Deak Joseph (Orcid ID: 0000-0002-0540-6080)
Liu Mengzhen (Orcid ID: 0000-0001-6550-6959)

Alcohol and Nicotine Polygenic Scores are Associated with the Development of Alcohol and Nicotine Use Problems from Adolescence through Young Adulthood

Joseph D. Deak, Ph.D.[1, 2], D. Angus Clark, Ph.D.[3], Mengzhen Liu. Ph.D.[4], Jonathan D., Schaefer, Ph.D.[4], SeonKyeong Jang, M.A.[4], C. Emily Durbin, Ph.D.[5], William G. Iacono, Ph.D.[4], Matt McGue, Ph.D.[4], Scott I. Vrieze, Ph.D.[4], & Brian M. Hicks, Ph.D.[3]

[1] Yale University

[2] VA Connecticut Healthcare System

[3] University of Michigan

[4] University of Minnesota

[5] Michigan State University

Acknowledgement: This work was supported by United States Public Health Service grants R37 AA09367 (McGue), R01 AA024433 (Hicks), T32 AA007477 (Blow), and T32 AA028259 (Vasilious) from the National Institute of Alcohol Abuse and Alcoholism and R01 DA034606 (Hicks), R37 DA005147 (Iacono), R01 DA013240 (Iacono), R01 DA044283 (Vrieze), R01

DA037904 (Vrieze), R01 DA042755 (McGue/Vrieze) and U01 DA046413 (Vrieze) from the National Institute on Drug Abuse.

Word Count: 3,848

Conflict of Interest Statement: None

Correspondence: Address correspondence to Joseph D. Deak, Department of Psychiatry, Human Genetics Division, Yale University School of Medicine, New Haven, CT. Electronic mail may be sent to joseph.deak@yale.edu.

Abstract

Background and Aims: Molecular genetic studies of alcohol and nicotine use have identified many genome-wide association study (GWAS) loci. We measured associations between drinking and smoking polygenic scores (PGS) and trajectories of alcohol and nicotine use outcomes from late childhood to early adulthood, substance-specific versus broader-liability PGS effects, and if PGS performance varied for consumption versus problematic substance use.

Design, setting, participants, and measurements: We fit latent growth curve models with structured residuals to scores on measures of alcohol and nicotine use and problems from age 14 to age 34. We then estimated associations between the intercept (initial status) and slope (rate of change) parameters and PGSs for drinks per week (DPW), problematic alcohol use (PAU), cigarettes per day (CPD), and ever being a regular smoker (SMK), controlling for sex and genetic principal components. All data were analyzed in the United States. PGSs were calculated for participants of the Minnesota Twin Family Study (*N*=3225) using results from the largest GWAS of alcohol and nicotine consumption and problematic use to date.

Findings: Each PGS was associated with trajectories of use for their respective substances (i.e., DPW [β_{mean}=0.08; β_{range}=0.02-0.12] and PAU [β_{mean}=0.12; β_{range}=-0.02-0.31] for alcohol; CPD [β_{mean}=0.08; β_{range}=0.04-0.14] and SMK [β_{mean}=0.18; β_{range}=0.05-0.36] for nicotine). The PAU and SMK PGSs also exhibited cross-substance associations (i.e., PAU for nicotine-specific intercepts, and SMK for alcohol intercepts and slope). All identified SMK PGS effects remained

as significant predictors of nicotine and alcohol trajectories (β_{mean}=0.15; β_{range}=0.02-0.33), even after adjusting for the respective effects of all other PGSs.

Conclusions: Substance use-related polygenic scores (PGSs) vary in the strength and generality versus specificity of their associations with substance use and problems over time. The regular smoker PGS appears to be a robust predictor of substance use trajectories and seems to measure both nicotine-specific and non-specific genetic liability for substance use, and potentially externalizing problems in general.

Introduction

Alcohol and nicotine use, respectively, contribute to 3 million (5.3%) and 7 million (12.3%) deaths worldwide each year, making both leading causes of global mortality[1,2]. Studies have demonstrated that genetic factors influence alcohol and nicotine use, and risk for alcohol use disorder (AUD) and nicotine use disorder (NicUD). Twin studies report heritability estimates of approximately 50% for AUD[3] and NicUD[4], and large-scale genome-wide association studies (GWAS) have identified hundreds of loci that exhibit genome-wide significant associations with alcohol and nicotine use phenotypes[5-7], providing new avenues for research on the genetic influences on substance use.

