Genetic imaging consortium for addiction medicine: From neuroimaging to genes


*Departments of Psychiatry and Psychology, University of Vermont, Burlington, VT, USA
†Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY, USA
‡Department of Psychiatry and Psychology, Hospital Clinic, IDIBAPS, CIBERSAM, University of Barcelona, Barcelona, Spain
§Department of Psychiatry and MRC Unit on Anxiety & Stress Disorders, University of Cape Town, Cape Town, South Africa
*Department of Psychiatry and MRC Unit on Anxiety & Stress Disorders, University of Cape Town, Cape Town, South Africa
jjMontreal Neurological Institute, McGill University, Montreal, QC, Canada
#Department of Psychosocial Research and Epidemiology, Netherlands Cancer Institute, Amsterdam, The Netherlands
**Institute of Psychiatry, King’s College London, London, UK
††Department of Psychiatry, Oregon Health & Science University, Portland, OR, USA
‡‡Department of Psychiatry, University of Michigan, Ann Arbor, MI, USA
§§Department of Psychology and Neuroscience, University of Colorado Boulder, Boulder, CO, USA
**Department of Psychiatry, Yale University School of Medicine, New Haven, CT, USA
||David Geffen School of Medicine, University of California at Los Angeles, Los Angeles, CA, USA
###School of Psychological Sciences, Monash Institute of Cognitive and Clinical Neurosciences and Monash Biomedical Imaging, Monash University, Melbourne, Australia
####Behavioural Science Institute, Radboud University, Nijmegen, The Netherlands
#####VA San Diego Healthcare System and Department of Psychiatry, University of California San Diego, La Jolla, CA, USA
####Laureate Institute for Brain Research, Tulsa, OK, USA

Progress in Brain Research, Volume 224, ISSN 0079-6123, http://dx.doi.org/10.1016/bs.pbr.2015.07.026
© 2016 Elsevier B.V. All rights reserved.
Abstract
Since the sample size of a typical neuroimaging study lacks sufficient statistical power to explore unknown genomic associations with brain phenotypes, several international genetic imaging consortia have been organized in recent years to pool data across sites. The challenges and achievements of these consortia are considered here with the goal of leveraging these resources to study addiction.

The authors of this review have joined together to form an Addiction working group within the framework of the ENIGMA project, a meta-analytic approach to multisite genetic imaging data. Collectively, the Addiction working group possesses neuroimaging and genomic data obtained from over 10,000 subjects. The deadline for contributing data to the first round of analyses occurred at the beginning of May 2015. The studies performed on this data should significantly impact our understanding of the genetic and neurobiological basis of addiction.

Keywords
Addiction, Genetic imaging, ENIGMA, Neuroimaging
1 INTRODUCTION

The emergence of numerous large-scale international genetic imaging consortia in recent years is the product of several rapidly evolving factors. The maturing field of neuroimaging has made significant progress toward adopting a widely accepted set of best practices which have been incorporated into several competing software distributions (e.g., SPM, AFNI, FreeSurfer, FSL) that are free to download and relatively easy to install. Adaptation of imaging software to the developing needs of the neuroimaging community and greater automation have been accompanied by tremendous efforts to annotate the software and to educate a large cadre of scientists who are now able to apply these methods to studies with increasingly large sample sizes. Evidence of these efforts can readily be found on the busy message boards of any of the major neuroimaging platforms. Furthermore, with the development of standard anatomical templates and coordinate-based reference systems, researchers worldwide can now relate their findings to previous results in a consistent way. In combination, these factors have facilitated the formation of several large-scale collaborations to overcome the limitation of small sample sizes in typical genetic imaging studies.

