Title

High Polygenic Risk Scores Are Associated With Early Age of Onset of Alcohol Use Disorder in Adolescents and Young Adults at Risk

Journal

Biological Psychiatry Global Open Science, 2(4)

ISSN

2667-1743

Authors

Nurnberger, John I
Wang, Yumin
Zang, Yong
et al.

Publication Date

2022-10-01

DOI

10.1016/j.bpsgos.2021.10.007

Peer reviewed

Archival Report

Biological Psychiatry: GOS

High Polygenic Risk Scores Are Associated With Early Age of Onset of Alcohol Use Disorder in Adolescents and Young Adults at Risk

John I. Nurnberger Jr., Yumin Wang, Yong Zang, Dongbing Lai, Leah Wetherill, Howard J. Edenberg, Fazil Aliev, Martin H. Plawecki, David Chorlian, Grace Chan, Kathleen Bucholz, Lance Bauer, Chella Kamarajan, Jessica E. Salvatore, Manav Kapoor, Victor Hesselbrock, Danielle Dick, Laura Bierut, Vivia McCutcheon, Jacquelyn L. Meyers, Bernice Porjesz, John Kramer, Samuel Kuperman, Sivan Kinreich, and Andrey P. Anokhin, for investigators from the Collaborative Study on the Genetics of Alcoholism

ABSTRACT

BACKGROUND: Genome-wide association studies have been conducted in alcohol use disorder (AUD), and they permit the use of polygenic risk scores (PRSs), in combination with clinical variables, to predict the onset of AUD in vulnerable populations.

METHODS: A total of 2794 adolescent/young adult subjects from the Collaborative Study on the Genetics of Alcoholism were followed, with clinical assessments every 2 years. Subjects were genotyped using a genome-wide chip. Separate PRS analyses were performed for subjects of European ancestry and African ancestry. Age of onset of DSM-5 AUD was evaluated using the Cox proportional hazard model. Predictive power was assessed using receiver operating characteristic curves and by analysis of the distribution of PRS.

RESULTS: European ancestry subjects with higher than median PRSs were at greater risk for onset of AUD than subjects with lower than median PRSs ($p = 3 \times 10^{-7}$). Area under the curve for the receiver operating characteristic analysis peaked at 0.88 to 0.95 using PRS plus sex, family history, comorbid disorders, age at first drink, and peer drinking; predictive power was primarily driven by clinical variables. In this high-risk sample, European ancestry subjects with a PRS score in the highest quartile showed a 72% risk for developing AUD and a 35% risk of developing severe AUD (compared with risks of 54% and 16%, respectively, in the lowest quartile).

CONCLUSIONS: Predictive power for PRSs in the extremes of the distribution suggests that these may have future clinical utility. Uncertainties in interpretation at the individual level still preclude current application.

https://doi.org/10.1016/j.bpsgos.2021.10.007

Alcohol use disorder (AUD) has its peak onset in late adolescence and early adulthood [ages 18–29 years (1)]. Family history of DSM-IV alcohol dependence is known to increase risk by at least twofold (2). Males are more likely than females to develop AUD (1–4), both within families of alcohol-dependent probands and in the general population (2).

In the United States, persons of African ancestry (AA) are less likely to develop AUD than persons of European ancestry (EA) (5–7). However, over recent years, there have been relatively greater increases of AUD in women and AA individuals compared with EA males (8).

Early age at first drink has been robustly associated with increased risk for adult drinking problems (9–11). There has been controversy over whether this is the best predictor of regular drinking and alcohol problems in adolescents/young adults (11–14). However, it does appear to be associated with genetic vulnerability (15–18).

There is a known risk relationship between other psychiatric disorders and AUD. Adolescents with a mood disorder are at increased risk for alcohol problems and disorders (19–21) and vice versa (21). Scores on an internalizing scale were positively correlated with risk for alcohol and other drug use disorders (22,23). There is an extensive literature supporting the relationship of externalizing disorders (especially conduct disorder) to subsequent development of AUD (22,24–27). There is a substantial overlap between genetic vulnerability to AUD and vulnerability to externalizing disorders in general (28).

Genetic vulnerability to common disorders is quantifiable using polygenic risk scores (PRSs), which combine the cumulative effects of many genetic loci into a single metric

ISSN: 2667-1743

(29–32). In practice, a PRS is calculated from a discovery genome-wide association study and applied to an independent target dataset. The PRS enumerates the number of copies of the risk allele carried by that individual at each locus (0, 1, or 2) and weights each risk allele by its effect size. Procedures for single nucleotide polymorphism (SNP) selection may generally be divided into 1) pruning and thresholding, which generates a set of scores based on different *p* value thresholds for the odds ratio estimates for disease association; or 2) Bayesian, which generates a single optimized score. Covariates for sex, ancestry, and other variables may also be added. Most of the literature on PRSs is based on thresholding methods, but Bayesian analysis is becoming more widely used because it obviates the need for multiple testing correction.

PRSs have been used in many complex medical conditions. Khera *et al.* (33) studied genetic risk for coronary artery disease, atrial fibrillation, type 2 diabetes mellitus, inflammatory bowel disease, and breast cancer. Using PRSs, they identified subgroups of subjects with 3 times, 4 times, and 5 times increased risk for each disorder. In a subsequent paper, Khera *et al.* (34) predicted severe obesity in a 20-year follow-up study of young adults, showing 1.3% severe obesity among subjects in the lowest PRS decile compared with 15.6% among subjects in the highest decile.

PRSs have been reported to discriminate between cases and controls with an area under the receiver operating characteristic (ROC) curve as high as 82% for schizophrenia, 65% for bipolar disorder, 58% for major depressive disorder, and 54% for anxiety, although there is substantial variability in estimates across studies (35). Recently Musliner *et al.* (36) showed a significant increase of conversion risk from major depressive disorder to bipolar disorder among subjects in different PRS subgroups in a sample of offspring of persons with bipolar disorder. Such studies have treatment implications (32).

In this paper, we explore the applicability of alcohol-related PRSs to AUD in an adolescent and young adult sample. We have previously reported in the Collaborative Study of the Genetics of Alcoholism (COGA) dataset effects of sex, ancestry, family type (case/comparison), and comorbid psychiatric disorders on risk for early onset of AUD (37). This report extends this analysis to include PRS, the age of first drink, and peer drinking. All of these variables (with the exception of peer drinking and comorbidity) are known to be fixed in value prior to the onset of AUD in any subject. COGA subjects have participated in separate multicenter studies of PRSs (38,39), but not in combination with clinical variables.

METHODS AND MATERIALS

Subjects

This analysis was based on the COGA Prospective Study Dataset (22). Ascertainment sites were University of Connecticut (Farmington, CT), Indiana University School of Medicine (Indianapolis, IN), University of Iowa (Iowa City, IA), Washington University in St. Louis (St. Louis, MO), State University of New York Downstate (Brooklyn, NY), Howard University (Washington, DC), and University of California San Diego (San Diego, CA). All participants provided informed consent for study procedures. All protocols were approved by institutional review boards at the various institutions.

