Polygenic Scores for Major Depressive Disorder and Risk of Alcohol Dependence

Allan M. Andersen, MD; Robert H. Pietrzak, PhD, MPH; Henry R. Kranzler, MD; Li Ma, PhD; Hang Zhou, PhD; Xiaoming Liu, PhD; John Kramer, PhD; Samuel Kuperman, MD; Howard J. Edenberg, PhD; John I. Nurnberger Jr, MD, PhD; John P. Rice, PhD; Jay A. Tischfield, PhD; Alison Goate, PhD; Tatiana M. Foroud, PhD; Jacquelyn L. Meyers, PhD; Bernice Porjesz, PhD; Danielle M. Dick, PhD; Victor Hesselbrock, PhD; Eric Boerwinkle, PhD; Steven M. Southwick, MD; John H. Krystal, MD; Myrna M. Weissman, PhD; Douglas F. Levinson, MD; James B. Potash, MD, MPH; Joel Gelernter, MD; Shizhong Han, PhD

IMPORTANCE Major depressive disorder (MDD) and alcohol dependence (AD) are heritable disorders with significant public health burdens, and they are frequently comorbid. Common genetic factors that influence the co-occurrence of MDD and AD have been sought in family, twin, and adoption studies, and results to date have been promising but inconclusive.

OBJECTIVE To examine whether AD and MDD overlap genetically, using a polygenic score approach.

DESIGN, SETTINGS, AND PARTICIPANTS Association analyses were conducted between MDD polygenic risk score (PRS) and AD case-control status in European ancestry samples from 4 independent genome-wide association study (GWAS) data sets: the Collaborative Study on the Genetics of Alcoholism (COGA); the Study of Addiction, Genetics, and Environment (SAGE); the Yale-Penn genetic study of substance dependence; and the National Health and Resilience in Veterans Study (NHRVS). Results from a meta-analysis of MDD (9240 patients with MDD and 9519 controls) from the Psychiatric Genomics Consortium were applied to calculate PRS at thresholds from \( P < .05 \) to \( P < .99 \) in each AD GWAS data set.

MAIN OUTCOMES AND MEASURES Association between MDD PRS and AD.

RESULTS Participants analyzed included 788 cases (548 [69.5%] men; mean [SD] age, 38.2 [10.8] years) and 522 controls (151 [28.9%] men; age [SD], 43.9 [11.6] years) from COGA; 631 cases (333 [52.8%] men; age [SD], 35.0 [7.7] years) and 756 controls (260 [34.4%] male; age [SD] 36.1 [7.7] years) from SAGE; 2135 cases (1375 [64.4%] men; age [SD], 39.4 [11.5] years) and 350 controls (126 [36.0%] men; age [SD], 43.5 [13.9] years) from Yale-Penn; and 317 cases (295 [93.1%] men; age [SD], 59.1 [13.1] years) and 1719 controls (1545 [89.9%] men; age [SD], 64.5 [13.3] years) from NHRVS. Higher MDD PRS was associated with a significantly increased risk of AD in all samples (COGA: best \( P = 1.7 \times 10^{-6} \), \( R^2 = 0.026 \); SAGE: best \( P = .001 \), \( R^2 = 0.01 \); Yale-Penn: best \( P = .035 \), \( R^2 = 0.0018 \); and NHRVS: best \( P = .004 \), \( R^2 = 0.0074 \)), with stronger evidence for association after meta-analysis of the 4 samples (best \( P = 3.3 \times 10^{-9} \)). In analyses adjusted for MDD status in 3 AD GWAS data sets, similar patterns of association were observed (COGA: best \( P = 7.6 \times 10^{-6} \), \( R^2 = 0.023 \); Yale-Penn: best \( P = .08 \), \( R^2 = 0.0013 \); and NHRVS: best \( P = .006 \), \( R^2 = 0.0072 \)). After recalculating MDD PRS using MDD GWAS data sets without comorbid MDD-AD cases, significant evidence was observed for an association between the MDD PRS and AD in the meta-analysis of 3 GWAS AD samples without MDD cases (best \( P = .007 \)).

