

# Whole-genome analyses of whole-brain data: working within an expanded search space

Sarah E Medland<sup>1,13</sup>, Neda Jahanshad<sup>2,3,13</sup>, Benjamin M Neale<sup>4-6</sup> & Paul M Thompson<sup>2,3,7-12</sup>

Large-scale comparisons of patients and healthy controls have unearthed genetic risk factors associated with a range of neurological and psychiatric illnesses. Meanwhile, brain imaging studies are increasing in size and scope, revealing disease and genetic effects on brain structure and function, and implicating neural pathways and causal mechanisms. With the advent of global neuroimaging consortia, imaging studies are now well powered to discover genetic variants that reliably affect the brain. Genetic analyses of brain measures from tens of thousands of people are being extended to test genetic associations with signals at millions of locations in the brain, and connectome-wide, genome-wide scans can jointly screen brain circuits and genomes; these analyses and others present new statistical challenges. There is a growing need for the community to establish and enforce standards in this developing field to ensure robust findings. Here we discuss how neuroimagers and geneticists have formed alliances to discover how genetic factors affect the brain. The field is rapidly advancing with ultra-high-resolution imaging and whole-genome sequencing. We recommend a rigorous approach to neuroimaging genomics that capitalizes on its recent successes and ensures the reliability of future discoveries.

Worldwide consortia offer new opportunities for discovery in neuroscience, vastly expanding the scale of brain imaging and genomics studies. The higher power to discover genetic factors associated with brain differences has led to key insights in neurology and psychiatry and helped us understand genetic pathways that shape the brain. Genes that contribute to mental illness and individual brain differences are now being reliably detected, mapped and quantified.

Large-scale genetic studies are uncovering common genetic variants associated with Alzheimer's disease<sup>1</sup>, schizophrenia<sup>2</sup> and many other disorders. These new discoveries often arise from genome-wide association studies (GWAS), which compare large samples of patients with healthy controls to detect differences in allele frequencies (or regress continuous traits on the number of risk alleles) at genetic markers across the genome.

What is perhaps less well known is that genomic screening efforts are also occurring in the brain imaging community. In the largest brain imaging studies yet performed, two consortia, Enhancing Neuroimaging Genetics through Meta-Analysis (ENIGMA) and the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE), recently amassed brain magnetic resonance imaging (MRI) scans and genome-wide data from 21,151 individuals<sup>3,4</sup>. These worldwide consortium efforts have increased the power to identify genetic variations that influence brain structure and function. The combined sample sizes are not unusual in genetics, but in neuroimaging they are vast. Neuroimaging and genotyping of large cohorts can be costly. However, consortium models draw on repositories of existing data, fueling collaborations and combining the expertise of scientists from around the world.

Twenty years of data collection in neuroimaging have led to rich and diverse data sets for genetic analysis. In particular, GWAS of neuroimaging measures can localize genetic effects within the brain, offering new information on how variants increase disease risk. Here we review some opportunities that arise when analyzing the genomics of brain images, including new statistical methods. As images and genomic data are collected at ever higher resolutions, data analyses now draw on large branches of work in mathematics, signal processing, informatics and basic neuroscience.

## GWAS of imaging

For neurological diseases, a few common genetic variants of large effect have been identified. Similarly, few copy number variants of large effect have been identified, but only a few common genetic risk factors, individually, substantially affect disease risk. The best-known genetic risk factor may be apolipoprotein E4 (ApoE4), which was discovered using a biologically informed approach<sup>5</sup>. Each copy of the E4 haplotype in the *APOE* gene boosts Alzheimer's disease risk by a factor of 3 or more; *APOE* explains around 6% of Alzheimer's disease risk.

<sup>1</sup>Quantitative Genetics, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia. <sup>2</sup>Imaging Genetics Center, Institute for Neuroimaging & Informatics, Keck School of Medicine of University of Southern California, Los Angeles, California, USA. <sup>3</sup>Department of Neurology, University of Southern California, Los Angeles, California, USA. <sup>4</sup>Program in Medical and Population Genetics, The Broad Institute of Harvard and MIT, Cambridge, Massachusetts, USA. <sup>5</sup>Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, Massachusetts, USA. <sup>6</sup>Psychiatric and Neurodevelopmental Unit, Massachusetts General Hospital, Boston, Massachusetts, USA. <sup>7</sup>Department of Neurology, University of Southern California, Los Angeles, California, USA. <sup>8</sup>Department of Psychiatry, University of Southern California, Los Angeles, California, USA. <sup>9</sup>Department of Engineering, University of Southern California, Los Angeles, California, USA. <sup>10</sup>Department of Radiology, University of Southern California, Los Angeles, California, USA. <sup>11</sup>Department of Pediatrics, University of Southern California, Los Angeles, California, USA. <sup>12</sup>Department of Ophthalmology, University of Southern California, Los Angeles, California, USA. <sup>13</sup>These authors contributed equally to this work. Correspondence should be addressed to P.M.T. ([pthomp@usc.edu](mailto:pthomp@usc.edu)).

Received 8 January; accepted 3 April; published online 27 May 2014;  
doi:10.1038/nn.3718

From an imaging perspective, ApoE4 carriers lose brain tissue faster as they age. As young adults, carriers have thinner cortical gray matter than noncarriers<sup>6</sup>. The first GWAS studies of brain images searched for associations between common variants and volumes of key brain regions, such as the hippocampus, and were well powered to detect the effects of *APOE* (ref. 7).

In large-scale GWAS meta-analyses, variants with large effects are exceptional; when found, they tend to be associated with later-onset diseases. The single nucleotide polymorphism (SNP) with the strongest evidence for association in the ENIGMA-CHARGE study<sup>3</sup> explained around 0.27% of the variance in hippocampal volume, although the differences between carriers and noncarriers had an effect comparable to that of several years of brain aging. In a related study of the overall volume of the brain, carriers of an *HMGA2* allele averaged 9 cm<sup>3</sup> larger brain volumes per allele.

Despite expectations that brain images might be more sensitive than clinical phenotypes to genetic effects, experience has indicated that, as with other complex traits, large sample sizes are needed to discover genetic variants that influence the brain. Significantly associated SNPs have an expected effect size <0.5% of the phenotypic variance<sup>8</sup> in quantitative traits, even for measures extracted from images. Work so far by imaging consortia such as ENIGMA and CHARGE suggests that the distribution of effect sizes for brain measures is similar to that seen when studying other quantitative traits, although only the simplest brain measures have so far been assessed with GWAS.

Even so, imaging studies can provide information that is not available in genetic comparisons of patients and controls. Imaging can identify

potential mechanisms and circuits promoting disease risk, making it easier to develop and evaluate treatments. A more mechanistic understanding should make it easier to discover disease-modifying factors.

For logistic reasons, the largest collaborative analyses of brain images have so far focused on simple measures such as regional volumes. Efforts are underway to examine whether larger gene effects may be found with diffusion tensor imaging (DTI) and functional MRI; a parallel effort (see below) seeks to prioritize brain measures and define new features to make genetic studies more efficient. Global harmonization efforts are underway to derive measures from MRI<sup>9–11</sup>, DTI<sup>12</sup> and electroencephalography (EEG) scans in a consistent way, making it easier to integrate data from studies initially designed with different goals. Global collaborations are also using meta-analysis to combine evidence from many cohorts, to discover the genetic variants that influence the brain<sup>13</sup>.

**Large-scale brain imaging consortia.** For the major imaging consortia that have published GWAS studies (Table 1)<sup>14,15</sup>, sample sizes are difficult to estimate and are increasing over time, as many studies are still enrolling participants. Published papers lag behind ongoing work: for instance, as of April 2014, collaborative work by ENIGMA and CHARGE has amassed MRI and GWAS data from 29,000 participants, but the largest published studies (from 2012) have around 11,000 participants per consortium and around 21,000 in aggregate<sup>3</sup>. Several new initiatives are planned (for example, the UK Biobank) that assess more cohorts with large-scale scanning and genotyping, offering ever-increasing power for genetic studies.

