Unravelling Polygenic Risk and Environmental Interactions in Adolescent Polysubstance Use: a U.S. Population-Based Observational Study

Dongmei Zhi, PhD[1,2,3], Brandon T. Sanzo, MS[3], Daniel H. Jung[4], Judit Cabana-Domínguez, PhD[5,6,7,8], Noelia Fernàndez-Castillo, PhD[5,9,10,11], Bru Cormand, PhD[5,8,9,10], Jing Sui, PhD[12], Rongtao Jiang, PhD[13], 23andMe Research Team[14], Eden A. Evins, MD[1,2], Scott E. Hadland, MPH[15,16], Joshua L. Roffman, MD[2], Richard T. Liu[2,17], Jodi Gilman, PhD[1,2a], Phil H. Lee, PhD[2,3,17a]

1. Center for Addiction Medicine, Massachusetts General Hospital, Boston, MA 02114, USA
2. Department of Psychiatry, Mass General Brigham and Harvard Medical School, Boston, MA 02115, USA
3. Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA 02114, USA
4. Department of Psychology and Behavioral Neuroscience, Cornell University, Ithaca, NY 14853, USA
5. Department of Genetics, Microbiology and Statistics, Faculty of Biology, Universitat de Barcelona, Barcelona, Catalunya, 08028, Spain
6. Psychiatric Genetics Unit, Group of Psychiatry, Mental Health and Addiction, Vall d'Hebron Research Institute (VHIR), Universitat Autònoma de Barcelona, Barcelona 08035, Spain.
7. Department of Mental Health, Hospital Universitari Vall d'Hebron, Barcelona 08035, Spain.
8. Biomedical Network Research Centre on Mental Health (CIBERSAM), Instituto de Salud Carlos III, Madrid, Spain
9. Centro de Investigación Biomédica en Red de Enfermedades raras (CIBERER), Instituto de Salud Carlos III, Madrid, Spain.
10. Institut de Biomedicina de la Universitat de Barcelona, Barcelona, Catalonia, Spain
11. Institut de Recerca Sant Joan de Déu, Esplugues de Llobregat, Catalonia, Spain
12. State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing, 100875, China
13. Department of Radiology and Biomedical Imaging, Yale School of Medicine, New Haven, CT 06510, USA
14. 23andMe, Inc.
15. Division of Adolescent and Young Adult Medicine, Mass General for Children, Boston, MA 02114, USA
16. Department of Pediatrics, Harvard Medical School, Boston, MA 02115, USA
17. Stanly Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA

*a: co-last author

***Correspondence:**
Phil H. Lee, PhD
Center for Genomic Medicine
Simches Research Building
185 Cambridge St.
Boston, MA 02114
Phone: 617-643-9304
Fax: 617-643-3080
Email: PLEE0@mgh.harvard.edu

Words: 3450 words
Abstract: 245 words
Tables and Figures: 5
References: 30

Keywords: Polysubstance use, polygenic scores, environmental factors, adolescents.

ABSTRACT

Background: Polysubstance use (PSU), defined as the use of multiple psychoactive substances, is associated with a heightened risk of subsequent health issues, including substance use disorders. However, the interplay between genetic susceptibility and environmental exposures in PSU initiation during adolescence remains understudied.

Methods: We examined associations of polygenic scores (PGSs) for general addiction risk, environmental factors, and their joint interactions with PSU initiation among 11,868 adolescents (aged 11-15 years) from the Adolescent Brain and Cognitive Development study. PSU status was assessed through interviews and toxicology screenings.

Results: Our sample included 7,898 adolescents (mean age 12.9 [0.6] years; 4,150 [53%] male). Of these, 541 (6.8%) had initiated single substance use (SSU), and 162 (2.1%) reported PSU). PGSs for general addiction risk were significantly associated with PSU (Odds Ratios [OR]=1.62, 95% CI=1.30-2.01) but not with SSU. Key environmental risk factors for PSU included prenatal substance use and peer victimization, whereas protective factors included planned pregnancy and positive family dynamics. Notably, gene-environment interaction analyses revealed that peer victimization (OR=2.4, 95% CI=1.4–4.2), prenatal substance use (OR=2.1, 95% CI=1.2–3.6), and substance availability (OR=2.3, 95% CI=1.3–3.9) substantially increased PSU risk among adolescents with high genetic susceptibility, while having minimal influence at low genetic risk levels (all $p < 0.05$ after multiple testing correction).

Conclusions: This study provides novel evidence linking polygenic risk to PSU in early adolescence and highlights PSU as a more severe manifestation of substance use liability driven by heightened genetic vulnerability and adverse environmental exposures.

INTRODUCTION

Adolescent polysubstance use (PSU), defined as the concurrent or sequential use of multiple psychoactive substances, is a pressing public health concern. In the United States (U.S.), roughly 20–25% of adolescents report using more than one substance in the past year.[1] PSU exposure during adolescence may lead to more severe and long-lasting consequences than single substance use (SSU), including persistent cognitive impairments, psychiatric comorbidities,[2] and elevated risk for substance use disorders (SUDs) in adulthood.[3]

Despite its significant public health impact, the underlying mechanisms of adolescent PSU remain poorly understood. Research on substance use, particularly in genetics, has predominantly focused on adult populations and individual substances[4,5], overlooking the developmental specificity of PSU during adolescence. This gap is critical, as adolescence represents a sensitive period when genetic predispositions interact with rapidly changing neurobiological and environmental factors. A nationwide survey revealed that initiating alcohol use before age 14 significantly increases the risk of earlier onset and more severe alcohol dependence.[6]

Notably, genetic influences play a significant role in substance use behaviors, with heritability estimates often exceeding 50%.[7] Genome-wide association studies (GWASs) have started to unveil the genetic architecture of substance use, identifying numerous loci associated with substance use and addiction vulnerabilities,[4,5] and highlighting extensive pleiotropy across SUDs, where the same genetic variants influence multiple disorders.[8] A particularly relevant discovery is the identification of a latent "addiction risk factor" that captures shared genetic liability across alcohol, tobacco, cannabis, and opioid use disorders,[8] demonstrating a common genetic vulnerability for SUD. This factor, which reflects a distinct cross-substance genetic risk independent of specific substance use behaviors, has been associated with neuropsychiatric traits such as risk-taking and executive dysfunction.[8,9] However, its role as a potential mechanism underlying adolescent PSU remains unexplored.

Environmental factors also play a pivotal role in shaping the trajectory of substance use during adolescence. Developmental, familial, peer, and school contexts, along with neighborhood-level influences, collectively contribute to early substance use.[10] Research increasingly supports an integrative model in which genetic predispositions and environmental factors interact dynamically. For instance, genetic risk levels might modulate the impact of environmental exposures on development. Adverse factors, such as high family conflict, excessive screen time, and disadvantaged school environments, may exacerbate risks for individuals with high genetic risk, whereas protective conditions, such as high socioeconomic status, can mitigate outcomes for those with lower genetic vulnerabilities.[11,12] Currently, research on gene-environment interactions affecting adolescent substance use risk is largely absent.

