



Published in final edited form as:

*Addiction*. 2023 September ; 118(9): 1675–1686. doi:10.1111/add.16210.

## Does Polygenic Risk for Substance-Related Traits Predict Ages of Onset and Progression of Symptoms?

Henry R. Kranzler, M.D.<sup>1,2</sup>, Richard Feinn<sup>3</sup>, Heng Xu<sup>1</sup>, Brendan L. Ho<sup>1</sup>, Divya Saini<sup>1</sup>, Olivia R. Nicastro<sup>1</sup>, Anya Jacoby<sup>1</sup>, Sylvanus Toikumo<sup>1,2</sup>, Joel Gelernter<sup>4</sup>, Emily E. Hartwell<sup>1,3</sup>, Rachel L. Kember<sup>1,2</sup>

<sup>1</sup>Department of Psychiatry, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA 19104

<sup>2</sup>Mental Illness Research, Education and Clinical Center, Veterans Integrated Service Network 4, Crescenz Veterans Affairs Medical Center, Philadelphia, PA 19104

<sup>3</sup>Department of Medical Sciences, Frank H. Netter School of Medicine at Quinnipiac University, North Haven, CT 06473

<sup>4</sup>Department of Psychiatry, Yale University School of Medicine, New Haven, CT 06511 and VA Connecticut Healthcare System, West Haven, CT 06516

### Abstract

**Background and aims:** Genetic risk can influence disease progression. We measured the impact of genetic risk for substance use disorders (SUDs) on substance use onset and progression of symptoms.

**Design:** Using findings from genome-wide association studies (GWAS) of alcohol use disorder (AUD), opioid use disorder (OUD), and smoking trajectory (SMK) we calculated polygenic risk scores (PRS) in deeply phenotyped independent samples.

**Setting:** Participants were recruited from 2000 through 2020 from U.S. inpatient or outpatient settings or through advertisements.

**Participants:** 5,692 European-ancestry individuals (EUR) (56.2% male) and 4,918 African-ancestry individuals (AFR) (54.9% male).

**Measurements:** Age of first substance use, regular use, reported problems, and dependence diagnosis, and progression from regular use to onset of problems and dependence for alcohol, opioids, and smoking. We examined the contribution of PRS to each milestone and progression measure.

**Findings:** EUR and males reported earlier onset and shorter progression times than AFR and females, respectively. Among EUR, higher AUD PRS predicted earlier onset and more rapid progression to alcohol-related milestones ( $p < 0.001$ ) and although a stronger moderator of problem onset among females ( $p = 0.017$ ), it was more predictive of the progression to problems among

\*Correspondence to Henry R. Kranzler, M.D., Center for Studies of Addiction, 3535 Market Street, Suite 500, Philadelphia, PA 19104; kranzler@penmedicine.upenn.edu; telephone: 215-746-1943.

males ( $p=0.005$ ). OUD and SMK PRS in EUR also predicted earlier onset of the respective milestones ( $p<0.001$ ). Among AFR, where power is lower, AUD PRS predicted age of regular alcohol use ( $p=0.039$ ) and dependence ( $p=0.001$ ) and progression from regular use to diagnosis ( $p=0.045$ ), while SMK PRS predicted earlier age of initiation ( $p=0.036$ ).

**Conclusions:** Genetic risk for substance use disorders appears to predict substance use milestones and symptom progression among European-ancestry individuals and, to a lesser extent, African-ancestry individuals.

### Keywords

Alcohol Use Disorder; Opioid Use Disorder; Smoking; Polygenic Risk Scores; Age of Onset; Milestones; Progression Measures; Prediction

---

## INTRODUCTION

Substance use disorders (SUDs), characterized by the chronic use of alcohol or other drugs, are common among adolescents and adults and result in clinically significant social impairments and medical and psychiatric disorders.<sup>1</sup> SUDs develop in stages following initial substance use, often progressing across a series of sequential transitions, which can be conceptualized as a continuous trajectory marked by milestones of escalating use or severity.<sup>2</sup> Charting the clinical course of SUDs using these developmental events can help to elucidate the factors that underlie symptom progression, a key example of which is the transition from regular substance use to substance dependence. The timing of milestones could also provide a personalized assessment of an individual's risk of symptom progression<sup>3</sup> and a more precise time window for a targeted intervention aimed at preventing the progression to a subsequent milestone.<sup>4</sup>

The age at which substance-related milestones occur and the rate of progression through them are influenced by both genetic and environmental factors.<sup>4</sup> These risk factors may be reflected in group differences in the initiation of substance use and the progression across substance-related milestones.<sup>5</sup> For example, Black individuals report first consuming alcohol and tobacco products and initiating regular drinking and binge drinking, and use of illicit drugs, later than White individuals.<sup>6,7,8</sup> Because the composition of these groups was not genetically determined, we use the terminology from the primary publications to describe them.

Although Black adolescents had a significantly lower risk of transitioning to regular use of alcohol and illicit drugs,<sup>8</sup> Black adults 30 years of age and older have also been shown to have a more rapid progression from regular alcohol use to regular drinking and intoxication than White adults.<sup>9</sup> These differences occur in the context of a general paucity of studies of population-group differences in substance-related symptom progression.<sup>10</sup> There is a similar dearth of findings on population group differences in age-of-onset and progression measures for opioids, though in one study Black individuals had a more rapid progression to opioid dependence (OD) than White individuals.<sup>11</sup>

Sex may also influence the developmental course of substance-related traits. In monozygotic and dizygotic twins, the initiation of alcohol and tobacco use occurred earlier among males than females with earlier initiation of use associated with a significantly increased likelihood of developing dependence on these substances<sup>4</sup>. In contrast, in another twin study, between-twin comparisons of the rate of progression through alcohol-related milestones showed no overall pattern of sex differences.<sup>2</sup> Thus, the relationship of sex to milestones and progression in SUDs is not fully understood. A controversial question in relation to sex differences is the validity of the phenomenon of telescoping, which posits that despite their later initiation of substance use, women manifest substance-related problems sooner than men.<sup>2</sup>

Genetic risk for SUDs is highly polygenic, involving potentially thousands of individual variants, each accounting for a very small proportion of trait variance. Polygenic risk scores (PRS) sum data from multiple genetic variants and account for greater proportions of trait variance than single polymorphisms.<sup>13</sup> PRS are useful in evaluating the risk for disease progression in diverse medical disorders, including breast cancer<sup>14</sup> and rheumatoid arthritis,<sup>15</sup> and the prediction of sudden death in individuals with coronary disease.<sup>16</sup> A recent large study of white European-ancestry (EUR) individuals in the UK Biobank showed the utility of a prostate cancer genetic risk score for triaging patients in primary care. Men in the highest quintile of risk had a prostate cancer incidence of 8.1% and could be fast-tracked for further investigation, while the incidence among those in the lowest risk quintile was <1% and they could more safely avoid invasive investigation.<sup>17</sup> Thus, there currently are clinical applications of PRS to differentiate individuals based on their genetic risk for a disease.

For SUDs, a PRS for alcohol dependence (AD) was associated with the progression from onset of regular drinking to AD in a EUR sample.<sup>18</sup> The ability to quantify the genetic risk of symptom progression could help to identify individuals at highest risk of developing more severe milestones (e.g., alcohol-related problems or AD) and who could benefit most from intensive interventions.