**Table 1** Imaging consortia that have published GWAS

Name of imaging consortium, URL	Sample size <sup>a</sup>	Goal	Representative papers
ENIGMA ( <a href="http://enigma.ini.usc.edu/">http://enigma.ini.usc.edu/</a> )	13,171; ~29,000 with CHARGE and replication cohorts	Genetic analysis of brain MRI, DTI, EEG and other brain measures, in cohorts across the lifespan; working groups on major brain diseases, including schizophrenia, bipolar illness, major depression and attention deficit–hyperactivity disorder	Stein <i>et al.</i> , 2012 (ref. 3) Jahanshad <i>et al.</i> , 2013 (ref. 12) Thompson <i>et al.</i> , 2014 (ref. 13) Hibar <i>et al.</i> , 2013 (ref. 11) Hibar <i>et al.</i> <sup>b</sup>
CHARGE ( <a href="http://web.chargeconsortium.com/">http://web.chargeconsortium.com/</a> ); includes several large cohorts, for example, the Framingham Heart study, the Rotterdam Study, the Icelandic AGES study, Atherosclerosis Risk in Communities (ARIC) and the Cardiovascular Health Study (CHS)	~29,000 when combined with ENIGMA	Genetic analysis of neuroimaging and cardiovascular health data, including stroke	Bis <i>et al.</i> , 2012 (ref. 4) Ikram <i>et al.</i> , 2012 (ref. 64) Psaty <i>et al.</i> , 2009 (ref. 65)
ADNI ( <a href="http://adni.loni.usc.edu/">http://adni.loni.usc.edu/</a> ) and related worldwide initiatives	~800 in phase 1 (ADNI1); ~850 additional participants in ADNI-GO and ADNI2	Longitudinal study of MRI, PET and CSF and plasma biomarkers from elderly people at ~50 North American sites; includes people with mild cognitive impairment and Alzheimer's disease	Weiner <i>et al.</i> , 2012 (ref. 14) Shen <i>et al.</i> , 2013 (ref. 39)
IMAGEN ( <a href="http://www.imagen-europe.com/">http://www.imagen-europe.com/</a> )	2,223 14-year-old children across Europe (in Pallière Martinot <i>et al.</i> 2013 (ref. 67))	European Research Project studying biological and environmental factors affecting mental health in adolescence	Nymberg <i>et al.</i> , 2012 (ref. 66) Pallière Martinot <i>et al.</i> , 2013 (ref. 67)
PING ( <a href="http://www.chd.ucsd.edu/research/ping-study.html">http://www.chd.ucsd.edu/research/ping-study.html</a> )	1,400 children, aged 3–20 years	Studies of the genetic basis of individual differences in brain structure and connectivity, cognition, and personality	Eicher <i>et al.</i> , 2013 (ref. 68)
IMAGEMEND ( <a href="http://www.imagemend.eu/">http://www.imagemend.eu/</a> )	13,000 participants	Primarily European study, focusing on schizophrenia, bipolar disorder and attention deficit–hyperactivity disorder	Project began October 2013
fBIRN ( <a href="http://www.birncommunity.org/">http://www.birncommunity.org/</a> )	690	Functional and structural imaging study of schizophrenia patients in the United States	Potkin and Ford, 2009 (ref. 69)

Many neuroimaging consortia were formed with goals other than genetic analysis (for example, to study disease effects on the brain), so here we focus on studies that have published GWAS of brain images. Some multisite consortia, such as IMAGEN, publish their own studies but also take part in larger consortia such as ENIGMA. ENIGMA also conducts studies of nine major brain diseases for which GWAS data are not necessary, so those numbers are not included here.

CSF, cerebrospinal fluid; ENIGMA, Enhancing Neuroimaging Genetics through Meta-Analysis; CHARGE, Cohorts for Heart and Aging in Genomic Epidemiology; ADNI, Alzheimer's Disease Neuroimaging Initiative; IMAGEN, Imaging Genetics; PING, Pediatric Imaging, Neurocognition, and Genetics Study; IMAGEMEND, Imaging Genetics For Mental Disorders; fBIRN: Functional Bioinformatics Research Network. In addition to the North American ADNI<sup>14</sup>, large-scale projects on the neuroimaging of Alzheimer's disease are underway in Europe, Japan, Australia, Korea, Taiwan and Argentina, with a new program in China and a planned program in Brazil<sup>15</sup>.

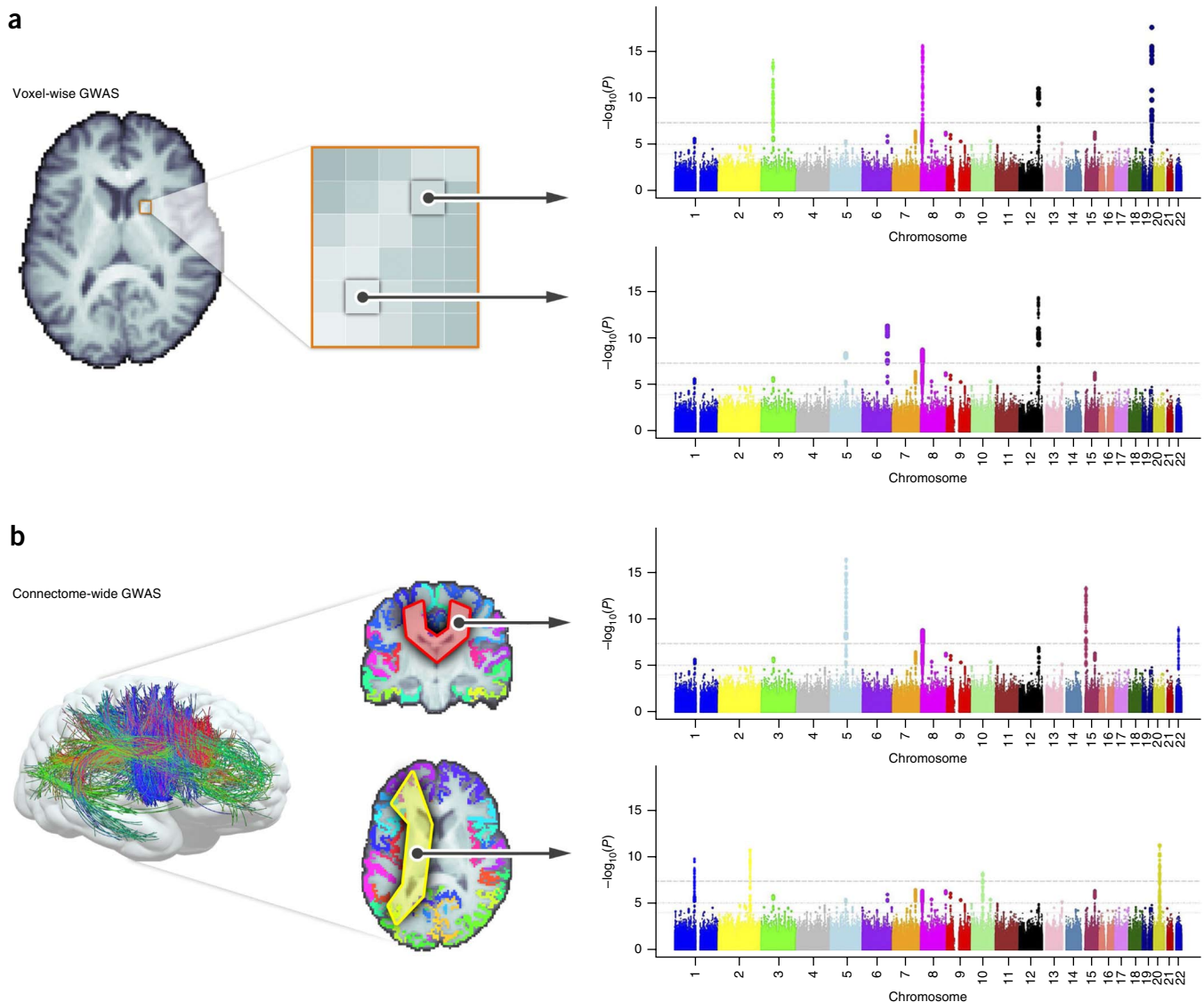
<sup>a</sup>Sample sizes are estimates as of April 2014. <sup>b</sup>Hibar, D.P., Stein, J.L., Renteria, M.E., Arias-Vasquez, A., Desrivieres, S. *et al.* (unpublished data).

One issue sometimes raised, when analyzing polygenic traits, is the fallacy of classical inference. Put simply, this means that, with a sufficient number of subjects, one will inevitably declare trivial effect sizes to be significant. While the effect sizes identified by GWAS are typically small, this does not indicate that the importance of the gene on the development of a trait or its effectiveness as a potential treatment target is trivial. For example, in studies of type 2 diabetes and cardiovascular diseases, the effect sizes of loci that are drug targets are small (<1% of the variation in disease risk), but the drugs themselves can be highly effective<sup>16–18</sup>. Thus, the effect size of a specific allele does not necessarily reflect the potential contribution of the pathway or mechanism to a trait.

