Reliability Issues in Imaging Genetics

By
Annika U. Eriksson

Capstone Project

Submitted in partial fulfillment of the requirements for the degree of
Master of Biomedical Informatics

Oregon Health and Science University

January 2017

School of Medicine

Oregon Health & Science University

CERTIFICATE OF APPROVAL

This is to certify that the Master's Capstone Project of

Annika U. Eriksson

"Reliability Issues in Imaging Genetics"

Has been approved

Shannon McWeeney, PhD
Capstone Advisor

Table of Contents

Acknowledgements

I would first like to thank Dr. Kemal Sönmez for encouraging me to take the first steps on this journey. I would also like to thank Dr. Eilis Boudreau and Dr. Beth Wilmot for their vast patience with me and the efforts that they have made to set me on the right path. Special thanks to Dr. Shannon McWeeney for her insight and for seeing me through to the end. Also thanks to the department for securing me the NLM Short Term Training grant which supported me in the early stages.

I am grateful to the many wonderful colleagues that have made my graduate school experience so memorable. My student experience was enriched by the moral support and professional guidance of Dr. Allison Fryer and Dr. Jackie Wirz. Most of all, I owe thanks to my family, for putting up with my absence and absent-mindedness and encouraging me every step of the way.

Exposition: Setting the stage

In behavioral neuroscience, questions of development and function are often pursued using either genetic or imaging approaches. The genetics of behavior has been studied in order to shed light on the processes that guide human psychological development. The underlying genetic basis for behavior has long been the subject of debate, but it has been accepted that genes play a role in both behavioral tendencies and susceptibility to disorders of the brain. Although heritability for certain brain-related disorders has been estimated at around 80-90%, as with other common traits this heritability is as yet mostly "missing", unexplained by single genetic variations.[1] Likewise, imaging has become an invaluable tool to researchers investigating the structure and function of the brain at the physiological level.[2,3] Detailed maps of the brain now describe the locations where particular functions are performed.[4] In this way, genetic and imaging data have complementary roles in guiding researchers closer to the biological mechanisms of behavior.

If known, the biological mechanisms underpinning mental and behavioral disorders could facilitate more precise treatment. However, modeling, predicting, or categorizing of human thoughts and decisions has limitations.[5,6] Any attempt that uses classical behaviorism or cognitive psychology results in subjective diagnoses.[7] Even modern understanding of behavior is based on clustered observations, with treatment options often dependent on guesswork.[8] As more empirical approaches, both genetics and brain imaging have been used to get closer to the underlying biology of psychiatric disorders and support the movement toward process-based, rather than symptom-based, classification.[9–12] While both fields have made progress independently, linking data from imaging and genetic studies can have even more power to reveal hidden biological mechanisms; for example, the differences in function observed in relation to variants in the COMT gene.[13,14] This approach can increase the potential for discovery while simultaneously presenting new computational challenges and amplifying existing issues of reliability for each of these data types.[15,16]

A challenge for behavioral geneticists lies with the heterogeneity of the traits in question confounded by the massive amounts of data being generated. While the study of genetics has undertaken dramatic advances including exponential increases in sample size due to the decreased cost of genotyping, the contributions of specific genes in such complex traits have not been adequately described.[17] At the same time, imaging has also proven useful in elucidating some biological mechanisms (for example, in cognition and memory),[12] but this field is likewise hampered by complexity. For example, some brain processes occur automatically at rest, while others are more active during a task; some involve the whole brain, while others are localized.[18] Also, the brain is intrinsically not stationary. Each neuron contains multiple types of channels in its membrane, exponentially increasing the pathways for signals to travel. In spite of these challenges, imaging and genetics both promise much insight into the biology of behavior.

Researchers in the emerging field of imaging genetics seek genes that are associated with both brain activity and behavioral traits. Combining imaging and genetics provides new challenges and opportunities. The traditional challenges of big data are amplified by the inherent

intricacy of both genetic and neural architecture.[19] Additionally, when modeling the relationships between genes, brain, and behavior, each step of the process is layered with unknowns and uncertainty. The inclusion of brain imaging data in genetic studies, on the other hand, is one way to reduce their complexity. This integration promises to lead to the demystification of behavior by placing its biological mechanisms in an anatomical context. This "endophenotype" approach, in which imaging is an intermediary between the genome and the mental state, provides an opportunity to probe the sometimes-indirect pathways involving genes, environment, and disrupted development.[20–22] Using the brain as an endophenotype may challenge the conceptual models of behavior, but actually increases the usefulness of the research.[23,24] Linking genetic and imaging data can model how genetic variables influence neural development, brain function, and ultimately behavior.[25] Understanding these risk factors for psychopathology is critical for intervention and prevention efforts.[26]

Though several consortia are actively working to collect large samples and develop the necessary methodological rigor for imaging genetics studies (Table 1), the existing literature shows a varied set of approaches with regard to study design, granularity of research questions, power enhancement approaches (Table 2), and analysis plans/methodological approaches. Also important to note is that when studies were not replicated, it was often due to lack of an appropriate sample. With this in mind, it seems apparent that replication is a more common practice in the more recent studies than the older ones, since more samples have become readily available (Table 3). In this paper, I will address issues of reliability facing the field of imaging genetics. Reliability (measurements and metrics of consistency or accuracy), in this context, is a time-sensitive matter requiring both excellent communication and sensitivity. I will examine the current approaches used to address these issues in each field separately, with examples from the large consortia that demonstrate both best practices and the computational expertise required. Specifically, I will attempt to delineate how appropriate strategies could be employed to produce more reliable imaging genetics research: research that is reproducible, replicable, rigorous, and robust.[27]

Foreshadowing: Importance of reliability

Science relies heavily on funding from public agencies and therefore the taxpayers, which in turn depend on and expect accurate results that can inform public policy and lead to scientific advances. Recent concerns about the lack of reliability across the sciences has damaged the perception of value and raised concern about scientific integrity. While the contributions of imaging genetics to behavioral neuroscience hold much promise, progress can only be made if a foundation is established for reliability. As these two high dimensional and complex data types (imaging and genetics) are integrated, it is critical to avoid the pitfalls researchers in each field have worked so diligently to overcome, and tackle novel problems that arise from the integration itself.

To avoid confusion, I will refer to the definitions put forth by statistician Jeff Leek when discussing measures of reliability (Box 1). In particular, Leek differentiated between

reproducibility and replicability, which have often been confounded in the literature but refer to specific procedures. *Reproducibility* improves confidence and demonstrates a lack of bias in the study, key for secondary analysis, while the key feature of a *replicable* study includes consistent results in independent studies. Without a replication step, there is no way to reliably make a claim about whether the results apply to the whole population. Each subsequent replication increases confidence in the reliability of the results. Consistent results from independent investigators are essential benchmarks for these procedures.[28,29]

Box 1. Leek's statistical definitions: reproducibility, replication[8]

Reproducibility. To reproduce a study is to arrive at exactly the same results using the same data, analysis plan, and code.

Replication. Drawing from the same population, using the same experimental design and analysis plan, but otherwise recreating the experiment and code, and arriving at consistent results.

Furthermore, I will borrow from esteemed professor and spinal cord regeneration specialist Oswald Steward,[30] who elaborates on the above two principles in application to neurobiological research and further discusses the concepts of robustness and rigor (Box 2).[27] When reviewing study results, one considers how broadly the claims can be applied. Decisions are made from the experimental design stage all the way through analysis that affect the interpretability and dependability of results. As a practical example of what it takes to achieve a high level of reliability, the Reproducibility Project used several sets of guidelines for evaluating rigor before choosing studies for full replication attempts.[31,32] These guidelines include Glenn Begley's "six red flags,"[33] Jason McDermott's two additional red flags,[34] Nature's reporting checklist for incoming submissions, and ARRIVE guidelines for reporting animal studies.[35] Most of these could be distilled into specific details at the experimental and analysis stages. The entirety of these criteria may be found in the Reproducibility Project: Cancer Biology.[31,36]

Box 2. Steward's definitions of robustness and rigor[17]

Robustness. Robust studies have far-reaching implications. This area also overlaps with scientific communication and affects the public opinion of how reliable — and valuable — the research appears.

Rigor. Rigorous research follows protocols that have governed the scientific method since the middle ages: proper blinding, using appropriate controls, and supplying enough detail that an outside source would be easily able to reproduce or replicate its findings.

Steward's "6 gold stars":

- Check submissions for "Begley's 6 red flags", and if present, require consideration of resulting caveats in Discussion sections.
- Require that papers report statistical power.
- Require statements about whether studies were done as "rolling experiments" and require information on timing of data collection.
- Require that all analyses be reported.
- Require a caveats/scientific rigor section in Discussions.
- Require specific indication of studies performed at the request of reviewers.

In a climate in which public confidence in science seems tenuous, reliability has become one of the most urgent areas of concern for many in the scientific community.[142-144] The need for investigators to be trained in best practices has motivated the creation of many Massive Online Open Courses (MOOCs), such as the Data Science Specialization sponsored by Johns Hopkins University.[28,29] The proponents of these data science educational outreach efforts attempt to lead the way to greater reliability by focusing on open data, data stewardship and transparency in research.