Polygenic scores (PGS) are one method for modeling aggregate genetic risk across the genome, and have provided valuable information about the unique and shared genetic influences on alcohol and nicotine use. PGS can be generated from a GWAS discovery sample by weighting genetic variants relative to the strength of their association with a given phenotype to calculate a measure of individual genetic risk in a target sample. For example, PGS calculated from GWAS-identified associations for alcohol use have demonstrated associations with alcohol-related outcomes in independent samples[8-10]. PGSs for alcohol and nicotine use have also been associated with use of other drugs (e.g., cannabis, cocaine, amphetamines, ecstasy, hallucinogens)[5,11], suggesting these PGS also index non-specific genetic influences on substance use.

While studies using PGS are beginning to trace the contours of the genetic architecture of substance use, they have yet to examine the influence of aggregate genetic risk on patterns of substance use over time. This is an important next step, because alcohol and nicotine use exhibit strong age-related mean-level trends, with typical initiation in adolescence followed by peak use in young adulthood and normative declines in heavy use and substance use disorders (SUDs) by age 30[12]. Understanding the etiology of substance use then requires accounting for these normative patterns of emergence, escalation, and decline, and there is some evidence that genetic influences for substance use varies across development[13,14].

Initial efforts using PGSs to examine associations with alcohol and nicotine use developmental trajectories have had modest success. A PGS for cigarettes smoked per day predicted later cigarette smoking and NicUD in early adulthood[15,16], but not alcohol use[15], suggesting the PGS measured substance-specific genetic influences on nicotine use. Evidence for longitudinal associations of alcohol-related PGS has been mixed. One study found that a PGS for AUD was associated with alcohol use in males at age 15.5 and greater increases of alcohol use at age 21.5[17], while other studies examining alcohol use in college student drinkers over a four-year timespan with an environment enriched for substance use have returned both positive[18] and null results[19]. Most prior studies were limited by smaller GWAS discovery samples relative to recent large-scale GWAS of alcohol and nicotine use outcomes.

We address these limitations using PGS measures derived from the large-scale GWAS of alcohol consumption (i.e., drinks per week; DPW)[5] and problematic alcohol use (PAU)[7],

given prior evidence suggesting differences in the polygenic architecture of alcohol use versus alcohol use problems[20,21], and PGS for cigarettes per day (CPD)[5] and for initiation of regular smoking (SMK)[5]. The nicotine-related PGSs have demonstrated varying degrees of genetic correlation (r_g) with nicotine dependence (r_g=0.42 and 0.95, for SMK and CPD, respectively)[6], and therefore, may index varying degrees of nicotine use versus NicUD, or nicotine-specific effects versus externalizing effects more broadly.

We examined respective PGS associations with trajectories of alcohol and nicotine use (quantity and frequency) and problems (AUD and NicUD criteria) from late childhood through young adulthood (ages 14-34), a period that covers normative initiation, peak use, and elevated risk for SUDs. Strengths of this approach include the ability to make stronger inferences about when in the developmental progression of substance use (e.g., initiation of use, escalation of use) these genetic influences have their effects, and the long follow-up period ensures that polygenic influences for alcohol and nicotine use are likely to have been expressed for most people. A second aim was to examine whether the associations between polygenic measures thought to index regular use versus problematic use differed for phenotypic measures of consumption (quantity/frequency) and substance use disorder (AUD and NicUD). A final aim was to examine whether the associations of the respective alcohol and nicotine-related PGS were limited to the specific substance or generalized to trajectories of both alcohol and nicotine outcomes, and determine if there was evidence of incremental predictive utility across the respective PGSs.

Methods

Participants

Participants were members of the Minnesota Twin Family Study (MTFS), a longitudinal study of 3762 (52% female) twins (1881 pairs) investigating the development of SUDs and related conditions[22-24]. All twin pairs were the same sex and living with at least one biological parent within driving distance to the University of Minnesota laboratories at the time of recruitment. Exclusion criteria included any cognitive or physical disability that would interfere with study participation. Twins were recruited the year they turned either 11-years old (*n*=2510; the younger cohort) or 17-years old (*n*=1252; the older cohort). Twins in the younger cohort were born from 1977 to 1984 and 1988 to 1994, while twins in the older cohort were born between 1972 and 1979. Families were representative of the area they were drawn from in terms of socioeconomic status (SES), mental health treatment history, and urban vs rural residence [22]. Consistent with the demographics of Minnesota for the target birth years, 96% of participants reported non-Hispanic White ethnicity and race.