**Data harmonization.** To amass the largest possible sample, data from different genotyping chips are typically prepared using standardized quality control procedures and then imputed using publicly available references from the HapMap (<http://www.hapmap.org/>) or the 1000

Genomes Project (<http://www.1000genomes.org/>). On the imaging side, automated algorithms can yield consistent measures from brain scans. For structural MRI, software packages such as FreeSurfer and FSL have been validated and widely used. For the main brain structures, these automated techniques give reproducible measures when a person is scanned repeatedly over time or on different scanners<sup>19</sup>. Further quality checks are possible when deriving measures from a large number of brain scans, such as spotting outliers from the distribution of values. Even so, genetic analyses have typically been more successful for the larger brain structures that tend to be measured more reliably.

**Searching maps of the brain.** A typical brain image—mapping brain activation with functional MRI, for example, or white matter microstructure with DTI—may have upwards of 2 million three-dimensional data points, or voxels (**Fig. 1a**). Statistical parametric maps have long been used by imagers to pinpoint regions where activation relates to behavior or to cognitive tasks performed in a scanner.



**Figure 1** Whole-brain GWAS. **(a)** Voxel-wise genetic association analysis. This kind of analysis involves a genome-wide search at each voxel in the brain, after aligning all subjects' images to a common template. **(b)** Extending this method to study brain connections, Jahanshad *et al.*<sup>30</sup> described connectome-wide searches. They combined diffusion-based MRI tractography and cortical parcellations to perform GWAS at all connections between cortical regions of interest. Artificial Manhattan plots are illustrated here, with thresholds shown based on a single GWAS. Despite the vast number of tests, promising findings emerged, even after correction, from these whole-connectome genetic screens.

Voxel-wise, genome-wide scans, or vGWAS<sup>20</sup>, extend this idea. They search the image, examining evidence for associations across the genome at each voxel in the image. Depending on the resolution adopted, vGWAS methods can aim to find genetic variants influencing up to 2 million locations in the brain. The sheer number of statistical tests requires sufficient computational power, as well as analytical rigor to avoid reporting false positive findings.

As a proof of principle, Stein *et al.*<sup>20</sup> screened 31,622 voxels and 448,293 SNPs to search for genetic variants related to regional brain volumes in around 700 participants scanned as part of the Alzheimer's Disease Neuroimaging Initiative (ADNI; <http://adni.loni.usc.edu/>). This initial study was underpowered, but later refinements focused on gene-based or pathway analyses and detected genetic effects that replicated across independent cohorts<sup>21</sup>.

Maps of brain circuitry can be screened as well (Fig. 1b). There is a wide range of methods to find fiber tracts in images, and we can extend the search to identify genetic effects on fiber bundles, one bundle at a time and one genetic variant at a time. In the first successful genome-wide screen of the human connectome, Jahanshad *et al.*<sup>22</sup> analyzed brain connections traced using tractography methods and found that a variant in the gene *SPON1* influenced brain connectivity *in vivo*. This gene's protein product is a key factor in the processing and cleavage of amyloid- $\beta$  precursor protein, APP—a key pathological hallmark of Alzheimer's disease. This finding linked risk for neurodegeneration in old age with brain integrity in young adulthood, around 50 years before Alzheimer's disease typically hits. Only a few weeks later, the same gene (though a different variant) was implicated in an independent study of cognitive decline<sup>23</sup>.

**Biological interpretation of GWAS results.** After identifying significant genetic effects, further work is needed to understand their biological mechanism and functional or behavioral significance. With a couple of notable exceptions (*APOE* and possibly a newly discovered Alzheimer's disease-associated variant in *TREM2* (refs. 24,25)), individual SNPs account for approximately 0.5% of the variance in imaging features tested so far. This effect size is comparable to that of SNPs affecting cognitive measures or psychiatric diagnosis. As with other GWAS, many imaging genetics studies follow up their top statistical findings with functional evaluations of the top SNPs.

SNPs identified in a GWAS of imaging data may affect gene expression, disease susceptibility or behavior—or they may have no detectable effects on any of these. Collaborative partnerships between genetics consortia in psychiatry or neurology (such as PGC or GERAD) and imaging consortia (such as ENIGMA) are beginning to show that some SNPs associated with brain volumes do indeed overlap with some SNPs that confer heightened risk for disease. Methods to test the overlap of two GWAS, such as rank-rank hypergeometric overlap (RRHO)<sup>26</sup>, can detect enrichment of SNPs affecting two traits, without relying on any one set of candidate genes selected in advance. They assess the rank order of SNPs in one GWAS and test whether it has any relation to the rank order of SNPs in another GWAS. Some brain-associated SNPs detected by screening images can predict histological measures from the same regions in mice when the homologous mouse gene is used as a predictor. Follow-up functional assessments of GWAS results may involve analyses of multiple kinds of images (MRI, functional MRI or DTI), as well as blood and plasma biomarkers. In the case of the *TREM2* variant, a potent risk factor for Alzheimer's disease<sup>24,25</sup>, carriers show faster brain atrophy and faster cognitive decline<sup>24,25,27</sup>. These multimodal imaging assessments can identify biological roles for the genes involved and their consequences for brain function.

## Multiple-testing issues

As genome-wide approaches to genetic analysis became possible, the genetics community derived statistical thresholds to limit the reporting of false positive discoveries that would fail to replicate. These thresholds were based on theoretical and empirical analyses of the number of independent tests that would be conducted if one were to examine evidence of association across the genome<sup>28,29</sup>. Basing these thresholds on the number of independent tests across the genome rather than the number of variants tested in any given study, as well as focusing on replication, has helped limit the reporting of type I errors in the literature.

In imaging also, decades of statistical work have been applied to evaluate whether a statistical association is likely robust. In imaging genetics, the significance threshold may depend on whether a study is a discovery study at the whole-brain and whole-genome levels (for which the threshold will be the strictest) or an attempt to replicate a previous finding. It may also change when the effects of disease-associated variants that already meet genome-wide significance are subsequently examined in neuroimaging data.

**How many tests?** Whole-genome analyses of whole-brain images involve running GWAS on multiple brain measures. For example, ENIGMA's second large-scale genetics study analyzed MRI scans from 29,000 people (Hibar, D.P., Stein, J.L., Renteria, M.E., Arias-Vasquez, A., Desrivieres, S. *et al.*, unpublished data). They performed a GWAS on 8 different brain measures. Analysis of correlations among the 8 measures revealed that the effective number of statistical tests was approximately 7. As such, the threshold for declaring a result significant was adjusted, by dividing the standard threshold for a genome-wide screen by 7, to  $P < (5 \times 10^{-8})/7 = 7.1 \times 10^{-9}$ .

Although neuroimaging data may appear very high dimensional (with millions of voxels), the effective number of degrees of freedom is generally much smaller. This is because imaging data have spatial contiguity properties that mean they are better treated as continuous functions of space. Thus, not all the tests are independent, as signals at neighboring voxels are correlated. Statistical parametric maps, the mainstay of voxel-based analyses, take into account the effective number of independent tests, which depends on, among other variables, the resolution of the scan and biological correlations of the signal<sup>30</sup>. If all the voxels in the brain were completely correlated, the effective number of tests would be 1, regardless of the scanner resolution. The correction for multiple comparisons is thus much less severe than might be anticipated at first glance, as we show in the next sections.

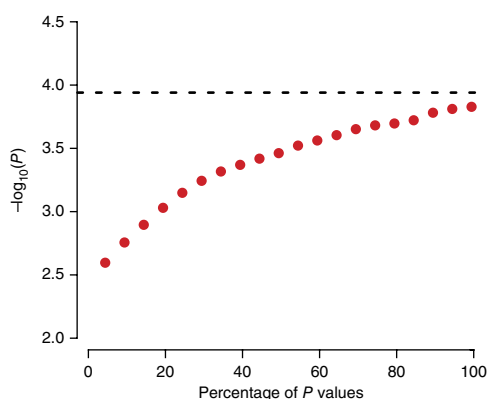
**Correcting for the number of voxels.** The history of brain mapping suggests why a strict voxel-based correction may not be widely adopted or used: genes are not likely to produce highly focal effects in images, and their effects are not likely to affect just a single voxel in isolation. Performing tests at each voxel without reference to what is going on elsewhere in the image is equivalent to a 'peak height' test or 'maximum statistic' test: it frames only one hypothesis about the most extreme statistic in the whole image. Neglecting information in the rest of the scan leads to tests that are not very sensitive to spatially extended effects. Imagers soon developed more complex, 'cluster-based' tests that considered the spatial extent of an effect in an image, in addition to its peak height<sup>31</sup>.