This sample (n = 2794) was a subset of the Prospective Study Dataset (N = 3286) (37); the subset included all dataset members with genotypic data. Adolescent and young adult subjects were assessed with a structured psychiatric interview [Semi-Structured Assessment for the Genetics of Alcoholism (40)] at 2-year intervals from 2004 to 2017. All subjects aged between 12 and 21 years at baseline assessment were invited to participate. Every subject with at least 1 complete interview (73% of those invited) was included in the dataset. The average age at first interview was 16.1 years (SD = 3.3), and the average age at last interview was 23.1 years (SD = 5.0). Most subjects (87%) came from extended families with a proband in treatment for an AUD (designated family type = case); 13% were from comparison families (family type = comparison), recruited from sources such as dental clinics and motor vehicle records. Subjects in comparison families were not prescreened and may be expected to have population rates of common disorders such as AUD (2). Subjects in the dataset were included in analyses regardless of initiation of alcohol use by the time of the baseline interview. More than 80% of subjects participated in at least one follow-up. The average number of interviews per subject was 4.0 (SD = 1.7). There was no evidence for selective attrition of affected subjects (22). All subjects had data on comorbid diagnoses, AUD diagnosis, peer drinking, and age of first drink. Genotyping was carried out and quality control procedures applied as previously described (41,42). Ancestry was assigned based on the first four genotypic principal components of population stratification (Figure S1). Individuals surrounding the CEU HapMap position were assigned to the EA sample, and those beyond a radius of about 0.002 units on the principal component 1 axis were assigned to the AA sample. To maintain power for family-based analyses, final ancestry assignment was based on the majority of individual family members. Families with an equal distribution were assigned to the most diverse group (AA). This analysis excluded subjects who were not assignable to either the EA or AA groups. There were 2794 subjects with genome-wide association study data available for this analysis, 67% EA (1872 subjects from 559 families) and 33% AA (922 subjects from 219 families). Among the subjects, 51% were female and 49% were male.

The dependent variables for the analyses were age of onset (AOO) for all AUDs and severe AUD as defined by DSM-5 (43). Externalizing disorders and internalizing disorders were defined as in Nurnberger *et al.* (37). Briefly, we considered DSM-IV (44) diagnoses of externalizing disorders (attention-deficit/hyperactivity disorder, conduct disorder, antisocial personality disorder, oppositional defiant disorder, and drug use disorder excluding alcohol or tobacco) and internalizing disorders (major depressive disorder, panic disorder, obsessive-compulsive disorder, social phobia, and agoraphobia) if they occurred before or at the same time as the onset of AUD. Comorbidity was scored as membership in one of four groups: externalizing (at least one disorder), internalizing (at least one disorder), both, or neither.

Age of first drink was defined using the Semi-Structured Assessment for the Genetics of Alcoholism question "How old were you the first time you had your very first whole drink?" Peer drinking was defined using the Semi-Structured Assessment for the Genetics of Alcoholism question "When you were 12–17, how many of your best friends used alcohol?" Answers were none, a few, most, or all. For analysis, we divided responses into two categories: most friends drink and most friends do not drink.

Construction of PRSs

We used the Million Veteran Program (MVP) (45,46) datasets as discovery samples and examined EA and AA PRSs for ICD 9/10 AUD, Alcohol Use Disorders Identification Test-Concise (AUDIT-C) scores, and MAX_ALC, defined as the highest number of drinks a subject reported drinking during a single day in a typical month (46). The MVP dataset included AUD cases (56,000, including 34,000 EA and 17,000 AA) and controls (219,000, including 167,000 EA and 39,000 AA). This is the largest sample of AUD cases available currently (although there are larger samples with data on consumption or problematic use). Variants located within 500 kb of the index variant and having $R^2 > 0.25$ with the index variant were clumped. PRSs were calculated as the sum of allele counts weighted by the sign of the log odds ratio and the negative log-transformed p value for each SNP. The weighting by p value was used because it is robust to variations in sample size from SNP to SNP. Each PRS was tested at nine thresholds. A p value of 3.3×10^{-4} was considered significant for a PRS after Bonferroni correction (although this is probably conservative, because PRSs are correlated). Calculations were performed with PLINK (47). For comparison, we also ran PRS-CS (48) for EA subjects and PRS-CSx (49) for AA subjects. For these analyses, the Bonferroni correction was 1.6×10^{-2}. All PRS results were standardized (mean = 0, SD = 1).

We applied PRS to the COGA Prospective Study EA dataset, controlling for sex, ancestry (using the first four principal components as indicated by the scree plot) (Figure S1), family history, and relatedness. We used the frailty model to capture relatedness for the COGA data; this is the most widely used method to handle correlated survival data (50). Specifically, the frailty model puts a random effect term in the Cox regression formula such that the members of the same family will share the same random effect.

PRS analyses were performed separately in AA subjects using PRSs derived from the MVP AA dataset. The relatively small number of subjects (4%) with non-EA, non-AA ancestries were not included in PRS analyses.

Kaplan-Meier curves were used to estimate the survival function of AOO of AUD. A Cox proportional hazards model was used to test the relationship of AOO with PRSs adjusted for sex, relatedness, and case/comparison family status in the original model. Other variables added into the model for specific analyses included comorbid disorders, age of first drink, and peer drinking. The proportional hazards assumption was tested in these analyses and was not violated.

Time-dependent ROC curves were created to examine predictive value for the diagnosis of any AUD or severe AUD at a given age (and an integrated value for the age range 15–27). Area under the curve (AUC) was calculated to assess the performance of the model. All statistical analyses (except calculation of PRS) were run using SAS/STAT 15.1 (51).

RESULTS

By the end of the follow-up period, 1544 of 2794 subjects were unaffected, and 606 were diagnosed with mild AUD, 365 with moderate AUD, and 279 with severe AUD. In total, 45% of subjects were diagnosed with AUD during the assessment period [for additional detail on outcomes, see (38)]. Table 1 shows the distribution of subjects in various clinical categories.

Results of the Cox proportional hazards model are presented in Tables 2 and 3. Risk was increased for males, subjects from case families, and subjects with increased PRS. PRS from MAX_ALC.P1 (which uses the threshold $p < .1$) was associated ($p < 6.1 \times 10^{-6}$ by Cox test) with AOO of AUD in EA subjects (Table 1). PRSs derived from AUD or AUDIT-C scores did not show a strong relationship with AOO in EA subjects ($p > .001$ at most thresholds). PRS variables were not significantly associated with AOO in AA subjects (Tables S1A and S1B).

A histogram of PRS in EA subjects from case and comparison families is shown in Figure 1. The distribution of PRS is shifted to the right in case subjects, although there is substantial overlap. The shift is not significant in the thresholded analysis ($p = 4.43 \times 10^{-4}$), but it is when calculated by PRS-CS ($p = 2.89 \times 10^{-6}$). A similar histogram for AA subjects is shown in Figure S2 (no significant shift by thresholded analysis or by PRS-CSx).

In general, recalculation of PRS by PRS-CS and PRS-CSx produced more significant p values (or closer to significance in the case of AA subjects) but did not change the results substantially, except as noted above.

Figure 2 illustrates the relationship between PRS and onset of AUD in EA subjects ($p = 5 \times 10^{-6}$ by the Cox model; $p = 5 \times 10^{-7}$ by log-rank test). The median AOO for AUD for subjects with a PRS in the top half of the distribution is 20, while the median AOO for subjects in the lower half of the PRS distribution is 22.

