Predicting Substance Use Disorders: A Multifactorial Risk Index Combining Clinical, Environmental, and Genetic Risk Factors

Peter B. Barr [1,2], Morgan N. Driver [3], Sally I-Chun Kuo [4], Mallory Stephenson [5], Fazil Aliev [5,6], Richard Karlsson Linnér [7], Andrey P. Anokhin [8], Kathleen Bucholz [8], Grace Chan [9,10], Howard J. Edenberg [11, 12], Alexis C. Edwards [5], Meredith W. Francis [8], K. Paige Harden [13, 14], Chella Kamarajan [1], Jaakko Kaprio [15,] Sivan Kinreich [1], John Kramer [10,] Samuel Kuperman [10], Antti Latvala [16], Jacquelyn L. Meyers [1, 2], Abraham A. Palmer [17, 18], Martin H. Plawecki [19], Bernice Porjesz [1], Richard J. Rose [20], Marc A. Schuckit [17], Jessica E. Salvatore [4], and Danielle M. Dick [3, 6]

[1] Department of Psychiatry and Behavioral Sciences, SUNY Downstate Health Sciences University, Brooklyn, NY, USA.
[2] VA New York Harbor Healthcare System, Brooklyn, NY, USA.
[3] Department of Human and Molecular Genetics, Virginia Commonwealth University, Richmond, VA, USA.
[4] Department of Psychiatry, Robert Wood Johnson Medical School, Rutgers University, Piscataway, NJ, USA.
[5] Virginia Institute for Psychiatric and Behavioral Genetics, Department of Psychiatry, Virginia Commonwealth University, Richmond, VA, USA.
[6] Department of Psychology, Virginia Commonwealth University, Richmond, VA, USA.
[7] Department of Economics, Leiden University, Leiden, The Netherlands.
[8] Department of Psychiatry, School of Medicine, Washington University in St. Louis, St Louis, MO, USA.
[9] Department of Psychiatry, School of Medicine, University of Connecticut, Farmington, CT, USA.
[10] Department of Psychiatry, Carver College of Medicine, University of Iowa, Iowa City, IA, USA
[11] Department of Medical and Molecular Genetics, School of Medicine, Indiana University, Indianapolis, IN, USA
[12] Department of Biochemistry and Molecular Biology, School of Medicine, Indiana University, Indianapolis, IN, USA
[13] Department of Psychology, University of Texas at Austin, Austin, TX, USA.
[14] Population Research Center, University of Texas at Austin, Austin, TX, USA.
[15] Institute for Molecular Medicine Finland, University of Helsinki, Helsinki, Finland
[16] Institute of Criminology and Legal Policy, University of Helsinki, Helsinki, Finland
[17] Department of Psychiatry, University of California San Diego, La Jolla, CA, USA.
[18] Institute for Genomic Medicine, University of California San Diego, La Jolla, CA, USA.
[19] Department of Psychiatry, School of Medicine, Indiana University, Indianapolis, IN, USA
[20] Department of Psychological and Brain Sciences, Indiana University, Bloomington, IN, USA

Corresponding author:
Peter B. Barr
Department of Psychiatry and Behavioral Sciences
SUNY Downstate Health Sciences University
450 Clarkson Ave, MSC 1203, Brooklyn, NY 11203
e-mail: peter.barr@downstate.edu.

ABSTRACT

Importance: Substance use disorders (SUDs) incur serious social and personal costs. Screening techniques that identify persons at risk before problems develop can improve prevention efforts.

Objective: To examine whether models that include polygenic scores (PGS) and a clinical/ environmental/risk index (CERI) are able to identify individuals as having a lifetime SUD.

Design: We tested the predictive power of PGS and CERI for lifetime diagnosis of DSM-IV substance dependence using four longitudinal cohorts.

Setting: The study included four samples: 1) the National Longitudinal Study of Adolescent to Adult Health (Add Health); 2) the Avon Longitudinal Study of Parents and Children (ALSPAC); 3) the Collaborative Study on the Genetics of Alcoholism (COGA); and 4) the Finnish Twin Cohort Study (FinnTwin12) for a combined sample of N = 15,134.

Participants: Participants in Add Health (N_{EUR} = 4,855; N_{AFR} = 1,605) and COGA (N_{EUR} = 1,878; N_{AFR} = 870) included individuals of both European and African ancestries. Participants in ALSPAC (N_{EUR} = 4,733) and FinnTwin12 (N_{EUR} = 1,193) were limited to individuals of European ancestries.

Exposures: A clinical/environmental risk index (CERI) composed of ten items and PGS for phenotypes with strong genetic overlap with SUDs (drinks per week, problematic alcohol use, externalizing problems, major depressive disorder, and schizophrenia).

Main Outcomes: Meeting lifetime criteria for DSM-IV: 1) alcohol dependence, 2) drug dependence, and 3) any substance dependence (alcohol, other drug, or nicotine).

Results: In the models containing the five PGS and CERI, the CERI was associated with all three outcomes (ORs = 1.35 – 1.64). PGS for problematic alcohol use was associated with alcohol dependence (OR = 1.14), PGS for externalizing was associated with drug dependence (OR = 1.14) and both were associated with any substance dependence (ORs = 1.11 – 1.19).

Including the five PGS, CERI, and covariates explained 6% - 13% of the variance in SUDs. Those in the top 10% of CERI and PGS had relative risk ratios of 3.82 - 9.13 for each SUD relative to the bottom 90%.

Conclusions and Relevance: Measures of clinical, environmental, and genetic, risk demonstrate modest ability to distinguish between affected and unaffected individuals for alcohol, drug, and any substance use disorders in young adulthood. These tools will continue to advance as we identify additional risk factors that can be incorporated into clinical practice and deliver on the goal of precision medicine.

INTRODUCTION

Substance use disorders (SUD) are associated with substantial costs to affected individuals, their families, and society. Approximately 73,000 Americans died as the result of an opioid overdose in the 12 months preceding May of 2021[1]. In 2016, alcohol use contributed 4.2% to the global disease burden and other drug use contributed 1.3%[2]. Excessive alcohol use and illicit drug use cost the United States an annual $250 billion[3] and $190 billion[4] respectively. Given the substantial human and economic costs of misuse and disorders, developing methods for identifying persons at heightened risk for SUD is a major public health concern.

There is evidence that targeting those at high risk is an effective strategy for interventions in substance misuse[5]. Ideally, screening tools for SUD risk would include measures of social, clinical, and genetic risk factors, each of which impacts the development of substance use disorders[6–10]. Prior research using an index of clinical and environmental risk factors (e.g. childhood disadvantage, family history of SUD, childhood conduct problems, childhood depression, early exposure to substances, frequent use during adolescence) found this risk index useful in identifying those with persistent SUDs[11]. Recent analyses evaluating the potential for polygenic scores (PGS), which aggregate risk for a trait across the genome using information from genome-wide association studies (GWAS), show current PGS do poorly in identifying individuals affected by SUDs[12]. Using individual genetic variants and clinical features outperformed clinical features alone[13], but individual variants have limited predictive power. Overall, there is limited work on the combined genetic, environmental, and clinical risk factors for SUDs. For other medical conditions, such as melanoma[14] or ischemic stroke[15], combining clinical and genetic risk factors showed improvement over models using individual risk factors.