Brains also vary in structure. Because of this, it is not expected that the same activation will occur at the exact same three-dimensional spatial coordinate (or voxel) in every person scanned. Although genomes can be aligned to a reference panel, registration of brain images to a standard three-dimensional template does not perfectly align functional areas of the brain across individuals.

‘Warping’ algorithms can now align structural imaging data with high precision, but function still varies relative to the anatomy, and relative to the three-dimensional coordinates of the reference template. To accommodate individual differences in locations of effects, tests of suprathreshold clusters (regions of images exceeding a statistical threshold) are much more widely used. These tests are much more sensitive to nonfocal signals that are not especially strong in any one voxel but are strong enough to reject the null hypothesis of no effect. Arguably, the effect of a gene on the brain, such as a growth factor, will be widespread and not focused on any one specific voxel.

If a correction is made for the total number of tests in an image, a second, more intriguing problem may arise: for some data, the image-wide threshold may depend on the magnetic field or the resolution of the scanner and the actual detail in the signal. Some measures extracted from an image—for example, the size of a tract or major nucleus—do not depend highly on the scanner resolution, and thus the effective number of tests is unchanged. But other signals—from diffusion MRI, for example—resolve finer scale fibers and connections when the scanner resolution or field strength is increased. This is addressed with methods that find features at multiple scales, prioritizing the most spatially extended signals<sup>32</sup>. In genetics, similar issues were observed as the coverage and depth of whole-genome sequencing improved. To a large extent, the sequencing of publicly available reference samples in the HapMap and 1000 Genome Projects has helped resolve many of these issues.

**Setting a statistical threshold.** To assess the dimensionality of imaging data and estimate a genome-wide significance level for the analysis of voxel level data, we extracted whole-brain voxel-level data (2,019,874 voxels) for 208 healthy, unrelated individuals for whom genotype data (Illumina Human610-Quad BeadChip) were available. These data were computed with methods previously described<sup>33</sup> using data from ADNI collected with 1.5-T scanners. We performed quantitative tests of association at each voxel using 100 independent markers. We then sampled these results to assess the average minimum  $P$  value. Following Dudbridge and Gusnanto<sup>28</sup>, we then estimated the dimensionality of the voxel-level analysis to be 9,122. Adopting a conservative correction for testing 10,000 independent traits, the genome-wide, voxel-wide significance threshold would then be  $(5 \times 10^{-8})/10,000 = 5 \times 10^{-12}$ . We plot the growth in the effective number of tests, based on the



**Figure 2** Testing genetic associations in an image. As a greater percentage of the voxels in the brain image are tested for associations ( $x$  axis), the minimum uncorrected  $P$  value will increase ( $y$  axis) even if the SNPs have no effect (that is, the null hypothesis is true). To correct for this, the genome-wide significance threshold can be divided by the effective number of independent tests on the image, based on  $P$  values that actually occur in simulated null data with the same spatial coherence.

minimum uncorrected  $P$  value observed in null data (Fig. 2). Encouragingly, the largest imaging genetics studies to date show effects on the order of  $P \approx 10^{-23}$  (Hibar, D.P., Stein, J.L., Renteria, M.E., Arias-Vasquez, A., Desrivieres, S. *et al.*, unpublished data), so such a threshold is within the reach of available sample sizes.

Other considerations apply when testing effects of a single, GWAS-significant disease locus across the whole brain, or testing its effects on a single region of interest. If a candidate gene is well validated at the genome-wide level—as is, for example, *APOE*—one could perform a legitimate hypothesis test for its effects on a region of interest, such as the hippocampal volume (at  $P < 0.05$ ) or across the whole brain—either using the conservative Bonferroni correction of  $0.05/1$  or  $0.05/10,000$ , respectively, or less conservative cluster- or set-based tests, which are more commonly used in neuroimaging.

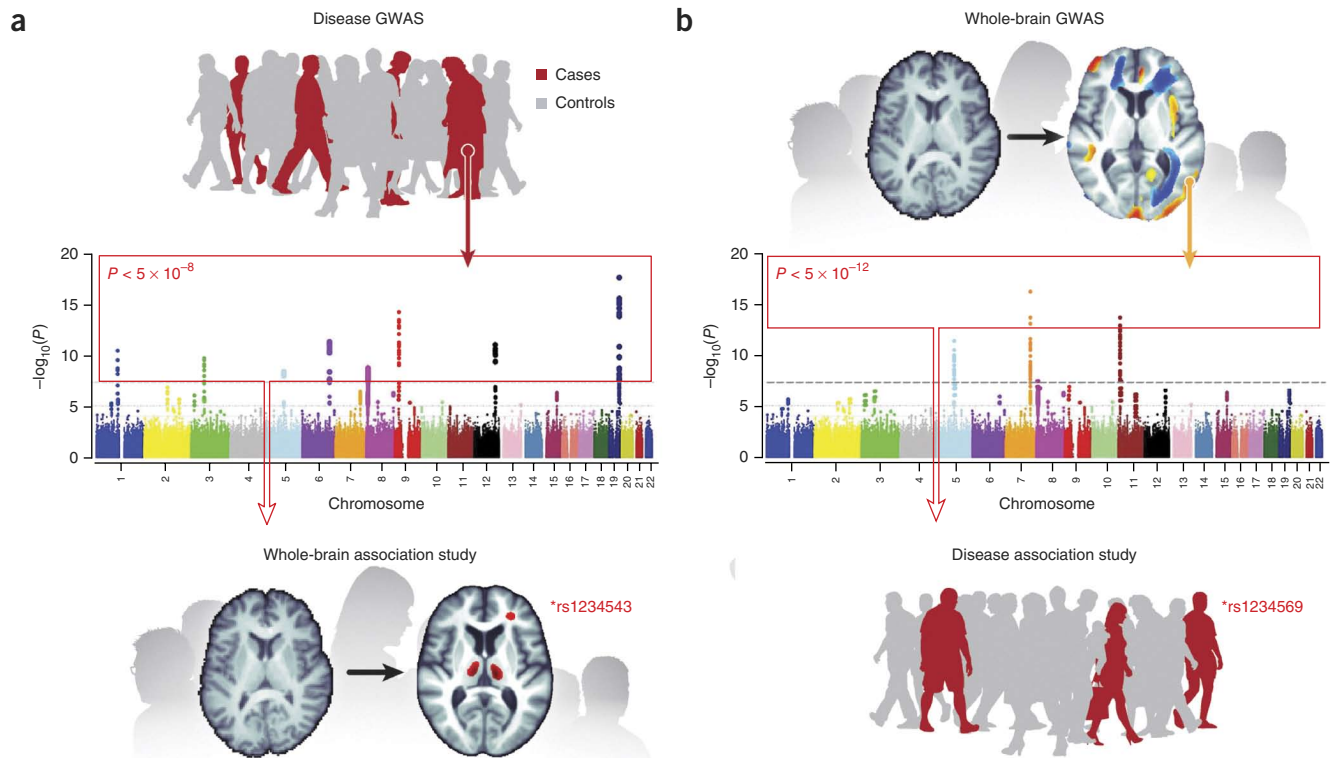
### Power and replication

Replication is vital to verify association signals and minimizes type I error in the literature. This push for replication comes from both the genetics and neuroimaging communities. If brain imaging studies are pooled after the data have been collected, it is unlikely that many similar cohorts will be scanned with exactly the same methods; if anything, the push for innovation has led to studies with different and ever-evolving designs. Even so, recent large-scale international imaging consortia galvanized the field to identify genetic associations with worldwide support. This offers a credible basis for prioritizing findings for follow-up analyses.

**How much power do we have?** Methods exist to compute how many subjects are needed to secure a given level of statistical power in a GWAS of a single trait, such as Alzheimer’s disease diagnosis, or the size of a specific region of the brain. The power to identify a given genetic effect, at a given voxel, can be determined if we can estimate the proportion of variance in the signal at that voxel due to the variant. One commonly used tool to calculate power is the Genetic Power Calculator (GPC)<sup>34</sup>.

For example, if we ran a univariate GWAS to find SNPs influencing a single voxel, the sample size required to detect a SNP effect as strong as ApoE4’s effect on Alzheimer’s disease (namely, 6% of the variance) with 80% power would be 660 participants. For a brain-wide voxel-wise search, we would apply an additional multiple testing correction. While work is currently underway to determine a suitably robust correction factor that can be used for both voxel- and connectome-level analyses, for the purposes of this illustration we will correct for 10,000–15,000 effectively independent voxels. Using this additional correction factor the threshold to declare significance increases from  $5 \times 10^{-8}$  to  $3 \times 10^{-12}$ . With this threshold, a sample of 1,015 individuals is needed to detect this same SNP effect in a voxel-level analysis with 80% power. For a more realistic SNP effect explaining around 0.5% of the variance, these sample sizes increase to 7,915 and 12,200 respectively—around the size of the first successful meta-analytic studies. These sample sizes are typically well beyond the reach of any single MRI study but well within the reach of collaborative consortia.