CONCLUSIONS AND RELEVANCE These results suggest that shared genetic susceptibility contributes modestly to MDD and AD comorbidity. Individuals with elevated polygenic risk for MDD may also be at risk for AD.
The association between alcohol dependence (AD) and major depressive disorder (MDD) has been studied by clinicians, epidemiologists, and geneticists for more than 100 years. Bleuler and Brill described “alcoholic melancholia,” characterized by brief periods of depression after alcohol binges, which could precipitate a suicide attempt. States of depressed mood associated with both alcohol use and withdrawal were documented by later researchers, while others noted that some individuals who increased their drinking during periods of depression tended to progress to AD.

More recently, the rate of AD and MDD comorbidity was assessed in cross-sectional epidemiologic studies. The National Comorbidity Study linked AD to 3.7-fold higher odds of experiencing a major depressive episode in the prior year. In the 2012-2013 National Epidemiologic Survey on Alcoholism and Related Conditions, individuals with a lifetime diagnosis of MDD had a 1.3-fold increased risk of an alcohol use disorder. Furthermore, both the National Comorbidity Study and Epidemiological Catchment Area study found that AD worsened the symptoms of MDD and vice versa, which may explain why people with comorbid AD and MDD are at greater risk for suicide.

This pattern of comorbidity has prompted hypotheses regarding causal relationships between AD and MDD. One hypothesis is that AD causes MDD either directly, through alcohol’s pharmacologic effects on the brain, or indirectly, through negative social consequences that increase the risk of MDD. A second hypothesis is that MDD can lead to AD. Longitudinal studies support an internalizing pathway from MDD to AD, with some individuals using alcohol to self-medicate negative mood states.

A third hypothesis, that shared genetic factors predispose to both AD and MDD, has had mixed support. A longitudinal study of Finnish twins showed a shared genetic liability between alcohol use and depressive symptoms across adolescence. Researchers from the Collaborative Study on the Genetics of Alcoholism (COGA) reported aggregation of MDD in relatives of probands with AD, whether or not the proband had MDD or the relative had AD. Other family studies have also suggested a genetic link between AD and MDD, including both sex-specific effects and environmental influences on their comorbidity.

Molecular genetic studies have implicated specific genomic regions and genes for both AD and MDD risk. Linkage analysis by COGA identified a region on chromosome 1 associated with both MDD and AD. Candidate gene studies reported a number of genes possibly underlying both disorders, including CHRM2, SLC6A4, COMT, and DRD2, although other studies failed to replicate these associations. Genome-wide association studies (GWAS) have produced significant findings for both AD and MDD, although not for comorbid AD-MDD.

In the present study, we examined whether AD and MDD overlap genetically using polygenic risk scores (PRSs), which are quantitative measures of the cumulative effects of common genetic variations across the genome on risk for a disorder. Using PRSs derived from a Psychiatric Genomics Consortium (PGC) meta-analysis of MDD, we tested whether PRSs for MDD were associated with AD case-control status in samples from 4 independent GWAS data sets: (1) COGA; (2) the Study of Addiction, Genetics, and Environment (SAGE); (3) the Yale-Penn genetic study of substance dependence; and (4) the National Health and Resilience in Veterans Study (NHRVS).

### Methods

### AD GWAS Data Sets

Basic information on the AD GWAS data sets is summarized in Table 1. All participants in the COGA, Yale-Penn, and NHRVS studies gave informed consent for participation, and all procedures and protocols used in each sample were approved by the institutional review board at the relevant institution. Participants received financial compensation. COGA involved multiple centers in the United States: Indiana University; State University of New York Health Science Center, Brooklyn; University of Connecticut; University of Iowa; University of California, San Diego; Washington University; and Howard University. The institutional review boards of all of these participating institutions approved the study. The Yale-Penn study was approved by the Yale University Human Investigation Committee. The NHRVS procedures were approved by the human subjects subcommittee of the Veterans Affairs (VA) Connecticut Health-care System and the VA Office of Research and Development.