In the current study, we examine the joint association of early life clinical/environmental risk factors and polygenic scores with SUDs in early adulthood across four longitudinal cohorts: the National Longitudinal Study of Adolescent to Adult Health (Add Health); the Avon

Longitudinal Study of Parents and Children (ALSPAC); the Collaborative Study on the Genetics of Alcoholism (COGA); and the youngest cohort of the Finnish Twin Cohort Study (FinnTwin12). These samples include population-based cohorts from three countries (United States, England, and Finland) and a predominantly high-risk sample. Two of the samples (COGA and Add Health) are ancestrally diverse. We focus on early adulthood as this is a critical period for the development and onset of SUDs[16]. Our research questions are guided by the understanding that risk factors for SUDs range from broader social conditions to genetic influences, and we must acknowledge all of these influences if we hope to deliver on the goal of an equitable precision medicine.

METHODS

Samples

Add Health is a nationally-representative longitudinal study of adolescents followed into adulthood in the United States[17]. Data has been collected from Wave I when respondents were between 11-18 (1994-1995) to Wave V (2016-2018) when respondents were 35-42. The current analysis uses data from Waves I, II, and Wave IV.

ALSPAC is an ongoing, longitudinal population-based study of a birth cohort in the (former) Avon district of South West England[18–21]. Pregnant female residents with an expected date of delivery between April 1, 1991, and December 31, 1992, were invited to participate (N = 14,541 pregnant women, 80% of those eligible). This analysis uses data up to the age 22 assessment (details of all the data that is available through a searchable, web-based tool: http://www.bristol.ac.uk/alspac/researchers/our-data/).

COGA is a family-based sample consisting of alcohol dependent individuals (identified through treatment centers across the United States), their extended families, and community controls (N ~16,000). We use a prospective sample of offspring of the original COGA

participants (baseline ages 12-22, N = 3,573) and have been assessed biennially since recruitment (2004-2019)[22].

FinnTwin12 is a population-based study of Finnish twins born 1983–1987 identified through Finland's Central Population Registry. A total of 2,705 families (87% of all identified) returned the initial family questionnaire late in the year in which twins reached age 11[23]. Twins were invited to participate in follow-up surveys when they were ages 14, 17, and approximately 22. Each cohort includes a wide range of social, behavioral, and phenotypic data measured across the life course. The measures of SUDs were derived from the corresponding young adult phases of data collection in each cohort (mean ages ~ 22 - 28). A full description of each sample is presented in the supplementary information.

Measures

Lifetime Diagnosis of Substance Use Disorder.

We constructed measures of lifetime SUD diagnosis based on the data that were available in each of the samples, defined as meeting criteria for three, non-mutually exclusive categories of substance dependence: 1) alcohol dependence; 2) drug dependence (inclusive of cannabis, cocaine, opioids, sedatives, and other substances); and 3) any substance dependence (alcohol, nicotine, or illicit drug). Our analyses focused on DSM-IV as this diagnostic system was most consistently used across all samples. There were two exceptions: (1) The ALSPAC included measures of likely alcohol dependence based on scores from the Alcohol Use Disorder Identification Test (AUDIT); and (2) for all samples, nicotine dependence was measured using a cutoff of 7 or higher on the Fagerstrom Test for Nicotine Dependence[24]. Where possible, we drew measures of substance dependence from the young adult waves of data collection to try and maintain temporal ordering between SUD diagnoses and measured risk factors.

Clinical/Environmental Risk Index

We created a clinical/environmental risk index (CERI) considering a variety of established risk factors for SUD (Table 1). The CERI included ten validated early life risk factors associated with later development of SUDs, including: low childhood socioeconomic status (SES), family history of SUD, early initiation of substance use, childhood internalizing problems, childhood externalizing problems, frequent drinking in adolescence, frequent smoking in adolescence, frequent cannabis use in adolescence, peer substance use, and exposure to trauma/traumatic experiences[11,25,26]. We dichotomized each risk factor (present vs not present) and summed them into an index for each person ranging from 0 to 10, providing a single measure of aggregate risk. A full list of how each measure is defined within each of the samples is available in the supplementary information.

Polygenic Scores

We constructed polygenic scores (PGS), which are aggregate measures of the number of risk alleles individuals carry weighted by effect sizes from GWAS summary statistics. We derived PGS from five recent GWAS of SUDs and comorbid conditions including: 1) externalizing problems (EXT) [27]; 2) major depressive disorder (MDD)[28]; 3) problematic alcohol use[29] (ALCP); 4) alcohol consumption (drinks per week, ALCC)[30]; and 5) schizophrenia (SCZ)[31]. We focused on these PGS because: 1) SUD show strong genetic overlap with other externalizing[32–34], internalizing[28,35], and psychotic disorders[29,36,37]; 2) both shared and substance-specific genetic risk are associated with later SUDs[38–40]; and 3) substance use and SUDs have only partial genetic overlap[41,42].

To date, GWAS have been overwhelmingly limited to individuals of European ancestries[43,44]. Importantly, polygenic scores derived from GWAS of one ancestry do not always transport into other ancestral populations[45,46]. We therefore used PRS-CSx[47], a new method that combines information from well powered GWAS (typically of European ancestries) and

ancestrally matched GWAS to improve predictive power of polygenic scores in the African ancestry samples from Add Health and COGA. PRS-CSx integrates GWAS summary statistics across multiple input populations and employs a Bayesian approach to correct GWAS summary statistics for the non-independence of SNPs in linkage disequilibrium (LD) with one another[47]. See the supplementary information for a detailed description.

Analytic Strategy

We pooled all the data for analysis using an integrative data analytic approach[48]. In order to account for population stratification in the polygenic scores, we first regressed each PGS on age, age^2, sex, sex*age, sex*age^2, and the first 10 ancestral PC's and saved the standardized residuals (Z-scores) as our PGS. Next, we pooled all of the data, including cohort as a fixed effect for each of the six cohorts (4 samples, of which two were split by ancestry) in subsequent analyses. Age of last observation and sex were also included as covariates. Finally, because COGA and FT12 included a large number of related individuals, we adjusted for familial clustering using cluster-robust standard errors[49].

We estimated a series of nested logistic regression models with the pooled data: 1) a baseline model (sex, age, and cohort), 2) a genetic risk model (baseline + PGS), 3) a clinical/environmental risk model (baseline + CERI), and 4) a combined risk model (baseline + PGS + CERI) to assess the predictive accuracy of each model using the difference in pseudo-R^2 ($\Delta Pseudo\text{-}R^2$) [50], between the baseline and corresponding models. We also calculated the discriminatory power of the combined model using the area under the curve (AUC) from a receiver operating characteristic (ROC) curve. Our analytic strategy was preregistered on the Open Science Framework (https://osf.io/etbw8). Deviations from the preregistration are described and outlined in the supplementary information.