The high dimensionality of genetic imaging datasets poses a difficult set of challenges. Human DNA consists of approximately 3 billion nucleotide base pairs. Variation in the population at any individual base is called a single-nucleotide polymorphism (SNP) and may contribute to the differential expression of phenotypic traits. Genomic studies have become a medical research priority because the identification of the genetic variation associated with a disease helps to clarify its molecular basis which, in turn, should lead to improved diagnostic categorization and more effective treatments (Sullivan et al., 2012). One way to proceed in identifying such associations is to investigate the relationship of traits of interest with candidate SNPs that are suggested on the basis of previous research (e.g., to examine the association of smoking behavior with SNPs related to the expression of nicotinic receptor subtypes). However, such a targeted approach is unlikely to expose the full range of SNPs involved in complex traits, such as addiction. To discover unknown trait-SNP associations, an unbiased search across the whole genome, known as a genome-wide association study (GWAS), is necessary. This latter strategy commonly involves testing hundreds of thousands to millions of SNPs and requires a strict multiple comparisons correction threshold, conventionally $p \leq 5 \times 10^{-8}$, to avoid reporting spurious results. Furthermore, findings must be replicated in at least one independent cohort before they are considered credible or at least generalizable. To meet these stringent thresholds, sharing data across multiple sites has become necessary.

There are now many successful examples of genetic imaging consortia, including ADNI (Alzheimer’s disease), IMAGEN (mental health and risk-taking behavior in teenagers), EPIGEN (epilepsy), the Saguenay Youth Study (development), fBIRN (schizophrenia), and CHARGE (heart and aging). These groups have pioneered
the use of multisite data sharing protocols and have demonstrated that analyses using shared data produce meaningful findings. The purpose of this review is to discuss how these resources can be leveraged to study addiction.

2 GENETIC BASIS OF ADDICTION

It is clear that addiction has a genetic component (Maes et al., 2004; Prescott and Kendler, 1999; Tsuang et al., 1998) although the specific set of genes involved remains obscure. Several GWAS of alcohol addiction have been published (Bierut et al., 2010; Edenberg et al., 2010; Heath et al., 2011; Treutlein et al., 2009) which have confirmed the risk of alcoholism associated with a number of SNPs, such as the ADH and ALDH2 genes, previously identified through the candidate gene approach. These studies have also identified some additional but as yet unreplicated variants that may contribute to alcohol dependence (Rietschel and Treutlein, 2013). However, results have largely differed from one GWAS to another with later studies providing only modest evidence of replication of previous findings. A similar situation exists with regard to cannabis in that published GWAS have not reproduced previous findings (Agrawal et al., 2011; Han et al., 2012; Hopfer et al., 2007). The genetic basis of nicotine dependence has been more closely examined than other substance addictions although again only a handful of results have been replicated across studies (Berrettini et al., 2008; Drgon et al., 2009; Gelernter et al., 2015; Thorgerirsson et al., 2008; Uhl et al., 2008b; Wang et al., 2012a; Zuo et al., 2013). Only a few published GWAS have examined the genetic basis of other drug use (e.g., Uhl et al., 2008a). In summary, most of the genetic variation underlying addiction remains to be explained.