Here, we examined the association of an AUD PRS with the progression from onset of regular drinking to onset of AD in a EUR sample selected using genetic principal components analysis (PCA). We extended the findings from a prior report<sup>18</sup> to include additional milestones and measures of progression of alcohol-related symptoms and conduct parallel analyses of opioid-related and smoking-related traits. Finally, we refine our understanding of the effects of sex and population group differences on these features by studying males and females and EUR and African-ancestry (AFR) participants, all of whom are well represented in our sample, which was recruited and deeply phenotyped for genetic studies of SUDs.<sup>19</sup>

## METHODS

### Discovery samples

We used genome-wide association study (GWAS) summary statistics for alcohol use disorder (AUD),<sup>20</sup> opioid use disorder (OUD),<sup>21</sup> and smoking trajectories (SMK)<sup>22</sup> as

discovery samples for calculating PRS in the Yale-Penn sample. GWAS provides a measure of effect for the association of each single nucleotide polymorphism with the respective phenotype. We chose discovery GWAS of SUDs because we thought that they would be most informative of the progression measures, which reflect the latency from regular substance use to either problematic substance use or a SUD diagnosis. The three discovery samples are described in detail in Supplemental Methods. Briefly, the discovery sample for AUD comprised 296,989 EUR and 80,764 AFR from the Million Veteran Program (MVP).<sup>20</sup> The discovery sample for OUD<sup>21</sup> included 302,585 EUR and 88,498 AFR from MVP. The SMK discovery sample<sup>22</sup> included 209,915 EUR and 54,867 AFR from MVP.

To compare the power among the discovery samples, we calculated the genomic inflation factor ( $\lambda_{gc}$ ). In the absence of population structure,  $\lambda_{gc}$  is a function of sample size and the number of causal variants and thus reflects power to detect significant SNP-trait associations.<sup>23</sup> We calculated  $\lambda_{gc}$  using the expected median of SNP test statistics from each set of GWAS summary statistics, with greater  $\lambda_{gc}$  denoting higher predictive power.

### Target sample

We calculated PRS in the Yale-Penn sample, which was recruited at five U.S. sites for genetic studies of dependence on cocaine, opioids, and alcohol.<sup>19</sup> The study was approved by the institutional review board at each site and participants gave written informed consent for data collection. Cases were identified through addiction treatment facilities, inpatient and outpatient psychiatric services, and posters and advertisements in local media. Although some individuals were recruited for family studies, data from family members are excluded in analyses presented here. Unaffected controls were recruited from non-psychiatric medical settings and through advertisements. The designation of ancestry in the target sample, as in the discovery samples, used principal components analysis of GWAS data.

### Semi-Structured Assessment for Drug Dependence and Alcoholism (SSADDA)

The SSADDA is a comprehensive psychiatric interview that comprises 24 modules assessing the physical, psychological, social, and psychiatric manifestations of SUDs, psychiatric disorders, and environmental covariates considered likely to have an impact on SUDs. The SSADDA's semi-structured format, accompanied by the rigorous training and quality control procedures used in the Yale-Penn sample,<sup>24</sup> allow a carefully trained non-clinician interviewer to assess diagnostic criteria and disorders and their ages of onset, to yield DSM-IV diagnoses of AD, OD, and nicotine dependence (ND). Ascertainment of ages of onset of the different milestones is done using questions that elicit estimated ages (e.g., of initiation of substance and regular substance use) and clustering of criteria within a 12-month period (for DSM-IV diagnoses).

### Genotyping, imputation, and polygenic risk scores

Yale-Penn samples were genotyped in three batches using the Illumina HumanOmni1-Quad microarray, the Illumina HumanCoreExome array, or the Illumina Multi-Ethnic Global array. Genotype data were filtered for individual call rates and excessive heterozygosity using PLINK v1.9 and were imputed using the Michigan Imputation Server<sup>25</sup> and the Haplotype Reference Consortium Panel.<sup>26</sup>

Using effect size estimates from the discovery samples, PRS were calculated in the target sample for AUD,<sup>20</sup> OUD,<sup>21</sup> and SMK<sup>22</sup> using Polygenic Risk Scores–Continuous Shrinkage software (PRS-CS)<sup>27</sup> and the 1000 Genomes Project phase 3 AFR and EUR samples for estimates of linkage disequilibrium. Global shrinkage parameters were obtained from each set of summary statistics by the PRS-CS package and effective sample sizes were used to calculate the final PRS. We used matched, genetically determined, ancestral summary statistics (e.g., an AFR GWAS for AUD was used to calculate AUD PRS in AFR Yale-Penn individuals). We report findings using effective sample sizes.

**Statistical Analysis**—We used PROC PHREG in SAS v9.4 to run Cox proportional hazards models for each of three substances (alcohol, opioids, and smoking) and the following four measures: age of first use, age of regular use, age of first bringing up problems with a healthcare professional, and age of diagnosis of the disorder (AD, OD, and ND). Using these milestones, we also examined two measures of progression: time from age of regular use to age of first bringing up problems and time from age of regular use to age of diagnosis. Analyses were conducted separately for EUR and AFR. Similarly, analyses were conducted separately for females and males within each ancestry group.

All models included the respective PRS, age, and the first 10 principal ancestral components as covariates. In analyses that did not examine sex as a factor, sex was included as a covariate. The models for the progression outcomes also included the age of regular use as a covariate. We tested the proportionality assumption of the Cox models by including an interaction term comprising the age of regular use and each of the two progression measures. Whereas the interaction terms in all progression models were significant, reflecting a lack of proportionality, we assessed the impact of age of onset of regular use on the two progression outcomes by analyzing the effects separately for participants with early onset (< 18 years) and late onset (>18 years) of regular use.<sup>18</sup>

For individuals who reported never having experienced a specific event, data were censored for the event in the Cox models and the age at interview was substituted for the missing age of onset. Thus, the age-of-onset outcomes for all substances have the same sample size. For the two progression outcomes, only individuals who reported regular use were included in the analysis, as censoring that age of onset would distort the analysis. Thus, the sample sizes for the two progression outcomes vary by substance.

For effect sizes, we report the hazard ratio (HR) with 95% confidence intervals, reflecting the change in hazard for a one-standard-deviation increase in PRS, with a HR>1.0 reflecting a greater likelihood that the event will occur as PRS increases. We also report an incremental pseudo-R<sup>2</sup> for the Cox models,<sup>28</sup> where models including the PRS are compared to models that include only the other covariates. We report p-values and adjusted p-values using the Hommel correction for multiple testing, with adjustments made separately by population group and substance. We chose the Hommel correction over a Bonferroni correction to increase power given the lack of independence of the milestones and progression measures.<sup>29</sup> We consider PRS as a significant predictor when the p-value, adjusted for multiple comparisons, is <0.05.

Whereas the analysis described here was not pre-registered the results should be considered exploratory.

## RESULTS

### Genomic inflation

Calculation of  $\lambda_{gc}$  showed that for all three traits EUR samples had better predictive power than AFR samples. For AUD, the  $\lambda_{gc}$  was 1.100 for EUR and 1.034 for AFR, while for OUD the respective values were 1.112 and 1.028, and for SMK they were 1.336 and 1.094.

### Rates of endorsement of substance use milestones

The target sample comprises 16,715 individuals, with genome-wide genotype data available for 10,610 individuals, including 5,692 EUR (56.2% male) and 4,918 AFR (54.9% male) (Table 1). More than 95% of both population groups reported ever having used alcohol and over 80% reported ever having used alcohol regularly. Lifetime opioid use was less common among AFR (33%) than EUR (51%), as was regular opioid use (21% vs. 41%). In both groups, nearly 90% of individuals reported having smoked more than 100 cigarettes lifetime, with over 65% endorsing regular smoking. In both AFR and EUR, the event endorsed least commonly for all three substances was having brought up problems to a healthcare professional, which ranged from 15%–31% in AFR and 31%–38% in EUR.

A majority (51.7%) of EUR and 53.5% of AFR met criteria for a DSM-IV diagnosis of AD. Similarly, 55.4% of EUR and 54.5% of AFR had an ND diagnosis. The prevalence of OD was lower in both groups: 39% among EUR 18% among AFR. The high rate of dependence diagnoses in the Yale-Penn sample reflects its ascertainment for studies of addiction genetics.<sup>19</sup> The prevalence of all three dependence diagnoses is lower and the onset of all milestones later among women and AFR than among men and EUR, respectively. In addition, EUR and males had shorter progression times for all substances than AFR and women, respectively (Table 2).