RESULTS

Table 2 presents the descriptive statistics and sample sizes across each of the cohorts and ancestries. All of the cohorts had similar proportions of females (~51% - 56%). The mean ages ranged from ~22 to ~29 years of age. The COGA cohorts (both European and African ancestries) reported the highest rates of SUD. These elevated rates are expected given the nature of the COGA sample (highly selected for substance use disorders). Add Health participants generally had higher rates of SUD than ALSPAC or FinnTwin12, but lower than COGA. Finally, ALSPAC and FinnTwin12 reported similar levels of alcohol, drug, and any substance dependence. COGA participants reported higher mean values on the CERI. The remainder of the cohorts report relatively similar rates of exposure to risk factors.

Table 3 presents the results from the *PGS only*, *CERI only*, and *combined* models for each outcome. Two of the five PGS were associated with the SUD outcomes in the *PGS only* model. PGS for externalizing (EXT OR = 1.19 – 1.35) and problematic alcohol use (ALCP OR = 1.08 – 1.15) were both associated with all SUD outcomes, though the association between ALCP and drug dependence was not significant after correcting for multiple testing. In the *CERI only* models, the CERI was consistently associated across each of the categories of SUD (ORs = 1.37 – 1.67). When we combined the PGS and risk index into the same model, the CERI remained significant across SUDs and was largely unchanged in magnitude (ORs = 1.35 – 1.64). EXT remained associated with drug dependence (OR = 1.14), ALCP remained associated alcohol dependence (OR = 1.14), and both remained associated with any substance dependence diagnosis (ORs = 1.11 – 1.19). Overall, the combined model explained 6.0%, 13.2%, and 12.8% of the variance in alcohol dependence, drug dependence, and any substance dependence, respectively.

To ensure the robustness of our results, we also performed: 1) a leave-one-out (LOO) analysis; 2) sex-stratified analyses, and 3) ancestry-specific analysis. The results from the LOO

and sex-stratified analyses were mostly identical to those from the full model. Results in the European ancestry cohorts mirrored the main results, while only the MDD PGS was associated with any of the SUDs in the African ancestry cohorts (see eTables 1-3).

Figure 1 (Panel A) presents the raw prevalence for each outcome across counts of the CERI. Those reporting 3 or more risk factors for drug dependence and 5 or more risk factors for both alcohol and any substance dependence, have a prevalence above lifetime prevalence estimates from nationally representative samples[51]. Panel B depicts the prevalence of each category of SUD across several mutually exclusive categories: 1) those in the bottom 90% of both the CERI and all PGS (averaged across the five scores); 2) those in the top 10% of the CERI but the bottom 90% of the PGS distribution; 3) those in the top 10% of the PGS distribution and the bottom 90% of the CERI; and 4) those in the top 10% of both PGS and the CERI. There is a steady increase in risk across those with elevated genetic risk, clinical/environmental risk, and both. Those in the top 10% of both PGS and CERI had the highest prevalence of each of the SUDs, though the error bars overlap with the estimates from those in the top 10% of the risk index, alone. Compared to those in the bottom 90% on both, those in the to the top 10% of both have a relative risk of 3.82 (95% CI = 3.53, 4.14) for alcohol dependence, 9.13 (95% CI = 8.22, 10.15) for drug dependence, and 4.17 (95% CI = 3.63, 4.80) for any substance dependence.

Finally, we consider the AUC for the combined model for each of the SUD categories. Figure 2 presents the ROC curves for the full (CERI and PGS) and baseline (covariates only) models for each SUD category. The full AUC in each model ranged from 0.74 for alcohol dependence, 0.86 for drug dependence, and 0.78 for any substance dependence. The overall change in AUC (from the baseline to the full model) that we achieve when adding the risk index and PGS was modest (ΔAUC = 0.05 – 0.08), and this improvement was due in large part to the explanatory power of the CERI.

DISCUSSION

Substance use disorders remain a serious threat to public health. Developing screening protocols that can identify those at greater risk of developing problems has the potential to improve prevention efforts. Prior work on early, targeted interventions for substance misuse suggests that those at highest risk stand to benefit the most from these prevention efforts[5]. In the current analysis, we examined the combination of clinical, environmental, and genetic risk factors for determining who is more likely to develop an SUD in early adulthood. We used previously validated measures of environmental and clinical risk[11,25,26] and polygenic scores for externalizing problems[27], major depressive disorder[28], problematic alcohol use[29], alcohol consumption[30], and schizophrenia[31]. The combination of genetic and social-environmental measures was significantly associated with the development of SUDs. The overall prediction was best for drug dependence, followed by any substance dependence, and lastly, alcohol dependence.

As expected, the CERI was the strongest association with each outcome. The proportion of those meeting criteria for each SUD in persons with 3 or more risk factors for drug dependence, and 5 or more risk factors for alcohol or any substance dependence, surpassed lifetime estimates of these SUD, respectively. Overall, the discriminatory power of the full model (AUC = .74 - .86) was similar to AUC estimates published in the original paper from which many of the risk index items were derived (AUC ~ 0.80)[11]. Interestingly, this risk index was originally developed for identifying persons with persistent SUD through early mid-life (~age 40). In the current analysis we demonstrated that the CERI in conjunction with demographic covariates and polygenic scores does equally well for those who meet criteria for any SUD by young adulthood.

In terms of measures of genetic risk, the overall predictive power of the PGS alone was in the range of 1.3 – 2.4%. Only the PGS for externalizing problems and problematic alcohol use were consistently associated with SUD outcomes. The externalizing PGS was associated

drug dependence, the problematic alcohol use was associated with alcohol dependence, and both were associate with any substance dependence. Overall, these results support prior evidence that genetic risk for SUD consists of a both shared and substance-specific variance[27,34,40]. Interestingly, even though the effect sizes for the PGS were attenuated in the model, the PGS for externalizing and problematic alcohol use remained significantly associated even when we included the risk index in the model. Since the risk index also included many of the phenotypes each of the PGS measured (e.g., childhood conduct disorder for externalizing, childhood depression for major depressive disorder; and frequent alcohol use for alcohol consumption), part of this attenuation is likely due to the inclusion of the actual phenotypes through which risk for some of these disorders is expressed. PGS are also confounded, at least in part, by environmental variance[52] and the reduction in effect sizes could be accounting for some of that confounding. Regardless, PGS may add information beyond well-known risk factors, which could eventually prove useful in settings when patients are hesitant to reveal information about potentially stigmatized behaviors, or when information about patients' early life exposures is unavailable. Additionally, PGS may provide information about risk before behaviors manifest, allowing for earlier intervention.

This analysis has several important limitations. First, although we included individuals of diverse ancestries, the PGS for our African Ancestry samples were still severely underpowered due to the small size of the discovery sample. Large-scale GWAS in diverse cohorts are vital to ensuring that any benefit of precision medicine is delivered in an equitable manner[53]. Second, while distinct, ancestry is closely related to race-ethnicity, one of the most profound social determinants of health[54]. Our measure of environmental risk may not fully capture risk factors and other social determinants that contribute to SUDs in populations beyond non-Hispanic whites. Future studies should include racially relevant measures of risk (e.g., experiences of interpersonal racism/discrimination, racial residential segregation) as well as other social and environmental measures that are known risk factors for SUDs (e.g., neighborhood social

conditions, alcohol outlet density). Further refinement of known risk factors may allow for even greater prediction of those at risk of SUD. Finally, while we tried to ensure time order between risk factors and onset of disorder, some risk factors (particularly adolescent substance use) could have occurred concurrently with diagnosis. Future work in samples with risk factors measured before the initiation of substance use (such as the Adolescent Brain Cognitive Development Study) will be important for replication efforts.