This study addresses significant knowledge gaps using data from the Adolescent Brain Cognitive Development (ABCD) Study, a population-based longitudinal cohort involving 11,868 U.S. children with sequential assessments of substance use, genomics, and environmental factors. Our investigation has three primary objectives: (1) to examine the association between addiction genetic risk, quantified through polygenic scores (PGSs), and adolescent problem substance use (PSU), (2) to identify environmental risk and protective factors influencing PSU, such as prenatal substance exposure, family dynamics, peer influence, socioeconomic status, and traumatic events, and (3) to characterize gene-environment (G×E) interactions that modulate PSU risk and resilience during adolescence. By elucidating how genetic risk and environmental context converge to shape substance use trajectories during this critical developmental period, this research aims to enhance our understanding of adolescent substance use vulnerability and inform targeted prevention strategies.

METHODS

Study Design and Participants

This observational study used data from the ABCD Study (release 5.1), which included 11,868 participants aged 9 to 10 at enrollment, recruited from elementary schools across 21 U.S. sites. Data were downloaded from the NIMH Data Archive (NDA; https://nda.nih.gov). The ABCD study received central institutional review board (IRB) approval from the University of California, San Diego, along with written parental consent and verbal assent from participants. Analyses utilized data from baseline to follow-up Year 3, encompassing substance use outcome measures for the full cohort. This secondary analysis of de-identified ABCD data was exempt from IRB review at Mass General Brigham, Boston, U.S.

Substance Use Outcomes

Substance use outcomes were derived from the Substance Use Interview, Substance Use Phone Interview (Mid-Year), and Hair Drug Toxicology data. Detailed information on these instruments, including questions and response options, is provided in **Table S1**. Participants were classified as cases for four binary substance use outcomes based on lifetime use: ALC (at least one full cup of alcohol consumed), NIC (at least one full dose of nicotine products, including cigarettes, e-cigarettes, smokeless tobacco, cigars, hookah, pipe, and nicotine replacements), CAN (at least one full dose of cannabis products, including smoked or vaped flower, edibles, concentrates, and synthetic tetrahydrocannabinol products), and OTH (any use of other substances, including prescription drug misuse and illicit drugs). Additionally, we defined two aggregated lifetime outcomes of substance use: SSU (single substance use, using only one type of substance at least once) and PSU (polysubstance use, with initiation of at least two types). Unless otherwise stated, controls for aggregated outcomes were participants reporting no lifetime substance use.

Genotype Data, Quality Control, and Imputation

In the ABCD study, saliva and blood samples were collected at baseline for genotyping using the Affymetrix NIDA Smokescreen Array. Quality-controlled (QC) genotype and TOPMED-imputed data for 11,665 participants were downloaded from the NDA website. The procedures for genotyping, imputation, and QC in the ABCD study are detailed elsewhere.[13] Genetic principal components (PCs) and genetic relatedness were estimated, and participants were classified into five genetic ancestry groups—AFR (African), AMR (American), EUR (European), SAS (South Asian), and EAS (East Asian)—through clustering with 1000 Genomes Reference samples (**Figure S1**). Due to small sample sizes and few substance use cases in the SAS and EAS groups, subsequent analyses focused on the AFR, AMR, and EUR groups. For PGS analyses, we included imputed SNPs with imputation quality scores ≥ 0.8, minor allele frequency ≥ 1%, and Hardy-Weinberg Equilibrium *p*-value > $1x10^{-6}$.

Genomic Structural Equation Modeling of General Substance Use

Building on recent multivariate GWAS studies,[8] we used GenomicSEM[14] to construct a general addiction risk factor capturing common genetic vulnerability shared across five SUDs (**Table S2**): alcohol use disorder (AUD; sample N = 313,963, effective sample size [N_{eff}]= 188,039),[15] cannabis use disorder (CUD; N = 374,287, N_{eff} = 65,159),[16] cocaine use disorder (COD; N = 8,463, N_{eff} = 6,285),[17] opioid use disorder (OUD; N = 447,950, N_{eff} = 179,212),[18,19] and nicotine dependence (ND; N = 244,890, N_{eff} = 39,420).[20] Consistent with significant genetic correlations among these disorders (**Figure S2**; **Table S3**), the common factor model fit the data well (χ^2 = 4.10, p-value = 0.54, comparative fit index = 1, standardized root mean square residual = 0.03).

The latent factor loaded significantly on all SUDs (standardized loadings on AUD = 0.84, CUD = 0.80, COD = 0.94, OUD = 0.98, ND = 0.62, **Table S4**).

Generation of Polygenic Scores (PGS)

We used PRS-CS[21] to estimate posterior SNP effect weights. Parameters were set to a gamma-gamma prior (a = 1, b = 0.5), global shrinkage (phi) = auto, 1,000 MCMC iterations (500 burn-in), and thinning factor = 5. Ancestry-matched reference panels were derived from the 1000 Genomes Project. Using the estimated posterior SNP weights, we calculated PGSs for substance use disorders (PGS_{SUD}) based on the latent addiction risk factor as described above. Consistent with GenomicSEM analysis, PGSs for individual SUDs were highly intercorrelated with PGS_{SUD} (**Figure S3**; **Table S5**). We regressed out PGS_{SUD} from each disorder-specific PGS to index genetic risk specific to each SUD (PGS_{AUD}, PGS_{CUD}, PGS_{COD}, PGS_{OUD}, PGS_{ND}). PGSs were adjusted for population stratification using the top 10 genetic principal components within each ancestry group.

Demographics and Environmental Measures

Participants' biological sex, race/ethnicity (parent-reported), and age (in months) were recorded at enrollment. Following prior studie,[22] 96 environmental measures at the initial collection were categorized into eight domains, including perinatal/early developmental events (N=12), prenatal health exposures (N=6), life events/lifestyle (N=21), family (N=30), neighborhood (N=8), school (N=3), peer relationship (N=11), and cultural exposures (N=5). Detailed instruments and descriptive statistics for these variables are provided in **Table S6** and **Table S7**.

Statistical Analyses

Our cross-sectional analysis focused on 7,898 independent participants of AFR, AMR, and EUR groups, with twins/siblings randomly excluded to ensure independence. Generalized linear mixed-effects models (GLMM) with binomial error structure and logit link function were used to examine the associations between PGS and PSU, stratified by ancestry to account for genetic background, and the associations between each environmental measure and PSU. Within each model, PSU was the dependent variable, with each PGS or environmental measure as an independent predictor. Age, sex, and race/ethnicity (only for environmental measures) were included as covariates, and enrollment sites as a random intercept to account for site-level variability. Scores of environmental measures with significantly positive skew were log-transformed, and all numeric variables were standardized. From GLMM, we present odds ratios (OR) and associated two-tailed p values.

To assess how associations between environmental measures and PSU outcomes vary across different levels of genetic risk, we examined gene-environment (GxE) interactions in substance use by stratifying participants into high PGS_{SUD} (top 25th percentile) and low PGS_{SUD} (bottom 75th percentile) levels. The GLMM included fixed effects for environmental measures, genetic risk level, and their interaction to capture gene-environment interactions while controlling for age, sex, and site as covariates. For all analyses with multiple comparisons, a false discovery rate (FDR) of 0.05 was used to declare significant findings. Analyses were conducted in R (version 4.3.2) and MATLAB (version R2017b).