**COGA**

Samples were drawn from an ongoing, family-based study of AD including participants from 6 sites across the United States. Patient cases had a lifetime diagnosis of AD by DSM-IV criteria. Individuals serving as controls had consumed alcohol but did not meet criteria for AD, alcohol abuse, or other substance use disorders in their lifetime. A sample of 1945 participants from COGA was genotyped (Illumina Human1M microarray; Illumina). After data quality control, 1310 European American individuals (788 with AD and 522 controls) were included.

**SAGE**

Samples were selected from 3 studies: COGA, the Family Study of Cocaine Dependence, and the Collaborative Genetic Study of Nicotine Dependence. Full details are available elsewhere. Alcohol dependence was defined based on DSM-IV criteria.

### Polygenic Risk Score Analysis for Depression and Alcohol Dependence

<table>
<thead>
<tr>
<th>Key Points</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Question</strong></td>
</tr>
<tr>
<td><strong>Findings</strong></td>
</tr>
<tr>
<td><strong>Meaning</strong></td>
</tr>
</tbody>
</table>
typed (Illumina Human1M microarray). After quality control and excluding participants overlapping with COGA, 1387 European American individuals (631 with AD and 756 controls) were retained for analysis.

Yale-Penn
The Yale-Penn samples included individuals recruited for genetic studies of cocaine dependence, opioid dependence, and AD, as described previously.50,51 Alcohol dependence was defined using DSM-IV criteria. Control participants did not meet the DSM-IV criteria for lifetime AD or other substance use disorders. DNA samples from 9459 individuals were genotyped using 1 of 2 platforms: (1) the Illumina HumanOmniQuad, version 1.0 microarray containing 988 306 autosomal single-nucleotide polymorphisms (SNPs) or (2) the Illumina Infinium Human Core Exome microarray, containing 265 919 exome-focused SNPs and approximately 240 000 tagging SNPs. Genome-wide association studies quality control and imputation were conducted for the merged GWAS samples based on a common set of autosomal SNPs (230 447). After quality control, the final analysis included 2485 European American individuals (2135 with AD and 350 controls).

NHRVS
The NHRVS data set included US military veterans who participated in a nationally representative cohort study in 2011.52,53 Alcohol dependence was assessed according to DSM-IV diagnostic criteria. Controls were defined as individuals without lifetime AD or other substance dependence. A total of 2825 samples were genotyped with the Infinium PsychArray, version 1.1. After quality control, 2036 European American individuals (317 AD and 1719 controls) were retained for analysis.

GWAS Quality Control
Each GWAS data set was cleaned via PLINK, version 1.07 before analysis.54 Samples and SNPs were excluded based on predetermined quality control metrics, including sample call rate of 95% or less, SNP call rate of 95% or less, minor allele frequency of 0.01 or less in controls, and P values for Hardy-Weinberg equilibrium ≤10−6 for controls. We estimated the genome-wide identity by descent sharing between all pairs of participants to detect duplicates (individuals genotyped twice within 1 sample or individuals in 1 sample also included in another) and then randomly removed 1 duplicate from each pair. We used EIGENSOFT, version 6.1.355 to compute principal components for each GWAS data set using pruned SNPs in low linkage disequilibrium (r2<0.2). Outliers, defined as participants with ancestry at least 3 SDs from the mean on 1 of the 2 largest principal components, were removed. After quality control, there were 843 188 (COGA), 837 142 (SAGE), 220 554 (Yale-Penn), and 287 923 (NHRVS) autosomal SNPs retained. To ensure that sample overlap did not contribute to inflated estimates of genetic overlap between AD and MDD, λ meta statistics were calculated. The λ meta is a statistic that uses effect size concordance to detect sample overlap or heterogeneity.56 All λ meta values were larger than 1 (1.02<λ meta <1.06), indicating no significant overlap between AD and PGC MDD samples.

Imputation
Following the best practice guidelines of IMPUTE2, version 2.3.2,57 we imputed 1000 Genomes variants into each GWAS sample. Prephasing was first performed with SHAPEIT, version 2.79058 to infer haplotypes for samples based on autosomal SNPs with minor allele frequency greater than 0.01. Imputation was carried out on prephased haplotypes using IMPUTE2 against reference data from the 1000 Genomes Phase 1 integrated variant set for all GWAS data except for the NHRVS and Yale-Penn samples, which were imputed against the 1000 Genomes Phase 3 haplotype reference. After postimputation (SNP missing rate <0.05, minor allele frequency >0.05, imputation quality score >0.5, and Hardy-Weinberg equilibrium >10−6), 5 960 232 (COGA), 5 955 937 (SAGE), 7 049 271 (Yale-Penn), and 7 055 726 (NHRVS) autosomal variants were retained. Imputed SNPs were used for calculating the PRS, as described below.