2.1 BRAIN ENDOPHENOTYPES

The failure to identify a greater proportion of risk genes is disappointing given the high heritability of addiction. Recent estimates of the heritability of dependence on different addictive substances include: 56% for alcohol, 72% for cocaine, 40% for other stimulants, 48% for cannabis, and 51% for sedatives (Bienvenu et al., 2011). The intermediate “endophenotypes” approach may be a more sensitive way to determine how genes influence addiction vulnerability (Glahn et al., 2007, 2014). An intermediate endophenotype is a quantifiable biomarker (e.g., regional brain volume or activity) that is genetically correlated with disease liability and observed to a greater degree in affected individuals and their relatives than in unaffected nonrelatives. Since these biomarkers are arguably more proximal to the molecular expression of DNA than the related complex trait, it may be possible to generate simpler models of single aspects of the disorder to effectively bridge the gap in understanding between genotype and phenotype. In addition, the statistical power to detect genetic associations may be greater than using diagnostic categories because intermediate endophenotypes represent a continuous scale on which individuals can be ranked.
At least three lines of evidence suggest that genetic neuroimaging may produce useful intermediate endophenotypes of addiction. First, 20 years of neuroimaging data amply demonstrate that brain structure and function interact with the use of addictive substances. For example, brain structure differences compared to healthy controls have been observed in cocaine-dependent individuals (Alia-Klein et al., 2011; Barros-Loscertales et al., 2011; Connolly et al., 2013; Hanlon et al., 2011; Ide et al., 2014; Mackey and Paulus, 2013; Matochik et al., 2003), cigarette smokers (Brener et al., 1995; Kuhn et al., 2010; Sutherland et al., 2013; Zhang et al., 2011), alcoholics (Cardenas et al., 2007; Jernigan et al., 1991; Rando et al., 2011), cannabis users (Batalla et al., 2013; Lorenzetti et al., 2014; Schacht et al., 2012; Yucel et al., 2008), and opiate users (Lyoo et al., 2006; Upadhyay et al., 2010; Wang et al., 2012b). These effects are widespread and likely reflect a mixture of preexisting differences that either confer vulnerability to addiction or are the cumulative effects of chronic exposure.

A second line of evidence suggesting that neuroimaging will generate useful intermediate phenotypes are twin- and SNP-based heritability studies which indicate a high heritability for structural brain measures, such as total amount of gray and white matter, overall brain volume, and addiction-relevant subcortical regions. Heritability estimates for brain measures \((h^2)\) are as high as 0.89 (Kremen et al., 2010) or even 0.96 (van Soelen et al., 2012) and subcortical regions appear to be moderately to highly heritable. One recent study reported high heritability estimates for the thalamus (0.80) and caudate nucleus (0.88) compared to a lower heritability for the left nucleus accumbens (0.44) (den Braber et al., 2013).

Third, biomarkers of addiction which are present to a greater degree in affected individuals and their relatives compared to unaffected nonrelatives have been reported. For example, a recent neuroimaging study acquired anatomical MRI and diffusion tensor image (DTI) scans in 50 biological sibling pairs and a group of nonrelated control subjects (Ersche et al., 2012). One sibling in each pair was dependent on cocaine or amphetamine. Fractional anisotropy in the DTI scans, an index of axonal integrity, was lower in dependent subjects and their nondependent siblings compared to the control subjects. Also, voxel-based morphometry indicated that gray matter volume in both dependent subjects and their siblings was lower in left posterior Sylvian fissure including parts of the postcentral gyrus, insula, and superior temporal gyrus and higher in the left putamen and left amygdala. The discovery of biomarkers that are quantifiably different in drug-dependent individuals and their siblings compared to nonrelated controls underscores the potential for neuroimaging to detect intermediate brain endophenotypes that will be useful in genomic research.

### 2.2 CHALLENGES

The search for robust genetic and brain structural correlates of drug use and dependence faces a number of substantial challenges. The inability to find extensive significant genome-wide associations might be attributable to the large degree of heterogeneity due to polydrug use and the high incidence of mental health comorbidities among drug users. It will be necessary to disambiguate several sources of
genetic variation. Epidemiological studies indicate that there will be genetic variation associated with a general vulnerability to addiction and to a lesser extent drug-specific associations as well as gene–environment interactions (Tsuang et al., 1998). Furthermore, lifetime drug use can be decomposed into a number of qualitatively different stages (e.g., initial experimentation, occasional use, transition to abuse and dependence, risk of relapse) that current research indicates will exhibit different sets of genetic associations (Belin and Deroche-Gamonet, 2012; Everitt and Robbins, 2013; Montigny et al., 2013). GWAS and candidate gene analyses also have their own unique shortcomings. While GWAS searches the whole genome for unknown associations, it will miss variants with small effect sizes that would pass the less stringent probability threshold of the candidate gene approach (Gizer and Ehlers, 2015). With the candidate gene approach, however, there is no way to verify whether published candidate gene studies are systematically biased toward reporting successes. To correct for this latter problem, it has even been suggested that candidate gene associations should be held to the same significance criterion as GWAS (Flint and Munafo, 2013). The solution will likely require a combination of the two search strategies to iteratively approximate the genetic polymorphisms involved in addiction using both intermediate endophenotypes and well-defined behavioral traits.