The younger cohort was assessed at ages 11 (M_{age}=11.78 years; SD=0.43 years) and 14 (M_{age}=14.90 years; SD=0.31 years), and all twins were assessed at target ages 17 (M_{age}=17.85 years; SD=0.64 years), 21 (M_{age}=21.08 years; SD=0.79 years), 24 (M_{age}=24.87 years; SD=0.94 years), and 29 (M_{age}=29.43 years; SD=0.67 years). A subgroup of twins from the younger cohort were also assessed at age 34 (M_{age}=34.62 years; SD=1.30 years). Supplemental Table 1 provides the number of participants for each assessment and descriptive statistics for the study measures. Analyses were conducted both with and without participants that consistently abstained from

substance use across time. Conclusions were similar across these models, so we report models using the full sample. Participation rates ranged from 80% to 93% among those recruited for a given follow-up assessment. The total sample included 1205 monozygotic (51.5% female) and 676 dizygotic (52.8% female) twin pairs[22,25].

Alcohol Use and AUD. All alcohol and nicotine variables were assessed during structured clinical interviews, while the use variables were also assessed using a computerized self-report questionnaire at ages 11, 14, and 17 that was completed in private. Alcohol variables included the average number of drinks per occasion in the past 12 months (i.e., alcohol quantity) and *DSM-III-R* symptoms of alcohol abuse and dependence (the diagnostic system when the study began, hereafter referred to as AUD symptoms). Free responses to alcohol quantity and the number of alcohol abuse and dependence symptoms were converted to scales that ranged from 0 to 8 (corresponding integer values were used for lower values [e.g., 1=1 symptom] with AUD symptoms capped at 8 and drinks per occasion coded as 7=7-9 drinks and 8=10 or more to reduce skew due to a small number of high values). The lifetime prevalence of *DSM-III-R* AUD (≥3 symptoms of abuse or dependence) was 26%.

Nicotine Use and NicUD**.** Nicotine variables included average cigarettes per day (or equivalent form of tobacco, e.g., chewing tobacco) and *DSM-III-R* symptoms of nicotine dependence (hereafter referred to as NicUD). Free responses were converted to a 0 to 6 scale for nicotine quantity (0=0 , 1=1, 2=2, 3=3, 4=4-6, 5=7 or more) and NicUD symptoms (number of symptoms capped at 6 to reduce skew). The lifetime prevalence of *DSM-III-R* NicUD was 33%.

PGS Methods. PGS were generated using the largest GWAS of nicotine and alcohol use traits to date. The DPW (average number of drinks per week; $N = 941,280$), CPD (average number of cigarettes per day; $N = 337,334$), and SMK (ever smoked regularly in lifetime; $N = 1,232,091$) PGSs were calculated using results of the GWAS & Sequencing Consortium of Alcohol and Nicotine use (GSCAN)[5], and the PAU PGS was calculated using results from the largest GWAS of AUD and the problems subscale of the Alcohol Use Disorder Identification Test (AUDIT-P) to date ($N = 435,563$)[7]. The MTFS sample was removed from the original GSCAN discovery sample to avoid overlap with the target sample[5]. PGS were created for participants of European ancestry, confirmed via principal components analysis[26], in the MTFS target sample following imputation to the most recent Haplotype Reference Consortium reference panel[27] and restricted to autosomal HapMap3 variants with a minor allele frequency ≥ 0.01 and an imputation quality > 0.7. The resulting filtered variants (~1 million variants) were then submitted to LDpred[28] to generate beta weights in the MTFS sample including variants of all significance levels (p-value ≤ 1). Individual PGS were then calculated in PLINK[26] for all individuals with phenotypic and genotypic data for the present study ($n = 3225$).