Combining PRSs With Demographic and Clinical Variables

As noted above, PRS was significantly associated with AOO covarying for sex, family type, and ancestry (Table 2). Adding in comorbid disorders, age of first drink, and peer drinking, PRS MAX_ALC.P1 remained associated with AOO of AUD in EA subjects (hazard ratio = 1.14, CI = 1.06−1.22, $p = 2.11 \times 10^{-4}$) (Table 3). For data in AA subjects, see Tables S1A and S1B. There is no significant relationship between PRS and AOO in AA subjects.

Risk Calculation Using ROC Curves

AUC peaks at 88% (all AUD) and 95% (severe AUD) for EA subjects (Table S2). Using a cut-point of 0.5, sensitivity is

Table 1. Description of the Sample by Clinical/Demographic Variables

Variable	European Ancestry, *n* = 559 Families			African Ancestry, *n* = 218 Families			Total
	n	Column%[a]	Row%[b]	*n*	Column%[a]	Row%[b]	*n* (%)
Sex							
Female	953	50.91%	66.74%	475	51.52%	33.26%	1428 (51.11%)
Male	919	49.09%	67.28%	447	48.48%	32.72%	1366 (48.89%)
Total	1872	100.00%	67.00%	922	100.00%	33.00%	2794 (100.00%)
Comorbidity							
Internalizing only	89	4.75%	80.18%	22	2.39%	19.82%	111 (3.97%)
Externalizing only	508	27.14%	59.42%	347	37.64%	40.58%	855 (30.60%)
Both	170	9.08%	66.93%	84	9.11%	33.07%	254 (9.09%)
Neither	1105	59.03%	70.20%	469	50.87%	29.80%	1574 (56.34%)
Total	1872	100.00%	67.00%	922	100.00%	33.00%	2794 (100.00%)
Peer Drinking							
Peer drinking	858	45.83%	70.04%	367	39.80%	29.96%	1225 (43.84%)
No peer drinking	1012	54.06%	64.58%	555	60.20%	35.42%	1567 (56.08%)
Not available	2	0.11%	100.00%	0	0.00%	0.00%	2 (0.07%)
Total	1872	100.00%	67.00%	922	100.00%	33.00%	2794 (100.00%)
Case/Comparison Status							
Subjects from case families	1559	83.28%	64.24%	868	94.14%	35.76%	2427 (86.86%)
Subjects from comparison families	313	16.72%	85.29%	54	5.86%	14.71%	367 (13.14%)
Total	1872	100.00%	67.00%	922	100.00%	33.00%	2794 (100.00%)
Age of First Drink							
Has onset age of first drink	1674	89.42%	67.61%	802	86.98%	32.39%	2476 (88.62%)
No first drink reported	198	10.58%	62.26%	120	13.02%	37.74%	318 (11.38%)
Total	1872	100.00%	67.00%	922	100.00%	33.00%	2794 (100.00%)
AUD Diagnosis							
Mild	423	22.60%	69.80%	183	19.85%	30.20%	606 (21.69%)
Moderate	279	14.90%	76.44%	86	9.33%	23.56%	365 (13.06%)
Severe	223	11.91%	79.93%	56	6.07%	20.07%	279 (9.99%)
No AUD	947	50.59%	61.33%	597	64.75%	38.67%	1544 (55.26%)
Total	1872	100.00%	67.00%	922	100.00%	33.00%	2794 (100.00%)

AUD, alcohol use disorder.
[a]Column% = % of total in that column.
[b]Row% = % of total in that row.

72% and specificity is 87% at age 15. Sensitivity is 73% and specificity is 67% at age 20. In most models, peak AUC is achieved with inclusion of all variables. AUC power is primarily driven by the variable age of first drink, and the independent effects of other variables (including PRS) are generally small. PRS adds 8.6% to AUC by itself and 0.7% after all clinical variables in the integrated EA all AUD analysis; in the integrated EA severe AUD analysis, PRS adds 10.4% by itself but adds no additional variance following all clinical variables. AUC peaks at 88% (all AUD) and 92% (severe AUD) for AA subjects (Table S2). In the AA analyses, PRS adds 5.5% initially to all AUD and 13.5% initially to severe AUD but <1% when added after all clinical variables.

Table 2. Cox Proportional Hazards Model for Sex, PRSs (EA_MAX_ALC.P1), and Family Type (Case/Comparison) in EA Subjects: Analysis of Maximum Likelihood Estimates

Parameter	Parameter Estimate	Standard Error	χ^2_1	*p* Value	HR	95% Confidence Limits
Sex	−0.28030	0.06852	16.7344	4.30×10^{-5}	0.756	0.661–0.864
EA_MAX_ALC.P1 (With PC1—PC4)	0.16451	0.03600	20.8803	4.89×10^{-6}	1.179	1.099–1.265
Family Type	0.54092	0.11717	21.3110	3.90×10^{-6}	1.718	1.365–2.161

The model is also adjusted for ancestry using the first four PCs of population stratification (PC1–PC4). All HRs show the effect of a particular variable after accounting for the effects of all other variables.

EA, European ancestry; HR, hazard ratio; PC, principal component; PRS, polygenic risk score.

Table 3. Cox Proportional Hazards Model for Sex, PRSs (EA_MAX_ALC.P1), Family Type (Case/Comparison), Age of First Drink, Peer Drinking, and Comorbidity in EA Subjects: Analysis of Maximum Likelihood Estimates

Parameter	Parameter Estimate	Standard Error	χ^2_1	*p* Value	Hazard Ratio	95% Confidence Limits
Sex	−0.20281	0.06854	8.7550	3.09×10^{-3}	0.816	0.714–0.934
EA_MAX_ALC.P1 (With PC1–PC4)	0.13171	0.03554	13.7311	2.11×10^{-4}	1.141	1.064–1.223
Family Type	0.16013	0.11287	2.0128	1.56×10^{-1}	1.174	0.941–1.464
Age of First Drink	−0.15056	0.01239	147.6970	5.53×10^{-34}	0.860	0.840–0.881
Peer Drinking	0.68893	0.07471	85.0399	2.92×10^{-20}	1.992	1.720–2.306
Comorbidity						
Both	0.30369	0.10633	8.1578	4.29×10^{-3}	1.355	1.100–1.669
Externalizing only	0.32966	0.07749	18.0974	2.10×10^{-5}	1.390	1.195–1.619
Internalizing only	−0.11756	0.15937	0.5442	4.61×10^{-1}	0.889	0.651–1.215

The model is also adjusted for ancestry using the first four PCs of population stratification (PC1–PC4). All hazard ratios show the effect of a particular variable after accounting for the effects of all other variables.

EA, European ancestry; PC, principal component; PRS, polygenic risk score.

Subgroups Based on Divisions of the PRS Distribution

Because AUC calculations do not capture the power of PRS at the extremes of the distribution, we plotted AOO for AUD in quartiles of EA PRS (Figure 3). Median AOO for subjects in the four quartiles were 20, 20, 22, and 24, respectively. By age 25, 72% of subjects in the highest quartile manifested AUD compared with 54% in the lowest quartile.

Figure 4 shows an analysis of AOO for severe AUD in EA subjects using PRS quartiles. By the age of 25, 35% of subjects in the top quartile were diagnosed with severe AUD as compared with 16% of subjects in the bottom quartile. Subgroup analyses in AA subjects are presented in Figures S2–S5.