3 ENHANCING NEUROIMAGING GENETICS THROUGH META-ANALYSIS

In 2009, researchers from large-scale neuroimaging and genetics consortia, including IMAGEN, EPIGEN, SYS, FBIRN, and ADNI, formed the Enhancing Neuroimaging Genetics through Meta-Analysis (ENIGMA) project to work through the challenges of bringing together data from multiple samples and sites worldwide in a single meta-analytic framework (http://enigma.ini.usc.edu/) (Thompson et al., 2014). The first published ENIGMA meta-analysis reported that the mean bilateral volume of the hippocampus was significantly associated with the intergenic variant rs7294919 (Stein et al., 2012). This proof-of-principle study established the feasibility of combining imaging and genomic data collected across multiple sites to investigate statistically significant effects of single-letter genomic differences in brain data. In addition, a follow-up study discovered eight genetic loci in which common variants were associated with the volumes of several subcortical structures, including the putamen, caudate, and hippocampus (Hibar et al., 2015). The SNPs associated with subcortical brain volumes were supported across 50 cohorts worldwide, suggesting the power to identify genetic effects that account for as little as 1% of the variance in regional brain volumes. Functional characterization of these genetic loci, in outbred mice, was consistent with possible effects on cell number and links to degenerative disease risk (Ashbrook et al., 2014). The protocol developed by the ENIGMA network to harmonize the data from multiple sites has been made freely available to collaborators and a support structure based in Dr. Paul Thompson’s
Imaging Genetics Center at the University of Southern California has been created to facilitate the application of the protocol to other projects.

The ENIGMA protocol contains several innovations to deal with special issues arising from multisite analyses, notably imputation of genomic data to a common reference panel and a pathway to harmonize neuroimaging data with standardized quality control procedures. For the initial ENIGMA study, all data were imputed to the HapMap3 reference panel because SNP data at the various sites were genotyped on different gene chips. The imputation protocol adds substantial power to the overall meta-analysis by creating a genomic dataset that is comparable across sites and by employing state-of-the-art approaches to account for hidden structure (e.g., ancestry) and relevant quality control variables. More recently, the ENIGMA imputation protocol implemented in MaCH (http://csg.sph.umich.edu/abecasis/MaCH/) has been updated to use the 1000 Genomes reference, a more in-depth analysis of the genome. To control for population stratification, multidimensional scaling (MDS) is applied to the genotyped data and the first four components are included as nuisance covariates in subsequent GWAS analyses (Hibar et al., 2015; Stein et al., 2012).

To process the neuroimaging data efficiently, one of two highly automated neuroimaging software packages (FSL’s FIRST and FreeSurfer) was used for the initial ENIGMA publications although in future studies, including those undertaken by the Addiction working group, only FreeSurfer (Fischl et al., 2002) will be employed (Fig. 1). The use of these standard software programs ensures the comparability of neuroimaging results across sites. Despite the automation of FreeSurfer, considerable time is still required to test for statistical outliers, inspect distributions of brain structure volumes, genomic inflation factors, and other statistical summaries at each site.