Rising Action: History of both fields independently

Both the study of genetics and techniques for brain imaging have come a long way in a relatively short time. The search for genetic links to disease, for example, evolved rapidly during the decades between the discovery of the structure of DNA and the successful completion of the Human Genome Project.[37] Since the advent of DNA sequencing, geneticists have sought to link particular variants to human traits and diseases by creating maps relative to known genetic markers.[37] Genetic mapping has advanced from genotyping whole families at large intervals to

extrapolate genetic distance,[38] to physical mapping using assembled whole genome sequences.[39] Likewise, even before the advent of brain imaging, a functional map of the brain was meticulously created by studying the effects of lesions on different regions of the brain.[40] Eventually, more detailed functional maps emerged from modern imaging technology and the ability to observe common areas of activation and deactivation.[41] With imaging genetics, integrating these functional brain maps with genetic maps creates a more informative model: a comprehensive map of the pathways from genetic factors and brain function to behavior.[26,42]

Imaging

To isolate these pathways from brain to behavior, a variety of imaging modalities have been employed. The choice of modality depends on the desired aspect of brain physiology. Cross-sectional X-ray views of the brain can be obtained by computerized tomography (CT) scans, often the modality used in an emergency to assess physical signs of trauma. Magnetic resonance imaging (MRI) uses a magnetic field and radio-frequency pulsation to produce high resolution three-dimensional data that can be used to measure structural components (e.g., brain size, feature size, relative dimensions within the organ). Brain structure is usually evaluated one dependent variable at a time by univariate analysis.[15] Complexity increases, however, with the study of function. Diffusion tensor imaging (DTI) allows observation of fiber tracts connecting the brain's regions. This more closely monitors the biochemical states of the brain by mapping the water diffusion properties of the tissues using MRI.[42] The structural information from DTI can be combined with functional data from another modality.[43]

Function, then, can be inferred from imaging blood flow during tasks with biologically active radioactive tracers using positron emission tomography (PET).[44] Similarly, the increase in blood oxygen level after a neuron fires, represented by the hemodynamic response function, is imaged by functional MRI (fMRI).[45] This is the ideal modality for recording brain activity when spatial resolution is desired and non-invasiveness preferred. If temporal resolution is a priority, electroencephalography (EEG), which records electrical activity on the scalp, can be used to augment the imaging data. The choice of imaging modality also depends on the level of granularity desired. The smallest unit of the MRI image is the voxel, a numerical representation of the signal at one location within the three-dimensional space of the brain. Therefore, one way to analyze MRI is to examine images voxel-by-voxel, using multivariate methods or employing dimension reduction strategies to increase statistical power. One simple such dimension reduction method is to examine relationships between multi-voxel regions of interest (ROIs), usually corresponding to brain structure. On a more global scale, graphical or network models are used to model whole brain connectivity. These methods apply to both structural (sMRI) and functional (fMRI) data, and can model the brain either at rest or during tasks. The signal from fMRI is noisy, and therefore the methods for statistically extrapolating activity from fMRI data include correcting for the hemodynamic response function as well as for movement and other technical artifacts.[46,47]

Genetics

Just as the challenge of imaging is to reduce complexity while capturing as much of the meaningful information in the data as possible, a variety of tools and approaches have been developed to address the same problem for genetic investigations. When it became cost effective to do so, genome-wide association studies (GWAS) were performed. GWASs scan the whole genome for statistically significant associations between each of thousands of single nucleotide polymorphisms (SNPs) and the trait in question.[48] This is challenging given the genome-wide search, as each statistical test increases the probability that a false positive will arise. Correcting for this bias eliminates all but the strongest associations, and leading to few significant and replicable SNP that emerge (note: this does not address causality).

In fact, most highly heritable traits have been shown to be associated with multiple small-effect variants. Added together, these only account for a fraction of the overall heritability of these traits. This is possibly due to limitations of the technology used in GWAS and the statistical challenges of high dimensional data. It appears likely that much of the heritability could be hidden in small-to moderate-effect variants that are somewhat common.[49] Larger sample sizes and full genetic sequencing may, in the future, reveal this to be the case.[50] Complex traits may also be passed on by inheritance through genetic architecture such as copy number variants, or by epistatic relationships in which two or more genes create a synergistic effect.[51] Other explanations for "missing heritability" include environment and epigenetics — alterations to the genome that result in changes in gene expression and functionality.[1,52] Finally, there may be more than one pathway to what is seen as a single common disease. No matter what is actually behind the heritability, genetic factors related to brain disorders have, in large part, continued to elude us.[1]

As the cost of whole-genome sequencing decreases, the ability to detect statistically significant associations in a case-control GWAS theoretically increases (as more individuals can be sequenced). In an attempt to increase both statistical power and interpretive meaning, candidate gene studies isolate sets of genes that have previously demonstrated a relationship to the trait or to a similar trait. However, candidate genes have accounted for only 15% of autism cases.[25]

In a landmark study, it was shown that for complex traits, testing for association of one million SNPs population-wide was more powerful than smaller scans of related individuals (noting this is highly dependent on phenotyping quality, among other variables).[53] Within five years of the first major GWAS, clinically relevant discoveries had been made, although concerns were also raised about the meaning of the sparse associations that were revealed.[54] As much of these issues were postulated to be due to study design, new strategies for dealing with replication issues in GWAS were developed and employed.[55] Over 100 biologically relevant genetic loci have been associated with schizophrenia (leveraging the Psychiatric Genomics Consortium and over 150,000 subjects).[56]

Whole-genome studies employ a variety of strategies to deal with the computational challenges of multiple testing. Many analysis plans include computing a polygenic risk score, or

weighted sum of associated SNPs, for individuals with and without the trait.[57] Another strategy for enhancing the power of these investigations is to utilize aggregate information (using a unit of analysis such as pathway or network). Gene-environment interactions may help explain the hierarchical and continuous nature of psychopathology. Finally, integrating imaging into a GWAS has the potential to resolve the overlapping, or comorbidity, of phenotypes that often appears through observational diagnosis.[58]

Imaging genetics

The first techniques for combining genetic and imaging data emerged around the year 2000.[15] In their simplest form, these studies chose a specific gene and predicted its effect in the brain based on its known biological function.[59] The next step in the evolution of imaging genetics was relating either multiple genes to a single measure of brain physiology, or a single genetic variant to multiple aspects of brain structure.[60,61] For example, one could look at a structural measurement in the brain, such as gray matter volume, and ask whether a particular gene or set of candidate genes is associated with this physical trait.[62–64] Genetic features range from a single gene to GWAS, and imaging features range from single generalized measure to whole-brain studies. Various such combinations of genetic features and imaging features have been used to analyze data from imaging genetics consortia (Table 3). [65,66]

Alternatively, one could ask how the entire genome is related to the whole brain by selecting features are the most highly differentiated between cases and controls using a multimodal fusion method.[67] Most of the studies in imaging genetics fit somewhere between two extremes, either on the many genes to few brain features end of the spectrum,[68,69] or the few genes to many brain features end. Integration strategies for imaging genetics can be primarily data driven or strongly hypothesis based, and all studies use at least one of several power enhancement approaches, each with its own unique benefits and drawbacks (Table 2). It is not uncommon to use more than one approach; however, a hypothesis-driven analysis would be more likely to use an *a priori* data reduction approach, whereas a hypothesis-free study would use a data-driven method. Within this research space, the goal is to capture as much of the complexity of psychiatric genetics as possible without losing the statistical power to see meaningful connections or being drowned out by technological noise and individual variability.[70]

This kind of multimodal integration comes with certain requirements. The following sections discuss the crisis of reliability in science and how it manifests at the intersection of the imaging and genetics fields. Finally, I will address the issues of robustness, reproducibility, replication, and rigor, using examples from the field.[27]

Crisis: Reliability in science

The problem of replication and reproducibility is a well-documented and constant presence in the era of big data, and one the scientific community at large can hardly afford to ignore.[71,72] According to one survey, most scientists have experienced replication failure first

hand.[71] Concerns about reliability could spell disaster for public and governmental support of scientific progress. Therefore, across disciplines it is important to address these issues preemptively, before money is spent on unreliable science and the collective reputation of biomedical research is compromised.

After selecting 100 promising papers, the Reproducibility Project: Psychology performed replications on each one and compared the results to the original studies. They reported "reproducibility" (they meant replicability) by assessing the statistical significance of the results, and reported effect sizes of the replication study compared to the original. They reported that replications reduced the number of significant findings by almost a third, particularly where there was weaker evidence in the original study. These findings show the importance of performing replication studies, as most of the replications returned smaller effect size estimates than the originals did in spite of adequate power.[31,32]

Reliability in genetics

The field of genetics has addressed reliability head-on with a number of initiatives, in response to the many reasons GWASs may fail to replicate. DbGAP addresses reliability by providing a platform for open data, addressing privacy issues, and providing guidelines for data sharing. Particularly important for reproducibility has been its emphasis on sharing analysis plans, documenting primary and secondary analyses and keeping a provenance of data types.[73] By using a diverse sample from multiple geographic areas and combining case-control analysis of seven disease states, The Wellcome Trust Case Control Consortium and Gene Environment Association Studies (GENEVA) multi-site GWAS consortium established many of the standards used in GWAS data analysis, underlining several important steps critical for reliability in subsequent genetic studies.[48,74,75] Namely, these studies described and accounted for the heterogeneity of its samples, justified and documented their data cleaning procedures, and demonstrated that replication of previous GWAS results in numerous cases can be obtained with sufficient statistical power.[48] Additionally, GWAS replication should always include testing the same markers and using the same analysis methods. With association studies of a complex trait, the definition and measurement, or "harmonization," of the phenotype must also remain consistent for the replication.[74,76]

Reliability in imaging

The reliability crisis in neuroscience, on the other hand, is just beginning.[77,78] Neuroinformatics, as a field, appears to lag behind its counterparts in bioinformatics. It encompasses a broad multi-disciplinary range of activities related to organizing data for studying structure-function relationships in the brain. The International Neuroinformatics Coordinating Facility (INCF), which oversees the development of community standards, has identified three areas of concern: computational methods, databases and sharing, and analysis tools.[79] Expertise is needed in each area, and reliability depends on efficient communication and up-to-date education on the latest developments.