### Effects of polygenic risk scores on alcohol-related milestones and symptom progression

In all figures, the PRS values are divided into tertiles, which are labeled low, medium, and high. Table 3 and Figure 1 (upper two panels) show the Cox regression model of the effects of the AUD PRS on alcohol-related milestones by population group. Significant associations reflect a younger age of onset and a shorter latency between milestones as a function of increasing PRS. Among EUR, AUD PRS was a significant predictor (adjusted p-value [ $p_{adj}$ ] $<0.001$ ) of all four milestones, with HRs ranging from 1.06 (age of first use and age of regular use) to 1.19 (age at which alcohol-related problems were brought up to a healthcare professional). The AUD PRS was also a significant predictor of the progression from age of regular alcohol use both to bringing up alcohol-related problems (HR=1.14,  $p_{adj}<0.001$ ) and age of AD diagnosis (HR=1.10,  $p_{adj}<0.001$ ) (Figure 2).

Among AFR, the AUD PRS significantly predicted both age of regular alcohol use (HR=1.05,  $p_{adj}=0.039$ ) and age of AD diagnosis (HR=1.09,  $p_{adj}=0.001$ ) (Table 3). It was



also a significant predictor of the progression from age of regular alcohol use to age of AD diagnosis (HR=1.06,  $p_{adj}=0.045$ ) (Figure 2).

In analyses stratified by age-of-onset of regular drinking (Supplemental Table 1), in both population groups, early onset of regular drinking (< 18 years) was associated with more rapid progression both to age of first reported alcohol-related problems and to age of initial AD diagnosis. Among EUR, the AUD PRS, irrespective of the age of onset of regular drinking, significantly predicted the progression to first reported alcohol-related problems and to an AD diagnosis (HRs = 1.09–1.18). Among AFR, the AUD PRS was positively associated with both progression measures (HRs=1.02–1.11), though the effect was significant only for progression to an AD diagnosis and only among individuals with a late onset of regular drinking (>18 years).

A sex-stratified analysis among EUR (Table 4 and Figure 3) showed that the association of AUD PRS with milestones was greater among women (HRs=1.06–1.30) than men (HRs=1.06–1.15). There was also a significant interaction effect of sex by AUD PRS on the age at which alcohol-related problems were raised with a health professional ( $p_{adj}=0.0165$ ). The genetic risk for AUD was a stronger moderator in females than males despite men having brought up problems earlier than women. This is evidenced by greater separation between survival curves for low, medium, and high PRS tertiles among women than men (Figure 3). Similarly, although the effects of the AUD PRS on the two measures of progression were significant in both sexes, the effects were greater among women (HR=1.27 for progression to reporting problems and 1.14 for progression to AD) than men (HR=1.09 for both). The progression from regular drinking to first reported alcohol-related problems in EUR differed significantly by sex ( $p_{adj}=0.0054$ ), with men bringing up alcohol-related problems sooner after beginning regular drinking than women. As can be seen in Figure 4, there is overlap between the survival curves of females in the highest PRS tertile and males in the lowest PRS tertile.

Among AFR, the only significant effect of the alcohol PRS when stratified by sex was on age of AD diagnosis. The effect was significant in both men (HR=1.09) and women (HR=1.11) (Supplemental Table 2). The only milestone on which the alcohol PRS had a significant effect was age of AD diagnosis (HR=1.09 in men and 1.11 in women; Supplemental Table 3). When stratified on both sex and age of onset of regular drinking (Supplemental Table 4), the only significant effect of the AUD PRS on progression measures among AFR was the time from onset of regular drinking to an AD diagnosis among women with late onset of regular drinking.

### Opioid-related milestones and progression

Among EUR, the OUD PRS significantly predicted all four milestones (HRs=1.14–1.19) but neither of the progression outcomes. Among AFR, the OUD PRS was not associated with any opioid-related milestones or progression outcomes (Table 5). Figure 1 (middle two panels) shows the difference between population groups in the survival curves for age of first OD diagnosis as a function of OUD PRS strata.

When stratified by sex, among EUR, the OUD PRS was significantly associated in both sexes with all four of the milestones but neither of the progression measures (Supplemental Table 3).

Among AFR, sex-stratified analyses of the effects of the OUD PRS show that the only milestone that was significant was age of onset of opioid use, an effect limited to women (Supplemental Table 4).

### Smoking-related milestones and progression

Among EUR, as with OUD, the SMK PRS significantly predicted all four age-of-onset measures (HRs=1.15–1.25), while among AFR it predicted only age of first use (HR=1.05) (Table 6). Survival curves for the age of diagnosis of ND by PRS strata (Figure 1, lower two panels), show the population differences on this key milestone. In neither population group was SMK PRS a significant predictor of the progression from regular smoking to reporting smoking-related problems or to ND.

Among EUR, sex-stratified analyses showed that, in both sexes, the SMK PRS is associated with all four milestones and, among females, with the progression from regular smoking to onset of ND (Supplemental Table 5). Similar analyses among AFR (Supplemental Table 6) yielded only one significant effect on smoking-related traits: among males SMK PRS predicted the age of smoking onset.

## DISCUSSION

There is growing interest in the clinical utility of PRS for identifying individuals at high risk for a variety of disorders. In addition to case identification, estimates of genetic risk are increasingly being used to predict disease progression. We examined the effects of polygenic risk on the age of onset of substance-related traits and the progression from regular substance use to substance-related problems and dependence. We used summary statistics from large GWAS of AUD,<sup>20</sup> OUD,<sup>21</sup> and SMK<sup>22</sup> to calculate PRS in a sample of deeply phenotyped individuals with alcohol or drug use disorders or screened controls. We compared these effects by population group and by sex within population group. While the effects of the PRS were statistically significant for many of the outcomes the effect size was generally small. As a measure of explained variance above the model covariates the incremental pseudo-r-square of the PRSs were in the range of 0.2% (HR=1.05) to 2.4% (HR=1.25).

Our most consistent findings were for alcohol-related traits. Among EUR, an AUD PRS predicted the age of all four alcohol-related milestones and both measures of progression from the onset of regular drinking. The findings replicate a previous observation that an AD PRS predicted the progression from onset of regular drinking to AD diagnosis in a sample of EUR.<sup>18</sup> We also extended the analysis to AFR, among whom an AUD PRS significantly predicted the age of regular alcohol use, age of an AD diagnosis, and the progression from regular use to an AD diagnosis.



Stratifying the analyses on the age of onset of regular drinking did not substantially alter the findings. Sex-stratified analyses showed that for some outcomes, the effects of PRS were greater among women than men. These findings could be relevant to the phenomenon of telescoping, in which women who, despite initiating substance use at a later age than men have been reported to have a more rapid progression to developing problems and presenting for treatment than men.<sup>30,31</sup> However, prior findings supporting telescoping are inconsistent.<sup>12,32</sup> Here we found that among EUR the AUD PRS predicted a significantly shorter time from onset of regular alcohol use to bringing up alcohol-related problems among men than women, while among AFR the effect of the AUD PRS on the progression from regular drinking to onset of AD was comparable for men and women.

Among EUR, there were also robust effects of OUD and SMK PRS on opioid- and tobacco-related milestones, respectively. However, neither PRS predicted the progression from age of onset of regular use either to bringing up problems related to these substances or dependence diagnoses. Among AFR, the only opioid-related milestone that was significantly associated with OUD PRS was the age of onset of opioid use among women. In this population group, the SMK PRS was associated with an earlier age of smoking initiation, a finding that was significant in men only.

Overall, we found consistently greater associations of PRS with substance-related milestones and symptom progression in EUR than AFR, attributable to the greater predictive value of the EUR summary statistics for all three substances, evidenced by greater genomic inflation factor values among EUR. Further, despite comparable numbers of EUR and AFR individuals with AD and ND in the Yale-Penn (i.e., target) sample, the number of AFR subjects in the target sample that endorsed opioid-related milestones and that met criteria for OD was about one-third the number among EUR. Thus, for opioids, differences in the target sample also likely contributed to the population-group difference in PRS effects. Differences in the sizes of the discovery and target samples by population group underscore the need, particularly in non-EUR populations, for larger GWAS samples and additional deeply phenotyped samples for more granular studies of genetic risk for substance use milestones and progression.