Data Analytic Strategy

Latent growth models with structured residuals (LGM-SR; see **Figure 1**) were used to model developmental trends in the alcohol and nicotine use outcomes[29,30]. These models include intercept factors that reflect status at the first time point (age 14 as there was almost no substance use at age 11), and slope factors that reflect the rate of change over the course of the

study. Slope factors were specified using a latent basis approach. That is, the first and last basis coefficient were fixed to 0 and 1, respectively, and the intervening coefficients were estimated, which provides a parsimonious way of capturing non-linear trajectories[31][1]. Intercept and slope factors were allowed to vary to capture individual differences in growth. The residual structure included occasion-specific latent factors that account for deviations from the intercept and slope implied trajectories. The autoregressive paths linking adjacent residual factors capture associations between variables over time after accounting for general growth trend (**Figure 1**) and were included because not accounting for residual autoregressive effects can lead to biased variance estimates in the growth factors[32,33].

Unconditional LGM-SR models were first fit to each outcome (**Figure 1-Panel a**). Conditional models were then fit in which the growth factors were regressed on a single PGS and the control variables (1 PGS Conditional Latent Growth Model; **Figure 1-Panel b**). The control variables included participant sex and the first five genetic principal components[34] to adjust for underlying ancestral substructure (associations between the control variables and the intercept and slope factors with and without PGS included can be found in the supplemental material). Sensitivity checks adjusting for birth cohort were also completed; the inclusion of birth cohort as a covariate did not affect the main conclusions. Conditional models were then estimated in which the growth factors were regressed on two PGSs simultaneously along with

[1] Alternative specifications of the growth model (e.g., piecewise models) were considered, and lead to the same conclusions as reported here

the control variables (2 PGS Conditional Latent Growth Model; **Figure 1-Panel c**). The 2 PGS models included both PGSs associated with a specific substance as predictors of the latent growth factors. That is, the 2 PGS models either included the DPW and PAU PGSs, or the CPD and SMK PGSs. Finally, conditional models were estimated in which the growth factors were regressed on all four PGSs simultaneously along with the control variables (4 PGS Conditional Latent Growth Model; **Figure 1-Panel d**). All major analyses were conducted using Mplus v8.4[35] with full information maximum likelihood estimation[36]. Every model included both twins from a given twin pair; as the focus was on the general growth trends over time across the sample, models were not stratified by zygosity. 95% confidence intervals (CIs) were derived using clustered (by family) percentile bootstrapping (with 1000 draws). This procedure performs well when estimating CIs for skewed variables such as substance use, and accounts for the family-based clustering of the observations (i.e., both twins from a given family of origin being included in the models)[37]. The present analysis plan was not pre-registered on a publicly available platform and should be considered largely exploratory in nature.

Results

Descriptive information for the study variables are reported in **Table 1**. Mean-levels of the alcohol and nicotine use outcomes increased from age 11 to age 20, and then decreased from age 20 to age 34. The rank-order stability of the alcohol and nicotine use outcomes between adjacent time points ranged from r=0.33-0.79.

The univariate models for alcohol and nicotine use related outcomes all fit the data well by conventional standards (Supplemental Table 1)[38]. Parameter estimates were consistent with the observed trajectories, suggesting a rise in alcohol and nicotine use throughout adolescence, and then a gradual decline in values after age 20. There was a statistically significant degree of variability in all of the slope factors, and all of the intercept factors except the alcohol quantity intercept factor (i.e., there was too little variability in alcohol quantity at age 14 to effectively estimate the intercept factor variance; Supplemental Table 2). Covariances between the intercept and slope factors were generally small and nonsignificant (likely in part due to the low rates of endorsement in early adolescence); on the other hand, the autoregressive coefficients were typically statistically significant and positive in sign, indicating that substance use beyond that predicted by the growth model at one time point was associated with similarly elevated substance use at subsequent time points. Correlations across the respective PGSs in the MTFS sample are reported in **Table 2**.

Standardized path coefficients from the single (1) PGS, two (2) PGS (adjusting for the other within substance PGS; e.g., covarying for DPW in the PAU 2 PGS analysis, and vice versa), and four (4) PGS (adjusting for all other PGS simultaneously; e.g., covarying for CPD, DPW, and PAU in the SMK 4 PGS analysis) models can be found in **Table 3**.

1 PGS Model Results

Drinks Per Week (DPW)

In the single PGS models, the DPW PGS had significant associations with the intercept and slope factors of AUD, and the slope factor for alcohol quantity. The DPW PGS also had significant cross-substance associations with the intercept factors for NicUD and nicotine quantity. DPW PGS effect sizes were small (β_{mean}=0.08; β_{range}=0.02-0.12) and comparable to the PAU PGS.