**Do discoveries generalize?** Replicated genome-wide association results are expected to generalize, but they may not if factors such as age or ethnicity influence the measures analyzed. Replication studies can determine whether discoveries generalize across populations—and if not, why not. This is especially important in genetic studies where a seemingly significant discovery could lead to further useful biological validations, including mouse knockout models and



**Figure 3** Approaches to imaging genetics. (a) Genome-wide association scans are performed in large-scale case-control studies to discover genetic risk factors for neurological diseases, implied by variants that surpass the genome-wide significance threshold of a single GWAS ( $P < 5 \times 10^{-8}$ ). A mechanistic understanding of how these variants alter specific structures and pathways in the living human brain can be accomplished through unbiased whole-brain association studies. Our example highlights associations with bilateral thalamic volume. (b) Through imaging genetic consortia, the vast accumulation of brain scans has allowed sample sizes to surpass that which is necessary to detect genome-wide significant loci associated with brain structure, even when taking into account a search across 10,000 independent tests in the image ( $P < 5 \times 10^{-12}$ ). Loci surpassing the brain-wide, genome-wide threshold not only reveal a genetic link to neuroanatomical deficits but also may help discover a genetic risk factor for disease through the affected neuroanatomical pathway. Numbers prefixed by “\*rs” represent unique SNPs within the genome.

genetic pathway analyses. If results are not robust, these efforts may be a waste of time.

### Tactics to reduce the search space

Perhaps, in these early days of imaging genomics, it would be prudent to take a conservative approach for reducing the search spaces. One could imagine a series of hierarchical tests, first performing whole-genome association analysis with brain-wide measures—structural and functional, as well as measures of brain connectivity and networks. This would ensure an unbiased screen at both the genomic and imaging levels. Subsequent follow up studies might include, for example, screening all significant disease GWAS hits with whole-brain data and then screening significant whole-brain hits, in independent samples, in specific regions of interest (Fig. 3). In all cases, the critical factor will be selecting the appropriate statistical threshold and ensuring sufficient power considering the expected effect size.

**Genomic efforts.** As so many tests can be performed on an image, many attempts have been made, in study design and analysis, to reduce the search space for an association. This includes selecting specific regions of interest, connections, pathways or genes. However, GWAS has typically shown little support for candidate genes studied in this way. For instance, the largest schizophrenia GWAS so far<sup>2</sup> did not support previous claims about several candidate genes involved

in the disorder. This is often the case for candidate genes selected on the basis of biological plausibility. Many polymorphic growth factor genes, such as *BDNF*, were initially expected to be strong candidates for affecting regional brain volumes. Perhaps surprisingly, in the ENIGMA-CHARGE analysis of 21,151 people, the only well-known candidate gene affecting hippocampal volume was *APOE*. A more nuanced interpretation is that some gene effects may be context dependent—they may depend on the age or diagnostic status of the cohort or interact with other factors that diminish their aggregate effect in a meta-analysis.

Depending on the goal at hand, rather than performing GWAS, it may be of interest to assess effects on the brain of candidate genes that have been previously identified in GWAS as exploratory follow-ups of the initial discoveries. For example, when a strong genome-wide supported variant or mutation has been found to be highly significant in a neurological disease on the basis of GWAS, then mapping its pathways of action in the brain is a high priority. While the rare variant in the *TREM2* gene was recently found to greatly increase disease risk in a genome-wide analysis of Alzheimer’s disease<sup>24,25</sup>, the neuroanatomical pathways remained unknown. Owing to the low prevalence of this variant (minor allele frequency  $\sim 0.005$ ), even current imaging sample sizes may not be able to detect these drastic effects if a genome-wide image-wide threshold of  $P \approx 5 \times 10^{-12}$  is applied. Instead, a candidate gene–type study of this variant was able to determine that elderly carriers of the mutation exhibit greatly

enhanced rates of temporal lobe atrophy<sup>27</sup>, offering a plausible explanation for the reduced cognitive functioning presented in the original discovery paper.

In addition to single SNP candidate analyses, the genetic search space can be narrowed by searching only genes or pathways of interest. However, without strong and robust evidence for such analyses, such approaches may be unwarranted.

**Dimension reduction.** As genomes and images are both high-dimensional, new mathematical techniques may be useful to pick out weak signals in these vast computational searches. One promising data reduction method uses penalized regression to detect weak signals in large data sets, using sparse regression or independent components analysis (ICA), among other data reduction models<sup>35</sup>. The mathematically related field of compressive sensing<sup>36</sup> has achieved great success in engineering and applied mathematics. Sparse methods try to model signals using as few predictors as possible: they penalize the number of predictors and favor those that explain the most variance in the data. These 'heavy-tailed priors' embody the prior belief that associations or effects have a sparse distribution. In other words, there are a small number of strong associations and a large number of weak associations. These data reduction methods boost the power and efficiency of the search when all predictors are weak. Even so, they often make effective predictions when aggregated<sup>37</sup>.

Whatever methods are used for dimensionality reduction, they need to be validated. In genetics, robust association studies are designed using simulations with real genomic data, and similar methods are being applied to different imaging tests to ensure that they are validated before becoming widely adopted.

**Results so far.** It is too early to predict what may be learned from genome-wide analyses of voxel-level data. The published studies that have addressed this research question are underpowered, with <1,000 participants (see **Supplementary Table 1**). Initial image-wide screens by Stein *et al.*<sup>20</sup> and Hibar *et al.*<sup>38</sup> on voxel-based GWAS failed to find any results after appropriate correction for multiple comparisons. There are some promising findings using dimension reduction, but these await replication.

While systematic comparisons are not yet possible, there are mixed results regarding the effects on the brain of popular candidate genes in psychiatry. As noted by ref. 39, many of the main Alzheimer's disease risk genes do indeed have robust effects on brain structure and pathology assessed using MRI and positron emission tomography (PET). By contrast, candidate genes encoding growth factors and genes in classical neurotransmitter pathways have not appeared among the top hits in the larger imaging GWAS. However, we must bear in mind that only the simplest brain measures have been tested in samples large enough to produce findings with genome-wide significance.

**Should we rank brain measures?** A more neuroscientifically driven approach is to rank features in brain images for genetic analysis if they are already implicated in a disease of interest. Meta-analyses of brain differences in different disorders can help to rigorously define these. ENIGMA's Schizophrenia<sup>10</sup> and Bipolar Disorder<sup>9</sup> Working Groups, for example, meta-analyzed regional brain volumes in the largest psychiatric MRI studies yet (2,047 patients and 2,059 controls for schizophrenia and 1,747 patients and 2,615 controls for bipolar disorder). Several basal ganglia components and the hippocampus show robust case-control differences in cohorts worldwide, and their volumes are highly heritable<sup>40</sup>.

Glahn *et al.*<sup>41</sup> advocated selecting brain measures on the basis of both their heritability and their relevance for a disease being studied. They defined the endophenotype ranking value (ERV) on the basis of the square root of the heritability ( $h_i^2$ ) of the disorder, the square root of the heritability ( $h_e^2$ ) of the brain measure, and their genetic correlation ( $\rho_g$ ):

$$ERV_{ie} = \left| \sqrt{h_i^2} \sqrt{h_e^2} \rho \right|$$

In schizophrenia, for example, decades of studies report morphometric differences in patients versus controls, for a range of different structures. The ENIGMA Schizophrenia Working Group is now ranking brain measures in order of their effect sizes for case-control differences; weighting these effects on the basis of their genetic correlation with the illness would give another ranking. This takes into account that there is not complete overlap between the variants influencing a disease-related trait and the disease itself. Clearly, a disease will not affect the brain identically in all patients. But the variety of cohorts in ENIGMA makes it possible to dig deeper into medication-related, geographic or demographic, and genetic factors to explain why brain differences vary so drastically across studies.

One intriguing tactic is to use information from one very widely available imaging modality to inform the genetic analysis of another. For example, worldwide, there are many more structural brain MRI scans from people who have been genotyped than functional imaging scans or scans of pathology (such as amyloid imaging). To empower the search for variants affecting data that is in shorter supply, methods such as seemingly unrelated regression (SUR) can exploit information from all data types, such as clinical and imaging features combined, boosting power to detect associations<sup>42</sup>. Other multivariate methods combine data from multiple scan types; however, they often require a complete data set, where all subjects have all sets of information. Because it is less commonly collected, there will never be as much PET, DTI or functional MRI data as standard anatomical MRI data; therefore, methods are helpful to impute or explicitly model these complex missing data structures<sup>43,44</sup>.