Among EUR, the effects of the AUD PRS were more consistent and robust than were the OUD or SMK PRS. This could be due to there being approximately one-third fewer participants in the SMK GWAS than either the AUD or OUD GWAS and approximately one-quarter fewer Yale-Penn participants with an OD diagnosis than either an AD or ND diagnosis. Substance-specific differences have also been shown to exist in symptom progression. National survey data showed that the cumulative probability of progressing to dependence was 67.5% for nicotine users and 22.7% for alcohol users.<sup>33</sup> A prospective study of adolescents showed that the shortest progression times (i.e., greatest addictive liabilities) were seen with opioids; tobacco and alcohol had the lowest liabilities.<sup>34</sup> Despite these findings, there are countervailing biological effects. For example, the rate of absorption of nicotine from smoking is much higher than gastrointestinal absorption of alcohol given the extensive surface area of pulmonary alveoli.<sup>35,36</sup> Thus, pharmacologic or other features specific to individual substances could add to or interact with genetic risk for dependence on them.

Although here we focus principally on genetic risk, environmental factors are also relevant to the age at which substance-related milestones occur.<sup>37</sup> Among both AFR and EUR, we found the lowest HRs for the age of first alcohol use (1.02 and 1.06, respectively). This is consistent with the notion that the initiation of substance use is strongly influenced by social and environmental factors, whereas the progression from first use to heavy use and from heavy use to problematic use or dependence is influenced more by neurobiological, including genetic, factors.<sup>35</sup> This is further supported by the finding that the age at first alcohol use is only modestly genetically correlated with AD ( $r_g=18-29\%$ ),<sup>38</sup> while for age of onset of regular drinking and AD the genetic correlation is moderate ( $r_g=0.54$ ).<sup>39</sup>

This study has limitations. First, we conducted analyses in only two population groups—EUR and AFR—as the other population groups in both the discovery and target samples are not large enough to support analyses. Secondly, the sample was recruited over 20 years at five sites in the eastern United States through multiple studies, thus it is not possible to specify the exact sources of recruitment. Whereas it is not a population sample, the generalizability of the findings is limited. Because all three discovery samples were from the MVP, which is preponderantly male, the effect size estimates from the GWAS could bias the PRS and their associations with symptoms, particularly in the context of existing sex differences. Third, we used DSM-IV substance dependence diagnoses in the target sample to ensure consistency across the substances, as we lacked some criteria required for a DSM-5 tobacco use disorder diagnosis. Fourth, we used a trajectory phenotype in the discovery GWAS for smoking, as it was the largest available GWAS for smoking in AFR. Nonetheless, the phenotype differs from the ICD-9/10 codes used in the AUD and OUD GWAS. Whereas the trajectories are a probabilistic categorization rather than a binary diagnosis, the trajectory-based groups are potentially more heterogeneous than AUD or OUD cases and controls. This variability could have diminished the association of the SMK PRS with smoking-related milestones or latency outcomes. The moderate or greater genetic correlations with widely used smoking-related traits (e.g., smoking initiation ( $r_g=0.52$ ), smoking cessation ( $r_g=0.85$ ), cigarettes per day ( $r_g=0.44$ ), and time from waking to the first cigarette ( $r_g=-0.49$ )) support the validity of the SMK trait.<sup>22</sup> Finally, the proportions of substance dependence diagnoses overall and by population in the target sample reflect the strategy used to recruit the sample, and therefore no conclusions may be drawn about these proportions *per se*. Finally, although effects of the PRSs were statistically significant for many of the outcomes the effect size was generally small, with the largest incremental pseudo- $R^2$  of the PRS being 2.7%.

Larger discovery samples are needed to increase the predictive power in both population groups and account for greater variance in the progression to problematic substance use. The goal of this effort is to augment non-genetic predictors with PRS to identify individuals at greatest risk to experience progression to more serious substance-related consequences and permit secondary preventive efforts.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Competing Interests:

Dr. Kranzler is a member of advisory boards for Dicerna Pharmaceuticals, Sophrosyne Pharmaceuticals, and Enthion Pharmaceuticals; a consultant to Sobrera Pharmaceuticals; the recipient of research funding and medication supplies for an investigator-initiated study from Alkermes; and a member of the American Society of Clinical Psychopharmacology's Alcohol Clinical Trials Initiative, which was supported in the last three years by Alkermes, Dicerna, Ethypharm, Lundbeck, Mitsubishi, and Otsuka. Drs. Gelernter and Kranzler hold U.S. Patent 10,900,082: Genotype-guided Dosing of Opioid Receptor Agonists, 26 Jan. 2021. The other authors have no disclosures to make.

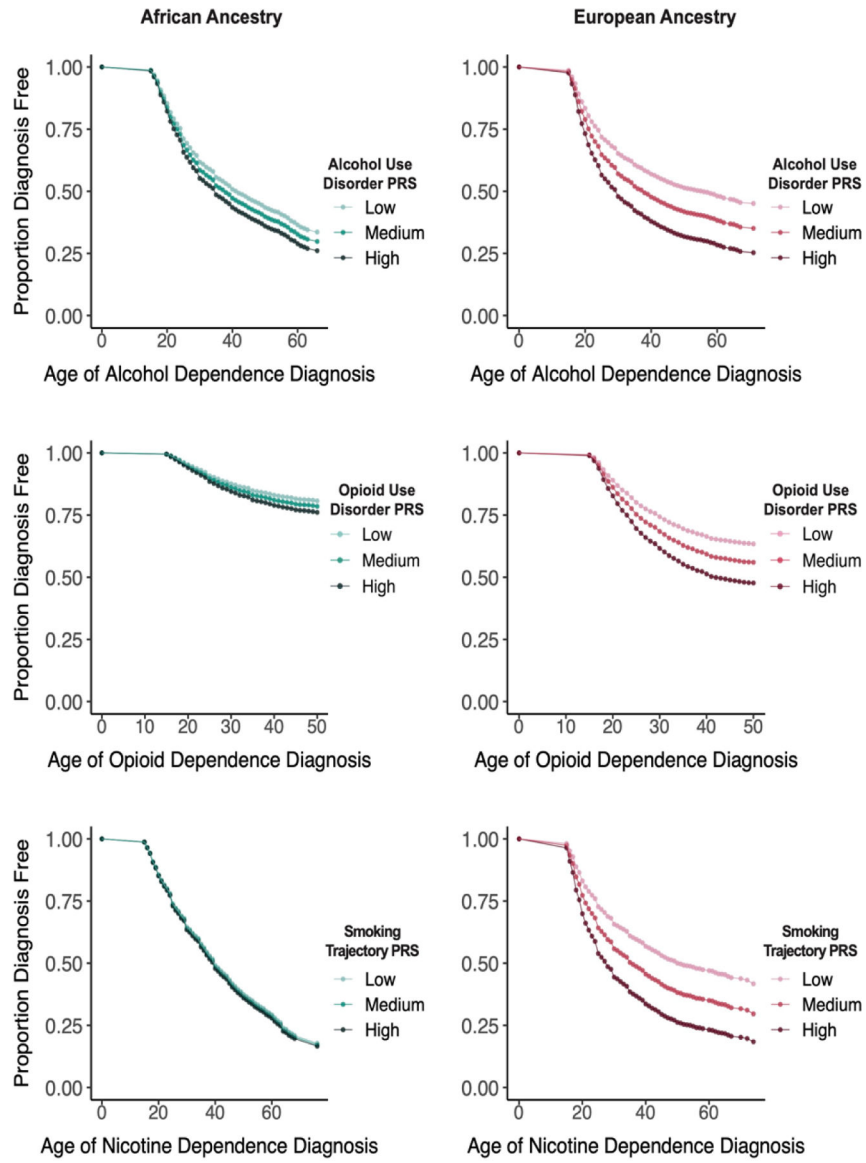
Supported by the Veterans Integrated Service Network 4 Mental Illness Research, Education and Clinical Center and NIH grants P30 DA046345, R01 AA026364, and K01 AA028292 (to RLK).