Polygenic Score Analysis
Each PRS was calculated as a weighted sum of reference alleles across independent SNPs on a genome-wide scale. Weights were log10 (odds ratio) of SNPs estimated from a meta-analysis of MDD GWASs (9240 MDD cases and 9519 controls of European ancestry) from the PGC.47 Summary results for 1235 109 autosomal SNPs were downloaded from the PGC website (https://www.med.unc.edu/pgc). Downloaded SNPs were pruned using P value–informed clumping in PLINK, with a cutoff of P = 0.2 within a 200-kilobase window. Polygenic risk scores were calculated for each AD GWAS sample using the PLINK score option, based on sets of pruned SNPs at successive P value thresholds (<.05 to ≤.99).

Table 1. Sample Demographics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>COGA</th>
<th>SAGE</th>
<th>Yale-Penn</th>
<th>NHRVS</th>
</tr>
</thead>
<tbody>
<tr>
<td>AD</td>
<td>788</td>
<td>631</td>
<td>2135</td>
<td>317</td>
</tr>
<tr>
<td>Men</td>
<td>548</td>
<td>333</td>
<td>1375</td>
<td>295</td>
</tr>
<tr>
<td>Age, mean (SD), y</td>
<td>38.2</td>
<td>35.0</td>
<td>39.4</td>
<td>59.1</td>
</tr>
<tr>
<td>MDD</td>
<td>149</td>
<td>35.0</td>
<td>414</td>
<td>511</td>
</tr>
<tr>
<td>Controls</td>
<td>522</td>
<td>756</td>
<td>350</td>
<td>1719</td>
</tr>
<tr>
<td>Men</td>
<td>151</td>
<td>260</td>
<td>126</td>
<td>1545</td>
</tr>
<tr>
<td>Age, mean (SD), y</td>
<td>43.9</td>
<td>36.1</td>
<td>43.5</td>
<td>64.5</td>
</tr>
<tr>
<td>MDD</td>
<td>65</td>
<td>NA</td>
<td>22</td>
<td>180</td>
</tr>
</tbody>
</table>

Abbreviations: AD, alcohol dependence; COGA, Collaborative Study on the Genetics of Alcoholism; MDD, major depressive disorder; NA, not available; NHRVS, National Health and Resilience in Veterans Study; SAGE, Study of Addiction, Genetics, and Environment.
We used GEMMA, version 0.96 software\(^59\) to evaluate associations between the PRS and AD in a mixed-model framework accounting for related samples and cryptic relationships. Association analyses were corrected for age, sex, and the top 10 principal components to control for population stratification. Each PRS was scaled from 0 to 2 so that association strength was comparable across GWAS data sets and the same samples.

We conducted 2 additional analyses to further evaluate bidirectional relationships between MDD and AD: (1) to determine whether the association of the PRS with AD was driven by comorbid MDD cases in AD GWAS data sets, we repeated the analyses adjusting for MDD status in the 3 AD GWAS data sets (COGA, NHRVS, and Yale-Penn) for which MDD status was available; and (2) to examine whether the association of the PRS with AD was driven by comorbid AD cases in MDD PGC data sets, we recalculated the MDD PRS for AD GWAS data sets using weights from meta-analysis of the 2 imputed data sets using METAL.\(^60\) The portion of the variance in AD case-control status explained by the PRS was assessed by the Nagelkerke $R^2$ included as a covariate (COGA: best $P < .99$, $R^2 = 0.023$; Yale-Penn: best $P < .90$, $R^2 = 0.018$; NHRVS: best $P < .60$, $R^2 = 0.017$).