Recognizing that multiple social, clinical, and genetic factors contribute to risk for SUDs is important as we move towards the goal of an equitable precision medicine that benefits all segments of the population. The results of this integrative data analysis provide initial evidence that each of these risk factors contribute unique information for SUDs in early adulthood. Expanding our sources of information (such as electronic health records, census data from home of record) and making use of increasingly well-powered PGS will continue to improve our ability to identify and intervene for those who have the greatest risk of developing SUDs.

ACKNOWLEDGEMENTS

Research reported in this publication was supported by the National Institute on Alcohol Abuse and Alcoholism and the National Institute of Drug Abuse of the National Institutes of Health under award numbers R01AA015416, R01DA050721, and K02AA018755; the Academy of Finland (grants 100499, 205585, 118555, 141054, 265240, 308248, 308698 and 312073); and the Scientific and Technological Research Council of Turkey (TÜBİTAK) under award number 114C117 (FA); and the Sigrid Juselius Foundation. The content is solely the responsibility of the authors and does not necessarily represent the official views of any of the funding bodies. This research also used summary data from the Psychiatric Genomics Consortium (PGC), the Million Veterans Program (MVP), the GWAS and Sequencing Consortium for Alcohol and Nicotine (GSCAN), UK Biobank, the Genomic Psychiatry Cohort (GPC) and 23andMe, Inc. We would like to thank the many studies that made these consortia possible, the researchers involved, and the participants in those studies, without whom this effort would not be possible. We would also like to thank the research participants and employees of 23andMe.

The Externalizing Consortium: Principal Investigators: Danielle M. Dick, Philipp Koellinger, K. Paige Harden, Abraham A. Palmer. Lead Analysts: Richard Karlsson Linnér, Travis T. Mallard, Peter B. Barr, Sandra Sanchez-Roige. Significant Contributors: Irwin D. Waldman. The Externalizing Consortium has been supported by the National Institute on Alcohol Abuse and Alcoholism (R01AA015416 - administrative supplement), and the National Institute on Drug Abuse (R01DA050721). Additional funding for investigator effort has been provided by K02AA018755, U10AA008401, P50AA022537, as well as a European Research Council Consolidator Grant (647648 EdGe to Koellinger). The content is solely the responsibility of the authors and does not necessarily represent the official views of the above funding bodies. **Add Health:** Add Health is directed by Robert A. Hummer and funded by the National Institute on Aging cooperative agreements U01 AG071448 (Hummer) and U01AG071450 (Aiello and

Hummer) at the University of North Carolina at Chapel Hill. Waves I-V data are from the Add Health Program Project, grant P01 HD31921 (Harris) from *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD), with cooperative funding from 23 other federal agencies and foundations. Add Health was designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill. **ALSPAC:** We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists, and nurses. The UK Medical Research Council and Wellcome (Grant ref: 217065/Z/19/Z) and the University of Bristol provide core support for ALSPAC. This publication is the work of the authors, and Peter Barr and Danielle Dick will serve as guarantors for the contents of this paper. A comprehensive list of grants funding is available on the ALSPAC website (http://www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf); This research was specifically funded by the Medical Research Council (MRC) under grants MR/L022206/1, MR/M006727/1, and G0800612/86812; the Wellcome Trust under grant 086684; and the National Institute on Alcohol Abuse and Alcoholism under 5R01AA018333-05. GWAS data was generated by Sample Logistics and Genotyping Facilities at Wellcome Sanger Institute and LabCorp (Laboratory Corporation of America) using support from 23andMe. **COGA:** We thank 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 eleven different centers: University of Connecticut (V. Hesselbrock); Indiana University (H.J. Edenberg, T. Foroud, Y. Liu, M. Plawecki); University of Iowa Carver College of Medicine (S. Kuperman, J. Kramer); SUNY Downstate Health Sciences University (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, R. Hart, J. Salvatore);

The Children's Hospital of Philadelphia, University of Pennsylvania (L. Almasy); Virginia Commonwealth University (D. Dick); Icahn School of Medicine at Mount Sinai (A. Goate, P. Slesinger); and Howard University (D. Scott). Other COGA collaborators include: L. Bauer (University of Connecticut); J. Nurnberger Jr., L. Wetherill, X., Xuei, D. Lai, S. O'Connor, (Indiana University); G. Chan (University of Iowa; University of Connecticut); D.B. Chorlian, J. Zhang, P. Barr, S. Kinreich, G. Pandey (SUNY Downstate); N. Mullins (Icahn School of Medicine at Mount Sinai); A. Anokhin, S. Hartz, E. Johnson, V. McCutcheon, S. Saccone (Washington University); J. Moore, Z. Pang, S. Kuo (Rutgers University); A. Merikangas (The Children's Hospital of Philadelphia and University of Pennsylvania); F. Aliev (Virginia Commonwealth University); H. Chin and A. Parsian are the NIAAA Staff Collaborators. We continue to be inspired by our memories of Henri Begleiter and Theodore Reich, founding PI and Co-PI 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. This national collaborative study is supported by NIH Grant U10AA008401 from the National Institute on Alcohol Abuse and Alcoholism (NIAAA) and the National Institute on Drug Abuse (NIDA). All code necessary to replicate this study is available upon request.

REFERENCES

1. Ahmad FB, Rossen LM SP. *Provisional Drug Overdose Death Counts*.; 2021. https://www.cdc.gov/nchs/nvss/vsrr/drug-overdose-data.htm
2. Degenhardt L, Charlson F, Ferrari A, et al. The global burden of disease attributable to alcohol and drug use in 195 countries and territories, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet Psychiatry*. 2018;5(12):987-1012. doi:10.1016/S2215-0366(18)30337-7
3. Sacks JJ, Gonzales KR, Bouchery EE, Tomedi LE, Brewer RD. 2010 National and State Costs of Excessive Alcohol Consumption. *Am J Prev Med*. 2015;49(5):e73-e79. doi:10.1016/j.amepre.2015.05.031
4. National Drug Intelligence Center. *National Drug Threat Assessment*. Vol 2019. United States Department of Justice; 2011. http://www.justice.gov/archive/ndic/pubs44/44849/44849p.pdf
5. Conrod PJ, O'Leary-Barrett M, Newton N, et al. Effectiveness of a Selective, Personality-Targeted Prevention Program for Adolescent Alcohol Use and Misuse: A Cluster Randomized Controlled Trial. *JAMA Psychiatry*. 2013;70(3):334-342. doi:10.1001/JAMAPSYCHIATRY.2013.651
6. Verhulst B, Neale MC, Kendler KS. The heritability of alcohol use disorders: a meta-analysis of twin and adoption studies. *Psychol Med*. 2015;45(5):1061-1072. doi:10.1017/S0033291714002165
7. Verweij KJH, Zietsch BP, Lynskey MT, et al. Genetic and environmental influences on cannabis use initiation and problematic use: A meta-analysis of twin studies. *Addiction*. 2010;105(3):417-430. doi:10.1111/j.1360-0443.2009.02831.x
8. Burt SA. Are there meaningful etiological differences within antisocial behavior? Results of a meta-analysis. *Clin Psychol Rev*. 2009;29(2):163-178.