## REFERENCES

1. McLellan AT, Koob GF & Volkow ND Preaddiction—A Missing Concept for Treating Substance Use Disorders. *JAMA Psychiatry* 79, 749 (2022). [PubMed: 35793096]
2. Deutsch AR et al. From alcohol initiation to tolerance to problems: Discordant twin modeling of a developmental process. *Dev. Psychopathol* 29, 845–861 (2017). [PubMed: 27417028]
3. Rhee SH et al. Genetic and Environmental Influences on Substance Initiation, Use, and Problem Use in Adolescents. *Arch. Gen. Psychiatry* 60, 1256 (2003). [PubMed: 14662558]
4. Huggett SB, Hatoum AS, Hewitt JK & Stallings MC The Speed of Progression to Tobacco and Alcohol Dependence: A Twin Study. *Behav. Genet* 48, 109–124 (2018). [PubMed: 29427143]
5. Lewis B, Hoffman L, Garcia CC & Nixon SJ Race and socioeconomic status in substance use progression and treatment entry. *J. Ethn. Subst. Abuse* 17, 150–166 (2018). [PubMed: 28846065]
6. Malone PS, Northrup TF, Masyn KE, Lamis DA & Lamont AE Initiation and persistence of alcohol use in United States Black, Hispanic, and White male and female youth. *Addict. Behav* 37, 299–305 (2012). [PubMed: 22136874]
7. Pérez A, Kuk AE, Bluestein MA, Sia HMS & Chen B Age of Initiation of Dual Tobacco Use and Binge Drinking among Youth (12–17 Years Old): Findings from the Population Assessment of Tobacco and Health (PATH) Study. *Int. J. Environ. Res. Public Health* 18, 12985 (2021). [PubMed: 34948595]
8. Burstein M Use and Abuse of Alcohol and Illicit Drugs in US Adolescents: Results of the National Comorbidity Survey–Adolescent Supplement. *Arch. Gen. Psychiatry* 69, 390 (2012). [PubMed: 22474107]
9. Johnson PB, Richter L, Kleber HD, McLellan AT & Carise D Telescoping of Drinking-Related Behaviors: Gender, Racial/Ethnic, and Age Comparisons. *Subst. Use Misuse* 40, 1139–1151 (2005). [PubMed: 16040374]
10. Sartor CE et al. Progression from First Drink, First Intoxication, and Regular Drinking to Alcohol Use Disorder: A Comparison of African American and European American Youth. *Alcohol. Clin. Exp. Res* 40, 1515–1523 (2016). [PubMed: 27256613]
11. Sartor CE, Kranzler HR & Gelernter J Rate of progression from first use to dependence on cocaine or opioids: A cross-substance examination of associated demographic, psychiatric, and childhood risk factors. *Addict. Behav* 39, 473–479 (2014). [PubMed: 24238782]
12. Alvanzo AAH et al. Race/ethnicity and sex differences in progression from drinking initiation to the development of alcohol dependence. *Drug Alcohol Depend.* 118, 375–382 (2011). [PubMed: 21652154]
13. Lewis CM & Vassos E Polygenic risk scores: from research tools to clinical instruments. *Genome Med.* 12, 44 (2020). [PubMed: 32423490]
14. Lakeman IMM et al. The predictive ability of the 313 variant-based polygenic risk score for contralateral breast cancer risk prediction in women of European ancestry with a heterozygous BRCA1 or BRCA2 pathogenic variant. *Genet. Med* 23, 1726–1737 (2021). [PubMed: 34113011]
15. Honda S et al. Association of Polygenic Risk Scores With Radiographic Progression in Patients With Rheumatoid Arthritis. *Arthritis Rheumatol.* 74, 791–800 (2022). [PubMed: 35048562]

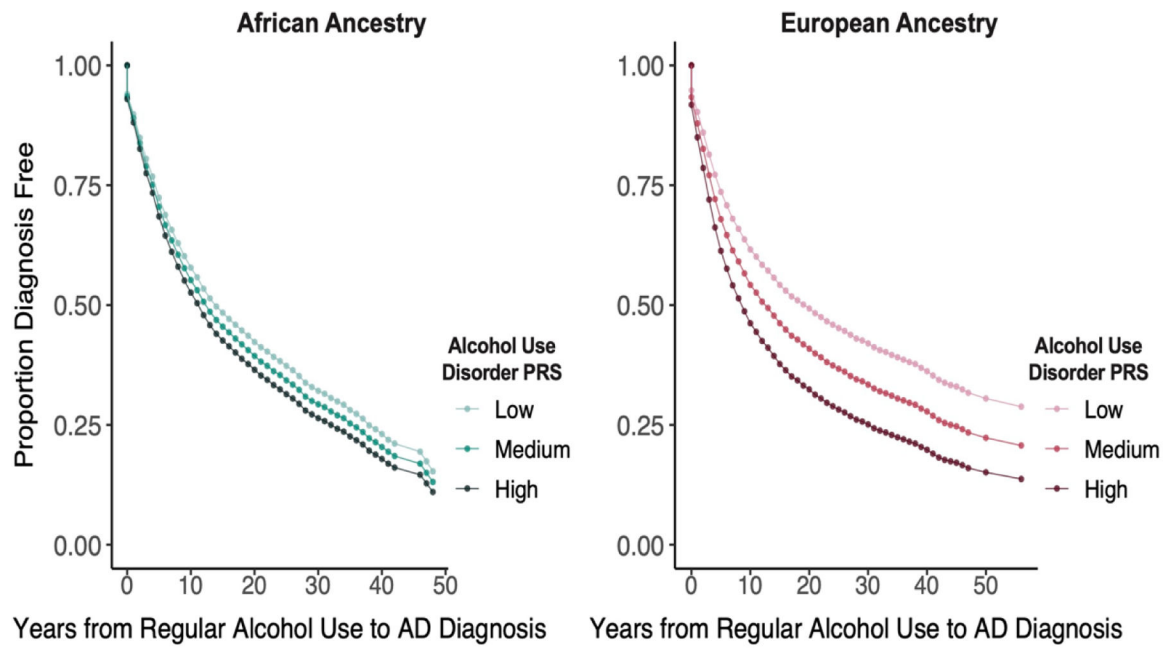
16. Sandhu RK et al. Polygenic Risk Score Predicts Sudden Death in Patients With Coronary Disease and Preserved Systolic Function. *J. Am. Coll. Cardiol* 80, 873–883 (2022). [PubMed: 36007985]
17. Green HD et al. Applying a genetic risk score for prostate cancer to men with lower urinary tract symptoms in primary care to predict prostate cancer diagnosis: a cohort study in the UK Biobank. *Br. J. Cancer* (2022) doi:10.1038/s41416-022-01918-z.
18. Yeung EW et al. Effects of genetic risk for alcohol dependence and onset of regular drinking on the progression to alcohol dependence: A polygenic risk score approach. *Drug Alcohol Depend.* 230, 109117 (2022). [PubMed: 34844060]
19. Kember RL et al. Phenome-wide association analysis of substance use disorders in a deeply phenotyped sample. *Biol Psychiatry.* 2022 Aug 18:S0006–3223(22)01515–3. doi: 10.1016/j.biopsych.2022.08.010. Epub ahead of print.
20. Kember RL et al. Genetic underpinnings of the transition from alcohol consumption to alcohol use disorder: shared and unique genetic architectures in a cross-ancestry sample. <http://medrxiv.org/lookup/doi/10.1101/2021.09.08.21263302> (2021) doi:10.1101/2021.09.08.21263302.
21. Kember RL et al. Cross-ancestry meta-analysis of opioid use disorder uncovers novel loci with predominant effects in brain regions associated with addiction. *Nat. Neurosci* (2022) doi:10.1038/s41593-022-01160-z.
22. Xu K et al. Genome-wide association study of smoking trajectory and meta-analysis of smoking status in 842,000 individuals. *Nat. Commun* 11, 5302 (2020). [PubMed: 33082346]
23. Yang J et al. Genomic inflation factors under polygenic inheritance. *Eur. J. Hum. Genet* 19, 807–812 (2011). [PubMed: 21407268]
24. Pierucci-Lagha A et al. Diagnostic reliability of the Semi-structured Assessment for Drug Dependence and Alcoholism (SSADDA). *Drug Alcohol Depend.* 80, 303–312 (2005). [PubMed: 15896927]
25. Das S et al. Next-generation genotype imputation service and methods. *Nat. Genet* 48, 1284–1287 (2016). [PubMed: 27571263]
26. The 1000 Genomes Project Consortium et al. . A global reference for human genetic variation. *Nature* 526, 68–74 (2015). [PubMed: 26432245]
27. Ge T, Chen C-Y, Ni Y, Feng Y-CA & Smoller JW Polygenic prediction via Bayesian regression and continuous shrinkage priors. *Nat. Commun* 10, 1776 (2019). [PubMed: 30992449]
28. Allison PD *Survival Analysis Using the SAS System: A Practical Guide*, 2nd ed. Cary, NC: SAS Institute Inc. (2010)
29. Blakesley RE et al. Comparisons of methods for multiple hypothesis testing in neuropsychological research. *Neuropsychology.* 23, 255–64 (2009). [PubMed: 19254098]
30. Hernandez-Avila CA, Rounsaville BJ & Kranzler HR Opioid-, cannabis- and alcohol-dependent women show more rapid progression to substance abuse treatment. *Drug Alcohol Depend.* 74, 265–272 (2004). [PubMed: 15194204]
31. Cheng HG, Chandra M, Alcover KC & Anthony JC Rapid transition from drinking to alcohol dependence among adolescent and young-adult newly incident drinkers in the United States, 2002–2013. *Drug Alcohol Depend.* 168, 61–68 (2016). [PubMed: 27620346]
32. Lewis B & Nixon SJ Characterizing Gender Differences in Treatment Seekers. *Alcohol. Clin. Exp. Res* 38, 275–284 (2014). [PubMed: 23930695]
33. Lopez-Quintero C et al. Probability and predictors of transition from first use to dependence on nicotine, alcohol, cannabis, and cocaine: Results of the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC). *Drug Alcohol Depend.* 115, 120–130 (2011). [PubMed: 21145178]
34. Ridenour TA, Lanza ST, Donny EC & Clark DB Different lengths of times for progressions in adolescent substance involvement. *Addict. Behav* 31, 962–983 (2006). [PubMed: 16677774]
35. Fattinger K, Benowitz NL, Jones RT & Verotta D Nasal mucosal versus gastrointestinal absorption of nasally administered cocaine. *Eur. J. Clin. Pharmacol* 56, 305–310 (2000). [PubMed: 10954344]
36. Norberg A, Jones AW, Hahn RG & Gabrielsson JL Role of Variability in Explaining Ethanol Pharmacokinetics: Research and Forensic Applications. *Clin. Pharmacokinet* 42, 1–31 (2003). [PubMed: 12489977]