### Integrating diverse imaging methods and genomic variation

**Functional neuroimaging.** Genetic analysis of functional neuroimaging data is an active field<sup>45,46</sup>, as functional MRI has been widely used for over two decades. Data standardization across protocol designs and consortia has not yet been fully established. So far, researchers have mainly used structural MRI measures to examine the genetic architecture of the brain. Future studies are likely to develop large databases of EEG and functional MRI measures. It is too early to say whether these measures will offer more explanatory power as mediators of genetically driven variability in behavior and related risk for psychopathology. Some groups advocate the adoption of standardized functional probes across large, federally funded studies, as is being done for behavioral phenotypes (for example, PhenX).

Functional imaging measures may be more challenging to harmonize than structural measures of regional brain volumes. For anatomical MRI, automated analysis software, such as FreeSurfer (<http://freesurfer.net/>) and FSL (<http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/>), is widely used. This has led to an understanding of the statistical norms and reliability of structural brain measures that informs subsequent GWAS. Ongoing efforts in data harmonization for genetic analysis include functional MRI, DTI and EEG<sup>12,47</sup>.

Once processing steps have been harmonized, and reliable and genetically influenced phenotypes have been extracted from the functional scans of multiple groups, then tests of genetic associations can

be performed. As with structural imaging GWAS, functional neuroimaging associations with the genome will likely begin with a handful of phenotypes of interest. When studies are sufficiently powered, a whole-brain approach will ultimately supplant the need for candidate regions based on incomplete knowledge of disease circuits. In either case, multiple-comparisons corrections are needed that take into account the number of independent tests performed.

**Neurogenetic syndromes.** A somewhat more consistent picture emerges from neuroimaging studies of neurogenetic syndromes. These include 22q deletion syndrome, fragile X syndrome, Williams syndrome, Turner syndrome and others. Neuroimaging shows generally consistent patterns of abnormalities in these disorders. In general, their effects on the brain and behavior are far greater than the effects of common genetic variants (see our review in ref. 48). Some neurogenetic disorders result from deletion of chromosomal material important for development, leading to brain abnormalities and characteristic behavioral and cognitive effects. Williams syndrome, for example, arises as a result of a contiguous deletion of about 26–28 genes from chromosome 7 and is characterized by abnormal social and emotional processing and abnormal brain structure and function.

In Williams syndrome, cortical folding is more complex in some regions, with distinctive patterns of greater and lesser gray matter thickness in different cortical areas; polymicrogyria and gyral anomalies are common. A related neurodevelopmental disorder, fragile X syndrome<sup>49</sup> involves an abnormal expansion of the CGG trinucleotide repeat affecting the fragile X mental retardation 1 (*FMR1*) gene on the X chromosome. This leads to impaired executive function and visuo-spatial skills. In affected females, brain volume abnormalities correlate with reduction in systemically measured levels of the fragile X mental retardation protein (FMRP), the gene product of *FMR1*, supporting theories that FMRP is required for normal dendritic pruning in fronto-striatal-limbic pathways. The neuroimaging findings in these and other neurogenetic disorders have led to a growing understanding of many genetic factors crucial for axonal guidance and migration, cortical folding and dendritic pruning. Aberrations in these processes can lead to widespread macroscopic effects in neuroimaging data. Consortia (such as the ENIGMA-22q initiative) are also studying neurogenetic syndromes to assess brain differences in children with these neurodevelopmental disorders, and what factors promote better brain and cognitive outcomes.

**Structural and rare variation.** For traits such as psychiatric diagnosis, genetic analyses have assessed an assortment of genetic variants. At this point, few brain imaging studies have evaluated copy number variants (CNVs), rare variants or other kinds of genetic variation; common variants are more widely studied because standardized protocols are available and because sufficiently large sample sizes needed to power a study are only now becoming available for other variants. When examining the impact of rare variation on imaging phenotypes, several different analytic approaches may be adopted. However, as explored by Zuk *et al.*<sup>50</sup>, the sample sizes required for rare variant analyses are as large or larger than those required for common variant analyses.

Current SNP chip technologies genotype around 250,000 to 5 million variants across the genome. However, with the advent of next generation sequencing (NGS) technology it is now theoretically possible to evaluate not just some, but all genetic variants. Two common modalities for NGS experiments are whole-exome sequencing (WES), which aims to capture all genetic variation, common and rare,

in the coding regions of the human genome (with some additional content such as noncoding RNAs), and whole-genome sequencing (WGS), which aims to capture the entire genome. In practice, the proportion of the variants obtained from WES and WGS depend on several factors, including the depth of the coverage (how many times each section of DNA is sequenced), error rates and the sequence complexity of the region (repetitive sequences, for example, are still difficult to sequence). As discussed elsewhere in this issue<sup>51</sup>, WES and WGS can evaluate rare and *de novo* (newly arisen) variation. Insights into neuropsychiatric disorders, such as epilepsy<sup>52,53</sup> and autism<sup>54–57</sup>, are already emerging from WGS studies. These discoveries explain a modest percentage of the variance in disease risk.

## Recommendations

In summary, some imaging studies are now large enough to discover reproducible effects of single SNPs on simple brain measures—for example, the volume of the hippocampus<sup>3,4</sup> and other subcortical regions. Some neuroimaging projects, such as ADNI, began by harmonizing the image acquisition across multiple centers, achieving samples of over a thousand individuals. A far larger sample, in the tens of thousands, can be achieved by combining data from multiple centers, even if the data were originally collected with other goals in mind. Consortium efforts such as ENIGMA and CHARGE show, without doubt, that there is sufficient harmony in the acquisition and analysis of brain scans to identify consistent genetic effects across cohorts worldwide.

Effect sizes of individual SNPs are small for all imaging traits examined so far. There is a need for rigor in declaring genome-wide significance, especially when several traits are assessed in the same study. Arguably, this issue is not avoided by focusing on a small set of candidate genes or biological pathways.

When searching the whole brain for an effect of genetics, consortia now have sufficient power to correct for the effective number of independent tests performed on the image. The history of neuroimaging has shown that tests for focal effects are not as powerful in detecting widespread effects, and other tests of cluster size or image-derived features can be more powerful. Even so, given the vast number of association tests undertaken, approaches are needed to correct for type I error.

Dimension reduction methods for ‘big data’ abound, and there is no substitute for replication. More efficient generative models of genomic and imaging variation will be created in the future, seeking the most relevant features to test and thus leading to tests performed on a more restricted set of parameters<sup>58</sup>. Such models have been successful in modeling natural languages, where the latent structure in the signals is modeled using a generative grammar of much lower dimension than the signals used to represent or record the data.

In the future, mutational load or burden tests of rare mutations might be applied to imaging data, as well as any other genetic analyses amenable to quantitative traits. A brain structural pathway or functional network might be treated similarly to a cellular pathway or protein interaction network; a burden test for rare loss-of-function mutations could be applied in a voxel-wise or connectome-wise manner to aspects of brain networks. In statistical analyses of the connectome, for example, we can test for effects on pathways or networks at various levels<sup>59</sup>. Often we test for associations element-wise—at single connections or hubs—and then we move to a more global level, aiming to detect more systematic relationships in whole pathways or organizational properties. Future studies are likely to identify specific genetic associations for higher order properties of brain organization.



Finally, imaging studies using sequencing-based genotypes can effectively be limited to study genomic regions already implicated in standard GWAS. One may then use data reduction methods such as least absolute shrinkage and selection operator (LASSO) or ridge regression to estimate the variance explained in the neighborhood of a sequenced locus<sup>60–62</sup>. Bayesian methods can also target causal disease associated SNPs in densely genotyped regions<sup>63</sup>. The experience of global consortia, at least for the limited structural measures so far subjected to GWAS, suggests that measures extracted from brain images do not show larger effect sizes than other quantitative traits. With rapid innovations in image acquisition and processing, the pool of phenotypes available from imaging is essentially infinite.

## Conclusion

Ongoing collaborations between brain imagers and geneticists have led to worldwide consortia dedicated to finding factors that influence brain imaging measures. This review has focused on GWAS, which have accelerated consortium studies worldwide. As sample sizes and power increase in imaging studies, other types of genetic variation—rare variants and CNVs—may also be assessed. As geneticists and neuroimagers work together, there have been several surprises along the way. Genetic analyses of brain images still require thousands of individuals to find credible genetic associations that replicate across cohorts.