## Results

Basic sample demographics are reported in Table 1. In the COGA sample, 18.9% of individuals with AD had MDD and 12.5% of controls had MDD. In the Yale-Penn sample, 19.4% of those with AD had MDD and 6.3% of controls had MDD. The NHRVS sample had the highest rate of comorbid MDD, with 35.0% in AD cases; the MDD rate in controls was 10.5%. Major depressive disorder status was unavailable for SAGE.

We observed significant associations between the MDD PRS and AD case-control status across all AD samples at various PRS $P$ value thresholds (COGA: best $P = 1.7 \times 10^{-6}$, $R^2 = 0.026$; SAGE: best $P = .001$, $R^2 = 0.01$; Yale-Penn: best $P = .035$, $R^2 = 0.0018$; NHRVS: best $P = .004$, $R^2 = 0.0074$) (Table 2). Meta-analysis of the 4 AD samples showed stronger evidence for association with the smallest value ($P = 3.3 \times 10^{-3}$; $P$ value threshold $= .4$) (Table 2). Regression coefficients ($\beta$) were positive at each PRS $P$ value threshold, suggesting that an elevated MDD PRS is linked to an increased risk of AD. However, the proportion of variance in AD explained by the MDD PRS was small, ranging from $R^2 = 0.0018$ in the Yale-Penn sample to $R^2 = 0.026$ in the COGA sample (eTable 1 in the Supplement and Figure). To investigate potential sex-specific effects, the same analyses were repeated stratified by sex. No consistent difference in the strength or significance of association between MDD PRS and AD status by sex was observed across the 4 samples (eTable 2 and eFigure in the Supplement).

We observed similar patterns of association between the MDD PRSs and AD case-control status when MDD status was included as a covariate (COGA: best $P = 7.6 \times 10^{-6}$, $R^2 = 0.023$; SAGE: best $P = 1.7 \times 10^{-6}$, $R^2 = 0.026$).

## Table 2. Meta-analysis of Associations Between MDD Polygenic Risk Scores and AD

<table>
<thead>
<tr>
<th>Thresholds</th>
<th>Adjusting for MDD in AD Samples</th>
<th>No</th>
<th>Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P &lt; .05$</td>
<td>$\beta$ 0.075, SE 0.016</td>
<td>$P = 2.8 \times 10^{-6}$, SE 0.064</td>
<td>$P = 1.3 \times 10^{-4}$, SE 0.017</td>
</tr>
<tr>
<td>$P &lt; .10$</td>
<td>$\beta$ 0.075, SE 0.016</td>
<td>$P = 3.1 \times 10^{-6}$, SE 0.061</td>
<td>$P = 2.5 \times 10^{-4}$, SE 0.017</td>
</tr>
<tr>
<td>$P &lt; .20$</td>
<td>$\beta$ 0.084, SE 0.016</td>
<td>$P = 1.0 \times 10^{-7}$, SE 0.070</td>
<td>$P = 2.0 \times 10^{-5}$, SE 0.017</td>
</tr>
<tr>
<td>$P &lt; .30$</td>
<td>$\beta$ 0.088, SE 0.016</td>
<td>$P = 2.6 \times 10^{-6}$, SE 0.074</td>
<td>$P = 8.4 \times 10^{-6}$, SE 0.017</td>
</tr>
<tr>
<td>$P &lt; .50$</td>
<td>$\beta$ 0.095, SE 0.016</td>
<td>$P = 3.3 \times 10^{-9}$, SE 0.079</td>
<td>$P = 3.3 \times 10^{-6}$, SE 0.017</td>
</tr>
<tr>
<td>$P &lt; .60$</td>
<td>$\beta$ 0.094, SE 0.016</td>
<td>$P = 3.5 \times 10^{-9}$, SE 0.080</td>
<td>$P = 2.5 \times 10^{-6}$, SE 0.017</td>
</tr>
<tr>
<td>$P &lt; .70$</td>
<td>$\beta$ 0.095, SE 0.016</td>
<td>$P = 4.2 \times 10^{-9}$, SE 0.079</td>
<td>$P = 3.4 \times 10^{-6}$, SE 0.017</td>
</tr>
<tr>
<td>$P &lt; .80$</td>
<td>$\beta$ 0.093, SE 0.016</td>
<td>$P = 3.8 \times 10^{-9}$, SE 0.079</td>
<td>$P = 3.2 \times 10^{-6}$, SE 0.017</td>
</tr>
<tr>
<td>$P &lt; .90$</td>
<td>$\beta$ 0.094, SE 0.016</td>
<td>$P = 6.5 \times 10^{-9}$, SE 0.078</td>
<td>$P = 4.4 \times 10^{-6}$, SE 0.017</td>
</tr>
<tr>
<td>$P &lt; .99$</td>
<td>$\beta$ 0.094, SE 0.016</td>
<td>$P = 5.3 \times 10^{-9}$, SE 0.079</td>
<td>$P = 3.7 \times 10^{-6}$, SE 0.017</td>
</tr>
</tbody>
</table>