Sensitivity Analysis

Sensitivity analyses tested the robustness of environmental effects on PSU by additionally adjusting for puberty, socioeconomic factors (family income and caregiver education), and family history of substance use.

RESULTS

Cohort Description

Table 1 summarizes the demographic characteristics of the study participants. Of 11,868 individuals recruited at baseline, 7,898 independent participants (67%) with PSU outcome data at Year 3 were included in the analysis. At Year 3, the mean age was 12.9 years (standard deviation = 0.6), and 3,748 participants were female (47.5%). From baseline to Year 3, we observed a consistent increase in substance use initiation among children, with 6.8% reporting lifetime SSU (N = 541) and 2.1% reporting PSU (N = 162) by Year 3 (**Figure 1A**). Among the four substance categories examined, nicotine use was the most frequently reported, with 3.6% of participants by Year 3 (**Figure 1B**). Within the PSU group, nicotine and cannabis were the substances most commonly used together (**Figure 1C**). All four substance use outcomes were positively correlated, with tetrachoric correlations ranging between 0.45 and 0.64 (**Table S8**; **Figure 1D**).

Associations of Polygenic Score with Polysubstance Use

We examined the associations between six PGSs related to addiction risk and lifetime PSU initiation at Year 3. These included PGS_{SUD}, which captures shared genetic liability across SUDs, as well as five disorder-specific PGSs: PGS_{AUD}, PGS_{COD}, PGS_{CUD}, PGS_{ND}, and PGS_{OUD} (see **Methods**). To account for genomic population structure, analysis was conducted separately within each genetic ancestry group.

In the EUR group, $PGS_{SUD,}$ representing shared genetic liability across SUDs, was significantly associated with PSU, after covarying for age, sex, genetic PCs, and enrollment sites (**Figure 2A; Table S9**). The adjusted OR for PGS_{SUD} was 1.62 (95% confidence interval (CI) = 1.30–2.01, $P_{FDR} = 2.52x10^{-4}$), and accounted for 5.68% of the phenotypic variation in PSU (likelihood ratio test, p-value = $1.35x10^{-5}$, **Figure 2B**). For the AFR and AMR groups, we found no significant findings, reflecting the limited cross-ancestry transferability of European-derived substance use PGSs.

In a secondary analysis of SSU, individuals with SSU showed a trend of heightened PGS_{SUD} compared to those with no substance use (NSU). However, this association did not retain significance after multiple-testing correction (adjusted OR = 1.18; 95% CI = 1.05–1.32; $P_{FDR} = 6.79x10^{-2}$; Pseudo R^2 = 0.78%; **Table S10**).

Associations of Environmental Exposures with Polysubstance Use

To identify potential risk and protective factors for PSU, we examined associations between 96 environmental measures and PSU compared to NSU. Of these, 67 measures (70%) showed significant associations with PSU ($P_{FDR} < 0.05$), with effect sizes ranging from OR = 0.56 to OR = 1.75 (**Figure 3A; Table S11**). The strongest risk factors for PSU included R-rated screen use (OR = 1.75; 95% CI = 1.57–1.96; $P_{FDR} = 1.48x10^{-21}$), peer reputation victimization (OR = 1.74; 95% CI = 1.55–1.96; $P_{FDR} = 2.09x10^{-19}$), prenatal substance use (OR = 1.71; 95% CI = 1.50–1.94; $P_{FDR} = 1.75x10^{-15}$), and area deprivation index (OR = 1.65; 95% CI = 1.39–1.95; $P_{FDR} = 7.29x10^{-8}$). Conversely, the strongest protective factors included family income (OR = 0.56; 95% CI = 0.47–0.66; $P_{FDR} = 3.98x10^{-10}$) and planned pregnancy (OR = 0.61; 95% CI = 0.52–0.72; $P_{FDR} = 7.29x10^{-8}$; **Figure 3B**). Among all environmental categories, peer relationship factors exhibited the largest mean effect sizes, with 9 of 12 measures reaching significance (**Figure S4**).

To better understand environmental factors associated with the escalation from single to polysubstance use, we examined exposures distinguishing PSU from SSU. Of 96 environmental measures, 20 (21%) showed significant associations ($P_{FDR} < 0.05$), with prenatal exposures and peer relationship factors showing the largest effect sizes (OR range: 0.59–1.58; **Figure S5; Table S12**).

Notably, all environmental factors significantly associated with PSU versus SSU were also significant for PSU versus NSU, suggesting that risk and protective factors operate along a continuum rather than being unique to the transition from SSU to PSU. The SSU group exhibited intermediate exposure levels between NSU and PSU, reinforcing a dose-response pattern (**Figure 3B**). Peer relationships, particularly peer reputation victimization, emerged as the strongest and most consistent predictors across all comparisons, highlighting their dominant role in distinguishing PSU from both SSU and NSU. Even after adjusting for puberty, socioeconomic factors, and family history of substance use, the primary associations remained statistically significant, although the number of significant associations decreased (**Figure S6**).

Interaction Effects of PGS and Environments on Polysubstance Use

We next examined whether high genetic risk amplifies the effects of environmental exposures on PSU. Among exposures significantly associated with PSU (**Tables S11 and S12**), five gene-environment interactions were nominally significant when comparing PSU to NSU (p-value < 0.05; uncorrected; **Figure 4A; Table S13**). Although these findings did not survive multiple testing correction, they suggest potential gene-environment interactions that may warrant further investigation.

In contrast, when comparing PSU to SSU, four interaction effects remained statistically significant after multiple testing corrections ($P_{FDR} < 0.05$; **Figure 4A; Table S14**). Environmental factors had a significantly stronger impact on substance use in the high genetic risk group compared to those with low genetic risk. Specifically, for PSU versus SSU, the effects of peer overt victimization, substance availability, prenatal substance use, and prenatal tobacco use were more pronounced among individuals with high genetic risk, whereas their influence was weak or negligible in those with low genetic risk (**Figure 4B**).

DISCUSSION

In this study, we examined how genetic susceptibility ("nature") and environmental exposures ("nurture") together influence adolescent PSU. To our knowledge, this is the first population-based study to explore both polygenic and environmental factors related to PSU in early adolescence. Our findings suggest that increased genetic liability for common addiction risk is linked to a higher risk of PSU during this developmental stage. Additionally, several environmental factors—including prenatal substance use, peer victimization, and substance availability—were more strongly associated with PSU among adolescents with greater genetic susceptibility.

Our findings provide novel insights into the early emergence of genetic risk for PSU. The association between increased genetic liability for adulthood SUDs and PSU risk in preadolescents underscores the relevance of genetic factors even at early developmental stages. This result extends previous research primarily focused on adults,[4,5] suggesting that genetic predispositions to substance use may manifest before the onset of diagnosable SUDs. Notably, SSU did not show a significant association with genetic risk, whereas PSU did, indicating that PSU may represent a subgroup with heightened genetic vulnerability for substance use. This distinction underscores the potential role of genetic predisposition in the escalation from SSU to PSU, reinforcing PSU as a high-risk phenotype within adolescent substance use trajectories. Taken together, our findings suggest that genetic risk factors contribute not only to the development of SUDs but also to the earlier initiation of multiple substances during critical developmental periods.