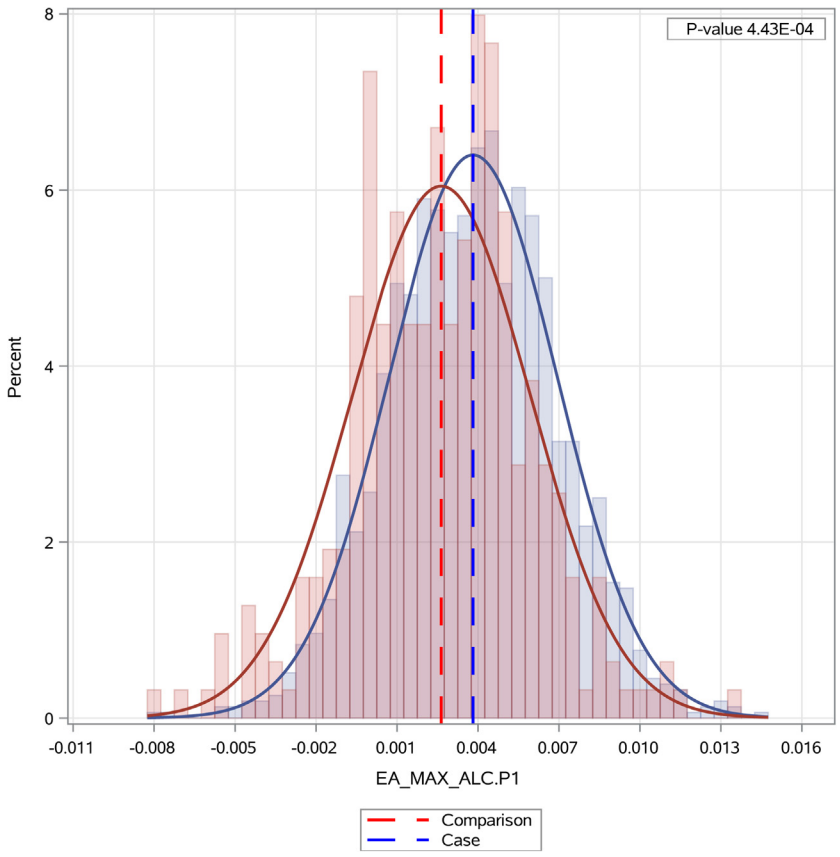

Figure 1. Polygenic risk score distribution in subjects from European ancestry (EA) case families (blue) and comparison families (red). The polygenic risk score distribution is shifted to the right in the case families, adjusting for sex and principal component 1 through principal component 4. The shift is not significant in the thresholded analysis (shown here for consistency) but is significant when calculated by the PRS-CS method (48) ($p = 2.89 \times 10^{-6}$). EA_MAX_ALC.P1, PRS based on MAX_ALC in European ancestry subjects using the threshold $p < .1$.

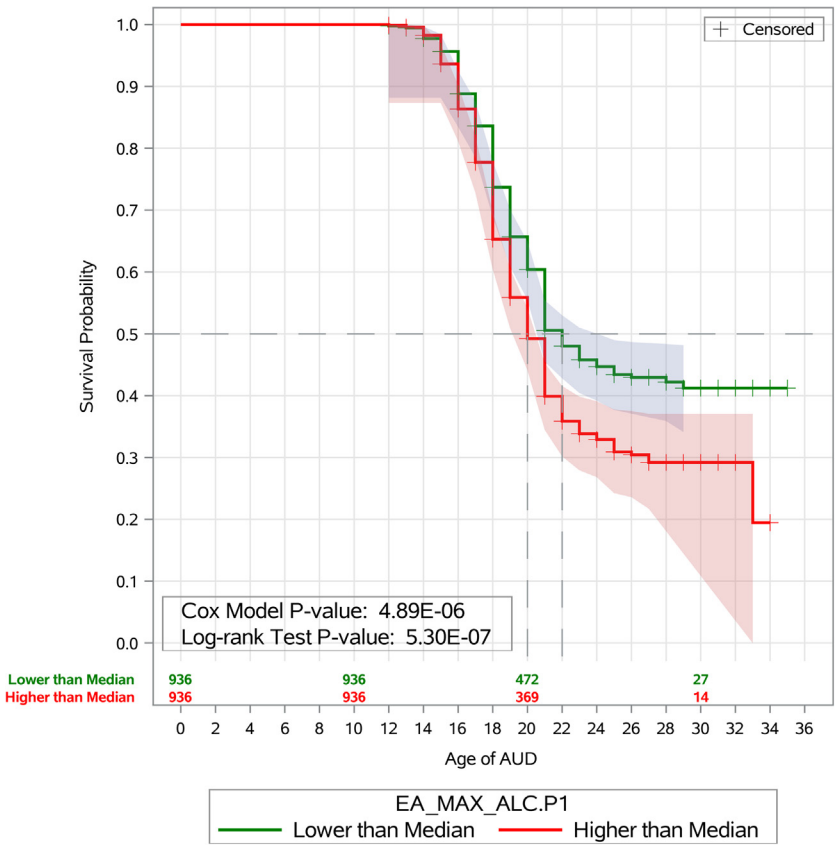

Figure 2. Age of onset for first diagnosis of alcohol use disorder (AUD) in European ancestry subjects with high polygenic risk scores (using a median split) is lower than age of onset for AUD for subjects with low polygenic risk scores. The Cox model shows the p value incorporating sex and principal components for ancestry ($p = 4.89 \times 10^{-6}$). The log-rank test is not adjusted for covariates ($p = 5.30 \times 10^{-7}$). EA_MAX_ALC.P1, PRS based on MAX_ALC in European ancestry subjects using the threshold $p < .1$.

DISCUSSION

Association of PRSs With Onset of AUD

In our analysis, a PRS derived from MAX_ALC was associated with AOO of AUD in our high-risk sample. Clinical prediction that includes easily measurable variables, such as comorbid conditions, age of first drink, and peer drinking, shows efficacy (0.88–0.95) that approaches the range of clinical utility (52), with or without PRS. The predictive value of PRS is most evident when the extremes of the PRS distribution are compared, as illustrated by a 1.3-times increase in the probability of developing any AUD and a 2.2-times increase in the probability of developing severe AUD by age 25 among EA individuals from the top and bottom PRS quartiles.

PRSs may be presumed to increase in predictive accuracy with the size of the discovery sample. We used the largest sample available at this time (MVP). Because AUD is a common disorder in the population [>29% lifetime prevalence (8)] and is only moderately heritable [approximately 55% (53)], we may expect that very large samples will be necessary to achieve optimal prediction (54).

PRSs and Ancestry

In our Cox model analyses, PRS was not significantly associated with AOO of AUD in the AA sample (although predictive ability using ROC was similar in the two groups using primarily clinical variables). Recent studies (55) show that PRSs for bipolar disorder derived from EA subjects are reasonably predictive when applied to East Asian subjects but less so when applied to AA populations. It is well established that linkage disequilibrium blocks (units of correlated genetic markers) are smaller in AA subjects than in other populations. This is related to the history of the human species, which extends an order of magnitude longer on the African continent than in other areas of the world (56). Thus, there are different allele frequencies and greater variation in allele frequency in AA subjects compared with those of other ancestries. The reliability of PRS developed from other ancestral groups is consequently less in AA subjects. Because PRS would currently be less useful for AA subjects than for EA subjects, questions arise regarding inequities in clinical application (57). One solution is increased emphasis on collection of genetic samples from diverse populations, especially those of AA. We should note that the sample sizes in both the discovery (MVP) and target (COGA) datasets in these analyses are considerably smaller for AA subjects compared with EA subjects.