Problematic Alcohol Use (PAU)

The PAU PGSs had significant associations with the intercept and slope factors of AUD and alcohol quantity. The PAU PGSs also had significant cross-substance associations with the intercept factors for NicUD and nicotine quantity. Effect sizes were small and comparable to the DPW PGS, with no statistically-significant differences between the two PGS, but slightly larger effects for the PAU (β_{mean}=0.12; β_{range}=-0.02-0.31).

Cigarettes Per Day (CPD)

The CPD PGS had significant associations with the intercept and slope factors for NicUD, and the intercept for nicotine quantity (β_{mean}=0.08; β_{range}=0.04-0.14).

Regular Smoking (SMK)

The SMK PGS had significant associations with the intercept and slope factors of NicUD and nicotine quantity. The SMK PGS also had significant cross-substance associations with the intercept and slope factors for AUD, and the intercept for alcohol quantity. Effect sizes were

small to medium, and slightly larger for the SMK PGS (β_{mean}=0.18; β_{range}=0.05-0.36) relative to the CPD PGS.

2 PGS Model Results

Alcohol Use PGSs (DPW and PAU)

When the DPW and PAU PGSs were included in the same model (2 PGS model), most PAU PGS effects observed in the single PGS model remained significant (β_{mean}=0.10; β_{range}=-0.03-0.30); however, the DPW effects were reduced to the point that most CIs included zero (β_{mean}=0.05; β_{range}=0.01-0.09). Adjusting for the effects of DPW, the PAU PGS had significant associations with the intercept and slope factors for AUD, and the intercept factors for alcohol quantity, NicUD, and nicotine quantity. Only the association between the PAU PGS and the slope factor for alcohol quantity was no longer significant (□=0.03, 95% CI: -0.04,0.11) after adjusting for the DPW PGS. In contrast, only the association between the DPW PGS and the intercept factor for AUD remained significant (□=0.09, 95% CI: 0.03,0.19) after adjusting for the PAU PGS.

Nicotine Use PGSs (CPD and SMK)

When the CPD and SMK PGSs were included in the same model (2 PGS model), the SMK PGS (β_{mean}=0.17; β_{range}=0.04-0.35) exhibited notably stronger associations than the CPD PGS (β_{mean}=0.05; β_{range}=0.02-0.09). Adjusting for the effects of CPD, the SMK PGS continued to

have significant associations with intercept and slope factors for NicUD, nicotine quantity, and AUD, and the intercept factor for alcohol quantity, and the effect sizes nearly the same as those in the one PGS model (**Table 3**). In contrast, none of the associations between the CPD PGS and the growth factors for the nicotine and alcohol use measures remained significant after adjusting for the SMK PGS.

Combined Alcohol and Nicotine PGS (4 PGS) Model Results

When the four alcohol and nicotine PGSs were included in the same model (i.e., examining respective PGS effects while adjusting for the other three PGSs simultaneously), the SMK PGS exhibited the most robust associations across both the alcohol and nicotine outcomes (β_{mean}=0.15; β_{range}=0.02-0.33). After adjusting for all other PGSs simultaneously (i.e., covarying for CPD, DPW, and PAU), the SMK PGS continued to have significant associations with the intercept and slope factors for NicUD, nicotine quantity, and AUD, and the intercept factor for alcohol quantity with only a modest decline in effect sizes (**Table 3**). The PAU PGS continued to have significant associations with the slope factor for AUD (□=0.08, 95% CI: 0.02,0.13) and the intercept factor for NicUD (□=0.08, 95% CI: 0.01,0.16) when covarying for the other PGSs. The DPW continued to have a significant association with the intercept factor for AUD (□=0.08, 95% CI: 0.02,0.16) after adjusting for the other PGSs. None of the associations for the CPD PGS remained significant after adjusting for the other PGSs. Figure 2 depicts exemplar substance use growth trajectories from ages 14 to 34 years for persons with high and low scores (i.e., ±1.5

standard deviation around the mean) on one of the four respective PGS, those with high or low scores across all four PGS, and people with average scores on each PGS.