The identification of significant environmental factors—including prenatal substance exposure, excessive adult-rating screen exposures, and peer victimization—is consistent with previous studies, reinforcing the crucial role of environmental influence in shaping PSU vulnerability.[23,24] For instance, exposures to R-rated movies may introduce adolescents to depictions of substance use, such as alcohol drinking and smoking, thereby increasing likelihood of substance use initiation. However, our study is the first to demonstrate significant GxE interactions specific to PSU. Emerging literature on G×E interactions highlights that neither genes nor environment alone can fully account for the complexity of adolescent substance use; instead, it is their interplay that shapes individual risk profiles.

Our study revealed that prenatal substance use, particularly prenatal tobacco use, may interact with a high genetic predisposition for addiction, thereby amplifying the risk of PSU initiation. Preclinical studies in animal models demonstrate that prenatal nicotine exposure induces long-lasting alterations in brain circuits—particularly within the mesolimbic dopamine system—thereby increasing offspring's propensity for drug self-administration.[25]

Adolescence is a critical developmental period during which peer relationships and experiences of victimization can significantly shape substance use behaviors. Here, we demonstrate that a higher genetic predisposition for substance use interacts with severe peer victimization, amplifying the risk of substance use. This finding aligns with previous research indicating that adolescents with heightened genetic vulnerability are particularly susceptible to negative peer influences, increasing their likelihood of engaging in substance use[11] Furthermore, prior studies suggest that a high genetic risk can intensify the effects of peer victimization on behavioral risk, thereby elevating the risk of addiction.[26] These earlier results, combined with the current findings, emphasize the importance of supportive social environments in mitigating substance use behaviors.

We also identified a significant GxE interaction between genetic risk and substance availability. Research has consistently shown that greater substance availability is associated with an increased likelihood of use; for instance, higher alcohol outlet densities are linked to elevated alcohol consumption and related problems,[27] and greater tobacco retailer density correlates with increased tobacco use prevalence, particularly among youth.[28] These studies emphasize the importance of regulating substance availability as a strategy to prevent and mitigate substance use, especially among vulnerable adolescent populations. Our findings indicate that when substances are readily accessible, this environment may act as a catalyst that amplifies underlying genetic vulnerabilities. This conclusion has important implications for regulations regarding the minimum legal age for purchasing nicotine and alcohol (e.g., ≥21 years) and cannabis legalization for recreational use, thus reducing the risk of substance misuse. This issue is particularly concerning during adolescence—a developmental stage characterized by ongoing neurodevelopment and immature self-regulation.

Our findings have significant implications for designing prevention and intervention strategies aimed at reducing PSU among youth. The increased effect of adverse environmental factors on children with greater genetic susceptibility highlights the need to address both genetic and environmental risk factors in early intervention initiatives. With the early manifestation of genetic risk, there is a rising interest in genetic screening tools to identify at-risk youth before substance use behaviors become ingrained. Although still in its nascent phase, understanding the genetic foundations of PSU in adolescents could aid in developing targeted prevention programs that integrate individual genetic profiles with environmental and behavioral risk assessments, ensuring a comprehensive and culturally sensitive approach.

Several limitations should be considered when interpreting our findings. First, although the ABCD study was designed to match the racial composition of U.S.-based 9-to 10-year-olds in 2015,[29] it represents a largely self-selected sample of higher-functioning families, with socioeconomic status factors exceeding the U.S. average, potentially limiting generalizability to populations with socioeconomic adversity. Second, adolescent self-reporting of substance use may introduce bias, as children might underreport or inaccurately report their substance use behaviors. However, the study was thoughtfully designed to maximize adolescents' truthful disclosure of substance use, including the use of hair toxicology data.[30] Furthermore, the relative rarity of substance use in the ABCD sample at this young age precluded meaningful examination of severity, patterns, and contexts. Third, despite statistically significant associations, the variance in PSU explained by PGS remains modest. As GWAS continues to expand and more genetic variants are identified, future analyses may uncover more subtle associations between genetic risk and PSU. However, current findings underscore the polygenic and multifactorial nature of adolescent PSU, suggesting that genetic risk must be considered alongside environmental and developmental influences.

In summary, our study demonstrates that the integration of genetic data with detailed assessments of environmental factors enhances our understanding of PSU risk during early adolescence. These findings underscore the importance of early identification of at-risk children through comprehensive evaluations of both genetic and environmental factors. As the ABCD cohort matures, we anticipate that more nuanced insights into the genetic and environmental contributions to PSU will emerge, ultimately informing more effective prevention and intervention strategies.

Acknowledgements

PL was supported by grants from the U.S. National Institute of Health (R01MH119243, R01GM148494) and by MassGeneral Brigham Department of Psychiatry. JMG is supported by K02DA052684 and R01DA051540. JLR is supported by R01MH124694. The ABCD study is supported by the National Institutes of Health (NIH) and additional federal partners under award numbers U01DA041048, U01DA050989, U01DA051016, U01DA041022, U01DA051018, U01DA051037, U01DA050987, U01DA041174, U01DA041106, U01DA041117, U01DA041028, U01DA041134, U01DA050988, U01DA051039, U01DA041156, U01DA041025, U01DA041120, U01DA051038, U01DA041148, U01DA041093, U01DA041089, U24DA041123, U24DA041147. Additional support for this work was made possible from NIEHS R01-ES032295 and R01-ES031074. Funding supporting this study was provided by the Spanish 'Ministerio de Ciencia, Innovación y Universidades' (PID2021-1277760B-I100 and PID2024-158634OB-I00, to BC and NFC), 'Generalitat de Catalunya/AGAUR', (2021-SGR-01093, to BC and NFC), ICREA Academia 2021 (to BC), 'Fundació La Marató de TV3′ (202218-31, to BC) and 'Ministerio de Sanidad, Servicios Sociales e Igualdad/Plan Nacional Sobre Drogas' (PNSD-2020I042 and PNSD-2024|056, to NFC). A full list of supporters is available online. A list of participating sites and a complete list of the study investigators can be found online. ABCD consortium investigators designed and implemented the study and provided data but did not necessarily participate in analysis or writing of this report. This manuscript reflects the views of the authors and does not reflect the opinions or views of the NIH or ABCD consortium investigators. The ABCD data repository grows and changes over time. The ABCD data used in this report came from http://dx.doi.org/10.15154/z563-zd24. We would like to thank the research participants and employees of 23andMe, Inc. for making this work possible.

Author Contributions

PHL, DZ, and BS have directly accessed and verified the ABCD data reported in the manuscript. PHL and JG conceptualized the idea and designed the study. PHL, DZ, JG, DJ, and BS were involved in data curation, formal analysis, investigation, visualization, and validation of the findings. PHL and JG were involved in funding acquisition, supervision, and project administration. PHL, JG, DZ, and BS wrote the original draft. All authors were involved in the critical review and editing. All authors approved the final version of the manuscript.