Clinical Utility of Predictive Algorithms

Effective treatments for AUD are now available, such as cognitive behavioral therapy, oral naltrexone, and long-acting

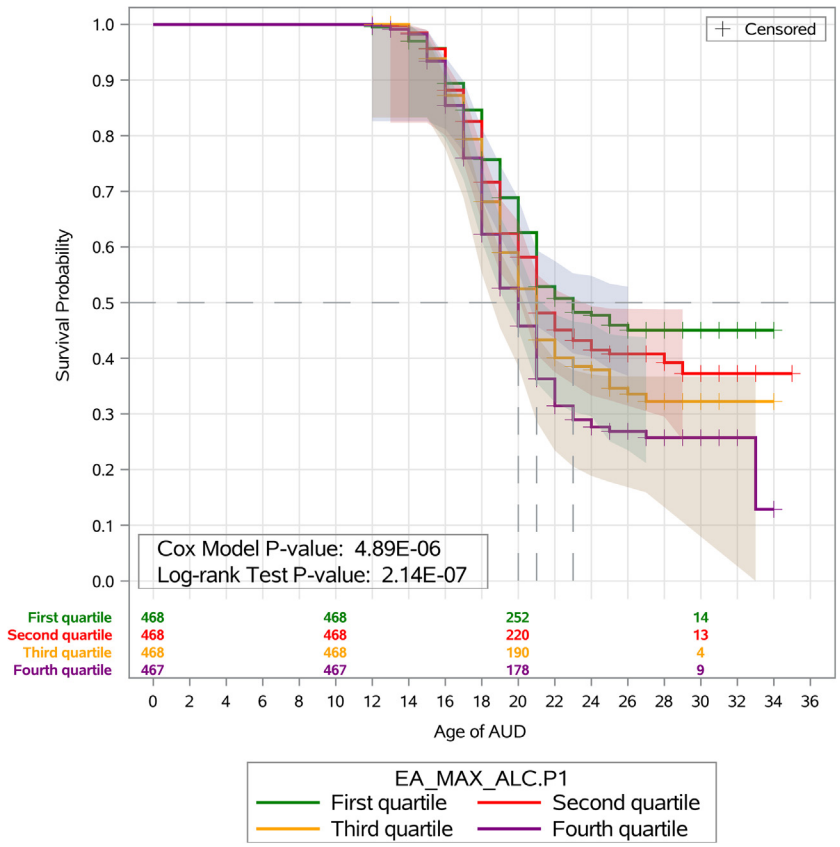

Figure 3. Kaplan-Meier curves (survival analysis) for age of first alcohol use disorder (AUD) diagnosis for European ancestry subjects in quartiles of the polygenic risk score distribution. Shaded areas around each curve represent 95% confidence intervals. The Cox model incorporates sex and ancestry principal components. The log-rank test does not incorporate covariates. EA_MAX_ALC.P1, PRS based on MAX_ALC in European ancestry subjects using the threshold $p < .1$.

injectable naltrexone, among others (58,59). These treatments are capable of saving lives, relationships, and careers, as well as minimizing medical comorbidities. Specific treatment and/or monitoring might be considered for adolescents and young adults in high-risk subgroups who are already drinking and manifesting problematic alcohol use patterns, such as binge drinking.

We note the importance of age of first drink in the AUC analyses. Early exposure to alcohol may have causal effects on later addictive behaviors. Adolescent exposure to alcohol preferentially increases alcohol drinking during adulthood in rat models (60), along with brain changes in important neurotransmitter systems and brain areas for appetitive behavior (61,62). Similar mechanisms in humans are a plausible hypothesis (63,64). Prevention strategies to delay exposure to alcohol might receive additional attention in public health efforts at harm reduction.

Limitations

The discovery sample differed from the target sample in several important respects. MVP is a study of older adult veterans, predominantly male. The COGA sample is mixed male and female adolescents and young adults and is a sample of subjects at risk for AUD. Perhaps one consequence of this is the observation that PRS for MVP AUD was less effective in the COGA sample than PRS for MAX_ALC; types of AUD diagnoses in the two samples may differ more than this quantitative lifetime measure. Replication of these results in independent cohorts would help establish their generalizability.

When discussing predictive testing, it is important to consider the issue of stigmatization, especially in young people. We are not developing algorithms for clinical use in individuals below drinking age at this time. However, we would advocate consideration of algorithms for prediction of early-onset AUD when problematic drinking is identified in adolescents. The fact that the algorithm includes genetic information is no reason to consider it additionally problematic in terms of stigmatization because heritable disorders may be successfully treated. Concerns regarding genetic information should be addressed as aspects of public health education, directed at consumers, providers, and ultimately the general population (65). The larger issue at present is difficulties in interpretation of individual PRS values, especially in mixed ancestry populations.

Methodologic improvements of PRS [e.g., PRS-CS, included here for comparison (48)], which obviates correction for multiple testing, and the use of SNP weights based on gene expression in PRS (66) might provide improved predictive capacity.

It is important to realize that physicians and scientists are no longer the only gatekeepers for individual genetic information.

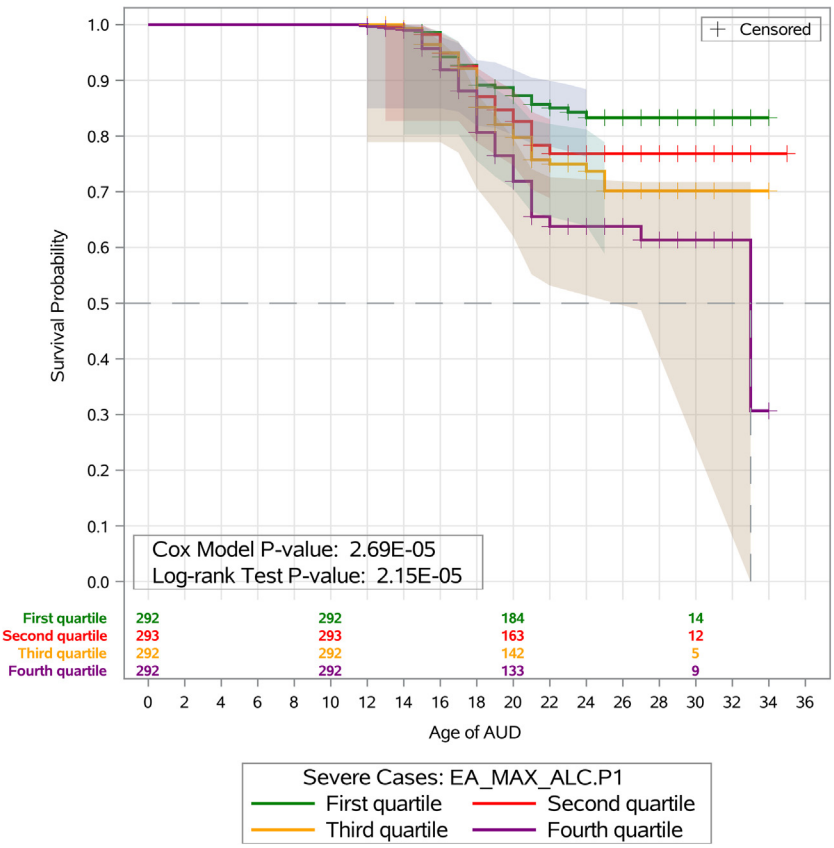

Figure 4. Kaplan-Meier curves (survival analysis) for age of first severe alcohol use disorder (AUD) diagnosis for European ancestry subjects in quartiles of the polygenic risk score distribution. Shaded areas around each curve represent 95% confidence intervals. The Cox model incorporates sex and ancestry principal components. The log-rank test does not incorporate covariates. EA_MAX_ALC.P1, PRS based on MAX_ALC in European ancestry subjects using the threshold $p < .1$.