Discussion

Using a longitudinal design, we extended prior studies investigating genetic influences on substance use by showing PGSs for alcohol and nicotine use phenotypes were each associated with problem use in middle adolescence, and a greater rate of increase in alcohol and nicotine use problems through young adulthood. Our findings show that alcohol and nicotine use-related PGS are statistically-significant predictors of trajectories for problematic alcohol and nicotine use, and extend prior cross-sectional studies by demonstrating that polygenic liability for alcohol and nicotine use are informative about the developmental progression of alcohol and nicotine use.

We also examined whether associations with PGSs differed for regular use versus SUD symptoms. For alcohol, DPW and PAU PGSs exhibited similar predictive power independently, though most DPW PGS effects decreased below significance when both PAU and DPW were included in the same model. Multiple factors may explain differences in performance between the PAU and DPW PGSs. The PAU phenotype was largely defined by lifetime AUD diagnosis, while DPW was the average number of weekly drinks over shorter reporting periods (past week/past 12 months). The greater severity and broader reporting period might account for some of the differential performance of the PAU and DPW PGS in the MTFS sample. Sample characteristics of the discovery samples (e.g., treatment-seeking vs. population-based) and levels

of substance use problems in the target sample[9,10] might also account for differences between the PAU and DPW PGSs. PGS may perform better in target samples with similar degrees of problematic substance use found in the PGS discovery sample. The PAU PGS was derived from a sample with elevated rates of drinking and AUD diagnosis (i.e., Million Veteran Program), while DPW contains large samples (e.g., 23andMe; n= ~404,000) that are less representative of the population at-large (e.g., high SES, relatively healthy, lower alcohol-related problems). The MTFS sample is a community-based sample that *is* representative of its target population, including prevalence of SUDs (e.g., 26% and 33% meeting criteria for AUD and NicUD, respectively). Also, the multiple assessment schedule of the MTFS increased the detection of positive diagnoses and peak substance use, which aids in accurately capturing problem use, similar to the longitudinal assessment of AUD in the PAU discovery sample.

For nicotine use, we found that the CPD PGS was only associated with problematic nicotine use, consistent with prior estimates of a high genetic correlation between CPD and nicotine dependence (r_g = 0.95)[6]. In contrast, the SMK PGS was a statistically-significant predictor of all alcohol and nicotine use outcomes, even after adjusting for the CPD, PAU, and DPW PGSs, indicating the SMK PGS indexes genetic influences that have general effects on misuse of multiple substances. The SMK PGS has also been associated with the use of multiple other substances (e.g., alcohol, cannabis, cocaine)[5], and has been found to load strongly onto a latent factor composed of externalizing (EXT) traits[39]. Using the same sample as in this report, we have also shown that these non-specific effects extend beyond substance use to include

externalizing problems (rule breaking and aggression) from ages 11 to 17 years old (i.e., prior to peak substance use), even after adjusting for contemporaneous nicotine use[40]. The notion of common genetic influences across substances is also consistent with multivariate twin studies that posited a common genetic etiology to account for the co-occurrence and family transmission of substance use problems, antisocial behavior, and disinhibited personality traits[41-43]. Taken together, these findings are consistent with the interpretation that the SMK PGS measures genetic influences on nicotine use and behavioral disinhibition more broadly, while the other substance use-related PGSs are relatively substance-specific. Efforts to extend these findings to additional substances (e.g., cannabis), and to directly compare the effects of the SMK PGS to the EXT PGS[39], are underway.

While the ability to use PGS to demonstrate associations with longitudinal trajectories of substance use is promising, substantial advancement is needed before PGSs have clinical relevance for SUDs. Stratifying individuals based upon "polygenic risk" has been shown to aid in mitigating adverse health outcomes for some health conditions (e.g., coronary disease)[44]; similar success has yet to be demonstrated for SUDs. Concerns related to clinical utility include that current SUD PGS account for a relatively small proportion of variance for clinical phenotypes (~5%), especially in comparison to other risk factors (e.g., SES, family history), and the potential for patient discrimination based upon genetic information[45-47].

Limitations of the study include that the PGS were generated from GWAS of European ancestry. The degree to which our results generalize to other ancestral groups is uncertain[48].

This limitation has the potential to proliferate health disparities if these findings are only applicable to individuals of European ancestry, further prioritizing the importance of extending efforts to diverse ancestry groups[49]. Additionally, genetic influences on substance use are influenced by environmental factors, and genetic and environmental influences vary developmentally[15,50]. Studies examining how PGSs interact with environmental influences longitudinally are needed.