Brain imaging can measure signals at over 2 million locations in the brain; if each signal is surveyed while screening the genome as well, the resulting statistical search is so vast that collaboration is essential to secure enough power to detect associations. The future is likely to bring ever more complex imaging data into the hands of geneticists, for genetic screening of the elaborate networks derived from brain connectivity analyses<sup>30</sup> and other deep phenotyping efforts. These worldwide efforts will reveal new factors that affect the brain and mechanistic factors that drive resilience and risk for disease, as well as normal variations in brain function and structure over the lifespan.

*Note: Any Supplementary Information and Source Data files are available in the online version of the paper.*

## ACKNOWLEDGMENTS

S.E.M. is supported by an Australian Research Council Future Fellowship, 110100548. N.J. and P.M.T. were supported, in part, by US National Institutes of Health R01 grants NS080655, AG040060, EB008432, MH097268, MH085667, MH089722 and MH094343, and grants U01 AG024904 and P41 EB015922. B.M.N. was supported in part by R01 MH101244.

## COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>.

- Lambert, J.C. *et al.* Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat. Genet.* **45**, 1452–1458 (2013).
- Ripke, S. *et al.* Genome-wide association analysis identifies 13 new risk loci for schizophrenia. *Nat. Genet.* **45**, 1150–1159 (2013).  
**This is one of the largest and most successful GWAS aiming at identifying genetic risk factors for psychiatric disease.**
- Stein, J.L. *et al.* Identification of common variants associated with human hippocampal and intracranial volumes. *Nat. Genet.* **44**, 552–561 (2012).  
**This paper, along with three others published simultaneously, reported the first large-scale imaging genetics studies. Analyzing scans from over 20,000 individuals in aggregate, measures derived from brain images were shown to be viable and promising traits for genome-wide search. Successful GWAS discoveries were replicated among the collaborating consortia.**
- Bis, J.C. *et al.* Common variants at 12q14 and 12q24 are associated with hippocampal volume. *Nat. Genet.* **44**, 545–551 (2012).

- Strittmatter, W.J. *et al.* Apolipoprotein E: high-avidity binding to beta-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer disease. *Proc. Natl. Acad. Sci. USA* **90**, 1977–1981 (1993).  
**Genetically informed priors were used to pinpoint ApoE4 as a risk haplotype associated with late-onset Alzheimer's disease. Twenty years later, this is perhaps the disease risk allele with the highest odds ratio for any common neurological disorder. Neuroimaging studies have since mapped the effects of ApoE4 on various brain traits (volume differences, cortical thinning patterns, shape variations, white matter pathology, etc.) in many cohorts of young and elderly individuals.**
- Shaw, P. *et al.* Cortical morphology in children and adolescents with different apolipoprotein E gene polymorphisms: an observational study. *Lancet Neurol.* **6**, 494–500 (2007).  
**This was one of the first papers to show that ApoE4 carriers had different brain morphology when they were young, around 50 years before the typical age of onset of Alzheimer's disease.**
- Potkin, S.G. *et al.* Hippocampal atrophy as a quantitative trait in a genome-wide association study identifying novel susceptibility genes for Alzheimer's disease. *PLoS ONE* **4**, e6501 (2009).
- Flint, J. & Munafò, M.R. Candidate and non-candidate genes in behavior genetics. *Curr. Opin. Neurobiol.* **23**, 57–61 (2013).
- Hibar, D.P. *et al.* Meta-analysis of structural brain differences in bipolar disorder: the ENIGMA-Bipolar Disorder. in *Proc. Organization of Human Brain Mapping* (2013).
- Turner, J.A. *et al.* A prospective meta-analysis of subcortical brain volumes in schizophrenia via the ENIGMA Consortium. in *Proc. Organization of Human Brain Mapping* (2013).
- Hibar, D.P. *et al.* ENIGMA2: genome-wide scans of subcortical brain volumes in 16,125 subjects from 28 cohorts worldwide. in *Proc. Organization of Human Brain Mapping* (2013).
- Jahanshad, N. *et al.* Multi-site genetic analysis of diffusion images and voxelwise heritability analysis: a pilot project of the ENIGMA-DTI working group. *Neuroimage* **81**, 455–469 (2013).  
**This paper showed how to combine DTI data from many sources, using a standardized protocol for uniform phenotypic extraction from images. The standardization of data analysis was then shown to result in mostly stable but some unreliable phenotypes. Robust phenotypes will be carried forward into future multisite imaging GWAS meta-analysis studies.**
- Thompson, P.M. *et al.* The ENIGMA Consortium: large-scale collaborative analyses of neuroimaging and genetic data. *Brain Imaging Behav.* doi:10.1007/s11682-013-9269-5 (8 January 2014).  
**This review describes the many projects that imaging consortia can take on and the challenges of merging data sets from a vast variety of sources. The successes and limitations of imaging consortia are discussed, in terms of standardizing protocols, distributed analysis of data, and meta-analysis.**
- Weiner, M.W. *et al.* The Alzheimer's Disease Neuroimaging Initiative: a review of papers published since its inception. *Alzheimers Dement.* **8**, S1–S68 (2012).
- Carrillo, M.C., Bain, L.J., Frisoni, G.B. & Weiner, M.W. Worldwide Alzheimer's Disease Neuroimaging Initiative. *Alzheimers Dement.* **8**, 337–342 (2012).
- Lehmann, J.M. *et al.* An antidiabetic thiazolidinedione is a high affinity ligand for peroxisome proliferator-activated receptor gamma (PPAR gamma). *J. Biol. Chem.* **270**, 12953–12956 (1995).
- Pearson, E.R. *et al.* Variation in *TCF7L2* influences therapeutic response to sulfonylureas: a GoDARTs study. *Diabetes* **56**, 2178–2182 (2007).
- Global Lipids Genetics Consortium. Discovery and refinement of loci associated with lipid levels. *Nat. Genet.* **45**, 1274–1283 (2013).
- Wonderlick, J.S. *et al.* Reliability of MRI-derived cortical and subcortical morphometric measures: effects of pulse sequence, voxel geometry, and parallel imaging. *Neuroimage* **44**, 1324–1333 (2009).
- Stein, J.L. *et al.* Voxelwise genome-wide association study (vGWAS). *Neuroimage* **53**, 1160–1174 (2010).  
**Despite null findings, this extended the field of imaging genetics, showing that it is feasible to perform a genome-wide search on tens of thousands of imaging phenotypes. Many researchers have since published on reducing the search space in images or genomes, allowing a less extreme correction for multiple comparisons across such a large number of tests.**
- Hibar, D.P. *et al.* Alzheimer's disease risk gene, *GAB2*, is associated with regional brain volume differences in 755 young healthy twins. *Twin Res. Hum. Genet.* **15**, 286–295 (2012).
- Jahanshad, N. *et al.* Genome-wide scan of healthy human connectome discovers SPON1 gene variant influencing dementia severity. *Proc. Natl. Acad. Sci. USA* **110**, 4768–4773 (2013).  
**The study searched the entire set of cortical connections for genetic influences, after reducing the dimensionality of the human connectome for GWAS by focusing on only heritable measures. A successful connectome-wide, genome-wide screen resulted in a significantly associated SNP within the APP pathway that also correlated with brain volumetric differences in the elderly. A SNP in the same gene was found only weeks later, through a GWAS, to associate with cognitive decline in an independent elderly sample.**
- Sherva, R. *et al.* Genome-wide association study of the rate of cognitive decline in Alzheimer's disease. *Alzheimers Dement.* **10**, 45–52 (2014).
- Guerreiro, R. *et al.* *TREM2* variants in Alzheimer's disease. *N. Engl. J. Med.* **368**, 117–127 (2013).