doi:https://doi.org/10.1016/j.cpr.2008.12.004

9. Rhee SH, Waldman ID. Genetic and environmental influences on antisocial behavior: A meta-analysis of twin and adoption studies. *Psychol Bull*. 2002;128(3):490-529. doi:10.1037/0033-2909.128.3.490

10. Kendler KS, Jacobson KC, Prescott CA, Neale MC. Specificity of Genetic and Environmental Risk Factors for Use and Abuse/Dependence of Cannabis, Cocaine, Hallucinogens, Sedatives, Stimulants, and Opiates in Male Twins. *Am J Psychiatry*. 2003;160(4):687-695. doi:10.1176/appi.ajp.160.4.687

11. Meier MH, Hall W, Caspi A, et al. Which adolescents develop persistent substance dependence in adulthood? Using population-representative longitudinal data to inform universal risk assessment. *Psychol Med*. 2016;46(4):877-889. doi:10.1017/S0033291715002482

12. Barr PB, Ksinan A, Su J, et al. Using polygenic scores for identifying individuals at increased risk of substance use disorders in clinical and population samples. *Transl Psychiatry*. 2020;10(1):196. doi:10.1038/s41398-020-00865-8

13. Kinreich S, Meyers JL, Maron-Katz A, et al. Predicting risk for Alcohol Use Disorder using longitudinal data with multimodal biomarkers and family history: a machine learning study. *Mol Psychiatry 2019 264*. 2019;26(4):1133-1141. doi:10.1038/s41380-019-0534-x

14. Gu F, Chen TH, Pfeiffer RM, et al. Combining common genetic variants and non-genetic risk factors to predict risk of cutaneous melanoma. *Hum Mol Genet*. Published online 2018. doi:10.1093/hmg/ddy282

15. O'Sullivan JW, Shcherbina A, Justesen JM, et al. Combining clinical and polygenic risk improves stroke prediction among individuals with atrial fibrillation. *medRxiv*. Published online January 1, 2020:2020.06.17.20134163. doi:10.1101/2020.06.17.20134163

16. Kessler RC, Berglund P, Demler O, et al. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Arch*

Gen Psychiatry. 2005;62(6):593. doi:10.1001/archpsyc.62.6.593

17. Harris KM, Halpern CT, Haberstick BC, Smolen A. The National Longitudinal Study of Adolescent Health (Add Health) sibling pairs data. *Twin Res Hum Genet*. 2013;16(01):391-398. doi:10.1017/thg.2012.137

18. Boyd A, Golding J, Macleod J, et al. Cohort profile: The 'Children of the 90s'-The index offspring of the avon longitudinal study of parents and children. *Int J Epidemiol*. Published online 2013. doi:10.1093/ije/dys064

19. Fraser A, Macdonald-wallis C, Tilling K, et al. Cohort profile: The avon longitudinal study of parents and children: ALSPAC mothers cohort. *Int J Epidemiol*. 2013;42(1):97-110. doi:10.1093/ije/dys066

20. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)-A metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform*. 2009;42(2):377-381. doi:10.1016/j.jbi.2008.08.010

21. Northstone K, Lewcock M, Groom A, et al. The Avon Longitudinal Study of Parents and Children (ALSPAC): an update on the enrolled sample of index children in 2019 [version 1; peer review: 2 approved]. *Wellcome Open Res*. 2019;4. doi:10.12688/wellcomeopenres.15132.1

22. Bucholz KK, McCutcheon V V., Agrawal A, et al. 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*. 2017;41(2):359-368. doi:10.1111/acer.13293

23. Rose RJRJ, Salvatore JEJE, Aaltonen S, et al. *FinnTwin12 Cohort: An Updated Review*. Vol 22.; 2019. doi:10.1017/thg.2019.83

24. Heatherton TF, Kozlowski LT, Frecker RC, Fagerstrom KO. The Fagerstrom Test for Nicotine Dependence: a revision of the Fagerstrom Tolerance Questionnaire. *Br J Addict*.

1991;86(9):1119-1127.

25. Hughes K, Bellis MA, Hardcastle KA, et al. The effect of multiple adverse childhood experiences on health: a systematic review and meta-analysis. *Lancet Public Heal*. Published online 2017. doi:10.1016/S2468-2667(17)30118-4

26. Sher KJ, Grekin ER, Williams NA. The development of alcohol use disorders. *Annu Rev Clin Psychol*. 2005;1(1):493-523. doi:10.1146/annurev.clinpsy.1.102803.144107

27. Karlsson Linner R, Mallard TT, Barr PB, et al. Multivariate genomic analysis of 1.5 million people identifies genes related to addiction, antisocial behavior, and health. *Nat Neurosci*. Published online January . h

28. Levey DF, Stein MB, Wendt FR, et al. Bi-ancestral depression GWAS in the Million Veteran Program and meta-analysis in >1.2 million individuals highlight new therapeutic directions. *Nat Neurosci*. Published online May 27, 2021. doi:10.1038/s41593-021-00860-2

29. Zhou H, Sealock JM, Sanchez-Roige S, et al. Genome-wide meta-analysis of problematic alcohol use in 435,563 individuals yields insights into biology and relationships with other traits. *Nat Neurosci*. Published online 2020. doi:10.1038/s41593-020-0643-5

30. Liu M, Jiang Y, Wedow R, et al. Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. *Nat Genet*. 2019;51(2):237-244. doi:10.1038/s41588-018-0307-5

31. Ripke S, Walters JTR, O'Donovan MC. Mapping genomic loci prioritises genes and implicates synaptic biology in schizophrenia. *medRxiv*. Published online 2020. doi:10.1101/2020.09.12.20192922

32. Barr PB, Dick DM. The Genetics of Externalizing Problems. *Curr Top Behav Neurosci*. 2020;47:93-112. doi:10.1007/7854_2019_120

33. Krueger RF, Hicks BM, Patrick CJ, Carlson SR, Iacono WGWG, McGue M. Etiological connections among substance dependence, antisocial behavior and personality:

Modeling the externalizing spectrum. *J Abnorm Psychol*. 2002;111(3):411-424. doi:10.1037//0021-843x.111.3.411

34. Kendler KS, Myers J. The boundaries of the internalizing and externalizing genetic spectra in men and women. *Psychol Med*. 2014;44(3):647-655. doi:10.1017/S0033291713000585

35. Polimanti R, Peterson RE, Ong JS, et al. Evidence of causal effect of major depression on alcohol dependence: Findings from the psychiatric genomics consortium. *Psychol Med*. Published online 2019. doi:10.1017/S0033291719000667

36. Johnson EC, Demontis D, Thorgeirsson TE, et al. A large-scale genome-wide association study meta-analysis of cannabis use disorder. *The Lancet Psychiatry*. Published online 2020. doi:10.1016/S2215-0366(20)30339-4