Abbreviations: AD, alcohol dependence; MDD, major depressive disorder.

Positive and Negative Controls

We tested the ability of the PGC-derived MDD PRS to predict MDD status in 3 AD GWAS data sets (COGA, NHRVS, and Yale-Penn) for which MDD status was available as a positive control. Height was similarly tested as a negative control in the same samples.
Yale–Penn: best $P = .08, R^2 = 0.0013$; and NHRVS: best $P = .006, R^2 = 0.0072$ (eTable 3 in the Supplement). Meta-analysis of the 3 samples strengthened evidence of association with the smallest $P = 2.5 \times 10^{-6} (P$ value threshold $= .50)$ (Table 2). The proportion of variance in AD explained by the MDD PRS was nearly identical to that of the unadjusted analyses in the corresponding samples.

We recalculated the MDD PRSs for AD GWAS samples using weights derived from 2 MDD GWAS data sets in which comorbid MDD-AD cases were excluded. We then tested the new MDD PRSs with AD case–control status in 3 AD GWAS samples without MDD cases. The new MDD PRS was nominally significantly associated with AD in the COGA sample at a PRS $P$ value threshold of $0.05 (P = .04)$ (eTable 4 in the Supplement). We did not observe significant evidence for association in the Yale–Penn and NHRVS samples, although the new MDD PRSs tended to increase AD risk across both samples. Meta-analysis of the 3 AD samples yielded significant evidence for association between the MDD PRS and AD with the smallest $P = .007$ (PRS $P$ value threshold $= .20$ or .05) (eTable 5 in the Supplement).

We tested whether the MDD PRS would predict MDD status as a positive control. A significant association was seen in the Yale–Penn sample ($7.6 \times 10^{-4} P \leq .004$) but not in either COGA or NHRVS, although a trend toward increased MDD risk was observed in both samples (eTable 6 in the Supplement). In a meta-analysis, significant associations were seen at all thresholds ($2.5 \times 10^{-4} P \leq .05$) (eTable 7 in the Supplement). Height, our negative control, showed no significant association with the MDD PRS in any sample at any $P$ value threshold (eTable 8 in the Supplement).

**Discussion**

In each independent AD GWAS sample, we found that a higher MDD PRS was associated with an increased risk for AD, and this result was even more pronounced in a meta-analysis of the 4 data sets. The same effect was seen in analyses with MDD status as a covariate in 3 AD GWAS samples. We also observed meta-analytic evidence for an association between MDD PRSs and AD, using a PRS derived from MDD GWAS data sets without comorbid MDD-AD cases. Our results suggest that the bi-directional associations between AD and MDD may be partially accounted for by shared genetic factors.

Our study methodology is similar to that of other published studies examining the association between AD genetic risk and AD age of onset$^{4,5}$; between childhood trauma, MDD genetic risk, and adult MDD$^{24}$; and between schizophrenia genetic risk and cannabis use.$^{64}$ Consistent with those studies, the proportion of variance in AD explained by the MDD PRS in our samples was small, on the order of $10^{-3}$ to $10^{-2}$. Similar effect sizes were seen in our positive control, MDD status.