Declaration of Interests

All other authors have no interests to declare.

Data Sharing

The Adolescent Brain Cognitive Development data used in this report are available from the NIMH Data Achieve (https://nda.nih.gov) to Authorized Users. We used open-source software packages from R, and publicly available methods, GenomicSEM, PLINK, PRS-CS (**Table S2**).

REFERENCES

1. Abuse S, Administration MHS. Key substance use and mental health indicators in the United States: Results from the 2018 National Survey on Drug Use and Health. *HHS Publication No PEP19-5068, NSDUH Series H-54* 2019; **170**: 51-8.
2. Maslowsky J, Schulenberg JE, O'Malley PM, Kloska DD. Depressive symptoms, conduct problems, and risk for polysubstance use among adolescents: Results from US national surveys. *Ment Health Subst Use* 2013; **7**(2): 157-69.
3. Clark SL, Dodd CG, Taylor L, et al. Characterizing patterns of substance use in trauma exposed youth. *J Psychiatr Res* 2023; **167**: 1-9.
4. Saunders GRB, Wang X, Chen F, et al. Genetic diversity fuels gene discovery for tobacco and alcohol use. *Nature* 2022; **612**(7941): 720-4.
5. Cheng Y, Dao C, Zhou H, et al. Multi-trait genome-wide association analyses leveraging alcohol use disorder findings identify novel loci for smoking behaviors in the Million Veteran Program. *Transl Psychiatry* 2023; **13**(1): 148.
6. Hingson RW, Heeren T, Winter MR. Age at drinking onset and alcohol dependence: age at onset, duration, and severity. *Arch Pediatr Adolesc Med* 2006; **160**(7): 739-46.
7. Deak JD, Johnson EC. Genetics of substance use disorders: a review. *Psychol Med* 2021; **51**(13): 2189-200.
8. Hatoum AS, Johnson EC, Colbert SMC, et al. The addiction risk factor: A unitary genetic vulnerability characterizes substance use disorders and their associations with common correlates. *Neuropsychopharmacology* 2022; **47**(10): 1739-45.
9. Brick LA, Benca-Bachman CE, Johnson EC, Gustavson DE, Carper M, Palmer RH. Genetic associations among internalizing and externalizing traits with polysubstance use among young adults. *medRxiv* 2023.
10. Cambron C, Kosterman R, Catalano RF, Guttmannova K, Hawkins JD. Neighborhood, Family, and Peer Factors Associated with Early Adolescent Smoking and Alcohol Use. *J Youth Adolesc* 2018; **47**(2): 369-82.
11. Dick DM. Gene-environment interaction in psychological traits and disorders. *Annu Rev Clin Psychol* 2011; **7**: 383-409.
12. Rice F, Harold GT, Shelton KH, Thapar A. Family conflict interacts with genetic liability in predicting childhood and adolescent depression. *J Am Acad Child Adolesc Psychiatry* 2006; **45**(7): 841-8.
13. Fan CC, Loughnan R, Wilson S, Hewitt JK, Group AGW. Genotype Data and Derived Genetic Instruments of Adolescent Brain Cognitive Development Study((R)) for Better Understanding of Human Brain Development. *Behav Genet* 2023; **53**(3): 159-68.
14. Grotzinger A, Rhemtulla M, de Vlaming R, et al. Genomic SEM provides insights into the multivariate genetic architecture of complex traits. *Nat Hum Behav* 2019; **3**: 513-25.
15. Sanchez-Roige S, Palmer AA, Fontanillas P, et al. Genome-Wide Association Study Meta-Analysis of the Alcohol Use Disorders Identification Test (AUDIT) in Two Population-Based Cohorts. *Am J Psychiatry* 2019; **176**(2): 107-18.
16. Johnson EC, Demontis D, Thorgeirsson TE, et al. A large-scale genome-wide association study meta-analysis of cannabis use disorder. *Lancet Psychiatry* 2020; **7**(12): 1032-45.
17. Cabana-Dominguez J, Shivalikanjli A, Fernandez-Castillo N, Cormand B. Genome-wide association meta-analysis of cocaine dependence: Shared genetics with comorbid conditions. *Prog Neuropsychopharmacol Biol Psychiatry* 2019; **94**: 109667.
18. Sanchez-Roige S, Fontanillas P, Jennings MV, et al. Genome-wide association study of problematic opioid prescription use in 132,113 23andMe research participants of European ancestry. *Mol Psychiatry* 2021; **26**(11): 6209-17.

19. Polimanti R, Walters RK, Johnson EC, et al. Leveraging genome-wide data to investigate differences between opioid use vs. opioid dependence in 41,176 individuals from the Psychiatric Genomics Consortium. *Mol Psychiatry* 2020; **25**(8): 1673-87.
20. Watanabe K, Stringer S, Frei O, et al. A global overview of pleiotropy and genetic architecture in complex traits. *Nat Genet* 2019; **51**(9): 1339-48.
21. Ge T, Chen CY, Ni Y, Feng YA, Smoller JW. Polygenic prediction via Bayesian regression and continuous shrinkage priors. *Nat Commun* 2019; **10**(1): 1776.
22. Zhi D, Jiang R, Pearlson G, et al. Triple Interactions Between the Environment, Brain, and Behavior in Children: An ABCD Study. *Biol Psychiatry* 2024; **95**(9): 828-38.
23. Green R, Wolf BJ, Chen A, et al. Predictors of Substance Use Initiation by Early Adolescence. *Am J Psychiatry* 2024; **181**(5): 423-33.
24. Doran N, Gonzalez MR, Courtney KE, et al. Social cognitive influences associated with susceptibility to nicotine and tobacco use in youth in the ABCD Study. *J Stud Alcohol Drugs* 2024.
25. Slotkin TA. Nicotine and the adolescent brain: insights from an animal model. *Neurotoxicol Teratol* 2002; **24**(3): 369-84.
26. Brendgen M, Boivin M, Vitaro F, Girard A, Dionne G, Perusse D. Gene-environment interaction between peer victimization and child aggression. *Dev Psychopathol* 2008; **20**(2): 455-71.
27. Gruenewald PJ, Freisthler B, Remer L, Lascala EA, Treno A. Ecological models of alcohol outlets and violent assaults: crime potentials and geospatial analysis. *Addiction* 2006; **101**(5): 666-77.
28. Lee JG, Henriksen L, Rose SW, Moreland-Russell S, Ribisl KM. A Systematic Review of Neighborhood Disparities in Point-of-Sale Tobacco Marketing. *Am J Public Health* 2015; **105**(9): e8-18.
29. Garavan H, Bartsch H, Conway K, et al. Recruiting the ABCD sample: Design considerations and procedures. *Dev Cogn Neurosci* 2018; **32**: 16-22.
30. Sullivan RM, Wade NE, Wallace AL, et al. Substance use patterns in 9 to 13-year-olds: Longitudinal findings from the Adolescent Brain Cognitive Development (ABCD) study. *Drug Alcohol Depend Rep* 2022; **5**.