Rather than using an analysis strategy where all phenotypic and genotypic data are sent to one central site for processing, as for example in the Psychiatric Genomics Consortium (http://www.med.unc.edu/pgc), ENIGMA employs a meta-analytic

**FIGURE 1**
Illustration of a structural MRI brain scan processed with FreeSurfer. Left, example of automated parcellation of the cortex. Right, local cortical thickness projected onto inflated surface of the brain.
strategy in which GWAS are computed locally using agreed upon covariates. The advantages of this approach include the active involvement in the analysis of the researchers who collect and curate the data, and the ability to draw upon local computer infrastructure at each site to ease demand on central data processing. Site-level GWAS are performed with Mach2qtl, a statistical genetics algorithm developed by Goncalo Abecasis and colleagues (Li et al., 2010). Multiple linear regression is performed on each SNP using trait as the dependent variable and allelic dosage (i.e., 0, 1, or 2 alleles) as the independent variable of interest. Sites control for a set of basic nuisance factors, namely the first four MDS components, age, sex, age × sex interaction effects, and nonlinear effects of age, including age² and age² × sex, by adding them as covariates in the regression model. Site-specific covariates may also be added (e.g., if data are acquired on two different scanners). Following quality control, the regression coefficient, standard error, and p-value for each SNP are forwarded to the coordinating site which conducts a unifying meta-analysis that weights the SNP coefficients by their standard error. This approach circumvents barriers associated with data sharing across sites and countries and allows sites to maintain responsibility for the integrity of their data. The meta-analysis is performed with an inverse standard error-weighted meta-analysis protocol implemented in METAL (Willer et al., 2010). Genomic control of p-values undertaken at the site level is repeated on the output of the meta-analysis to provide an additional control for population stratification or cryptic relatedness not accounted for by the MDS components (Devlin and Roeder, 1999). Additionally, associations are verified in replication samples that have been acquired independently of the discovery dataset.

3.1 DISEASE WORKING GROUPS

From the time that the pilot project by Stein et al. was published in April 2012, several ENIGMA working groups have been formed to focus more closely on applying the ENIGMA meta-analysis protocols to case–control differences in various brain-related diseases. With such large studies comes the ability to perform high power association studies to identify biomarkers for monitoring disease state and targets for drug therapies. ENIGMA working groups have been formed to study ADHD, schizophrenia, OCD, HIV, PTSD, major depressive disorder, and bipolar disorder (Jahanshad et al., 2013; Schmaal et al., 2015; van Erp et al., 2015).

4 ENIGMA ADDICTION WORKING GROUP

The authors of this review have joined together to leverage the structure of the ENIGMA project to study addiction. The international membership represents research laboratories from four continents and nine different time zones (Fig. 2). An initial site survey has identified datasets, including both case/control and cohort studies, that collectively contain neuroimaging and genomic data on over 10,000 subjects. Table 1 provides a summary of the Addiction working group datasets.
FIGURE 2
World map of the current membership of the ENIGMA Addiction working group.

Table 1  Summary of ENIGMA Addiction Working Group Datasets as of February 2015

<table>
<thead>
<tr>
<th>Substance</th>
<th>Pattern of Use</th>
<th>Cases</th>
<th>Female Cases</th>
<th>Cases and Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alcohol</td>
<td>Occasional</td>
<td>150</td>
<td>75</td>
<td>150</td>
</tr>
<tr>
<td></td>
<td>Dependent</td>
<td>1695</td>
<td>560</td>
<td>2124</td>
</tr>
<tr>
<td></td>
<td>Abstinent</td>
<td>61</td>
<td>24</td>
<td>177</td>
</tr>
<tr>
<td>Nicotine</td>
<td>Occasional</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Current</td>
<td>1132</td>
<td>385</td>
<td>1797</td>
</tr>
<tr>
<td></td>
<td>Abstinent</td>
<td>10</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>Cannabis</td>
<td>Occasional</td>
<td>91</td>
<td>30</td>
<td>213</td>
</tr>
<tr>
<td></td>
<td>Dependent</td>
<td>238</td>
<td>33</td>
<td>348</td>
</tr>
<tr>
<td></td>
<td>Abstinent</td>
<td>17</td>
<td>7</td>
<td>34</td>
</tr>
<tr>
<td>Stimulants</td>
<td>Occasional</td>
<td>175</td>
<td>69</td>
<td>228</td>
</tr>
<tr>
<td></td>
<td>Dependent</td>
<td>906</td>
<td>182</td>
<td>1408</td>
</tr>
<tr>
<td></td>
<td>Abstinent</td>
<td>68</td>
<td>9</td>
<td>108</td>
</tr>
<tr>
<td>Gambling</td>
<td>Occasional</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Dependent</td>
<td>59</td>
<td>0</td>
<td>187</td>
</tr>
<tr>
<td></td>
<td>Abstinent</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Heroin</td>
<td>Occasional</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Dependent</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Abstinent</td>
<td>38</td>
<td>15</td>
<td>70</td>
</tr>
<tr>
<td>Cohort</td>
<td></td>
<td>–</td>
<td>–</td>
<td>6445</td>
</tr>
<tr>
<td>Totals</td>
<td></td>
<td>4640</td>
<td>1389</td>
<td>13,299</td>
</tr>
</tbody>
</table>
The strengths of the working group are currently found in several large developmental cohorts as well as case/control studies of dependent users of alcohol, nicotine, stimulants (cocaine and methamphetamine), and cannabis.