Despite these limitations, our findings are a successful extension of prior work by demonstrating substance-specific and generalized PGS effects on longitudinal trajectories of alcohol and nicotine use problems across late childhood and early adulthood. The results also provide initial evidence that the SMK PGS may index non-specific genetic risk for substance use and externalizing behaviors in general. This effort serves as a key step in demonstrating the influence of alcohol and nicotine PGS across developmental periods in which individuals initiate substance use, increase quantity and frequency of use, and begin to experience substance use problems. Our hope is that this work will aid in mitigating adverse health outcomes related to problematic substance use in the future.

References

1. World Health Organization, 2018. Global Status Report on Alcohol and Health 2018 Ed, World Health Organization, Geneva, Switzerland, 2018.

2. World Health Organization. WHO Report on the Global Tobacco Epidemic, 2017: Monitoring Tobacco Use and Prevention Policies. World Health Organization; 2017.

3. Verhulst B, Neale MC, Kendler KS. The heritability of alcohol use disorders: a meta-analysis of twin and adoption studies. Psychol Med. 2015 Apr;45(5):1061–72.

4. Agrawal A, Verweij KJH, Gillespie NA, et al. The genetics of addiction—a translational perspective. Translational psychiatry. 2012;2(7):e140–e140.

5. 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. Nature genetics. 2019;51(2):237–244.

6. Quach BC, Bray MJ, Gaddis NC, et al.. Expanding the genetic architecture of nicotine dependence and its shared genetics with multiple traits. Nature Communications. 2020;11(1). doi:10.1038/s41467-020-19265-z.

7. 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. Nature Neuroscience. Published online 2020:1–10.

8. Salvatore JE, Aliev F, Edwards AC, et al. Polygenic scores predict alcohol problems in an independent sample and show moderation by the environment. Genes. 2014;5(2):330–346.

9. Savage JE, Salvatore JE, Aliev F, et al. Polygenic risk score prediction of alcohol dependence symptoms across population-based and clinically ascertained samples. Alcoholism: Clinical and Experimental Research. 2018;42(3):520–530.

10. 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. Translational Psychiatry. 2020;10(1). doi:10.1038/s41398-020-00865-8.

11. Chang L-H, Couvy-Duchesne B, Liu M, et al. Association between polygenic risk for tobacco or alcohol consumption and liability to licit and illicit substance use in young Australian adults. Drug and alcohol dependence. 2019;197:271–279.

12. Jackson KM, Sartor CE. The natural course of substance use and dependence. In: Sher KJ, ed. The Oxford Handbook of Substance Use and Substance Use Disorders. New York, NY: Oxford University Press; 2016:67-134.

13. Malone SM, Taylor J, Marmorstein NR, McGue M, Iacono WG. Genetic and environmental influences on antisocial behavior and alcohol dependence from adolescence to early adulthood. Development and Psychopathology. 2004;16(4):943–966.

14. Bergen SE, Gardner CO, Kendler KS. Age-Related Changes in Heritability of Behavioral Phenotypes Over Adolescence and Young Adulthood: A Meta-Analysis. Twin Research and Human Genetics. 2007;10(3):423-433. doi:10.1375/twin.10.3.423

15. Vrieze SI, Hicks BM, Iacono WG, McGue M. Decline in genetic influence on the co-occurrence of alcohol, marijuana, and nicotine dependence symptoms from age 14 to 29. American Journal of Psychiatry. 2012;169(10):1073–1081.

16. Belsky DW, Moffitt TE, Baker TB, et al. Polygenic Risk and the Developmental Progression to Heavy, Persistent Smoking and Nicotine Dependence: Evidence From a 4-Decade Longitudinal Study. JAMA Psychiatry. 2013;70(5):534-542. doi:10.1001/jamapsychiatry.2013.736

17. Li JJ, Cho SB, Salvatore JE, et al. The impact of peer substance use and polygenic risk on trajectories of heavy episodic drinking across adolescence and emerging adulthood. Alcoholism: clinical and experimental research. 2017;41(1):65–75.

18. Ksinan AJ, Su J, Aliev F, et al. Unpacking genetic risk pathways for college student alcohol consumption: the mediating role of impulsivity. Alcoholism: Clinical and Experimental Research. 2019;43(10):2100–2110.