Figure Legends

Figure 1. (**A**) Lifetime prevalence of no substance use (NSU), single substance use (SSU), and polysubstance use (PSU) from baseline to Year 3. (**B**) Lifetime prevalence of alcohol (ALC), nicotine (NIC), cannabis (CAN), and other substance use (OTH) from baseline to Year 3. (**C**) PSU patterns—co-occurrence of lifetime use among the four substances. (**D**) Pairwise tetrachoric correlations of lifetime use among the four substances. *, $P_{FDR} < 0.05$.

Figure 2. (**A**) Associations between addiction risk polygenetic score (PGS) and lifetime polysubstance use (PSU) initiation by Year 3 within each genetic ancestry group, presented as the odds ratios (OR) with 95% confidence intervals. (**B**) Pseudo R^2 values explained by PGS for PSU within each genetic ancestry group, controlling for age, sex, genetic principal components, and enrollment sites. Statistical significance is indicated as follows: *, $P_{FDR} < 0.05$.
Abbreviations: SUD, substance use disorder; AUD, alcohol use disorder; CUD, cannabis use disorder; COD, cocaine use disorder; OUD, opioid use disorder; ND, nicotine dependence; EUR, European; AFR, African/Admixed African American; AMR, Hispanic/Admixed Latin American.

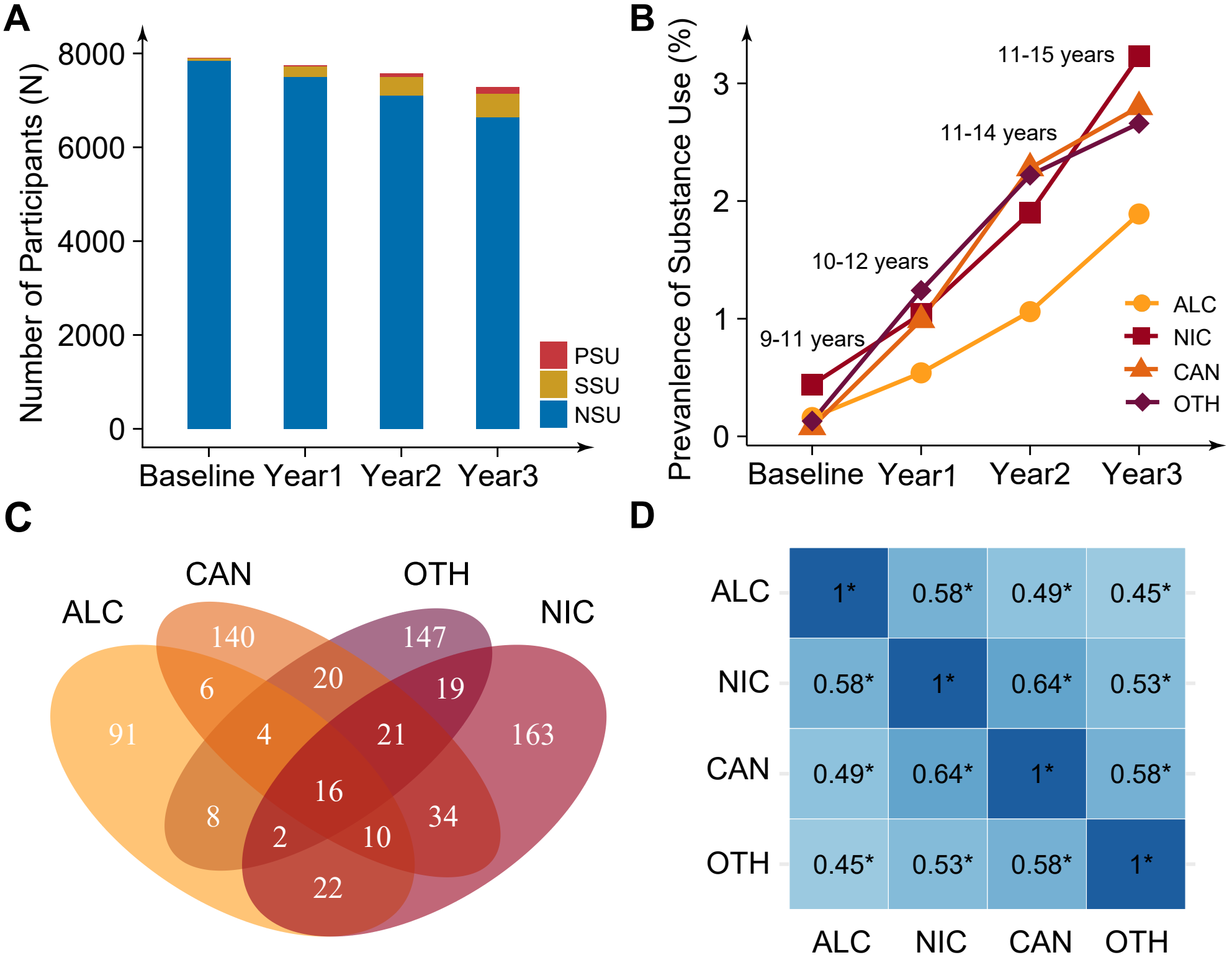
Figure 3. Associations between environmental measures and polysubstance use (PSU). (**A**) Among 96 environmental measures, 67 (70%) demonstrated significant associations with PSU when compared to no substance use (NSU) after controlling for age, sex, ethnicity, and site confounders. Environmental factors with non-significant associations are shown in lighter colors. The fifteen highest-ranking risk factors (including substance availability) and the five most protective factors with the strongest numerical associations with substance use are highlighted. (**B**) Mean and standard deviation (z-score normalized) of the top five risk factors showing the strongest numerical associations between PSU and NSU, as well as the top five protective factors. Aside from planned pregnancy and family intellectual culture, all other exposures indicated significant associations not only between PSU and NSU but also between PSU and single substance use (SSU). For planned pregnancy, family income, caregiver education, family cohesion, and family intellectual culture, higher values suggest less substance use, while for all other measures, higher values correlate with greater substance use. The Y-axis represents the mean of environmental exposure with z-score normalization, while the X-axis delineates the three substance use groups: NSU, SSU, and PSU, respectively.

