *Human Brain Mapping*, *34*(3), 713–725. doi:10.1002/hbm.21468.

Vermeer S. E., Hollander M., van Dijk E. J., Hofman A., Koudstaal P. J., Breteler M. M., et al. (2003). Silent brain infarcts and white matter lesions increase stroke risk in the general population: the Rotterdam Scan Study. *Stroke*, *34*(5), 1126–1129. doi:10.1161/01.STR.0000068408.82115.D2.

Verweij K. J., Mosing M. A., Zietsch B. P., & Medland S. E. (2012). Estimating heritability from twin studies. *Methods in Molecular Biology*, *850*, 151–170. doi:10.1007/978-1-61779-555-8_9.

Veyrieras J. B., Kudaravalli S., Kim S. Y., Dermitzakis E. T., Gilad Y., Stephens M., et al. (2008). High-resolution mapping of expression-QTLs yields insight into human gene regulation. *PLoS Genetics*, *4*(10), e1000214. doi:10.1371/journal.pgen.1000214.

Videbech P., & Ravnkilde B. (2004). Hippocampal Volume and Depression: A meta-analysis of MRI Studies. *The American Journal of Psychiatry*, *161*(11), 1957–1966.

Vinkhuyzen A. A., Wray N. R., Yang J., Goddard M. E., & Visscher P. M. (2013). Estimation and partition of heritability in human populations using whole-genome analysis methods. *Annual Review of Genetics*, *47*, 75–95. doi:10.1146/annurev-genet-111212-133258.

Visscher, P. M., Hill, W. G., & Wray, N. R. (2008). Heritability in the genomics era–concepts and misconceptions. Nature Reviews. Genetics, 9(4), 255–266, doi:10.1038/nrg2322.

Visscher P. M., Brown M. A., McCarthy M. I., & Yang J. (2012a). Five years of GWAS discovery. *American Journal of Human Genetics*, *90*(1), 7–24. doi:10.1016/j.ajhg.2011.11.029.

Visscher P. M., Goddard M. E., Derks E. M., & Wray N. R. (2012b). Evidence-based psychiatric genetics, AKA the false dichotomy between common and rare variant hypotheses. *Molecular Psychiatry*, *17*(5), 474–485. doi:10.1038/mp.2011.65.

Vounou M., Nichols T. E., Montana G., & Alzheimer's Disease Neuroimaging I. (2010). Discovering genetic associations with high-dimensional neuroimaging phenotypes: a sparse reduced-rank regression approach. *NeuroImage*, *53*(3), 1147–1159. doi:10.1016/j.neuroimage.2010.07.002.

Vuoksimaa E., Panizzon M. S., Chen C. H., Fiecas M., Eyler L. T., Fennema-Notestine C., et al. (2014). The genetic association between neocortical volume and general cognitive ability is driven by global surface area rather than thickness. *Cerebral Cortex*. doi:10.1093/cercor/bhu018.

Wallace G. L., Eric Schmitt J., Lenroot R., Viding E., Ordaz S., Rosenthal M. A., et al. (2006). A pediatric twin study of brain morphometry. *Journal of Child Psychology and Psychiatry*, *47*(10), 987–993. doi:10.1111/j.1469-7610.2006.01676.x.

Wallace G. L., Lee N. R., Prom-Wormley E. C., Medland S. E., Lenroot R. K., Clasen L. S., et al. (2010). A bivariate twin study of regional brain volumes and verbal and nonverbal intellectual skills during childhood and adolescence. *Behavior Genetics*, *40*(2), 125–134. doi:10.1007/s10519-009-9329-1.

Wang Q., Xiang B., Deng W., Wu J., Li M., Ma X., et al. (2013a). Genome-wide association analysis with gray matter volume as a quantitative phenotype in first-episode treatment-naive patients with schizophrenia. *PloS One*, *8*(9), e75083. doi:10.1371/journal.pone.0075083.

Wang Y., Li J., Chen C., Chen C., Zhu B., Moysis R. K., et al. (2013b). COMT rs4680 Met is not always the 'smart allele': Val allele is associated with better working memory and larger hippocampal volume in healthy Chinese. *Genes, Brain, and Behavior*, *12*(3), 323–329. doi:10.1111/gbb.12022.

Wellcome Trust Case Control C. (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3000 shared controls. *Nature*, *447*(7145), 661–678. doi:10.1038/nature05911.

Welter D., MacArthur J., Morales J., Burdett T., Hall P., Junkins H., et al. (2014). The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Research*, *42*(Database issue), D1001–D1006. doi:10.1093/nar/gkt1229.

Winkler A. M., Kochunov P., Blangero J., Almasy L., Zilles K., Fox P. T., et al. (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., Lee S. H., Mehta D., Vinkhuyzen A. A. E., Dudbridge F., & Middeldorp C. M. (2014). Research review: polygenic methods and their application to psychiatric traits. *Journal of Child Psychology and Psychiatry*, *55*(10), 1068–1087. doi:10.1111/jcpp.12295.

Wright I. C., Rabe-Hesketh S., Woodruff P. W., David A. S., Murray R. M., & Bullmore E. T. (2000). Meta-analysis of regional brain volumes in schizophrenia. *The American Journal of Psychiatry*, *157*(1), 16–25.

Wright I. C., Sham P., Murray R. M., Weinberger D. R., & Bullmore E. T. (2002). Genetic contributions to regional variability in human brain structure: methods and preliminary results. *NeuroImage*, *17*(1), 256–271.

Yang J., Benyamin B., McEvoy B. P., Gordon S., Henders A. K., Nyholt D. R., et al. (2010). Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics*, *42*(7), 565–569. doi:10.1038/ng.608.

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

Yoon U., Fahim C., Perusse D., & Evans A. C. (2010). Lateralized genetic and environmental influences on human brain morphology of 8-year-old twins. *NeuroImage*, *53*(3), 1117–1125. doi:10.1016/j.neuroimage.2010.01.007.

Zhang X., Lee M. R., Salmeron B. J., Stein D. J., Hong L. E., Geng X., et al. (2013). Prefrontal white matter impairment in substance users depends upon the catechol-o-methyl transferase (COMT) val158met polymorphism. *NeuroImage*, *69*, 62–69. doi:10.1016/j.neuroimage.2012.11.056.