Reliability Issues in Imaging Genetics

Neuroimaging has gradually moved toward standardizing and storing data in an effort to improve reliability, but functional imaging has presented new statistical challenges which are being actively discussed.[78,80–82] As many methods require further validation, empirical studies to determine the most reliable methods for neuroimaging are needed. This methodological vetting, along with education efforts in the realm of data science, has been cited as the key to improving robustness.[28,29]

In terms of rigor, it has been pointed out that many neuroimaging papers report only the statistical values of their findings.[82] Accepted guidelines for reporting fMRI studies require inclusion of estimated effect size along with statistical tests and claims, descriptions of subjects and tasks, quality control, multiple testing correction steps, and more.[80] Replication efforts appear to be hindered by the continuously evolving requirements for statistical rigor. For example, cluster-wise inference, a common method used in fMRI analysis to increase statistical power over the more accurate voxel-wise approach,[83] has been criticized for making assumptions about the smoothness of the underlying signals.[81]

On top of all this, neuroscience studies have been severely underpowered. Just as is has in GWAS, this lack of power directly affects the reliability of the results.[78] It also indirectly affects the perceived reliability of the field because these underpowered studies tend to be affected by publication bias. There must be awareness of the need for increased sample size as the neuroimaging field continues to pursue smaller effect targets. It is often argued that small imaging studies should only be for the low-hanging fruit that they are powered to detect.[78,80–82] Though sample size and power have been the subjects of much discussion,[84–87] one thing is clear: reliability is dependent on the appropriateness of the sample size to detect the expected effect size.[78] In general, pursuit of higher standards for reliability is critical across all branches of science, with direct effects on public perception and support of innovation and progress.[84]

Escalation: Reliability in imaging and genetics

More recent still are the problems faced when combining signals from neuroimaging and GWAS; studies in imaging genetics have been plagued with the same issues as those focused on its separate parts, (e.g., spurious signals, heterogeneity, small sample sizes).[17] Combining imaging with genetics compounds almost every statistical and methodological problem presented above.[88,89] Additionally, both fields acknowledge a large amount of individual variation, the effects of which are not yet fully known.[51,90] Additionally, experts in the fields of data management and processing pipelines have scrambled to keep up with the needs brought about by the integration of genetic and neuroimaging data.[91] Awareness of these issues is important, but action is essential. For the field to move forward, it must address the crisis with an appropriate level of urgency. Individuals with computational expertise can minimize its impact through education and collaboration, sharing of tools and pipelines, and communicating the results of experiments along with their caveats in a clear, precise, accessible voice.

One early success in imaging genetics was the association of a variant in the ZNF804A gene with psychosis schizophrenia and bipolar disorder. This imaging GWAS was aimed at the

connectivity endophenotype, which is quantifiably "disturbed" in subjects with psychosis.[92,93] Early imaging GWAS studies sparked a number of investigations into ZNF804A, including fine mapping of the gene,[94] association studies of other variants in the gene,[95] examination of structural and cognitive effects,[96, 97] and a successful replication.[98]

Imaging genetics consortia capitalize on the successes of GWAS; for example, the genetic risk of Alzheimer's disease was linked to variants in the APOE gene.[99] One of the largest consortia with both imaging and genetic data is Enhancing NeuroImaging Genetics through Meta-Analysis (ENIGMA), a worldwide effort of more than 70 institutions.[100] ENIGMA combines structural MRI and DTI, and arose from the broad availability of these technologies in the 1990's. The rapid pace of image collection required an organized effort to standardize data across the globe. Both structural and functional imaging required the community to develop "average" images on which to map results. Additionally, ENIGMA collaborators assert that imaging (particularly structural imaging) has generally reproduced well, especially as many analysis tools have become automated.[100] Combining genetics and imaging data from multiple sites to increase sample size has opened up the promising new frontier of using imaging as an endophenotype for brain disorders.[13]

The Alzheimer's Disease Neuroimaging Initiative (ADNI) sought to develop a better understanding of the course and disease model of Alzheimer's disease (AD) pathology and progression. The study included subjects with mild cognitive impairment (MCI) due to its overlap with, and sometimes progression into, AD in the general population at a rate of approximately 10-15%. This allowed them to look into risk factors and early biomarkers. Early results confirming the importance of APOE variants and candidate regions of interest were valuable in moving the field forward and providing directions for future research.[101] The consortium spawned publication of 200 papers in 6 years.[65]

The following two cases represent major attempts at maximizing the extent of both imaging and genetic data from ADNI:

Case 1. Stein et. al., 2010[102]

In 2010, the most exhaustive imaging genetic analysis to date was published: a voxelwise GWAS around 500,000 SNPs. As the MRI images each contained over 30,000 voxels, some sort of data reduction step was unavoidable due to the number of statistical tests that would be required. In this case, only the top voxel for each SNP was selected at the risk of losing information from less prominent associations. Yet even with this liberal a measure, when the total number of tests was corrected for, no SNP passed the significance threshold. The investigators didn't even try to replicate their study, but asserted that a comparatively small sample size would be required for replication. However, the most significant findings did map to genes with known biological functions, an encouraging revelation for imaging geneticists.[102]

Case 2. Hibar et. al., 2011[103]

A similar study added a gene-level data reduction step. Based on the fact that the gene is the ultimate unit of biological function (by way of its protein product), a multi-locus analysis was

performed. This statistical approach for high-dimensional data simply tests each group of SNPs in a gene as a unit. The hope is to increase the probability of locating a single gene with multiple weak-effect SNPs. This study still misses out on genes that are not top hits, due to an additional data reduction step of only selecting the top hit for each imaging measure. Nothing statistically significant was found using this approach, but again the results were deemed biologically reasonable.[103]

The landscape of ADNI-based studies includes imaging genetic methods of integration for both structural and functional imaging. The investigators pursued a range of genetic approaches from gene-based to genome-wide; from binary (case-control) studies to quantitative trait analysis (which increases statistical power, allowing for smaller sample size). The results from ADNI studies have confirmed a list of about 10 genes with verifiable associations with AD phenotypes, it is claimed. These promising results have potential to be used in treatment prediction within personalized medicine paradigms.[104] As an overall multi-site coordinated effort, ADNI focuses on the following issues central to reliability: making the process for data sharing easy, developing automated de-identification, and sustaining quality control.[65]

The additional requirements of functional imaging were addressed by the Function Biomedical Informatics Research Network (fBIRN), a case-control schizophrenia study across multiple sites. The initial analysis attempted to differentiate between activity during retrieval of memory as opposed to encoding of memory, two different neural processes.[62] Other participating studies found differences in connectivity in subjects with different clinical features, such as hallucinations.[105] Another examined auditory processing during task-based fMRI.[106] A conceptual meta-analysis of these studies indicates strong justification for continuing to probe these imaging indicators for associations with genetic factors.[107] Six genes out of a genome-wide scan were significant for efficiency in the dorsolateral prefrontal cortex network, measured as a quantitative trait. Interestingly, these genes are known to be involved in cortical development, supporting the hypothesis that schizophrenia is a developmental disorder.[62] The contributions of the fBIRN meta-analysis to the pursuit of robust research will be discussed below.

While there are many benefits to consortia, it is important to note their challenges. Namely, the differences in populations, geography, medical conditions and environmental exposures must all be accounted for. Most consortia do not collect longitudinal data on subjects, which would allow for the study of development and changes across the lifespan. The notable exception to this is the Avon Longitudinal Study of Parents and Children (ALSPAC), which began in the early 1990s and is still following subjects. ADNI also follows up longitudinally with subjects. Both Case 1 and Case 2, above, were affiliated with ADNI and considered moderately successful given that they generated sets of candidate genes for further study.[65]

The INCF, as stated previously, organizes its efforts in neuroimaging research around the areas of computational methods, databases and sharing, and analysis tools.[79] In the following sections, I will discuss some of the ways the imaging genetics community is addressing these concerns. Standardized formats for and increased access to data, along with development of

advanced methods and tools, are the rallying cries of the reliability movement. These endeavors work together to serve the interest of robustness, reproducibility, replication, and rigor.[27]

Resolution: Robustness

The scientific community needs robust methods to handle the size of genetic and imaging datasets and the inherent variability within. The foundation of imaging analysis is to make sense of noisy signals picked up by highly sensitive instruments; thus, true signal can be obscured by processing noise or by batch effects unwittingly created by elements of experimental design. Extracting interesting features from complex multi-dimensional data often relies on decomposing the data into its most important sources of variation using latent variable methods. Principle Component Analysis (PCA), Independent Component Analysis (ICA), and Canonical Correlation Analysis (CCA) are all variations on this theme. Between- and within-group component analysis can call attention to potential batch effects and sources of extraneous noise. In a hypothesis-free analysis, these methods can be used for discovering unanticipated sources of signal with biological meaning. Such techniques are routinely used to reduce the computational burden posed by large amounts of data.

Another way to combat the noise in large heterogenous cross-modality studies is by leveraging prior knowledge.[108–110] *A priori* methods of data reduction can be used on the imaging or genetic side, or both. One IMAGEMEND study examined the relationship between functional imaging and a genetic score, which was based solely on 14 SNPs previously found significant to memory in a GWAS study.[64] In this case, the genetic data was drastically reduced to a single measure, resulting in a univariate analysis.