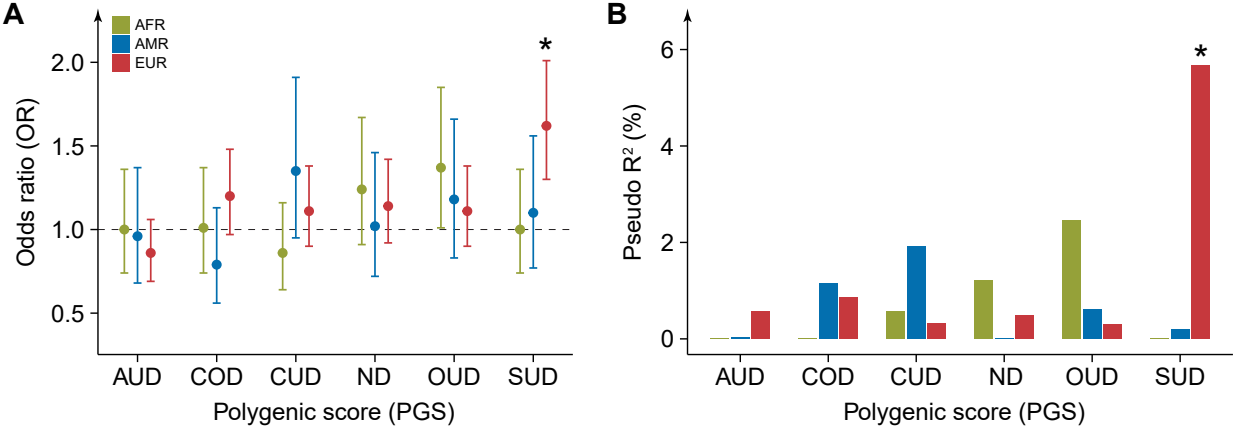
Figure 4. (**A**) Interaction effects between environmental exposures and polygenic risk scores (PGS) on polysubstance use (PSU) compared to no substance use (NSU), as well as PSU compared to single substance use (SSU). Data are presented as odds ratios (OR) with 95% confidence intervals (CI). The widths of the lines extending from the center points illustrate the 95% CI. (**B**) The four significant interactions for PSU versus SSU demonstrate the associations between environmental exposures and substance use at both high and low genetic risk levels. The red line indicates high genetic risk levels, while the blue line denotes low genetic risk levels, along with the 95% CI. *, $P_{FDR} < 0.05$.

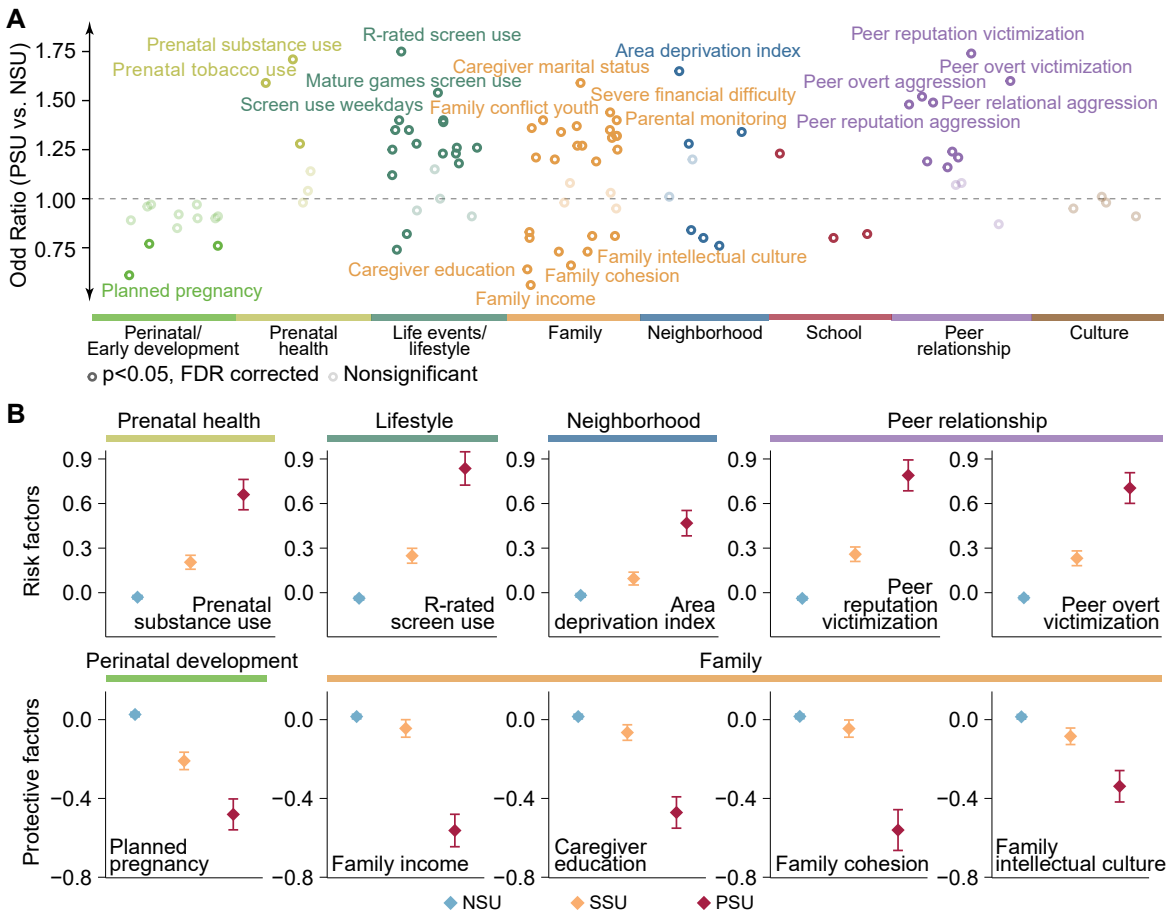
Table 1. Demographic and Substance Use Outcomes of Study Participants

	Total participants	Participants in main analysis	Statistics	EUR	AFR	AMR
Number	11,868	7,898		5,002	1,478	1,418
Male	6,188 (52.1%)	4,150 (52.5%)	χ^2 = 0.28; P = 0.6	2,646 (52.9%)	750 (50.7%)	754 (53.2%)
Female	5,677 (47.8%)	3,748 (47.5%)	NA	2,356 (47.1%)	728 (49.3%)	664 (46.8%)
Age, months	155 (7.8)	154.8 (7.7)	T = 1.23; P = 0.22	154.9 (7.7)	154.6 (7.6)	154.5 (7.9)
Substance use outcomes						
Alcohol use	241 (2.0%)	159 (2.0%)	NA	122 (2.4%)	19 (1.3%)	18 (1.3%)
Cannabis use	358 (3.0%)	251 (3.2%)	NA	137 (2.7%)	65 (4.4%)	49 (3.5%)
Nicotine use	414 (3.5%)	287 (3.6%)	NA	158 (3.2%)	72 (4.9%)	57 (4%)
Other substance use	350 (2.9%)	237 (3.0%)	NA	132 (2.6%)	52 (3.5%)	53 (3.7%)
No substance use	10,846 (91.4%)	7,195 (91.1%)	NA	4,574 (91.4%)	1,331 (90.1%)	1,290 (91%)
Aggregated outcomes						
Single substance use	777 (6.5%)	541 (6.8%)	NA	342 (6.8%)	104 (7%)	95 (6.7%)
Polysubstance use	245 (2.1%)	162 (2.1%)	NA	86 (1.7%)	43 (2.9%)	33 (2.3%)

Note: Chi-square tests were used to assess sex differences, and two-sample *t*-tests were used to assess age differences between the full and main analysis samples. AFR, African/Admixed African American; AMR, Hispanic/Admixed Latin American; EUR, European; NA, not applicable.