Direct-to-consumer companies offer genome-wide data and, in some cases, PRSs for multiple conditions, including psychiatric disorders. Individual patients may obtain this information and may bring it to their doctors. It is incumbent on clinical professionals to guide patients on conservative interpretation of such data.

Clinical trials may eventually be used to formally test the value of PRSs in combination with appropriate clinical variables. The model might be similar to that used in recent trials using pharmacogenetic testing (67), although appropriate follow-up would presumably be more extended. This would provide an opportunity for real-life examination of the feasibility and utility of such scores. Appropriate clinician training would be an essential part of such trials.

In conclusion, several variables had significant effects on AOO of AUD in this study: sex, family history, age of first drink, peer drinking, comorbidity, and PRS. Discriminatory power in the ROC model was maximized by using age of first drink along with other variables. PRS was useful in identifying subgroups at unusually high risk. Such algorithms might have a place in the future clinical practice of psychiatry. Larger samples, especially AA samples, will be necessary to support more effective use of PRS in diverse clinical populations. In combination with clinical variables, PRS may aid in prediction of outcome and clinical decision making. Following additional study, clinical trials may help to assess feasibility and utility.

ACKNOWLEDGMENTS AND DISCLOSURES

This national collaborative study is supported by the National Institutes of Health (Grant No. U10AA008401 from the National Institute on Alcohol Abuse and Alcoholism and Grant No. K01 AA024152 from the National Institute on Drug Abuse [to JES]).

The Collaborative Study on the Genetics of Alcoholism (COGA) (Principal Investigators B. Porjesz, V. Hesselbrock, T. Foroud; Scientific Director, A. Agrawal; Translational Director, D. Dick) includes 11 different centers: University of Connecticut (V. Hesselbrock); Indiana University (H.J. Edenberg, T. Foroud, J. Nurnberger Jr, Y. Liu); University of Iowa (S. Kuperman, J. Kramer); SUNY Downstate (B. Porjesz, J. Meyers, C. Kamarajan, A. Pandey); Washington University in St. Louis (L. Bierut, J. Rice, K. Bucholz, A. Agrawal); University of California at San Diego (M. Schuckit); Rutgers University (J. Tischfield, A. Brooks, R. Hart); The Children's Hospital of Philadelphia, University of Pennsylvania (L. Almasy); Virginia Commonwealth University (D. Dick, J. Salvatore); Icahn School of Medicine at Mount Sinai (A. Goate, M. Kapoor, P. Slesinger); and Howard University (D. Scott). Other COGA collaborators include L. Bauer (University of Connecticut); L. Wetherill, X. Xuei, D. Lai, S. O'Connor, M. Plawecki, and Y. Zang (Indiana University); L. Acion (University of Iowa); G. Chan (University of Iowa; University of Connecticut); D.B. Chorlian, J. Zhang, S. Kinreich, and G. Pandey (SUNY Downstate); M. Chao (Icahn School of Medicine at Mount Sinai); A. Anokhin, V. McCutcheon, and S. Saccone (Washington University); and F. Aliev and P. Barr (Virginia Commonwealth University). H. Chin and A. Parsian are the National Institute on Alcohol Abuse and Alcoholism Staff Collaborators.

We continue to be inspired by our memories of Henri Begleiter and Theodore Reich, founding Principal Investigator and Co-Principal Investigator of COGA, and also owe a debt of gratitude to other past organizers of COGA, including Ting-Kai Li, P. Michael Conneally, Raymond Crowe, and Wendy Reich, for their critical contributions.

JIN is an investigator for Janssen on an unrelated study. All other authors report no biomedical financial interests or potential conflicts of interest.

ARTICLE INFORMATION

From the Department of Psychiatry (JIN, MHP), Department of Biostatistics and Health Data Science (YW, YZ), Department of Medical and Molecular Genetics (JIN, DL, LW, HJE), Department of Biochemistry and Molecular Biology (HJE), and Stark Neurosciences Research Institute (JIN), Indiana University School of Medicine, Indianapolis, Indiana; Department of Psychology (FA, JES, DD) and Virginia Institute for Psychiatric and Behavioral Genetics (JES), Virginia Commonwealth University, Richmond, Virginia; Department of Psychiatry (GC, JK, SKu), University of Iowa Carver College of Medicine, Iowa City, Iowa; Henri Begleiter Neurodynamics Laboratory (DC, CK, JLM, BP, SKi), Department of Psychiatry, State University of New York Downstate Health Sciences University, Brooklyn; Department of Psychiatry (MK), Icahn School of Medicine at Mount Sinai, New York, New York; Department of Psychiatry (GC, LBa, VH), University of Connecticut, Farmington, Connecticut; and the Department of Psychiatry (KB, LBi, VM, APA), Washington University School of Medicine, St. Louis, Missouri.

Address correspondence to John I. Nurnberger Jr, M.D., Ph.D., at jnurnber@iu.edu.

Received Jun 2, 2021; revised Sep 30, 2021; accepted Oct 1, 2021.

Supplementary material cited in this article is available online at https://doi.org/10.1016/j.bpsgos.2021.10.007.