Data-driven methods performed on multimodal datasets have potential to detect disorder-related signatures, provide evidence of unexpected relationships, and generate new hypotheses. Other solutions have included dimension reduction methods such as downsampling, which has the advantage of allowing cross-modality and cross-disorder analyses.[111,112] An IMAGEMEND project demonstrated a data driven linked ICA method that related combinations of structural and functional imaging features to schizophrenia.[113] This exact approach has not yet been applied in an imaging genetic study; although Vince Calhoun, with his background in electrical engineering, has developed another variation on ICA for functional imaging combined with GWAS data.[112]

Systems-level analysis is advantageous if a robust solution is required. For example, a robust regression method outperforms the default mass univariate methods that result in a high level of false positives.[114] In the ENIGMA dataset, basic measures of network organization have been proven heritable. Using data from multiple sites to increase power, analysis of preliminary data detected associations between genes and structural connections in the brain. This structural connectivity measurement is a power enhancement strategy that increases the range of the phenotype and may explain other traits shared by individuals, such as behavior or disease risk.[115]

In addition, fBIRN reported on the variations caused by technical and individual differences in a small sample using multiple types of machines and locations. Phase I assessed

these differences by following a small cohort of subjects to ten different scanners across the country. Phase II and III centered on collecting data from schizophrenic subjects and healthy controls.[116] It is important to be aware of the expected baseline variation before making assumptions about larger studies. A meta-analysis of many different studies must control for demographic and location based sources of variation. In this way, preliminary data gathered from imaging genetic consortia may provide a starting point for future robust pursuits.

Resolution: Reproducibility

The first step in verifying the reliability of experimental results is to reproduce the experiment, so as to check that the results of the analysis are not affected by who performs it.[29,117,118] A large part of the pursuit of reproducibility is a need for increased data sharing. In a recent review, big data expert Martin Wiener asserted that open data is the key to overcoming the reproducibility challenge in neuroscience. His recommendations for improving data sharing may be summarized by incentives, discoverability, and sustainability (Box 3).[119]

Box 3. Wiener's reproducibility recommendations: data sharing[81]

Incentives. Data-centric efforts should be just as rewarding as traditional research endeavors, with metrics to determine the success of sharing efforts, a data sharing index identifier (similar to publishing index identifiers). An infrastructure will need to be easy and streamline the incentives, otherwise many will be discouraged from sharing.

Discoverability. Metadata needs to be useful. There has been much speculation about potential barriers to data sharing, and why it hasn't caught on in the culture of neuroscience.

Sustainability. Keeping pace with growth by safeguarding the existence of data management and sharing platforms.

From a funding and publication standpoint, there is little external motivation for an investigator to reproduce another's analysis, much less replicate an entire study. It seems much more rewarding to introduce a new hypothesis or develop a distinct approach to a slightly new problem, but this makes it all the more difficult to make direct comparisons between studies. Smaller studies that are deficient in statistical power are harder to reproduce, but perhaps preregistration at the experimental design stage would help reduce the waste of performing a non-reproducible study.[78,80–82]

Any improvement to open data access is a mechanism for success.[120] Long-term strategies depend on maintaining distributed systems with interoperable queries, workflows, and pipelines.[91] Russell Poldrack is the Director of the Stanford Center for Reproducible Neuroscience, which runs the OpenfMRI data repository and developed the Brain Imaging Data Structure (BIDS) for organizing and describing MRI datasets.[121,122] Poldrack has personally demonstrated how difficult it is to ensure full reproducibility of even a simple experiment on a single subject (himself).[123] His team has gone to great lengths to make their data and code available, even providing access to a computing platform where needed.

Other collaborative imaging genetics repositories emphasize open data and reproducibility as well.[124,125] fBIRN is committed to publicly sharing its data, which has yielded several important studies on schizophrenia.[126] The project put all of its imaging data (structural and functional) in a standardized format, along with neurocognitive testing data, into the Human Imaging Database (HID), logging about 80 downloads per month.[116] All the imaging data from fBIRN is available to the public for subsequent analyses,[62] and one of the goals of the IMAGEMEND consortium is to benchmark and test new computational methods for reproducibility.[127]

Resolution: Replication

A true replication draws from the same population, using the same experimental design and analysis plan, but otherwise recreating the experiment and code, and arriving at consistent results. Replication is essential for identifying and reducing false positives, both in GWAS and in imaging studies. While determining a suitable significance threshold ensures detection of the smallest true signal possible, the replication is the gold standard in verifying the reliability of previous results.

It has been reported that false positive rates in fMRI studies can be as high as 70%. Both meta-analyses and the rigorous validation of methods through permutation testing are needed to minimize false positives. Most methods for fMRI haven't been validated with real data, and experts are still in the process of determining the nature of the effects of spatial autocorrelation, technical artifacts, and noise, and their effects on false positive rates.[81] Recent imaging genetics studies divide their samples into independent "discovery" and "replication" datasets. One significant SNP was found, associated with connectivity between regions of interest, which in turn were associated with dementia. An exploratory analysis (in need of replication) showed promise for larger sample sizes able to make genetic discoveries.[115]

The effective number of statistical tests on a genetic-by-imaging analysis is an ongoing problem.[101] Imaging genetics presents an enhanced multiple comparison problem over either neuroimaging or genetic studies alone. Many strategies have been proposed for dealing with multiple comparisons.[128] One of these is data reduction, which theoretically limits the amount of noise interfering with signal. Data reduction methods overcome the lack of power in studies less than 1000 subjects. Study designs and methods are constantly evolving to meet the demand for innovation.[88] Lack of consistency in study design makes comparison difficult, and a thorough

analysis plan pooling data within consortia and across studies is even more statistically problematic when the methods are so diverse.

A common manifestation of the data reduction approach is to pool information from local brain regions of interest (ROIs) in order to reduce the number of statistical tests. It is important to note that the strategy of constraining one's study to specific brain regions does not lower the rate of false positives, according to Meyer-Lindenberg;[129] however, ROI studies have other benefits which include being computationally less expensive than a voxelwise study. Another data reduction strategy, ranking imaging features by heritability, has been a successful form of prioritization.[130,131] For data reduction, the advantage of multivariate methods is a proven increase in power; however, there is enhanced difficulty in interpreting results of multivariate methods due to their complexity.[104]

Reproducibility metrics can be used as a benchmark for optimal processing pipelines in imaging data. Traditionally, imaging analysts focused on pipelines that resulted in biologically interesting results, or reasonable results. This method is prone to bias.[132] Better to use metrics that don't depend on results as much as the ability to reproduce the results in secondary analyses. Another metric traditionally used for benchmarking was the p-value, but it turns out this does not indicate degree of reproducibility at all, and also introduces bias to the preprocessing pipeline.[89]

Resolution: Rigor

Traditional barriers to science, which included access to technology, are disintegrating in the era of cloud based services. Anyone with a laptop is able to participate in science, and it is easier than ever for someone with computational or analytical skills to make a contribution. In this ecosystem, open data can thrive, and along with it an implicit understanding that results may be objectively double-checked.[119] This growing emphasis on rigor can be seen in The Human Connectome Project, still in its data acquisition phase, making every effort to document its protocols as each phase of the project unrolls.[133]

According to Wiener, lack of rigorous standards are one of the many challenges to fostering a data sharing ecosystem in neuroscience (Box 4). He paints the landscape his field with the broad strokes of cultural, technical and practical challenges.[119] I propose that rigorous imaging genetic research also faces challenges in these three dimensions.

Box 4. Wiener's challenges to research: rigor edition[81]

Cultural. Overcoming the reliability crisis is even more critical now as science as a whole is under attack by anti-intellectualism. Failure to replicate results threatens to discredit and reduce support for progress in science. The complexity of the problem by nature demands that tackling it must be a team effort. One investigator cannot find the solution alone. (This can be compared to particle physics and astronomy). People with internal motivation are contributing much in this team building approach.

Technical. Numerous sub-communities have created various types of data and scales of analysis. This has led to a challenge for interoperability between labs and tools. Determining the statistical threshold for significance is not a simple task because not all tests are independent. Genetic markers are in linkage disequilibrium due to being inherited together, and there are varying degrees of this throughout the genome. Imaging measures have spatial correlations that are still not well understood, and functional imaging signals possess additional temporal autocorrelation along with spatial non-independence.

Practical. The movement needs experimentalists, architects, app engineers, data scientists, scientific users and educators. The levels of involvement can be distilled into three or four essential roles existing at the levels of data, infrastructure, apps, and algorithms.

Open data is a fundamentally cultural principle that is part of today's generation's way of thinking.[134] The Allen Institute for Brain Science is a prototype of social neuroscience that can be an example of science unperturbed by political tides, and captures the spirit of the team-based approach needed in these collaborative times.[134] Advocates are needed for governance of incentives, database standards, and communicating the general importance of data sharing and the reliability of science. As a community grassroots movement, experts are needed on ethics, privacy, and security concerns.[119]

Then there is communication with the non-scientist. Leveraging all of the technical data from various efforts against each other by making them available, compatible, and user friendly.[135]

Rigorous analysis steps are needed to evaluate potential pitfalls at every step of an experiment. Errors and assumptions waste time and money, and can lead to retractions. For

example, an oversight recently led to statistical issues when a gene set analysis contained a SNP that was tagged in several genes.[136] That this error was missed demonstrates the importance of rigor, particularly as this field is interdisciplinary and new. Another rigorous step to take if possible: use a separate dataset for analysis after selecting variables of interest to avoid bias.[137]

Conclusion: A call to action

The merging of GWAS with neuroimaging appears to be a natural pairing to explain the biological roots of complex disease. While genetics seeks explain the heritability of behavior, imaging illuminates its neural processes. The idea that linking genes and brain function might better describe mental processes provides hope for a frustratingly difficult-to-treat array of mental disorders. If the collaborative environment can continue to press forward with reliable science, great strides might be made in the areas of diagnosis, treatment, and prevention.