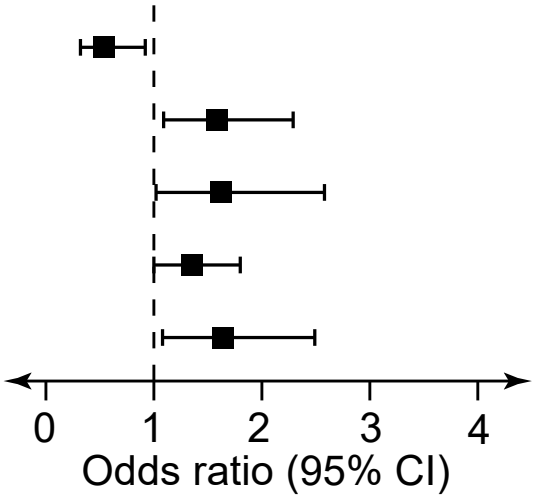

A
Number of Participants (N)
8000
6000
4000
2000
0
Baseline
Year1
Year2
Year3
PSU
SSU
NSU
B
Prevanlence of Substance Use (%)
3
2
1
0
11-15 years
11-14 years
10-12 years
9-11 years
ALC
NIC
CAN
OTH
Baseline
Year1
Year2
Year3
C
ALC
CAN
OTH
NIC
140
147
6
20
19
91
4
21
163
16
8
34
2
10
22
D
ALC
NIC
CAN
OTH
1*
0.58*
0.49*
0.45*
0.58*
1*
0.64*
0.53*
0.49*
0.64*
1*
0.58*
0.45*
0.53*
0.58*
1*
ALC
NIC
CAN
OTH

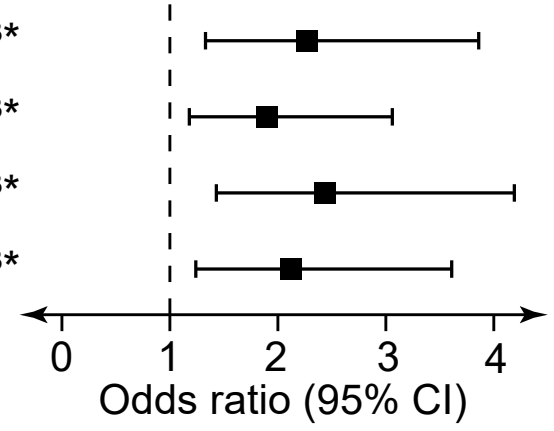
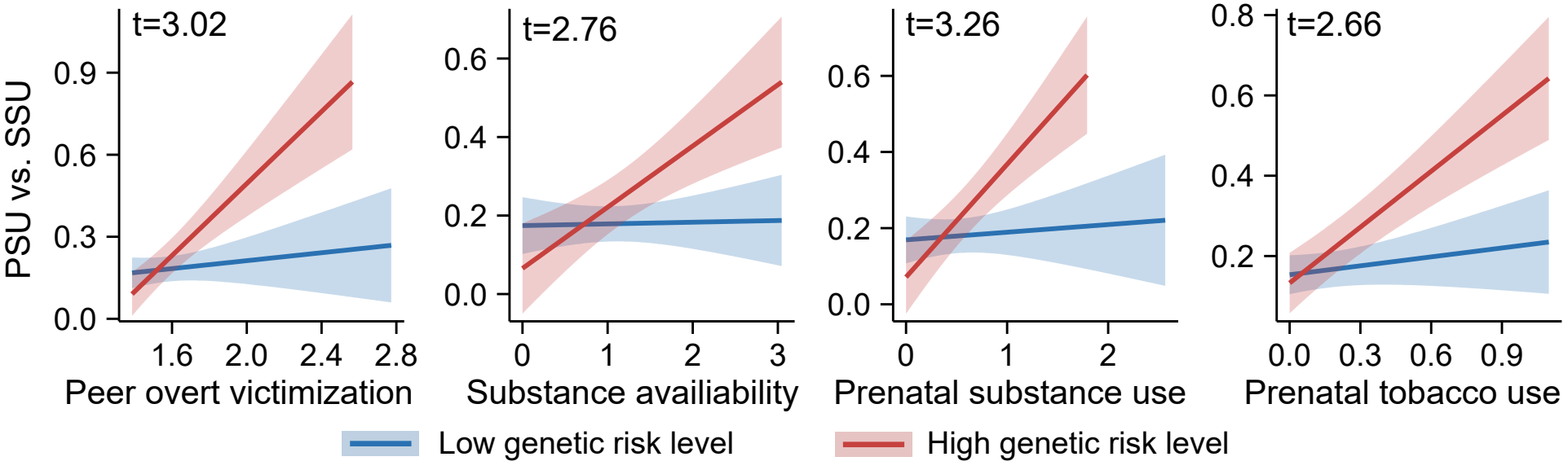

A
AFR
AMR
EUR
Odds ratio (OR)
2.0
1.5
1.0
0.5
*
AUD
COD
CUD
ND
OUD
SUD
Polygenic score (PGS)
B
Pseudo R² (%)
6
4
2
0
*
AUD
COD
CUD
ND
OUD
SUD
Polygenic score (PGS)


A
Odd Ratio (PSU vs. NSU)
1.75
1.50
1.25
1.00
0.75
Prenatal substance use
Prenatal tobacco use
R-rated screen use
Mature games screen use
Screen use weekdays
Caregiver marital status
Family conflict youth
Severe financial difficulty
Parental monitoring
Area deprivation index
Peer reputation victimization
Peer overt victimization
Peer overt aggression
Peer relational aggression
Peer reputation aggression
Planned pregnancy
Caregiver education
Family income
Family cohesion
Family intellectual culture
Perinatal/ Early development
Prenatal health
Life events/ lifestyle
Family
Neighborhood
School
Peer relationship
Culture
p<0.05, FDR corrected
Nonsignificant
B
Prenatal health
Lifestyle
Neighborhood
Peer relationship
Risk factors
0.9
0.6
0.3
0.0
Prenatal substance use
R-rated screen use
Area deprivation index
Peer reputation victimization
Peer overt victimization
Perinatal development
Family
Protective factors
0.0
−0.4
−0.8
Planned pregnancy
Family income
Caregiver education
Family cohesion
Family intellectual culture
NSU
SSU
PSU

A

Interaction term	Number of participants	OR (95% CI)	P value
PSU vs. NSU			
Sleep health*PGS	4552	1.35 (1.00, 1.80)	0.047
Peer relation victimization*PGS	3532	0.54 (0.32, 0.92)	0.023
Parental emotion issues*PGS	4361	1.62 (1.02, 2.58)	0.043
Maternal tobacco use*PGS	4658	1.64 (1.08, 2.49)	0.021
Substance availiability*PGS	4398	1.58 (1.09, 2.29)	0.015
PSU vs. SSU			
Peer overt victimization*PGS	425	2.44 (1.43, 4.19)	1.21×10^{-3}*
Prenatal tobacco use*PGS	413	1.90 (1.18, 3.06)	8.14×10^{-3}*
Prenatal substance use*PGS	428	2.12 (1.24, 3.61)	5.95×10^{-3}*
Substance availiability*PGS	408	2.27 (1.33, 3.86)	2.70×10^{-3}*





B