37. Kendler KS, Schmitt E, Aggen SH & Prescott CA Genetic and Environmental Influences on Alcohol, Caffeine, Cannabis, and Nicotine Use From Early Adolescence to Middle Adulthood. *ARCH GEN PSYCHIATRY* 65, 9 (2008).
38. Prescott CA & Kendler KS Age at First Drink and Risk for Alcoholism: A Noncausal Association. *Alcohol. Clin. Exp. Res* 23, 101–107 (1999). [PubMed: 10029209]
39. Grant JD et al. Adolescent alcohol use is a risk factor for adult alcohol and drug dependence: evidence from a twin design. *Psychol. Med* 36, 109–118 (2006). [PubMed: 16194286]

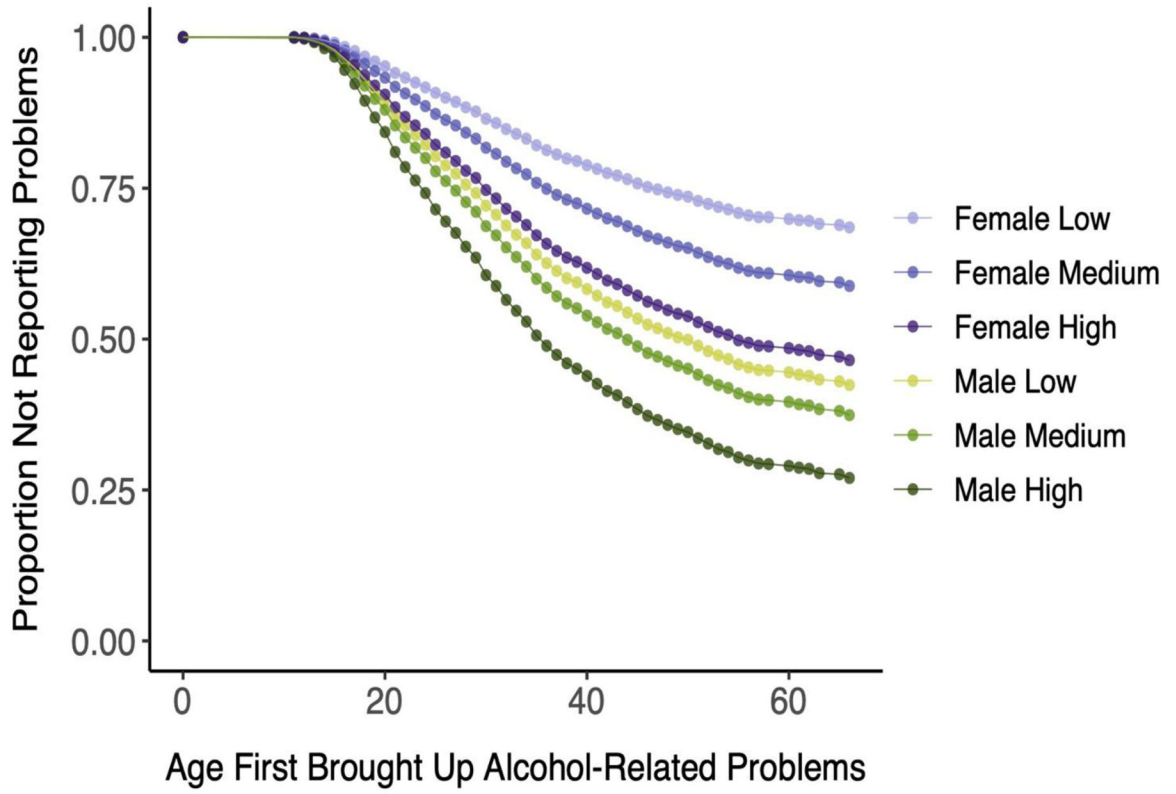


**Figure 1:** Age of Alcohol, Opioid, and Nicotine Dependence Diagnoses by European or African Ancestry and Low, Medium, and High Polygenic Risk Scores for Alcohol Use Disorder, Opioid Use Disorder, and Smoking Trajectory, Respectively