In preparing this paper, I reviewed the major publications related to the imaging genetics consortia in Table 1. All groups included in the table are GWAS studies that also have imaging data on some or all of their subjects, with two exceptions. In the Saguenay Youth Study, imaging was the focal point, with genetic conclusions drawn primarily from the close relatedness of the study participants: an isolated population with low genetic variation due to a strong founder effect.[138] When genotyping was required, it was done one SNP at a time.[139] The Human Connectome Project also does not currently have genetic data on its subjects, being a five-year project begun in 2012 with genotyping on the agenda for year five.[133]

Ease of access to consortium data serves several purposes. Meta-analyses increase statistical power while simultaneously exposing issues with combining data from multiple sites.[42,90] This has laid the foundation for harmonizing imaging genetics standards. In the case of ADNI in particular, it has inspired experimentation with multiple methods of integrating GWAS and imaging data that have yielded some interesting findings.[15,66] As new phases of these projects are added, they will likely incorporate continued refinement of both imaging and genetic preprocessing pipelines.

A crucial question for computational scientists is how to incorporate a whole genome analysis into a functional imaging study in order to detect (or to not rule out) a potential signal that has not been previously identified. Moving forward, researchers are looking at ever smaller effect sizes, and this requires larger samples.[78] These studies are also less likely to replicate, and when they do, the resulting effect sizes are expected to be even smaller upon secondary analysis.[32] This problem is not surprising, but being aware of the issues affecting reliability reveal several opportunities for improvement.

Next frontiers

A new dimension of complexity emerges with the study of time-varying connectivity, capturing information from the constantly changing brain while more accurately representing its function than static measurements.[140] Methods and standards will be needed to handle dynamic

paradigm. Other data modalities (epigenetic, proteomic, copy number variations, etc.) will also be important in the coming years.[67] Consideration must be given to the issues each data type brings to the table, and experts must learn to share their skills and knowledge to avoid costly assumptions. In addition, the methods that have been evolving for integrating imaging and genetics will be valuable in the movement toward multimodality studies.[15]

The process of science today, in which complex concepts are described with data and teased apart by statistical methods, then stretched out into associations to create publishable claims, has created a tangle of challenging opportunities. The robustness of these claims and their interpretation must be addressed.[141] Individuals must demand accountability from each other for open sharing of data and code to reproduce results;[142] they must give incentives for replication and encourage a culture of rigorous standards in both practice and reporting.[143] From the experimental design phase, to analysis and development of methods, to interpretation and communication of results to the press and the public, experts are needed. They are especially needed to address issues of reliability at each phase. Government regulatory agencies must understand the implications of reliability in research as they decide on important matters such as approving tests, treatments, and other products. The field of behavioral neuroscience and that of imaging genetics in particular have a critical need for this kind of computational expertise at every level.

I have not discussed such practices as p-hacking, the garden of forking paths, or the file drawer problem. These all have a role in reliability problems in science,[29] but in spite of their obvious contribution, I chose not to include fraud in my analysis. Instead, I propose that the scientific community should encourage secondary analysis. Improving overall reliability will prevent errors from propagating, regardless of their source. The next frontier is to create an ongoing system of checks, like the criteria used by the Reproducibility Project,[32] but global and proactive. Focusing the collective energies of imaging genetics on robustness, reproducibility, replication, and rigor will help us to avoid the pitfalls other fields have already met and overcome.

Bibliography

1. Maher B. Personal genomes: The case of the missing heritability. Nature 2008;456:18–21.

2. Poldrack RA, Congdon E, Triplett W, Gorgolewski KJ, Karlsgodt KH, Mumford JA, et al. A phenome-wide examination of neural and cognitive function. Sci Data 2016;3:160110.

3. Berman MG, Jonides J, Nee DE. Studying mind and brain with fMRI. Soc Cogn Affect Neurosci 2006;1:158–61.

4. Ogawa S, Tank DW, Menon R, Ellermann JM, Kim SG, Merkle H, et al. Intrinsic signal changes accompanying sensory stimulation: functional brain mapping with magnetic resonance imaging. Proc Natl Acad Sci U S A 1992;89:5951–5.

5. Lilienfeld SO. The Research Domain Criteria (RDoC): an analysis of methodological and conceptual challenges. Behav Res Ther 2014;62:129–39.

6. First MB, Kendler KS, Leibenluft E. The Future of the DSM: Implementing a Continuous Improvement Model. JAMA Psychiatry [Internet] 2016;Available from: http://dx.doi.org/10.1001/jamapsychiatry.2016.3004

7. Kendler KS. The transformation of American psychiatric nosology at the dawn of the twentieth century. Mol Psychiatry 2016;21:152–8.

8. Howland RH. Pharmacogenetic testing in psychiatry: not (quite) ready for primetime. J Psychosoc Nurs Ment Health Serv 2014;52:13–6.

9. Tosto G, Reitz C. Genomics of Alzheimer's disease: Value of high-throughput genomic technologies to dissect its etiology. Mol Cell Probes 2016;30:397–403.

10. Liu Y, Wang Y, Huang C, Zeng D. Estimating personalized diagnostic rules depending on individualized characteristics. Stat Med [Internet] 2016;Available from: http://dx.doi.org/10.1002/sim.7182

11. Fu CHY, Mourao-Miranda J, Costafreda SG, Khanna A, Marquand AF, Williams SCR, et al. Pattern classification of sad facial processing: toward the development of neurobiological markers in depression. Biol Psychiatry 2008;63:656–62.

12. Passingham D, Sakai K. The prefrontal cortex and working memory: physiology and brain imaging. Curr Opin Neurobiol 2004;14:163–8.

13. Meyer-Lindenberg A, Weinberger DR. Intermediate phenotypes and genetic mechanisms of psychiatric disorders. Nat Rev Neurosci 2006;7:818–27.

14. Meyer-Lindenberg A, Kohn PD, Kolachana B, Kippenhan S, McInerney-Leo A, Nussbaum R, et al. Midbrain dopamine and prefrontal function in humans: interaction and modulation by COMT genotype. Nat Neurosci 2005;8:594–6.

15. Liu J, Calhoun VD. A review of multivariate analyses in imaging genetics. Front Neuroinform 2014;8:29.

16. Roffman JL, Weiss AP, Goff DC, Rauch SL, Weinberger DR. Neuroimaging-genetic paradigms: a new approach to investigate the pathophysiology and treatment of cognitive deficits in schizophrenia. Harv Rev Psychiatry 2006;14:78–91.

17. McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, Ioannidis JPA, et al. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. Nat Rev Genet 2008;9:356–69.

18. Ramus F. Genes, brain, and cognition: a roadmap for the cognitive scientist. Cognition 2006;101:247–69.

19. McIntosh A, Deary I, Porteous DJ. Two-back makes step forward in brain imaging genomics. Neuron 2014;81:959–61.

20. Gottesman II, Gould TD. The endophenotype concept in psychiatry: etymology and strategic intentions. Am J Psychiatry 2003;160:636–45.

21. Preston GA, Weinberger DR. Intermediate phenotypes in schizophrenia: a selective review. Dialogues Clin Neurosci 2005;7:165–79.

22. Cannon TD, Keller MC. Endophenotypes in the genetic analyses of mental disorders. Annu Rev Clin Psychol 2006;2:267–90.

23. Flint J, Munafò MR. The endophenotype concept in psychiatric genetics. Psychol Med 2007;37:163–80.

24. Kendler KS, Neale MC. Endophenotype: a conceptual analysis. Mol Psychiatry 2010;15:789–97.

25. Abrahams BS, Geschwind DH. Connecting genes to brain in the autism spectrum disorders. Arch Neurol 2010;67:395–9.

26. Hyde LW, Bogdan R, Hariri AR. Understanding risk for psychopathology through imaging gene-environment interactions. Trends Cogn Sci 2011;15:417–27.

27. Steward O. A Rhumba of "R's": Replication, Reproducibility, Rigor, Robustness: What Does a Failure to Replicate Mean? eNeuro [Internet] 2016;3. Available from: http://dx.doi.org/10.1523/ENEURO.0072-16.2016

28. Leek JT, Peng RD. Opinion: Reproducible research can still be wrong: adopting a prevention approach. Proc Natl Acad Sci U S A 2015;112:1645–6.

29. Patil P, Peng RD, Leek J. A statistical definition for reproducibility and replicability [Internet]. 2016. Available from: http://biorxiv.org/lookup/doi/10.1101/066803

30. Steward O, Popovich PG, Dietrich WD, Kleitman N. Replication and reproducibility in spinal cord injury research. Exp Neurol 2012;233:597–605.

31. Errington TM, Iorns E, Gunn W, Tan FE, Lomax J, Nosek BA. An open investigation of the reproducibility of cancer biology research. Elife [Internet] 2014;3. Available from: http://dx.doi.org/10.7554/eLife.04333

32. Open Science Collaboration. PSYCHOLOGY. Estimating the reproducibility of psychological science. Science 2015;349:aac4716.

33. Begley CG. Six red flags for suspect work. Nature 2013;497:433–4.

34. McDermott JE. Reproducibility: two more red flags for suspect work. Nature 2013;499:284.

35. Kilkenny C, Browne WJ, Cuthill IC, Emerson M, Altman DG. Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. PLoS Biol 2010;8:e1000412.

36. Template for coding studies [Internet]. Google Docs [cited 2016 Dec 4];Available from: https://docs.google.com/spreadsheets/d/16ztw220fMW4fQhMdbriedn1qhfTbdGKr7xZrofZ0BC4/edit?usp=embed_facebook

37. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. Nature 2001;409:860–921.

38. Botstein D, White RL, Skolnick M, Davis RW. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet 1980;32:314–31.

39. Park ST, Kim J. Trends in Next-Generation Sequencing and a New Era for Whole Genome Sequencing. Int

Neurourol J 2016;20:S76–83.

40. Lashley KS. Brain mechanisms and intelligence: A quantitative study of injuries to the brain. Chicago: University of Chicago Press; 1929.

41. Lancaster JL, Rainey LH, Summerlin JL, Freitas CS, Fox PT, Evans AC, et al. Automated labeling of the human brain: A preliminary report on the development and evaluation of a forward-transform method. Hum Brain Mapp 1997;5:238–42.