Each site has committed to completing the site-level analyses through local personnel. ENIGMA provides detailed image analysis protocols that will be adopted at all sites (http://enigma.ini.usc.edu/protocols/imaging-protocols/). The fact that many sites already employ these protocols or very similar processing pipelines will minimize the time required for data preprocessing. Data analysis support will be provided by a postdoctoral associate (S.M.) at the University of Vermont and will also be available from the engineers and analysts in Dr. Thompson’s ENIGMA support team. Easy-to-use instructions on how to preprocess the neuroimaging and genomic data and check for data quality have also been prepared.

A multisite genetic neuroimaging meta-analysis will only be successful if careful attention is paid to the assessment of behavioral variables. The experience of the ENIGMA research consortium shows that the pooling of neuroimaging data requires the evaluation, and where possible standardization, of site effects on phenotypic characterizations and brain measures. The chosen phenotypes and brain measures must offer optimal sensitivity to disease effects, clinically relevant modulators of disease, and treatment effects. While each site possesses extensive phenotyping on its research participants, there are important differences across sites in the instruments and questionnaires used. The Addiction working group will develop common measures of quantity and frequency of use derived from the different instruments and assessments obtained at each site. Standardized addiction scores will be generated across the varying developmental and clinical profiles. This approach has been effective in harmonizing measures of alcohol consumption for the purpose of large genetics studies, such as the Gene–Environment Association Studies (GENEVA) consortium. For example, the GENEVA consortium was able to convert disparate alcohol measures into useful categories representing onset and safe compared with unsafe consumption (Holman and English, 1995; Holman et al., 1996). There are many methodological problems associated with measurement heterogeneity for alcohol consumption in the context of genomic studies. These include questions with regard to how abstention should be interpreted, the episodic nature of alcohol consumption, the coding of current drug use state at the time of scanning, the quantity and frequency of substance use across reference periods, differences in cultural norms, the standardization of drinking units, as well as recall and other respondent biases (see review, Agrawal et al., 2012). As recommended by Agrawal et al., the Addiction working group will use the guidelines and where possible attempt to align the addiction-related phenotypes with the NIH PhenX toolkit measures for alcohol and drug consumption (e.g., lifetime use, age at first use, and symptoms of dependence).

4.1 INITIAL PROJECT

The first analysis will examine the structural correlates of four simple drug use categories, no lifetime use, occasional use, abuse, and dependence. Data related to four substances, i.e., alcohol, nicotine, stimulants (cocaine and methamphetamine), and
cannabis, will be used to identify the neural substrates of core addiction processes as well as substance-specific factors. Performing GWAS on the identified brain regions will significantly reduce the dimensionality of the brain imaging data. This first analysis will establish relationships between the sites and bring to light any major difficulties that need to be addressed. Analysis will begin after the first data freeze which will occur at the beginning of May 2015. Likely, the most important early challenge for the group will be the development of an assessment instrument that harmonizes the different drug use measures at the various sites.