37. Zhou H, Rentsch CT, Cheng Z, et al. Association of OPRM1 Functional Coding Variant With Opioid Use Disorder: A Genome-Wide Association Study. *JAMA Psychiatry*. Published online June 2020. doi:10.1001/jamapsychiatry.2020.1206

38. Kendler KS, Gardner C, Dick DM. Predicting alcohol consumption in adolescence from alcohol- specific and general externalizing genetic risk factors, key environmental exposures and their interaction. *Psychol Med*. 2011;41(7):1507-1516. doi:10.1017/S003329171000190X.Predicting

39. Meyers JL, Salvatore JE, Vuoksimaa E, et al. Genetic Influences on Alcohol Use Behaviors Have Diverging Developmental Trajectories□: A Prospective Study Among Male and Female Twins. *Alcohol Clin Exp Res*. 2014;38(11):2869-2877. doi:10.1111/acer.12560

40. Barr PB, Mallard TT, Sanchez-Roige S, et al. Parsing Genetically Influenced Risk Pathways: Genetic Loci Impact Problematic Alcohol Use Via Externalizing and Specific Risk. *medRxiv*. Published online January 1, 2021:2021.07.20.21260861. doi:10.1101/2021.07.20.21260861

41. Sanchez-Roige S, Palmer AA, Clarke TK. Recent Efforts to Dissect the Genetic Basis of Alcohol Use and Abuse. *Biol Psychiatry*. Published online 2020. doi:10.1016/j.biopsych.2019.09.011

42. Walters RK, Polimanti R, Johnson EOECEO, et al. Trans-ancestral GWAS of alcohol dependence reveals common genetic underpinnings with psychiatric disorders. *Nat Neurosci*. 2018;21(12):1656-1669. doi:10.1038/s41593-018-0275-1

43. Dick DM, Barr P, Guy M, Nasim A, Scott D. Review: Genetic research on alcohol use outcomes in African American populations: A review of the literature, associated challenges, and implications. *Am J Addict*. 2017;26(5):486-493. doi:10.1111/ajad.12495

44. Mills MC, Rahal C. A scientometric review of genome-wide association studies. *Commun Biol*. 2019;2(1):9. doi:10.1038/s42003-018-0261-x

45. Martin AR, Gignoux CR, Walters RK, et al. Human Demographic History Impacts Genetic Risk Prediction across Diverse Populations. *Am J Hum Genet*. 2017;100(4):635-649. doi:10.1016/j.ajhg.2017.03.004

46. Duncan L, Shen H, Gelaye B, et al. Analysis of polygenic risk score usage and performance in diverse human populations. *Nat Commun*. Published online 2019. doi:10.1038/s41467-019-11112-0

47. Ruan Y, Anne Feng Y-C, Chen C-Y, et al. Improving Polygenic Prediction in Ancestrally Diverse Populations. *medRxiv*. Published online January 1, 2021:2020.12.27.20248738. doi:10.1101/2020.12.27.20248738

48. Curran PJ, Hussong AM. Integrative Data Analysis: The Simultaneous Analysis of Multiple Data Sets. *Psychol Methods*. 2009;14(2):81-100. doi:10.1037/a0015914

49. Cameron CA, Gelbach JB, Miller DL. Robust inference with multiway clustering. *J Bus Econ Stat*. 2011;29(2):238-249. doi:10.1198/jbes.2010.07136

50. Nagelkerke NJD. A note on a general definition of the coefficient of determination. *Biometrika*. 1991;78(3):691-692. doi:10.1093/biomet/78.3.691

51. Hasin DS, Grant BF. The National Epidemiologic Survey on Alcohol and Related Conditions (NESARC) Waves 1 and 2: review and summary of findings. *Soc Psychiatry Psychiatr Epidemiol*. 2015;50(11):1609-1640. doi:10.1007/s00127-015-1088-0

52. Kong A, Thorleifsson G, Frigge ML, et al. The nature of nurture: Effects of parental genotypes. *Science (80-)*. 2018;359(6374):424-428. doi:10.1126/science.aan6877

53. Martin AR, Kanai M, Kamatani Y, Okada Y, Neale BM, Daly MJ. Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat Genet*. 2019;51(4):584-591. doi:10.1038/s41588-019-0379-x

54. Williams DR, Mohammed SA, Leavell J, Collins C. Race, socioeconomic status, and health: complexities, ongoing challenges, and research opportunities. *Ann N Y Acad Sci*. 2010;1186(1):69-101. doi:10.1111/j.1749-6632.2009.05339.x

Table 1: Items included in the Clinical/Environmental Risk Index (CERI)

Measure	Definition
1) Low childhood SES	Parent(s) report having less than basic level of education [culturally dependent]; having a low-skill or menial occupation; income at or below the poverty line; or receipt of government assistance.
2) Family history of SUD	Biological parent self-reports history of SUD for themselves or other biological parent or meets criteria for SUD from clinical interview/AUDIT threshold of 8 or higher.
3) Childhood behavior problems	Respondent meets criteria for conduct disorder or oppositional defiant disorder from a clinical interview or computer-based prediction; or has a behavior problems score at or above the 90th percentile at 15 or younger.
4) Childhood internalizing problems	Respondent reports diagnosis of depression/anxiety or panic disorder; meets criteria for internalizing disorder in clinical interview/computer-based prediction; or has a CES-D score above a threshold of 16 at 15 or younger.
5) Early initiation of substance use	Respondent reports age of first whole alcoholic drink, smoked whole cigarette, or tried cannabis before the age of 15.
6) Adolescent alcohol use	Frequency of self-reported use 5 or more days per week at age 18 and below.
7) Adolescent tobacco use	Frequency of self-reported use at daily use at age 18 and below.
8) Adolescent cannabis use	Frequency of self-reported use 5 or more days per week at age 18 and below.
9) Peer substance use	Respondent reports the majority of their best friends use alcohol/tobacco/cannabis; their three best friends smoke daily/drink once a month/use cannabis once a month; or more than one friend smokes/drinks alcohol/has tried other drugs.
10) Traumatic events	Respondent reports exposure to any traumatic event.

Full description of sample specific definitions available in the supplementary information.