42. Thompson PM, Ge T, Glahn DC, Jahanshad N, Nichols TE. Genetics of the connectome. Neuroimage 2013;80:475–88.

43. Guye M, Bartolomei F, Ranjeva J-P. Imaging structural and functional connectivity: towards a unified definition of human brain organization? Curr Opin Neurol 2008;21:393–403.

44. Shulman GL, Fiez JA, Corbetta M, Buckner RL, Miezin FM, Raichle ME, et al. Common Blood Flow Changes across Visual Tasks: II. Decreases in Cerebral Cortex. J Cogn Neurosci 1997;9:648–63.

45. Ogawa S, Lee TM, Kay AR, Tank DW. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. Proc Natl Acad Sci U S A 1990;87:9868–72.

46. Logothetis NK, Pauls J, Augath M, Trinath T, Oeltermann A. 10.1038/35084005 [Internet]. Nature2001;412:150–7. Available from: http://www.nature.com/doifinder/10.1038/35084005

47. Lindquist MA. The Statistical Analysis of fMRI Data. Stat Sci 2008;23:439–64.

48. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007;447:661–78.

49. Gibson G. Rare and common variants: twenty arguments. Nat Rev Genet 2012;13:135–45.

50. Cirulli ET, Goldstein DB. Uncovering the roles of rare variants in common disease through whole-genome sequencing. Nat Rev Genet 2010;11:415–25.

51. Gelernter J. Genetics of complex traits in psychiatry. Biol Psychiatry 2015;77:36–42.

52. Cariaga-Martinez A, Saiz-Ruiz J, Alelú-Paz R. From Linkage Studies to Epigenetics: What We Know and What We Need to Know in the Neurobiology of Schizophrenia. Front Neurosci 2016;10:202.

53. Risch N, Merikangas K. The future of genetic studies of complex human diseases. Science 1996;273:1516–7.

54. Visscher PM, Brown MA, McCarthy MI, Yang J. Five years of GWAS discovery. Am J Hum Genet 2012;90:7–24.

55. Gorlov IP, Moore JH, Peng B, Jin JL, Gorlova OY, Amos CI. SNP characteristics predict replication success in association studies. Hum Genet 2014;133:1477–86.

56. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. Nature 2014;511:421–7.

57. Hamshere ML, Langley K, Martin J, Agha SS, Stergiakouli E, Anney RJL, et al. High loading of polygenic risk for ADHD in children with comorbid aggression. Am J Psychiatry 2013;170:909–16.

58. van den Heuvel MP, van Soelen ILC, Stam CJ, Kahn RS, Boomsma DI, Hulshoff Pol HE. Genetic control of functional brain network efficiency in children. Eur Neuropsychopharmacol 2013;23:19–23.

59. Hariri AR. The neurobiology of individual differences in complex behavioral traits. Annu Rev Neurosci 2009;32:225–47.

60. Whelan R, Conrod PJ, Poline J-B, Lourdusamy A, Banaschewski T, Barker GJ, et al. Adolescent impulsivity phenotypes characterized by distinct brain networks. Nat Neurosci 2012;15:920–5.

61. Eicher JD, Powers NR, Miller LL, Akshoomoff N, Amaral DG, Bloss CS, et al. Genome-wide association study of shared components of reading disability and language impairment. Genes Brain Behav 2013;12:792–801.

62. Potkin SG, Ford JM. Widespread cortical dysfunction in schizophrenia: the FBIRN imaging consortium. Schizophr Bull 2009;35:15–8.

63. Debette S, Bis JC, Fornage M, Schmidt H, Ikram MA, Sigurdsson S, et al. Genome-wide association studies of MRI-defined brain infarcts: meta-analysis from the CHARGE Consortium. Stroke 2010;41:210–7.

64. Luksys G, Fastenrath M, Coynel D, Freytag V, Gschwind L, Heck A, et al. Computational dissection of human episodic memory reveals mental process-specific genetic profiles. Proc Natl Acad Sci U S A 2015;112:E4939–48.

65. Weiner MW, Veitch DP, Aisen PS, Beckett LA, Cairns NJ, Green RC, et al. The Alzheimer's Disease Neuroimaging Initiative: a review of papers published since its inception. Alzheimers Dement 2012;8:S1–68.

66. Meda SA, Narayanan B, Liu J, Perrone-Bizzozero NI, Stevens MC, Calhoun VD, et al. A large scale multivariate parallel ICA method reveals novel imaging-genetic relationships for Alzheimer's disease in the ADNI cohort. Neuroimage 2012;60:1608–21.

67. Calhoun VD, Sui J. Multimodal fusion of brain imaging data: A key to finding the missing link(s) in complex mental illness. Biol Psychiatry Cogn Neurosci Neuroimaging 2016;1:230–44.

68. Xia Z, Chibnik LB, Glanz BI, Liguori M, Shulman JM, Tran D, et al. A putative Alzheimer's disease risk allele in PCK1 influences brain atrophy in multiple sclerosis. PLoS One 2010;5:e14169.

69. Wan J, Kim S, Inlow M, Nho K, Swaminathan S, Risacher SL, et al. Hippocampal surface mapping of genetic risk factors in AD via sparse learning models. Med Image Comput Comput Assist Interv 2011;14:376–83.

70. Casey BJ, Soliman F, Bath KG, Glatt CE. Imaging genetics and development: challenges and promises. Hum Brain Mapp 2010;31:838–51.

71. Baker M. 1,500 scientists lift the lid on reproducibility. Nature 2016;533:452–4.

72. Reality check on reproducibility. Nature 2016;533:437.

73. Mailman MD, Feolo M, Jin Y, Kimura M, Tryka K, Bagoutdinov R, et al. The NCBI dbGaP database of genotypes and phenotypes. Nat Genet 2007;39:1181–6.

74. Bennett SN, Caporaso N, Fitzpatrick AL, Agrawal A, Barnes K, Boyd HA, et al. Phenotype harmonization and cross-study collaboration in GWAS consortia: the GENEVA experience. Genet Epidemiol 2011;35:159–73.

75. Cornelis MC, Agrawal A, Cole JW, Hansel NN, Barnes KC, Beaty TH, et al. The Gene, Environment Association Studies consortium (GENEVA): maximizing the knowledge obtained from GWAS by collaboration across studies of multiple conditions. Genet Epidemiol 2010;34:364–72.

76. Kraft P, Zeggini E, Ioannidis JPA. Replication in genome-wide association studies. Stat Sci 2009;24:561–73.

77. Glatard T, Lewis LB, Ferreira da Silva R, Adalat R, Beck N, Lepage C, et al. Reproducibility of neuroimaging analyses across operating systems. Front Neuroinform 2015;9:12.

78. Button KS, Ioannidis JPA, Mokrysz C, Nosek BA, Flint J, Robinson ESJ, et al. Power failure: why small sample size undermines the reliability of neuroscience. Nat Rev Neurosci 2013;14:365–76.

79. Bjaalie JG, Grillner S. Global neuroinformatics: the International Neuroinformatics Coordinating Facility. J Neurosci 2007;27:3613–5.

80. Poldrack RA, Fletcher PC, Henson RN, Worsley KJ, Brett M, Nichols TE. Guidelines for reporting an fMRI study. Neuroimage 2008;40:409–14.

81. Eklund A, Nichols TE, Knutsson H. Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. Proc Natl Acad Sci U S A 2016;113:7900–5.

82. Chen G, Taylor PA, Cox RW. Is the Statistic Value All We Should Care about in Neuroimaging? [Internet]. 2016. Available from: http://biorxiv.org/lookup/doi/10.1101/064212

83. Woo C-W, Krishnan A, Wager TD. Cluster-extent based thresholding in fMRI analyses: pitfalls and recommendations. Neuroimage 2014;91:412–9.

84. Button KS, Ioannidis JPA, Mokrysz C, Nosek BA, Flint J, Robinson ESJ, et al. Confidence and precision increase with high statistical power. Nat Rev Neurosci 2013;14:585–6.

85. Quinlan PT. Misuse of power: in defence of small-scale science. Nat Rev Neurosci 2013;14:585.

86. Ashton JC. Experimental power comes from powerful theories - the real problem in null hypothesis testing. Nat Rev Neurosci 2013;14:585.

87. Bacchetti P. Small sample size is not the real problem. Nat Rev Neurosci 2013;14:585.

88. Medland SE, Jahanshad N, Neale BM, Thompson PM. Whole-genome analyses of whole-brain data: working within an expanded search space. Nat Neurosci 2014;17:791–800.

89. Strother S, La Conte S, Kai Hansen L, Anderson J, Zhang J, Pulapura S, et al. Optimizing the fMRI data-processing pipeline using prediction and reproducibility performance metrics: I. A preliminary group analysis. Neuroimage 2004;23 Suppl 1:S196–207.

90. Psaty BM, O'Donnell CJ, Gudnason V, Lunetta KL, Folsom AR, Rotter JI, 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 2009;2:73–80.

91. Dinov ID, Petrosyan P, Liu Z, Eggert P, Zamanyan A, Torri F, et al. The perfect neuroimaging-genetics-computation storm: collision of petabytes of data, millions of hardware devices and thousands of software tools. Brain Imaging Behav 2014;8:311–22.

92. Esslinger C, Walter H, Kirsch P, Erk S, Schnell K, Arnold C, et al. Neural mechanisms of a genome-wide supported psychosis variant. Science 2009;324:605.

93. O'Donovan MC, Craddock N, Norton N, Williams H, Peirce T, Moskvina V, et al. Identification of loci associated with schizophrenia by genome-wide association and follow-up. Nat Genet 2008;40:1053–5.