5 SUMMARY AND FUTURE DIRECTIONS

Several international consortia have been organized in recent years to improve the statistical power of genetic imaging association analyses by pooling data from multiple sites. The authors of this review have formed an Addiction working group within the framework of the ENIGMA project to leverage the acquired knowledge about data sharing across multiple sites to study the genetic and neurobiological mechanisms underlying addiction. The ENIGMA Addiction working group will attempt to identify brain endophenotypes starting with a volumetric investigation of the core neural substrates of addiction. The identification of core brain regions using structural MRI will reduce the number of dimensions in subsequent genomic analyses of problematic substance use. The Addiction working group will adopt the meta-analytic methods used successfully by the ENIGMA project. However, a mega-analysis approach, i.e., analysis of all pooled raw data at one location, may offer opportunities to conduct in-depth examinations of the neurobiology of drug use that are not possible in a meta-analysis. While practical concerns about sharing data were part of the motivation for the meta-analysis approach used by ENIGMA, more sensitive analyses may be possible by going beyond the pooling of effect sizes and the sharing of summary statistics (e.g., volume measurements of specific cortical and subcortical structures) to the sharing of complete, fully anonymized datasets, where available. We believe that the obstacles to this level of data sharing are surmountable. Depending on how the consortium grows (i.e., the addition of new members and of new datasets from current members), the Addiction working may in also decide to include multi-modal assessments of brain function including task-related and resting-state fMRI, DTI, and EEG (e.g., Jahanshad et al., 2015; Kochunov et al., 2015). At the present time, the working group is focused on resolving problems related to multisite data pooling with a manageable number of 18 sites.

Recent advances in the statistical analysis of genomic data present several promising new ways to investigate the combined datasets. We will explore the application of genome-wide complex trait analyses (Yang et al., 2011) to assess the heritability and genetic correlations among brain regions and phenotypic measures associated with alcohol and drug use. This method produces estimates of the variance explained by all SNPs over the whole genome for a complex trait and is suitable for large samples of nonrelated subjects. The working group will also investigate emerging statistical methods to detect significant associations in high dimensional data, such
as the parallel independent components analysis with a reference mask (Liu et al., 2012), meta-analysis of voxel-based data (Jahanshad et al., 2015), and novel applications of correspondence analysis (Cioli et al., 2014).

5.1 ADDICTION MEDICINE
There are multiple ways in which the progress of the working group could impact the practice of addiction medicine. Since there is strong evidence that addiction has a genetic component (Maes et al., 2004; Prescott and Kendler, 1999; Tsuang et al., 1998), a GWAS with sufficient power, such as the one envisaged by the working group, will likely detect novel genetic associations with behavioral features of addiction or with intermediate brain phenotypes. Not only will these novel associations drive future research aimed at understanding the neural processes involved in problematic substance use and potentially provide novel targets for pharmacological intervention, but they could also lead to the development of predictive genetic and neuroimaging biomarkers. Addiction medicine would benefit enormously from a set of predictive tools that could be used to estimate risk at various stages of the disorder, e.g., risk of transition from healthy to problematic patterns of use or risk of relapse after treatment (Paulus, 2015). Current research also points toward a heterogeneity of causes (Tsuang et al., 1998). If addictive behavior can be attributed to many small effects in a range of brain systems, it is possible that combined neuroimaging and genetic testing could identify differential vulnerabilities which could be used to customize treatment to address the specific challenges of the individual patient.

ACKNOWLEDGMENTS
This work was supported by a National Institute on Drug Abuse (NIDA) Grant 1R21DA038381 and by a National Institutes of Health (NIH) Grant U54 EB 020403 with funds provided for the trans-NIH Big Data to Knowledge (BD2K) initiative. Support was also provided by an NIH Grant 1P20GM103644-01A1 awarded to the Vermont Center on Behavior and Health.

REFERENCES