25. Jonsson, T. *et al.* Variant of *TREM2* associated with the risk of Alzheimer's disease. *N. Engl. J. Med.* **368**, 107–116 (2013).
26. Plaisier, S.B., Taschereau, R., Wong, J.A. & Graeber, T.G. Rank-rank hypergeometric overlap: identification of statistically significant overlap between gene-expression signatures. *Nucleic Acids Res.* **38**, e169 (2010).
27. Rajagopalan, P., Hibar, D.P. & Thompson, P.M. *TREM2* and neurodegenerative disease. *N. Engl. J. Med.* **369**, 1565–1566 (2013).
28. Dudbridge, F. & Gusnanto, A. Estimation of significance thresholds for genome-wide association scans. *Genet. Epidemiol.* **32**, 227–234 (2008).  
**This paper proposed methods to correct for surveying large number of SNPs in performing a GWAS.**
29. Pe'er, I., Yelensky, R., Altshuler, D. & Daly, M.J. Estimation of the multiple testing burden for genome-wide association studies of nearly all common variants. *Genet. Epidemiol.* **32**, 381–385 (2008).
30. Kiebel, S.J., Poline, J.B., Friston, K.J., Holmes, A.P. & Worsley, K.J. Robust smoothness estimation in statistical parametric maps using standardized residuals from the general linear model. *Neuroimage* **10**, 756–766 (1999).
31. Hayasaka, S., Phan, K.L., Liberzon, I., Worsley, K.J. & Nichols, T.E. Nonstationary cluster-size inference with random field and permutation methods. *Neuroimage* **22**, 676–687 (2004).
32. Worsley, K.J., Marrett, S., Neelin, P. & Evans, A.C. Searching scale space for activation in PET images. *Hum. Brain Mapp.* **4**, 74–90 (1996).  
**This paper was one of the first to statistically test features at multiple scales in a brain image, allowing powerful detection of signals in PET scans of the brain.**
33. Hua, X. *et al.* Unbiased tensor-based morphometry: improved robustness and sample size estimates for Alzheimer's disease clinical trials. *Neuroimage* **66**, 648–661 (2013).
34. Purcell, S., Cherny, S.S. & Sham, P.C. Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. *Bioinformatics* **19**, 149–150 (2003).  
**This paper reports methods for evaluating the statistical power attainable in a genome-wide association study.**
35. Hibar, D.P., Kohannim, O., Stein, J.L., Chiang, M.C. & Thompson, P.M. Multilocus genetic analysis of brain images. *Front. Genet.* **2**, 73 (2011).
36. Candès, E.J., Romberg, J.K. & Tao, T. Stable signal recovery from incomplete and inaccurate measurements. *Commun. Pure Appl. Math.* **59**, 1207–1223 (2006).
37. Inkster, B. *et al.* Histone deacetylase gene variants predict brain volume changes in multiple sclerosis. *Neurobiol. Aging* **34**, 238–247 (2013).
38. Hibar, D.P. *et al.* Voxelwise gene-wide association study (vGeneWAS): multivariate gene-based association testing in 731 elderly subjects. *Neuroimage* **56**, 1875–1891 (2011).
39. Shen, L. *et al.* Genetic analysis of quantitative phenotypes in AD and MCI: imaging, cognition and biomarkers. *Brain Imaging Behav.* doi:10.1007/s11682-013-9262-z (5 October 2013).
40. den Braber, A. *et al.* Heritability of subcortical brain measures: a perspective for future genome-wide association studies. *Neuroimage* **83**, 98–102 (2013).
41. Glahn, D.C. *et al.* High dimensional endophenotype ranking in the search for major depression risk genes. *Biol. Psychiatry* **71**, 6–14 (2012).  
**Theoretically, a good phenotype for a successful GWAS must be heritable. To serve as an endophenotype for disease, it must also be strongly implicated in disease. This paper developed an algorithm to rank complex quantitative phenotypes from imaging, as well as other markers, in terms of their potential for success in GWAS.**
42. Jahanshad, N. *et al.* Seemingly Unrelated Regression empowers detection of network failure in dementia. *Neurobiol. Aging*. (in the press).
43. Yuan, L., Wang, Y., Thompson, P.M., Narayan, V.A. & Ye, J. Multi-source feature learning for joint analysis of incomplete multiple heterogeneous neuroimaging data. *Neuroimage* **61**, 622–632 (2012).
44. Xiang, S. *et al.* Bi-level multi-source learning for heterogeneous block-wise missing data. *Neuroimage* doi:10.1016/j.neuroimage.2013.08.015 (27 August 2013).
45. Glahn, D.C. *et al.* Genetic control over the resting brain. *Proc. Natl. Acad. Sci. USA* **107**, 1223–1228 (2010).
46. Potkin, S.G. *et al.* A genome-wide association study of schizophrenia using brain activation as a quantitative phenotype. *Schizophr. Bull.* **35**, 96–108 (2009).  
**Using brain activation as a quantitative phenotype, this paper used functional imaging scans from a consortium to perform one of the first imaging GWAS in a psychiatric disorder.**
47. Kochunov, P. *et al.* Multi-site study of additive genetic effects on fractional anisotropy of cerebral white matter: comparing meta and mega analytical approaches for data pooling. *Neuroimage* doi:10.1016/j.neuroimage.2014.03.033 (18 March 2014).
48. Dennis, E.L. & Thompson, P.M. Typical and atypical brain development: a review of neuroimaging studies. *Dialogues Clin. Neurosci.* **15**, 359–384 (2013).
49. Lee, A.D. *et al.* 3D pattern of brain abnormalities in Fragile X syndrome visualized using tensor-based morphometry. *Neuroimage* **34**, 924–938 (2007).
50. Zuk, O. *et al.* Searching for missing heritability: designing rare variant association studies. *Proc. Natl. Acad. Sci. USA* **111**, E455–E464 (2014).
51. Gratten, J., Wray, N.R., Keller, M.C. & Visscher, P.M. Light after the long dark tunnel: elucidating the genetic architecture of psychiatric disorders through large-scale genomics. *Nat. Neurosci.* **17**, 791–800 (2014).
52. Carvill, G.L. *et al.* Targeted resequencing in epileptic encephalopathies identifies *de novo* mutations in *CHD2* and *SYNGAP1*. *Nat. Genet.* **45**, 825–830 (2013).
53. Epi4K Consortium & Epilepsy Phenome/Genome Project. *De novo* mutations in epileptic encephalopathies. *Nature* **501**, 217–221 (2013).
54. Neale, B.M. *et al.* Patterns and rates of exonic *de novo* mutations in autism spectrum disorders. *Nature* **485**, 242–245 (2012).  
**This paper highlighted the importance of *de novo* genetic mutations in the risk for autism.**
55. O'Roak, B.J. *et al.* Sporadic autism exomes reveal a highly interconnected protein network of *de novo* mutations. *Nature* **485**, 246–250 (2012).
56. Sanders, S.J. *et al.* *De novo* mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature* **485**, 237–241 (2012).
57. Iossifov, I. *et al.* *De novo* gene disruptions in children on the autistic spectrum. *Neuron* **74**, 285–299 (2012).
58. Chattopadhyay, I. & Lipson, H. Data smashing. Preprint at <http://arxiv.org/pdf/1401.0742v1.pdf> (2014).
59. Duarte-Carvajalino, J.M. *et al.* Hierarchical topological network analysis of anatomical human brain connectivity and differences related to sex and kinship. *Neuroimage* **59**, 3784–3804 (2012).
60. Tibshirani, R. Regression shrinkage and selection via the Lasso. *J. R. Stat. Soc. Series B Stat. Methodol.* **58**, 267–288 (1996).
61. Kohannim, O. *et al.* Boosting power to detect genetic associations in imaging using multi-locus, genome-wide scans and ridge regression. in *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on 1855–1859* (IEEE, 2011).
62. Kohannim, O. *et al.* Discovery and replication of gene influences on brain structure using LASSO regression. *Front. Neurosci.* **6**, 115 (2012).
63. Wellcome Trust Case Control Consortium *et al.* Bayesian refinement of association signals for 14 loci in 3 common diseases. *Nat. Genet.* **44**, 1294–1301 (2012).
64. Ikram, M.A. *et al.* Common variants at 6q22 and 17q21 are associated with intracranial volume. *Nat. Genet.* **44**, 539–544 (2012).
65. Psaty, B.M. *et al.* Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium. Design of prospective meta-analyses of genome-wide association studies from 5 cohorts. *Circ. Cardiovasc. Genet.* **2**, 73–80 (2009).
66. Nymberg, C., Jia, T., Ruggeri, B. & Schumann, G. Analytical strategies for large imaging genetic datasets: experiences from the IMAGEN study. *Ann. NY Acad. Sci.* **1282**, 92–106 (2013).
67. Paillère Martinot, M.L. *et al.* White-matter microstructure and gray-matter volumes in adolescents with subthreshold bipolar symptoms. *Mol. Psychiatry* **19**, 462–470 (2014).
68. Eicher, J.D. *et al.* Genome-wide association study of shared components of reading disability and language impairment. *Genes Brain Behav.* **12**, 792–801 (2013).
69. Potkin, S.G. & Ford, J.M. Widespread cortical dysfunction in schizophrenia: the FBIRN imaging consortium. *Schizophr. Bull.* **35**, 15–18 (2009).