Table 2: Descriptive statistics of SUDs risk factors across each cohort

Phenotype	Cohort	Ancestry	N*	Mean/%	SD
Female	Add Health	AFR	1,605	55.26%	-
	Add Health	EUR	4,855	53.59%	-
	ALSPAC	EUR	4,733	56.71%	-
	COGA	AFR	870	51.38%	-
	COGA	EUR	1,878	51.33%	-
	FinnTwin12	EUR	1,193	53.73%	-
Age (at last observation)	Add Health	AFR	1,605	28.89	1.69
	Add Health	EUR	4,855	28.84	1.70
	ALSPAC	EUR	4,733	22.47	2.20
	COGA	AFR	870	24.13	5.12
	COGA	EUR	1,878	24.24	5.26
	FinnTwin12	EUR	1,193	22.44	0.72
Alcohol dependence	Add Health	AFR	1,605	3.93%	-
	Add Health	EUR	4,855	12.75%	-
	ALSPAC	EUR	4,733	5.92%	-
	COGA	AFR	870	11.49%	-
	COGA	EUR	1,878	21.14%	-
	FinnTwin12	EUR	1,193	8.55%	-
Drug dependence	Add Health	AFR	1,605	6.73%	-
	Add Health	EUR	4,855	10.79%	-
	ALSPAC	EUR	4,733	0.78%	-
	COGA	AFR	870	26.44%	-
	COGA	EUR	1,878	23.59%	-
	FinnTwin12	EUR	1,193	1.34%	-
Any substance dependence†	Add Health	AFR	1,605	11.21%	-
	Add Health	EUR	4,855	25.81%	-
	ALSPAC	EUR	4,733	8.87%	-
	COGA	AFR	870	30.69%	-
	COGA	EUR	1,878	34.66%	-
	FinnTwin12	EUR	1,193	10.98%	-
CERI	Add Health	AFR	1,605	1.95	1.48
	Add Health	EUR	4,855	2.07	1.65
	ALSPAC	EUR	4,733	2.08	1.19
	COGA	AFR	870	3.98	2.24
	COGA	EUR	1,878	3.65	2.38
	FinnTwin12	EUR	1,193	2.62	1.27

* Available samples with genotypic, phenotypic, and environmental risk data
† Any substance dependence includes those who meet criteria for alcohol, nicotine, or drug dependence.
CERI = clinical/environmental risk index

Table 3: Estimates for PGS Only, CERI Only, and Combined Risk Models

		Alcohol Dependence			Drug Dependence			Any Substance Dependence		
		OR	(95% CI)		OR	(95% CI)		OR	(95% CI)	
PGS Only Model*	ALCC PGS	1.06	(1.00	1.11)	1.04	(0.97	1.10)	1.00	(0.96	1.05)
	ALCP PGS	**1.15**	**(1.10**	**1.21)**	1.08	(1.01	1.14)	**1.12**	**(1.07**	**1.16)**
	EXT PGS	**1.19**	**(1.13**	**1.25)**	**1.35**	**(1.28**	**1.41)**	**1.34**	**(1.29**	**1.39)**
	MDD PGS	0.98	(0.92	1.04)	1.05	(0.99	1.12)	1.02	(0.97	1.06)
	SCZ PGS	1.07	(1.01	1.13)	1.06	(0.99	1.13)	1.05	(1.00	1.10)
ΔPseudo-R^2			*0.013*			*0.018*			*0.024*	
CERI Only Model*	CERI	**1.37**	(**1.33**,	**1.41**)	**1.67**	(**1.61**,	**1.72**)	**1.58**	(**1.54**,	**1.63**)
ΔPseudo-R^2			*0.054*			*0.129*			*0.120*	
Combined Model*	CERI	**1.35**	**(1.32**	**1.38)**	**1.64**	**(1.61**	**1.67)**	**1.55**	**(1.52**	**1.58)**
	ALCC PGS	1.05	(0.99	1.11)	1.04	(0.97	1.11)	1.00	(0.95	1.04)
	ALCP PGS	**1.14**	**(1.09**	**1.20)**	1.06	(0.99	1.13)	**1.11**	**(1.06**	**1.16)**
	EXT PGS	1.08	(1.01	1.14)	**1.14**	**(1.07**	**1.22)**	**1.19**	**(1.14**	**1.24)**
	MDD PGS	0.95	(0.89	1.01)	0.99	(0.93	1.06)	0.97	(0.92	1.02)
	SCZ PGS	1.06	(0.99	1.12)	1.02	(0.95	1.09)	1.04	(0.99	1.09)
ΔPseudo-R^2			*0.06*			*0.132*			*0.128*	

* All models included age, sex, and cohort as covariates. PGS residualized on age, sex, and first 10 ancestral principal components.
Bolded estimates = $p < .05$ after correction for multiple testing ($p < .05/3 = .017$)
ΔPseudeo-R^2 denotes pseudo-R^2 above model including age, sex, and cohort. CI = confidence interval; PGS = polygenic score; CERI = clinical/environmental risk index