94. Williams HJ, Norton N, Dwyer S, Moskvina V, Nikolov I, Carroll L, et al. Fine mapping of ZNF804A and genome-wide significant evidence for its involvement in schizophrenia and bipolar disorder. Mol Psychiatry 2011;16:429–41.

95. Steinberg S, Mors O, Børglum AD, Gustafsson O, Werge T, Mortensen PB, et al. Expanding the range of ZNF804A variants conferring risk of psychosis. Mol Psychiatry 2011;16:59–66.

96. Donohoe G, Rose E, Frodl T, Morris D, Spoletini I, Adriano F, et al. ZNF804A risk allele is associated with relatively intact gray matter volume in patients with schizophrenia. Neuroimage 2011;54:2132–7.

97. Walters JTR, Corvin A, Owen MJ, Williams H, Dragovic M, Quinn EM, et al. Psychosis susceptibility gene ZNF804A and cognitive performance in schizophrenia. Arch Gen Psychiatry 2010;67:692–700.

98. Riley B, Thiselton D, Maher BS, Bigdeli T, Wormley B, McMichael GO, et al. Replication of association between schizophrenia and ZNF804A in the Irish Case-Control Study of Schizophrenia sample. Mol Psychiatry 2010;15:29–37.

99. Sadigh-Eteghad S, Talebi M, Farhoudi M. Association of apolipoprotein E epsilon 4 allele with sporadic late onset Alzheimer`s disease. A meta-analysis. Neurosciences 2012;17:321–6.

100. Thompson PM, Stein JL, Medland SE, Hibar DP, Vasquez AA, Renteria ME, et al. The ENIGMA Consortium: large-scale collaborative analyses of neuroimaging and genetic data. Brain Imaging Behav 2014;8:153–82.

101. Shen L, Kim S, Risacher SL, Nho K, Swaminathan S, West JD, et al. 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 2010;53:1051–63.

102. Stein JL, Hua X, Lee S, Ho AJ, Leow AD, Toga AW, et al. Voxelwise genome-wide association study (vGWAS). Neuroimage 2010;53:1160–74.

103. Hibar DP, Stein JL, Kohannim O, Jahanshad N, Saykin AJ, Shen L, et al. Voxelwise gene-wide association study (vGeneWAS): multivariate gene-based association testing in 731 elderly subjects. Neuroimage 2011;56:1875–91.

104. Shen L, Thompson PM, Potkin SG, Bertram L, Farrer LA, Foroud TM, et al. Genetic analysis of quantitative phenotypes in AD and MCI: imaging, cognition and biomarkers. Brain Imaging Behav 2014;8:183–207.

105. Wible CG, Lee K, Molina I, Hashimoto R, Preus AP, Roach BJ, et al. fMRI activity correlated with auditory hallucinations during performance of a working memory task: data from the FBIRN consortium study. Schizophr Bull 2009;35:47–57.

106. Lavigne KM, Menon M, Woodward TS. Impairment in subcortical suppression in schizophrenia: Evidence from the fBIRN Oddball Task. Hum Brain Mapp 2016;37:4640–53.

107. Lewis CM, Levinson DF, Wise LH, DeLisi LE, Straub RE, Hovatta I, et al. Genome scan meta-analysis of schizophrenia and bipolar disorder, part II: Schizophrenia. Am J Hum Genet 2003;73:34–48.

108. Kilaru V, Iyer SV, Almli LM, Stevens JS, Lori A, Jovanovic T, et al. Genome-wide gene-based analysis suggests an association between Neuroligin 1 (NLGN1) and post-traumatic stress disorder. Transl Psychiatry 2016;6:e820.

109. Chen J, Calhoun VD, Pearlson GD, Perrone-Bizzozero N, Sui J, Turner JA, et al. Guided exploration of genomic risk for gray matter abnormalities in schizophrenia using parallel independent component analysis with reference. Neuroimage 2013;83:384–96.

110. Ge T, Feng J, Hibar DP, Thompson PM, Nichols TE. Increasing power for voxel-wise genome-wide association studies: the random field theory, least square kernel machines and fast permutation procedures. Neuroimage 2012;63:858–73.

111. Vounou M, Nichols TE, Montana G, Alzheimer's Disease Neuroimaging Initiative. Discovering genetic associations with high-dimensional neuroimaging phenotypes: A sparse reduced-rank regression approach. Neuroimage 2010;53:1147–59.

112. Liu J, Pearlson G, Windemuth A, Ruano G, Perrone-Bizzozero NI, Calhoun V. Combining fMRI and SNP data to investigate connections between brain function and genetics using parallel ICA. Hum Brain Mapp 2009;30:241–55.

113. Brandt CL, Doan NT, Tønnesen S, Agartz I, Hugdahl K, Melle I, et al. Assessing brain structural associations with working-memory related brain patterns in schizophrenia and healthy controls using linked

independent component analysis. Neuroimage Clin 2015;9:253–63.

114. Fritsch V, Da Mota B, Loth E, Varoquaux G, Banaschewski T, Barker GJ, et al. Robust regression for large-scale neuroimaging studies. Neuroimage 2015;111:431–41.

115. Jahanshad N, Kochunov PV, Sprooten E, Mandl RC, Nichols TE, Almasy L, et al. Multi-site genetic analysis of diffusion images and voxelwise heritability analysis: a pilot project of the ENIGMA-DTI working group. Neuroimage 2013;81:455–69.

116. Keator DB, van Erp TGM, Turner JA, Glover GH, Mueller BA, Liu TT, et al. The Function Biomedical Informatics Research Network Data Repository. Neuroimage 2016;124:1074–9.

117. Patil P, Peng RD, Leek J. A statistical definition for reproducibility and replicability [Internet]. 2016. Available from: http://biorxiv.org/lookup/doi/10.1101/066803

118. Leek JT, Jager LR. Is most published research really false? [Internet]. 2016. Available from: http://biorxiv.org/lookup/doi/10.1101/050575

119. Wiener M, Sommer FT, Ives ZG, Poldrack RA, Litt B. Enabling an Open Data Ecosystem for the Neurosciences. Neuron 2016;92:617–21.

120. Teeters JL, Godfrey K, Young R, Dang C, Friedsam C, Wark B, et al. Neurodata Without Borders: Creating a Common Data Format for Neurophysiology. Neuron 2015;88:629–34.

121. Poldrack RA, Barch DM, Mitchell JP, Wager TD, Wagner AD, Devlin JT, et al. Toward open sharing of task-based fMRI data: the OpenfMRI project. Front Neuroinform 2013;7:12.

122. Gorgolewski KJ, Auer T, Calhoun VD, Craddock RC, Das S, Duff EP, et al. The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. Sci Data 2016;3:160044.

123. Poldrack RA, Laumann TO, Koyejo O, Gregory B, Hover A, Chen M-Y, et al. Long-term neural and physiological phenotyping of a single human. Nat Commun 2015;6:8885.

124. Zuo X-N, Anderson JS, Bellec P, Birn RM, Biswal BB, Blautzik J, et al. An open science resource for establishing reliability and reproducibility in functional connectomics. Sci Data 2014;1:140049.

125. Kennedy D, Haselgrove C, Grethe J, Wagner K, Buccigrossi R, Preuss N. The neuroimaging informatics tools and resources clearinghouse (NITRC). Front 2nd INCF Congr of Neuro [Internet] 1970;Available from: http://www.frontiersin.org/10.3389/conf.neuro.11.2009.08.024/event_abstract

126. van Erp TGM, Guella I, Vawter MP, Turner J, Brown GG, McCarthy G, et al. Schizophrenia miR-137 locus risk genotype is associated with dorsolateral prefrontal cortex hyperactivation. Biol Psychiatry 2014;75:398–405.

127. Frangou S, Schwarz E, Meyer-Lindenberg A, IMAGEMEND. Identifying multimodal signatures associated with symptom clusters: the example of the IMAGEMEND project. World Psychiatry 2016;15:179–80.

128. Hua W-Y, Nichols TE, Ghosh D, Alzheimer's Disease Neuroimaging Initiative. Multiple comparison procedures for neuroimaging genomewide association studies. Biostatistics 2015;16:17–30.

129. Meyer-Lindenberg A, Nicodemus KK, Egan MF, Callicott JH, Mattay V, Weinberger DR. False positives in imaging genetics. Neuroimage 2008;40:655–61.

130. Patel S, Park MTM, Alzheimer's Disease Neuroimaging Initiative, Chakravarty MM, Knight J. Gene Prioritization for Imaging Genetics Studies Using Gene Ontology and a Stratified False Discovery Rate Approach. Front Neuroinform 2016;10:14.

131. Glahn DC, Curran JE, Winkler AM, Carless MA, Kent JW Jr, Charlesworth JC, et al. High dimensional endophenotype ranking in the search for major depression risk genes. Biol Psychiatry 2012;71:6–14.

132. Skudlarski P, Constable RT, Gore JC. ROC analysis of statistical methods used in functional MRI: individual subjects. Neuroimage 1999;9:311–29.

133. Van Essen DC, Ugurbil K, Auerbach E, Barch D, Behrens TEJ, Bucholz R, et al. The Human Connectome Project: a data acquisition perspective. Neuroimage 2012;62:2222–31.

134. Koch C, Jones A. Big Science, Team Science, and Open Science for Neuroscience. Neuron 2016;92:612–6.

135. To the Cloud! A Grassroots Proposal to Accelerate Brain Science Discovery. Neuron 2016;92:622–7.

136. Dixson L, Walter H, Schneider M, Erk S, Schäfer A, Haddad L, et al. Retraction for Dixson et al., Identification of gene ontologies linked to prefrontal-hippocampal functional coupling in the human brain. Proc Natl Acad Sci U S A 2014;111:13582.