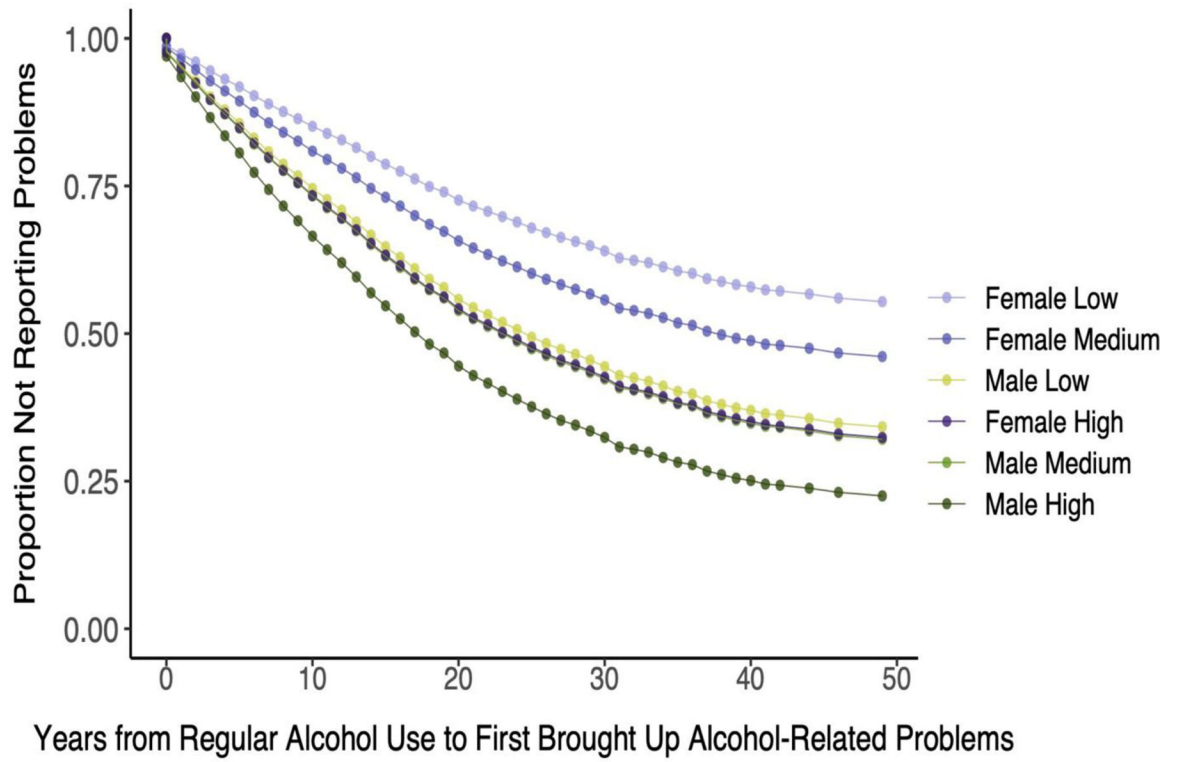




**Figure 2:**  
Years from Regular Alcohol Use to Alcohol Dependence Diagnosis by European or African Ancestry and Low, Medium, and High Alcohol Use Disorder Polygenic Risk Scores



**Figure 3:**  
Age First Brought Up Alcohol-Related Problems to a Healthcare Professional Among European-Ancestry Individuals by Sex Alcohol Use Disorder Polygenic Risk Score



**Figure 4:** Years from First Regular Alcohol Use to First Brought Up Alcohol-Related Problems to a Healthcare Professional Among European-Ancestry Individuals by Sex Alcohol Use Disorder Polygenic Risk Score. Note that the lines for Female High and Male Medium are nearly wholly overlapping and require careful inspection to differentiate them.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 1:**

Prevalence of Substance Use Disorder Milestones by Ancestral Group and Sex

	African Ancestry (n=4,918)			European Ancestry (n=5,692)		
	All (n=4,918)	Male (n=2,704)	Female (n=2,214)	All (n=5,692)	Male (n=3,200)	Female (n=2,492)
<i>Alcohol</i>						
First use	4,679 (95.1%)	2,582 (95.5%)	2,097 (94.7%)	5,476 (96.2%)	3,079 (96.2%)	2,397 (96.2%)
Regular use	4,034 (82.1%)	2,390 (88.4%)	1,644 (74.3%)	4,792 (84.2%)	2,848 (89.0%)	1,944 (78.0%)
Brought up problems	1,878 (38.2%)	1,219 (45.1%)	659 (29.8%)	2,174 (38.2%)	1,479 (46.2%)	695 (27.9%)
Dependence diagnosis	2,631 (53.5%)	1,693 (62.6%)	938 (42.4%)	2,945 (51.7%)	1,942 (60.7%)	1,003 (40.2%)
<i>Opioids</i>						
First use	1,642 (33.4%)	1,109 (41.0%)	533 (24.1%)	2,913 (51.2%)	1,927 (60.2%)	986 (39.6%)
Regular use	1,042 (21.2%)	682 (25.2%)	360 (16.3%)	2,307 (40.5%)	1,518 (47.4%)	789 (31.7%)
Brought up problems	759 (15.4%)	476 (17.6%)	283 (12.8%)	2,039 (35.8%)	1,338 (41.8%)	701 (28.1%)
Dependence diagnosis	889 (18.1%)	573 (21.2%)	316 (14.3%)	2,202 (38.7%)	1,437 (44.9%)	765 (30.7%)
<i>Nicotine</i>						
First use	4,269 (86.8%)	2,448 (90.5%)	1,821 (82.2%)	4,998 (87.8%)	2,895 (90.5%)	2,103 (84.4%)
Regular use	3,324 (67.6%)	2,015 (74.5%)	1,309 (59.1%)	3,709 (65.2%)	2,269 (70.9%)	1,440 (57.8%)
Brought up problems	1,017 (20.7%)	608 (22.5%)	409 (18.5%)	1,751 (30.8%)	1,073 (33.5%)	678 (27.2%)
Dependence diagnosis	2,679 (54.5%)	1,601 (59.2%)	1,078 (48.7%)	3,156 (55.4%)	1,966 (61.4%)	1,190 (47.8%)

**Table 2:** Age (Median [Q1–Q3]) of Substance Use Disorder Milestones by Ancestral Group and Sex

	African Ancestry			European Ancestry		
	All 42 [35–48]	Male 43 [36–49]	Female 41 [33–47]	All 39 [28–49]	Male 39 [28–48]	Female 40 [28–49]
<i>Alcohol</i>						
First use	16 [13–18]	15 [13–15]	16 [14–19]	15 [13–17]	14 [12–16]	16 [14–18]
Regular use	18 [16–21]	17 [15–20]	19 [16–22]	17 [15–19]	17 [15–18]	18 [16–21]
Brought up problems	30 [24–37]	30 [24–38]	30 [24–36]	26 [20–34]	25 [20–33]	27 [20–35]
Dependence diagnosis	23 [19–29]	23 [19–28]	25 [20–30]	21 [18–26]	21 [18–25]	21 [18–27]
<i>Opioids</i>						
First use	21 [17–26]	20 [17–25]	22 [18–28]	18 [16–23]	18 [16–22]	19 [16–25]
Regular use	22 [18–28]	21 [17–27]	23 [19–29]	21 [18–26]	20 [17–25]	22 [18–28]
Brought up problems	30 [24–36]	30 [24–37]	30 [25–36]	25 [21–32]	25 [21–32]	26 [21–33]
Dependence diagnosis	25 [20–31]	25 [19–31]	25 [20–31]	23 [19–29]	22 [19–28]	24 [19–31]
<i>Nicotine</i>						
First use	14 [12–17]	14 [12–17]	15 [12–17]	14 [12–16]	13 [12–16]	14 [12–16]
Regular use	18 [15–21]	18 [15–21]	18 [15–21]	16 [15–19]	16 [15–19]	16 [14–18]
Brought up problems	37 [29–44]	38 [30–45]	36 [29–43]	30 [23–38]	30 [23–38]	29 [22–38]
Dependence diagnosis	25 [20–32]	25 [19–31]	25 [20–33]	21 [18–27]	21 [18–26]	21 [18–28]

Note: The number of subjects in each cell is shown in Table 1.