REFERENCES

1. Hasin DS, Stinson FS, Ogburn E, Grant BF (2007): Prevalence, correlates, disability, and comorbidity of DSM–IV alcohol abuse and dependence in the United States: Results from the National Epidemiologic Survey on Alcohol and Related Conditions. Arch Gen Psychiatry 64:830–842.
2. Nurnberger JI Jr, Wiegand R, Bucholz K, O'Connor S, Meyer ET, Reich T, *et al.* (2004): A family study of alcohol dependence: Coaggregation of multiple disorders in relatives of alcohol-dependent probands [published correction appears in Arch Gen Psychiatry 2005; 62: 848]. Arch Gen Psychiatry 61:1246–1256.
3. Delker E, Brown Q, Hasin DS (2016): Alcohol consumption in demographic subpopulations: An epidemiologic overview. Alcohol Res 38:7–15.
4. Vasilenko SA, Evans-Polce RJ, Lanza ST (2017): Age trends in rates of substance use disorders across ages 18–90: Differences by gender and race/ethnicity. Drug Alcohol Depend 180:260–264.
5. Smith SM, Stinson FS, Dawson DA, Goldstein R, Huang B, Grant BF (2006): Race/ethnic differences in the prevalence and co-occurrence of substance use disorders and independent mood and anxiety disorders: Results from the National Epidemiologic Survey on Alcohol and Related Conditions [published correction appears in Psychol Med 2008; 38:606]. Psychol Med 36:987–998.
6. Huang B, Grant BF, Dawson DA, Stinson FS, Chou SP, Saha TD, *et al.* (2006): Race-ethnicity and the prevalence and co-occurrence of Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, alcohol and drug use disorders and Axis I and II disorders: United States, 2001 to 2002. Compr Psychiatry 47:252–257.
7. Grant BF, Goldstein RB, Saha TD, Chou SP, Jung J, Zhang H, *et al.* (2015): Epidemiology of DSM-5 alcohol use disorder: Results from the National Epidemiologic Survey on Alcohol and Related Conditions III. JAMA Psychiatry 72:757–766.
8. Grant BF, Chou SP, Saha TD, Pickering RP, Kerridge BT, Ruan WJ, *et al.* (2017): Prevalence of 12-month alcohol use, high-risk drinking, and DSM-IV alcohol use disorder in the United States, 2001–2002 to 2012–2013: Results from the National Epidemiologic Survey on Alcohol and Related Conditions. JAMA Psychiatry 74:911–923.
9. Grant BF, Dawson DA (1997): Age at onset of alcohol use and its association with DSM-IV alcohol abuse and dependence: Results from the National Longitudinal Alcohol Epidemiologic Survey. J Subst Abuse 9:103–110.
10. DeWit DJ, Adlaf EM, Offord DR, Ogborne AC (2000): Age at first alcohol use: A risk factor for the development of alcohol disorders. Am J Psychiatry 157:745–750.
11. Deutsch AR, Slutske WS, Richmond-Rakerd LS, Chernyavskiy P, Heath AC, Martin NG (2013): Causal influence of age at first drink on alcohol involvement in adulthood and its moderation by familial context. J Stud Alcohol Drugs 74:703–713.
12. Sartor CE, Bucholz KK, Nelson EC, Madden PAF, Lynskey MT, Heath AC (2011): Reporting bias in the association between age at first alcohol use and heavy episodic drinking. Alcohol Clin Exp Res 35:1418–1425.
13. Maimaris W, McCambridge JJ (2014): Age of first drinking and adult alcohol problems: Systematic review of prospective cohort studies. J Epidemiol Community Health 68:268–274.
14. Kuntsche E, Rossow I, Engels R, Kuntsche S (2016): Is 'age at first drink' a useful concept in alcohol research and prevention? We doubt that. Addiction 111:957–965.
15. Kuperman S, Chan G, Kramer JR, Bierut L, Bucholz KK, Fox L, *et al.* (2005): Relationship of age of first drink to child behavioral problems and family psychopathology. Alcohol Clin Exp Res 29:1869–1876.
16. Kuperman S, Chan G, Kramer JR, Wetherill L, Bucholz KK, Dick D, *et al.* (2013): A model to determine the likely age of an adolescent's first drink of alcohol. Pediatrics 131:242–248.
17. Kuperman S, Chan G, Kramer J, Wetherill L, Acion L, Edenberg HJ, *et al.* (2017): A GABRA2 polymorphism improves a model for prediction of drinking initiation. Alcohol 63:1–8.
18. Agrawal A, Sartor CE, Lynskey MT, Grant JD, Pergadia ML, Grucza R, *et al.* (2009): Evidence for an interaction between age at first drink and genetic influences on DSM-IV alcohol dependence symptoms. Alcohol Clin Exp Res 33:2047–2056.
19. Kessler RC, Avenevoli S, McLaughlin KA, Green JG, Lakoma MD, Petukhova M, *et al.* (2012): Lifetime co-morbidity of DSM-IV disorders in the US National Comorbidity Survey Replication Adolescent Supplement (NCS-A). Psychol Med 42:1997–2010.
20. Boschloo L, Vogelzangs N, van den Brink W, Smit JH, Veltman DJ, Beekman ATF, Penninx BWJH (2013): Depressive and anxiety disorders predicting first incidence of alcohol use disorders: Results of the Netherlands Study of Depression and Anxiety (NESDA). J Clin Psychiatry 74:1233–1240.
21. Kandel DB, Johnson JG, Bird HR, Weissman MM, Goodman SH, Lahey BB, *et al.* (1999): Psychiatric comorbidity among adolescents with substance use disorders: Findings from the MECA study. J Am Acad Child Adolesc Psychiatry 38:693–699.
22. Bucholz KK, McCutcheon VV, Agrawal A, Dick DM, Hesselbrock VM, Kramer JR, *et al.* (2017): Comparison of parent, peer, psychiatric, and cannabis use influences across stages of offspring alcohol involvement: Evidence from the COGA prospective study. Alcohol Clin Exp Res 41:359–368.
23. Acion L, Kramer J, Liu X, Chan G, Langbehn D, Bucholz K, *et al.* (2019): Reliability and validity of an internalizing symptom scale based on the adolescent and adult Semi-Structured Assessment for the Genetics of Alcoholism (SSAGA). Am J Drug Alcohol Abuse 45:151–160.
24. Cloninger CR (1987): Neurogenetic adaptive mechanisms in alcoholism. Science 236:410–416.
25. Babor TF, Hofmann M, DelBoca FK, Hesselbrock V, Meyer RE, Dolinsky ZS, Rounsaville B (1992): Types of alcoholics, I. Evidence for an empirically derived typology based on indicators of vulnerability and severity. Arch Gen Psychiatry 49:599–608.
26. Kuperman S, Schlosser SS, Kramer JR, Bucholz K, Hesselbrock V, Reich T, Reich W (2001): Developmental sequence from disruptive behavior diagnosis to adolescent alcohol dependence. Am J Psychiatry 158:2022–2026.
27. Groenman AP, Janssen TWP, Oosterlaan J (2017): Childhood psychiatric disorders as risk factor for subsequent substance abuse: A meta-analysis. J Am Acad Child Adolesc Psychiatry 56:556–569.
28. Kendler KS, Myers J (2014): The boundaries of the internalizing and externalizing genetic spectra in men and women. Psychol Med 44:647–655.
29. Wray NR, Goddard ME, Visscher PM (2007): Prediction of individual genetic risk to disease from genome-wide association studies. Genome Res 17:1520–1528.