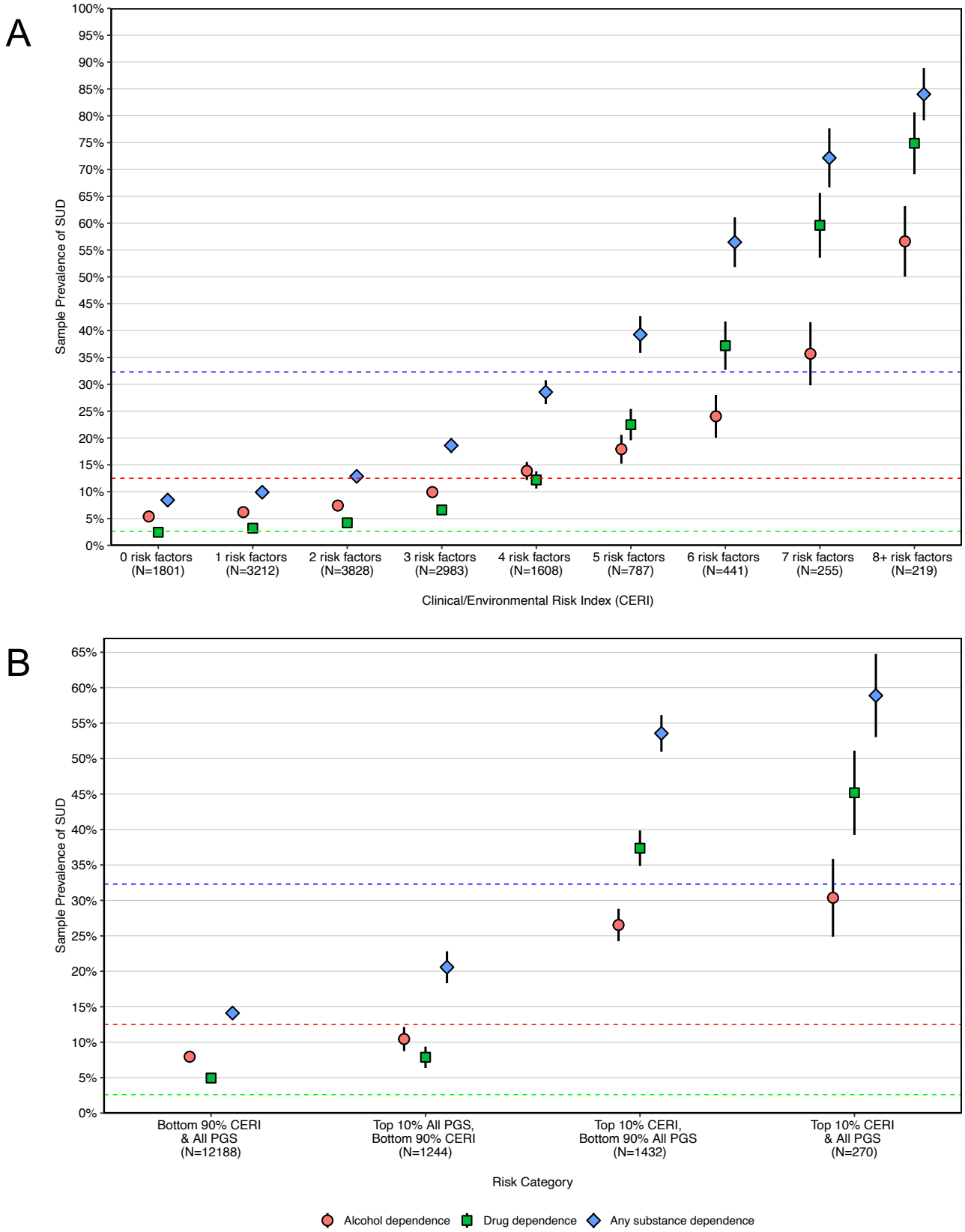
FIGURE CAPTIONS

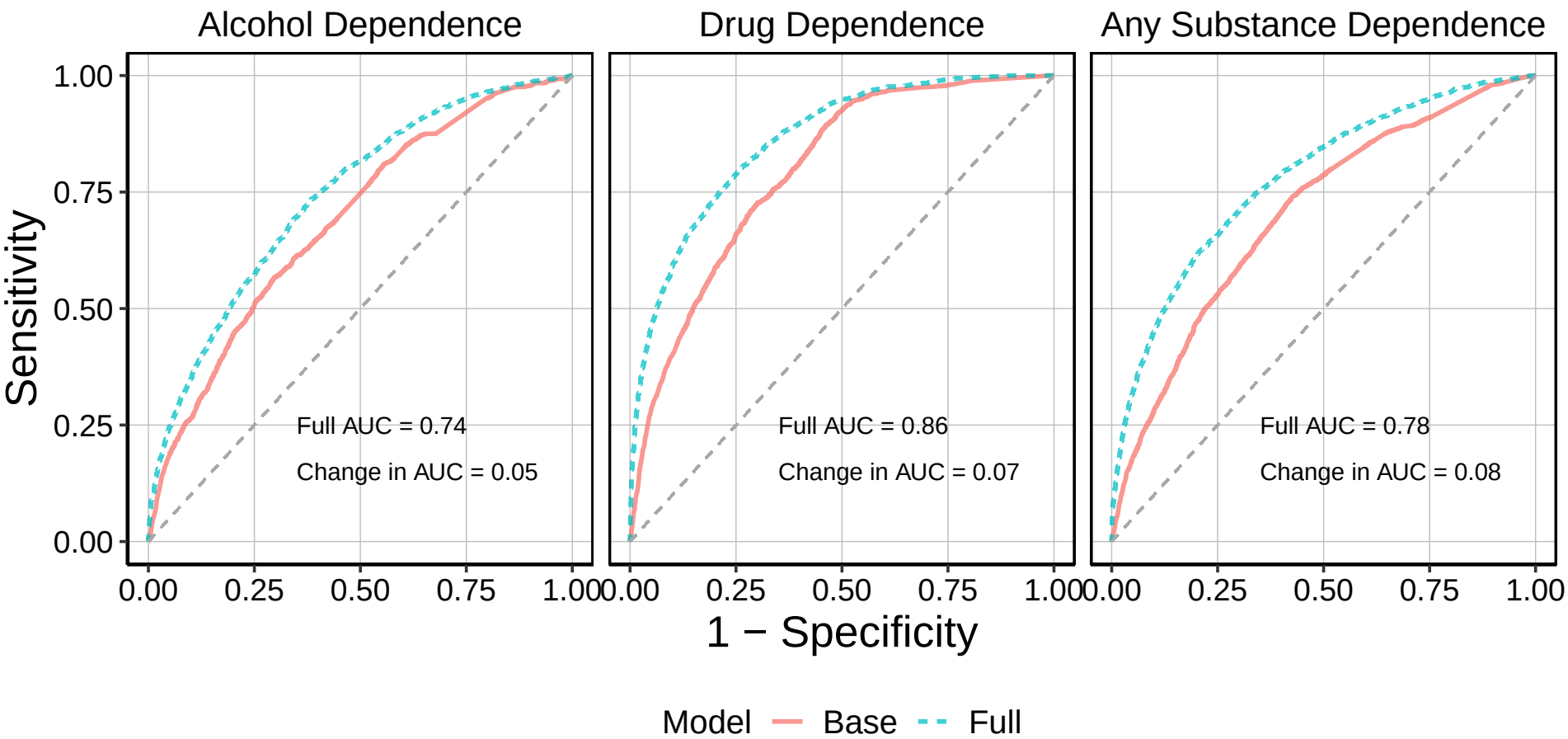
Figure 1: SUD Prevalence Across Genetic and Environmental Risk Factors

Panel A: Prevalence (and 95% confidence intervals) of those who meet criteria for alcohol, drug, any substance dependence across counts for items in the risk index. Panel B: Prevalence (and 95% confidence intervals) of those who meet criteria for alcohol, drug, or any substance dependence across four categories: 1) those below the 90th percentile for all PGS and the CERI; 2) those at or above the 90th percentile for the CERI; 3) those at or above the 90th percentile for all PGS; and 4) those at or above the 90th percentile for both the CERI and PGS. PGS and risk index were first residualized on sex, age, age^2, cohort, sex*age, sex*age^2, sex*cohort, cohort*age, cohort*age^2, sex*cohort*age, and sex*cohort*age^2. Dotted colored lines represent corresponding lifetime prevalence estimates for alcohol dependence (red), drug, dependence (green), and any substance use disorder (blue) from nationally representative data[51].

Figure 2: ROC Curves for Combined and Baseline Models

Receiver operating characteristic (ROC) curves for baseline models (red line, covariates only) and the full models (blue line, PGS + CERI + covariates) for each substance use disorder. Area under the curve (AUC) is presented for the PGS model in each cell. Change in AUC represents value of the difference between AUC from the full model and AUC from the base model.


A
100%
95%
90%
85%
80%
75%
70%
65%
60%
55%
50%
45%
40%
35%
30%
25%
20%
15%
10%
5%
0%
Sample Prevalence of SUD
0 risk factors (N=1801)
1 risk factors (N=3212)
2 risk factors (N=3828)
3 risk factors (N=2983)
4 risk factors (N=1608)
5 risk factors (N=787)
6 risk factors (N=441)
7 risk factors (N=255)
8+ risk factors (N=219)
Clinical/Environmental Risk Index (CERI)
B
65%
60%
55%
50%
45%
40%
35%
30%
25%
20%
15%
10%
5%
0%
Sample Prevalence of SUD
Bottom 90% CERI & All PGS (N=12188)
Top 10% All PGS, Bottom 90% CERI (N=1244)
Top 10% CERI, Bottom 90% All PGS (N=1432)
Top 10% CERI & All PGS (N=270)
Risk Category
Alcohol dependence
Drug dependence
Any substance dependence


Alcohol Dependence
Drug Dependence
Any Substance Dependence
Sensitivity
1 − Specificity
0.00
0.25
0.50
0.75
1.00
Full AUC = 0.74
Change in AUC = 0.05
Full AUC = 0.86
Change in AUC = 0.07
Full AUC = 0.78
Change in AUC = 0.08
Model
Base
Full