Although our results do not illuminate the mechanism by which this shared genetic liability for MDD and AD acts, several possibilities are worth investigating in future studies. Anxiety disorders and MDD share genetic liability,$^{24,66}$ and anxiety has been proposed as a significant risk factor linking AD and MDD via the internalizing pathway.$^{49}$ More broadly, personality traits, such as neuroticism, disinhibition, and sensation-seeking, are associated with a range of psychiatric disorders, including anxiety, MDD, and AD, and are potential mechanisms through which shared AD-MDD polygenic risk may exert its effects.$^{67}$ Inclusion of comorbid disorders and personality trait measures in future studies may help to clarify these pathways.

The PGC MDD GWAS samples are likely enriched for individuals with AD, which could drive the associations initially observed between MDD PRSs and AD. To examine this possibility, we recalculated the MDD PRS using 2 MDD GWAS data sets without comorbid MDD-AD cases and tested the association of the new PRS with AD in AD GWAS samples without MDD cases. We observed a significant association between this recalculated MDD PRS and AD status only in the COGA sample at a threshold of $P < .05 (P = .04)$, potentially due to the smaller sample size of this training data set. Nonetheless, meta-analysis of all samples showed significant associations between the recalculated MDD PRS and AD, suggesting that the observed associations are not driven solely by comorbid AD cases in the PGC MDD GWAS and providing further support for genetic overlap between MDD and AD.
Prior family and twin studies suggested a role for sex-specific transmission of shared AD-MDD risk. In post hoc analyses, we examined sex-specific patterns of association between MDD PRSs and AD risk. We did not find consistent evidence of sex-specific transmission of shared AD-MDD liability in sex-stratified analyses, although our study was likely underpowered to detect such an effect given the reduced sample sizes in sex-specific subgroups.

Limitations
This work should be viewed in light of 2 limitations. First, MDD GWASs with larger sample sizes will likely improve the predictive ability of the MDD PRS and thus further refine the associations that we observed. Second, further studies are needed to investigate whether the MDD PRS association is specific to AD or whether it generalizes to substance dependence more broadly. There is evidence that a large proportion of genetic risk for substance use disorders involves predisposition to addiction generally, and some data have demonstrated that PRSs for psychopathology predict general substance involvement.

Conclusions
From a public health standpoint, MDD and AD are individually and jointly associated with significant morbidity and mortality, so understanding their association is a priority. Our findings suggest that common genetic factors contribute to MDD-AD comorbidity and that some individuals carry a genetic predisposition for both disorders. The consistency of our findings across 4 independent samples suggests the feasibility and value of a meta-analysis of AD and MDD GWAS data sets to identify specific genetic variations underlying this shared predisposition.
Polygenic Risk Score Analysis for Depression and Alcohol Dependence

Genotyping was performed at the Johns Hopkins University Center for high throughput genotyping for studying the co-occurrence of addictive and mental disorders: implications for prevention and service utilization. Am J Orthopsychiatry. 1996;66(1):37-50.

Genetics of Alcoholism (COGA) sample.

The genetics of alcoholism: identifying specific genes through family studies. Addict Biol. 2006;11(3-4):386-396.


Assistance with data cleaning was provided by the Database of Genotype and Phenotype (dbGAP). Funding support for SAGE was provided through the NIH Genes, Environment and Health Initiative (GEI) grant U01 HG004422. SAGE is one of the GWASs funded as part of the Gene Environment Association Studies (GENEVA) under GEI. Assistance with phenotype harmonization and genotype cleaning, as well as with general study coordination, was provided by the GENEVA Coordinating Center grant U01 HG004446. Assistance with data cleaning was provided by the National Center for Biotechnology Information. Support for collection of data sets and samples was provided by COGA grant U01 AA084801, the Collaborative Genetic Study of Nicotine Dependence grant P01 CA083932, and the Family Study of Cocaine Dependence grant R01 DA03423. Funding support for genotyping, which was performed at the Johns Hopkins University Center for Inherited Disease Research, was provided by NIH GEI U01HG004438, NIAAA, and NIH and contract “high throughput genotyping for studying the genetic contributions to human disease” (HHSN268200782096C). The COGA case-control GWAS data set used for the analyses described in this article was obtained from dbGAP (accession number phs000012.v1.p1). Genotyping was performed at the Johns Hopkins University Center for Inherited Disease Research. Funding for data collection, phenotype harmonization, genotype cleaning, and analysis was through COGA grant U01 AA084801.

REFERENCES