30. International Schizophrenia Consortium, Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, *et al.* (2009): Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature 460:748–752.
31. Bogdan R, Baranger DAA, Agrawal A (2018): Polygenic risk scores in clinical psychology: Bridging genomic risk to individual differences. Annu Rev Clin Psychol 14:119–157.
32. Fullerton JM, Nurnberger JI (2019): Polygenic risk scores in psychiatry: Will they be useful for clinicians? F1000 Res 8:F1000 Faculty Rev-1293.
33. Khera AV, Chaffin M, Aragam KG, Haas ME, Roselli C, Choi SH, *et al.* (2018): Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. Nat Genet 50:1219–1224.
34. Khera AV, Chaffin M, Wade KH, Zahid S, Brancale J, Xia R, *et al.* (2019): Polygenic prediction of weight and obesity trajectories from birth to adulthood. Cell 177:587–596.e9.
35. So HC, Sham PC (2017): Exploring the predictive power of polygenic scores derived from genome-wide association studies: A study of 10 complex traits. Bioinformatics 33:886–892.
36. Musliner KL, Krebs MD, Albiñana C, Vilhjalmsson B, Agerbo E, Zandi PP, *et al.* (2020): Polygenic risk and progression to bipolar or psychotic disorders among individuals diagnosed with unipolar depression in early life. Am J Psychiatry 177:936–943.
37. Nurnberger JI Jr, Yang Z, Zang Y, Acion L, Bierut L, Bucholz K, *et al.* (2019): Development of alcohol use disorder as a function of age, severity, and comorbidity with externalizing and internalizing disorders in a young adult cohort. J Psychiatr Brain Sci 4:e190016.
38. Kapoor M, Chou YL, Edenberg HJ, Foroud T, Martin NG, Madden PAF, *et al.* (2016): Genome-wide polygenic scores for age at onset of alcohol dependence and association with alcohol-related measures. Transl Psychiatry 6:e761.
39. Barr PB, Ksinan A, Su J, Johnson EC, Meyers JL, Wetherill L, *et al.* (2020): Using polygenic scores for identifying individuals at increased risk of substance use disorders in clinical and population samples. Transl Psychiatry 10:196.
40. Bucholz KK, Cadoret R, Cloninger CR, Dinwiddie SH, Hesselbrock VM, Nurnberger JI Jr, *et al.* (1994): A new, semi-structured psychiatric interview for use in genetic linkage studies: A report on the reliability of the SSAGA. J Stud Alcohol 55:149–158.
41. Lai D, Wetherill L, Bertelsen S, Carey CE, Kamarajan C, Kapoor M, *et al.* (2019): Genome-wide association studies of alcohol dependence, DSM-IV criterion count and individual criteria. Genes Brain Behav 18:e12579.
42. Lai D, Wetherill L, Kapoor M, Johnson EC, Schwandt M, Ramchandani VA, *et al.* (2020): Genome-wide association studies of the self-rating of effects of ethanol (SRE). Addict Biol 25:e12800.
43. American Psychiatric Association (2013): Diagnostic and Statistical Manual for Mental Disorders, 5th ed. Washington, DC: American Psychiatric Publishing.
44. American Psychiatric Association (1994): Diagnostic and Statistical Manual for Mental Disorders, 4th ed. Washington, DC: American Psychiatric Publishing.
45. Kranzler HR, Zhou H, Kember RL, Vickers Smith R, Justice AC, Damrauer S, *et al.* (2019): Genome-wide association study of alcohol consumption and use disorder in 274,424 individuals from multiple populations [published corrections appear in Nat Commun 2019; 10: 2275 and Nat Commun 2019; 10:4050]. Nat Commun 10:1499.
46. Gelernter J, Sun N, Polimanti R, Pietrzak RH, Levey DF, Lu Q, *et al.* (2019): Genome-wide association study of maximum habitual alcohol intake in >140,000 U.S. European and African American veterans yields novel risk loci. Biol Psychiatry 86:365–376.
47. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, *et al.* (2007): PLINK: A tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81:559–575.
48. Ge T, Chen CY, Ni Y, Feng YA, Smoller JW (2019): Polygenic prediction via Bayesian regression and continuous shrinkage priors. Nat Commun 10:1776.
49. Ruan Y, Lin YF, Feng YA, Chen CY, Lam M, Guo Z, *et al.* (2022): Improving polygenic prediction in ancestrally diverse populations. Nat Genet 54:573–580.
50. Hougaard P (1995): Frailty models for survival data. Lifetime Data Anal 1:255–273.
51. SAS version 9.4. (2016). Cary, NC: SAS Institute Inc.
52. Eeltink E, van der Horst MZ, Zinkstok JR, Aalfs CM, Luykx JJ (2021): Polygenic risk scores for genetic counseling in psychiatry: Lessons learned from other fields of medicine. Neurosci Biobehav Rev 121:119–127.
53. Edwards AC, Gillespie NA, Aggen SH, Kendler KS (2013): Assessment of a modified DSM-5 diagnosis of alcohol use disorder in a genetically informative population. Alcohol Clin Exp Res 37:443–451.
54. Walters RK, Polimanti R, Johnson EC, McClintick JN, Adams MJ, Adkins AE, *et al.* (2018): Transancestral GWAS of alcohol dependence reveals common genetic underpinnings with psychiatric disorders. Nat Neurosci 21:1656–1669.
55. Mullins N, Forstner AJ, O'Connell KS, Coombes B, Coleman JRI, Qiao Z, *et al.* (2021): Genome-wide association study of more than 40, 000 bipolar disorder cases provides new insights into the underlying biology. Nat Genet 53:817–829.
56. Reich D (2018): Who We Are and How We Got Here: Ancient DNA and the New Science of the Human Past. New York: Random House.
57. Duncan L, Shen H, Gelaye B, Meijsen J, Ressler K, Feldman M, *et al.* (2019): Analysis of polygenic risk score usage and performance in diverse human populations. Nat Commun 10:3328.
58. Carroll KM, Kiluk BD (2017): Cognitive behavioral interventions for alcohol and drug use disorders: Through the stage model and back again. Psychol Addict Behav 31:847–861.
59. Kranzler HR, Soyka M (2018): Diagnosis and pharmacotherapy of alcohol use disorder: A review. JAMA 320:815–824.
60. McBride WJ, Bell RL, Rodd ZA, Strother WN, Murphy JM (2005): Adolescent alcohol drinking and its long-range consequences. Studies with animal models. Recent Dev Alcohol 17:123–142.
61. McClintick JN, McBride WJ, Bell RL, Ding ZM, Liu Y, Xuei X, Edenberg HJ (2016): Gene expression changes in glutamate and GABA-A receptors, neuropeptides, ion channels, and cholesterol synthesis in the periaqueductal gray following binge-like alcohol drinking by adolescent alcohol-preferring (P) rats. Alcohol Clin Exp Res 40:955–968.
62. McClintick JN, McBride WJ, Bell RL, Ding ZM, Liu Y, Xuei X, Edenberg HJ (2018): Gene expression changes in the ventral hippocampus and medial prefrontal cortex of adolescent alcohol-preferring (P) rats following binge-like alcohol drinking. Alcohol 68:37–47.
63. Koob GF, Le Moal M (2001): Drug addiction, dysregulation of reward, and allostasis. Neuropsychopharmacology 24:97–129.
64. Koob GF (2003): Alcoholism: Allostasis and beyond. Alcohol Clin Exp Res 27:232–243.
65. Nurnberger JI Jr, Austin J, Berrettini WH, Besterman AD, DeLisi LE, Grice DE, *et al.* (2018): What should a psychiatrist know about genetics? Review and recommendations from the Residency Education Committee of the International Society of Psychiatric Genetics. J Clin Psychiatry 80:17nr12046.
66. Wang D, Liu S, Warrell J, Won H, Shi X, Navarro FCP, *et al.* (2018): Comprehensive functional genomic resource and integrative model for the human brain. Science 362:eaat8464.
67. Greden JF, Parikh SV, Rothschild AJ, Thase ME, Dunlop BW, DeBattista C, *et al.* (2019): Impact of pharmacogenomics on clinical outcomes in major depressive disorder in the GUIDED trial: A large, patient- and rater-blinded, randomized, controlled study. J Psychiatr Res 111:59–67.