**Table 3:** Effect of Alcohol Use Disorder Polygenic Risk Score on Age of Onset and Progression of Alcohol-Related Measures by Population Group

Outcome Measure	European Ancestry						African Ancestry							
	B	SE	HR	95% CI	P-Value	Adjusted P-Value	R <sup>2</sup>	B	SE	HR	95% CI	P-Value	Adjusted P-Value	R <sup>2</sup>
Age of First Use	0.057	0.014	1.06	1.03–1.09	<0.001	<0.001	0.3%	0.015	0.017	1.02	0.98–1.05	0.382	0.382	0.0%
Age of Regular Use	0.060	0.015	1.06	1.03–1.10	<0.001	<0.001	0.3%	0.046	0.018	1.05	1.01–1.09	0.013	<b>0.039</b>	0.1%
Age First Brought Up Problems	0.173	0.023	1.19	1.14–1.24	<0.001	<0.001	1.0%	0.065	0.027	1.07	1.01–1.13	0.019	0.057	0.1%
Age of Diagnosis of AD	0.136	0.019	1.15	1.10–1.19	<0.001	<0.001	0.8%	0.086	0.023	1.09	1.04–1.14	0.0002	<b>0.001</b>	0.2%
Regular Use to Brought Up Problems	0.133	0.024	1.14	1.09–1.20	<0.001	<0.001	0.6%	0.054	0.029	1.04	0.98–1.09	0.236	0.382	0.0%
Regular Use to AD Diagnosis	0.096	0.021	1.10	1.06–1.15	<0.001	<0.001	0.4%	0.061	0.025	1.06	1.01–1.12	0.015	<b>0.045</b>	0.1%

Note: All models include sex, age, and the first 10 ancestry principal components. Progression models (Regular Use to ...) also include Age of Regular Use as a covariate and the interaction between time and Age of Regular Use because the proportional hazards assumption was violated. Events=number of individuals meeting the milestone or progression event, SE=standard error, HR=hazard ratio, 95%CI=95% confidence interval, R<sup>2</sup>=incremental pseudo R-squared, AD=alcohol dependence



**Table 4:**

Effect of Alcohol Use Disorder Polygenic Risk Score in Cox Proportional Hazard Models of Alcohol-Related Time-to-Event and Progression Measures by Sex among Individuals of European Ancestry

Outcome Measure	Male						Female							
	B	SE	HR	95% CI	P-Value	Adjusted P-Value	R <sup>2</sup>	B	SE	HR	95% CI	P-Value	Adjusted P-Value	R <sup>2</sup>
Age of First Use	0.057	0.019	1.06	1.02–1.10	0.003	<i>0.005</i>	0.3%	0.060	0.022	1.06	1.02–1.10	0.006	<i>0.006</i>	0.3%
Age of Regular Use	0.056	0.020	1.06	1.02–1.10	0.005	<i>0.005</i>	0.3%	0.070	0.024	1.07	1.02–1.13	0.004	<i>0.006</i>	0.4%
Age First Brought Up Problems	0.136	0.027	1.15	1.09–1.21	<0.0001	<i>0.0005</i>	0.8%	0.264	0.041	1.30	1.20–1.41	<0.0001	<i>0.0004</i>	1.7%
Age of AD Diagnosis	0.123	0.024	1.13	1.08–1.19	<0.0001	<i>0.0005</i>	0.8%	0.171	0.034	1.19	1.11–1.27	<0.0001	<i>0.0004</i>	1.0%
Regular Use to Brought Up Problems	0.089	0.029	1.09	1.03–1.16	0.002	<i>0.0045</i>	0.3%	0.234	0.044	1.27	1.16–1.38	<0.0001	<i>0.0004</i>	1.5%
Regular Use to AD Diagnosis	0.085	0.026	1.09	1.04–1.15	0.001	<i>0.004</i>	0.4%	0.127	0.039	1.14	1.05–1.23	0.001	<i>0.003</i>	0.6%

Note: All models include sex, age, and the first 10 ancestry principal components. Progression models (Regular Use to ...) also include Age of Regular Use as a covariate and the interaction between time and Age of Regular Use because the proportional hazards assumption was violated. Events=number of individuals meeting the milestone or progression event, SE=standard error, HR=hazard ratio, 95%CI=95% confidence interval, R<sup>2</sup> = incremental pseudo R-squared, AD=alcohol dependence

**Table 5:**

Effect of Opioid Use Disorder Polygenic Risk Score in Cox Proportional Hazard Models of Opioid-Related Time-to-Event and Progression Measures by Population Group

Outcome Measure	European Ancestry						African Ancestry							
	B	SE	HR	95% CI	P-Value	Adjusted P-Value	R <sup>2</sup>	B	SE	HR	95% CI	P-Value	Adjusted P-Value	R <sup>2</sup>
Age of First Use	0.135	0.019	1.14	1.10–1.19	<0.0001	<b>0.0002</b>	0.9%	0.069	0.027	1.07	1.02–1.13	0.010	0.060	0.1%
Age of Regular Use	0.150	0.021	1.16	1.11–1.21	<0.0001	<b>0.0002</b>	0.9%	0.034	0.033	1.03	0.97–1.10	0.312	0.936	0.0%
Age First Brought Up Problems	0.177	0.023	1.19	1.14–1.25	<0.0001	<b>0.0002</b>	1.1%	0.037	0.039	1.04	0.96–1.12	0.345	0.936	1.0%
Age of OD Diagnosis	0.173	0.022	1.19	1.14–1.24	<0.0001	<b>0.0002</b>	1.1%	0.052	0.036	1.05	0.98–1.13	0.150	0.600	0.0%
Regular Use to Problems	0.054	0.027	1.06	1.00–1.11	0.045	0.090	0.2%	0.003	0.047	1.00	0.92–1.10	0.942	0.942	0.0%
Regular Use to OD Diagnosis	0.035	0.034	1.04	0.97–1.11	0.316	0.316	0.0%	-0.028	0.058	0.97	0.87–1.09	0.624	0.942	0.0%

Note: All models include sex, age, and the first 10 ancestry principal components. Progression models (Reg Use to ...) also include Age of Regular Use as a covariate and the interaction between time and Age of Regular Use because the proportional hazards assumption was violated. Events=number of individuals meeting the milestone or progression event, SE=standard error, HR=hazard ratio, 95%CI=95% confidence interval, R<sup>2</sup> = incremental pseudo R-squared, OD=opioid dependence

**Table 6:** Effect of Smoking Polygenic Risk Score in Cox Proportional Hazard Models of Smoking-Related Time-to-Event and Progression Measures by Population Group

Outcome Measure	European Ancestry							African Ancestry						
	B	SE	HR	95% CI	P-Value	Adjusted P-Value	R <sup>2</sup>	B	SE	HR	95% CI	P-Value	Adjusted P-Value	R <sup>2</sup>
Age of First Use	0.139	0.015	1.15	1.12–1.18	<0.0001	<b>0.0002</b>	1.5%	0.044	0.016	1.05	1.01–1.08	0.006	<b>0.036</b>	0.2%
Age of Regular Use	0.217	0.017	1.24	1.20–1.29	<0.0001	<b>0.0002</b>	2.7%	0.027	0.018	1.03	0.99–1.07	0.125	0.446	0.0%
Age First Brought Up Problems	0.220	0.025	1.25	1.19–1.31	<0.0001	<b>0.0002</b>	1.3%	0.039	0.032	1.04	0.98–1.11	0.223	0.669	0.0%
Age of ND Diagnosis	0.223	0.019	1.25	1.21–1.30	<0.0001	<b>0.0002</b>	2.4%	0.026	0.020	1.03	0.99–1.07	0.188	0.564	0.0%
Regular Use to Problems	0.039	0.026	1.04	0.99–1.09	0.137	0.137	0.0%	0.023	0.033	1.02	0.96–1.09	0.479	0.958	0.0%
Regular Use to ND Diagnosis	0.041	0.020	1.04	1.00–1.08	0.040	0.080	0.1%	-0.000	0.021	1.00	0.96–1.40	0.980	0.980	0.0%

Note: All models include sex, age, and the first 10 ancestry principal components. Progression models (Reg Use to ...) also include Age of Regular Use as a covariate and the interaction between time and Age of Regular Use because the proportional hazards assumption was violated. Events=number of individuals meeting the milestone or progression event, SE=standard error, HR=hazard ratio, 95%CI=95% confidence interval, R<sup>2</sup> = incremental pseudo R-squared, ND=nicotine dependence