137. Kriegeskorte N, Simmons WK, Bellgowan PSF, Baker CI. Circular analysis in systems neuroscience: the dangers of double dipping. Nat Neurosci 2009;12:535–40.

138. Pausova Z, Paus T, Abrahamowicz M, Almerigi J, Arbour N, Bernard M, et al. Genes, maternal smoking, and the offspring brain and body during adolescence: design of the Saguenay Youth Study. Hum Brain Mapp 2007;28:502–18.

139. Pausova Z, Syme C, Abrahamowicz M, Xiao Y, Leonard GT, Perron M, et al. A common variant of the FTO gene is associated with not only increased adiposity but also elevated blood pressure in French Canadians. Circ Cardiovasc Genet 2009;2:260–9.

140. Calhoun VD, Miller R, Pearlson G, Adalı T. The chronnectome: time-varying connectivity networks as the next frontier in fMRI data discovery. Neuron 2014;84:262–74.

141. Loeb A. Good data are not enough. Nature 2016;539:23–5.

142. Baker M. Reproducibility: Seek out stronger science. njobs 2016;537:703–4.

143. Go forth and replicate! Nature 2016;536:373.

144. Niarchou M, Zammit S, Lewis G. The Avon Longitudinal Study of Parents and Children (ALSPAC) birth cohort as a resource for studying psychopathology in childhood and adolescence: a summary of findings for depression and psychosis. Soc Psychiatry Psychiatr Epidemiol 2015;50:1017–27.

145. Jahanshad N, Rajagopalan P, Hua X, Hibar DP, Nir TM, Toga AW, et al. Genome-wide scan of healthy human connectome discovers SPON1 gene variant influencing dementia severity. Proc Natl Acad Sci U S A 2013;110:4768–73.

146. Schumann G, Loth E, Banaschewski T, Barbot A, Barker G, Büchel C, et al. The IMAGEN study: reinforcement-related behaviour in normal brain function and psychopathology. Mol Psychiatry 2010;15:1128–39.

147. Nymberg C, Jia T, Ruggeri B, Schumann G. Analytical strategies for large imaging genetic datasets: experiences from the IMAGEN study. Ann N Y Acad Sci 2013;1282:92–106.

148. Holmes AJ, Hollinshead MO, O'Keefe TM, Petrov VI, Fariello GR, Wald LL, et al. Brain Genomics Superstruct Project initial data release with structural, functional, and behavioral measures. Sci Data 2015;2:150031.

149. Bakken TE, Roddey JC, Djurovic S, Akshoomoff N, Amaral DG, Bloss CS, et al. Association of common genetic variants in GPCPD1 with scaling of visual cortical surface area in humans. Proc Natl Acad Sci U S A

2012;109:3985–90.

150. Jernigan TL, Brown TT, Hagler DJ Jr, Akshoomoff N, Bartsch H, Newman E, et al. The Pediatric Imaging, Neurocognition, and Genetics (PING) Data Repository. Neuroimage 2016;124:1149–54.

151. Pausova Z, Paus T, Abrahamowicz M, Bernard M, Gaudet D, Leonard G, et al. Cohort Profile: The Saguenay Youth Study (SYS). Int J Epidemiol [Internet] 2016;Available from: http://dx.doi.org/10.1093/ije/dyw023

152. Ramanan VK, Kim S, Holohan K, Shen L, Nho K, Risacher SL, et al. Genome-wide pathway analysis of memory impairment in the Alzheimer's Disease Neuroimaging Initiative (ADNI) cohort implicates gene candidates, canonical pathways, and networks. Brain Imaging Behav 2012;6:634–48.

153. Weiner MW, Veitch DP, Aisen PS, Beckett LA, Cairns NJ, Cedarbaum J, et al. Impact of the Alzheimer's Disease Neuroimaging Initiative, 2004 to 2014. Alzheimers Dement 2015;11:865–84.

154. Potkin SG, Turner JA, Guffanti G, Lakatos A, Fallon JH, Nguyen DD, et al. A genome-wide association study of schizophrenia using brain activation as a quantitative phenotype. Schizophr Bull 2009;35:96–108.

155. Gollub RL, Shoemaker JM, King MD, White T, Ehrlich S, Sponheim SR, et al. The MCIC collection: a shared repository of multi-modal, multi-site brain image data from a clinical investigation of schizophrenia. Neuroinformatics 2013;11:367–88.

156. Cao H, Duan J, Lin D, Shugart YY, Calhoun V, Wang Y-P. Sparse representation based biomarker selection for schizophrenia with integrated analysis of fMRI and SNPs. Neuroimage 2014;102 Pt 1:220–8.

Tables

Table 1. Major collaborative projects that include both brain imaging and genome wide genotype data. Publications in italics are review papers describing the consortium.

Imaging Genetics Consortia (expanded from Medland 2014[88])					
Consortium	**Imaging Type**	**Sample Size**	**Composition**	**Population**	**Select Publications**
Avon Longitudinal Study of Parents and Children (ALSPC)	sMRI	14,000 women + families	subjects were pregnant when enrolled	Bristol, UK	Eicher 2013,*[61] *Niarchou 2015*[144]
IMAging GEnetics for MENtal Disorders (IMAGEMEND)	DTI, sMRI, fMRI	12,667 +relatives	SCZ, BPD, ADHD, healthy controls	Europe	Luksys 2015,[64] *Frangou 2016*[127]
Enhancing NeuroImaging Genetics through Meta-Analysis (ENIGMA)*	DTI, sMRI	12000+	various	worldwide	Stein 2010,*[102] Hibar 2011,*[103] Jahanshad 2013,*[145] *Thompson 2014*
Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE)*	DTI, sMRI	12000+	cohort	United States and Europe	*Psaty 2009*[90], Debette 2010*[63]
IMAGEN	sMRI, fMRI	2000+	healthy teenagers	Europe	*Schumann 2010*,[146] *Nymberg 2012*,[147] Whelan 2012[60]
Brain Genomics Superstruct Project (GSP)	sMRI, fMRI	1,570	healthy adults	Massachusetts	*Holmes 2015*[148]
Pediatric Imaging, Neurocognition, and Genetics (PING)	DTI, sMRI, fMRI	~1400	healthy children and teenagers	United States	Bakken 2012,*[149] *Jernigan 2016*[150]
Human Connectome Project (HCP)	sMRI, fMRI, DTI, MEG	1,200 +twins +siblings	healthy adults	United States	*Van Essen 2012*
Saguenay Youth Study (SYS)*	sMRI	1,000	French- Canadian teenagers	Quebec	*Pausova 2016*[151]
Alzheimer's Disease Neuroimaging Initiative (ADNI)	DTI, fMRI, PET	~822	Alzheimer's disease, MCI, normal adults	United States and Canada	Shen 2010,*[101] Meda 2012,*[66] Ramanan 2012,[152] *Weiner 2014*,[153] *Shen 2014*[104]
Function Biomedical Informatics Research Network (fBIRN)	fMRI	350+	SCZ, healthy controls	United States	Potkin 2009[154], Keator 2016[116]
MIND Clinical Imaging Consortium (MCIC)	DTI, sMRI, fMRI	331	SCZ and controls	New Mexico, Minnesota, Massachusetts, Iowa	*Gollub 2013*,[155] Cao 2014[156]

*= structural imaging only

Table 2. Strengths and weaknesses of common power enhancement approaches. Most imaging genetic studies use data reduction approaches (shaded).

Power enhancement approaches (Adapted from Thompson 2014[100])			
Power enhancement approach	**Examples**	**Strengths**	**Weaknesses**
Enhance the dataset	• Increase sample size • Increase genomic coverage • Increase range of phenotype • Meta-analysis	detects smaller effect sizes	can be expensive
Data reduction (*a priori*)	• By classical genetics (heritability) • By relevance to disease/trait • By prioritization	reduces noise	may miss unanticipated hits
Data reduction (data-driven)	• Multivariate statistics • Hierarchical clustering • Latent variable methods	fewer assumptions	difficult to interpret directionality and effect size of individual variables
Multimodality approaches	• Parallel ICA • Joint ICA	exploits joint information	computationally expensive, interpretation unclear

Table 3. A snapshot of 12 major publications from imaging genetics consortia, showing a diverse range of scope and methodology. Those that did attempt replication (shaded) were mostly successful, with one exception.

Scope and Methodology of Imaging Genetic Studies				
Study	**Genetic features**	**Imaging features**	**Primary power enhancement approach**	**Extent of replication**
Potkin 2009[154]	GWAS	activation in single region	enhance dataset by increasing range of phenotype	recommended
Debette 2010*[63]	GWAS	presence of infarcts	enhance dataset by meta-analysis	single-SNP replication failed, repeat recommended
Shen 2010*[101]	GWAS	ROIs and grey matter volume	data reduction by multivariate statistics	recommended
Stein 2010*[102]	GWAS	all voxels	data reduction by prioritization	estimate of sample size needed
Hibar 2011*[103]	gene-based GWAS	all voxels	data reduction by multivariate statistics	recommended
Bakken 2012,*[149]	GWAS	visual cortical surface areas	data reduction with multivariate statistics	replicated in two independent cohorts
Meda 2012*[66]	GWAS	ROIs	multimodality approach	recommended
Whelan 2012[60]	select SNPs	ROIs	data reduction by relevance to disease/trait	estimate of replicability
Eicher 2013*[61]	single gene	language-related fiber tracts	data reduction by relevance to disease/trait	replicated in PING dataset
Jahanshad 2013*[145]	GWAS	ROI connectivity	data reduction by classical genetics (heritability)	discovery and replication subsamples
Cao 2014[156]	GWAS	all voxels	multimodality approach	not mentioned
Luksys 2015[64]	GWAS	select voxels	data reduction by relevance to disease/trait	discovery and replication subsamples

*= structural imaging only