19. Su J, Kuo SI-C, Meyers JL, Guy MC, Dick DM. Examining interactions between genetic risk for alcohol problems, peer deviance, and interpersonal traumatic events on trajectories of alcohol use disorder symptoms among African American college students. Dev Psychopathol. 2018;30(5):1749-1761. doi:10.1017/S0954579418000962

20. Kranzler HR, Zhou H, Kember RL, et al. Genome-wide association study of alcohol consumption and use disorder in 274,424 individuals from multiple populations. Nature communications. 2019;10(1):1–11.

21. 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. American Journal of Psychiatry. 2019;176(2):107–118.

22. Iacono WG, Carlson SR, Taylor J, Elkins IJ, McGue M. Behavioral disinhibition and the development of substance-use disorders: findings from the Minnesota Twin Family Study. Development and psychopathology. 1999;11(4):869–900.

23. Keyes MA, Malone SM, Elkins IJ, Legrand LN, McGue M, Iacono WG. The enrichment study of the Minnesota twin family study: increasing the yield of twin families at high risk for externalizing psychopathology. Twin Research and Human Genetics. 2009;12(5):489–501.

24. Wilson S, Haroian K, Iacono WG, et al. Minnesota Center for Twin and Family Research. Twin Research and Human Genetics. Published online 2019:1–7.

25. McGue M, Zhang Y, Miller MB, et al. A genome-wide association study of behavioral disinhibition. Behavior genetics. 2013;43(5):363–373.

26. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience. 2015;4(1):s13742–015.

27. Das S, Forer L, Schönherr S, et al. Next-generation genotype imputation service and methods. Nature genetics. 2016;48(10):1284–1287.

28. Vilhjálmsson BJ, Yang J, Finucane HK, et al. Modeling linkage disequilibrium increases accuracy of polygenic risk scores. The american journal of human genetics. 2015;97(4):576–592.

29. Curran PJ, Howard AL, Bainter SA, Lane ST, McGinley JS. The separation of between-person and within-person components of individual change over time: A latent curve model with structured residuals. Journal of consulting and clinical psychology. 2014;82(5):879.

30. Berry D, Willoughby MT. On the practical interpretability of cross-lagged panel models: Rethinking a developmental workhorse. Child development. 2017;88(4):1186–1206.

31. Wu W, Selig JP, Little TD. Longitudinal data analysis. In: Little TD, editor. Oxford handbook of quantitative methods, Vol 2: Statistical Analysis. New York, NY: Oxford University Press; p. 387–410

32. Kwok O, West SG, Green SB. The impact of misspecifying the within-subject covariance structure in multiwave longitudinal multilevel models: A Monte Carlo study. Multivariate Behavioral Research. 2007;42(3):557–592.

33. Sivo S, Fan X, Witta L. The biasing effects of unmodeled ARMA time series processes on latent growth curve model estimates. Structural Equation Modeling. 2005;12(2):215–231.

34. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. Nature genetics. 2006;38(8):904–909.

35. Muthén LK, Muthén BO. Mplus user's guide. 8th ed. Los Angeles, CA

36. Allison, P. D. In Missing Data. Millsap RE, Maydeu-Olivares A. The SAGE Handbook of Quantitative Methods in Psychology. Sage Publications; 2009.

37. Falk CF. Are robust standard errors the best approach for interval estimation with nonnormal data in structural equation modeling? Structural Equation Modeling: A Multidisciplinary Journal. 2018;25(2):244–266.

38. West SG, Taylor AB, Wu W. Model fit and model selection in structural equation modeling. Handbook of structural equation modeling. 2012;1:209–231.

39. Linnér, R. K., Mallard, T. T., Barr, P. B., Sanchez-Roige, S., Madole, J. W., Driver, M. N., … Dick, D. M. (2020). Multivariate genomic analysis of 1.5 million people identifies genes related to addiction, antisocial behavior, and health. BioRxiv. doi:10.1101/2020.10.16.342501

40. Hicks BM, Clark DA, Deak JD, et al.. Polygenic Risk Score for Smoking is associated with Externalizing Psychopathology and Disinhibited Personality Traits but not Internalizing Psychopathology in Adolescence. 2020. doi:10.1101/2020.07.29.227405.

41. Hicks BM, Foster KT, Iacono WG, McGue M. Genetic and environmental influences on the familial transmission of externalizing disorders in adoptive and twin offspring. JAMA psychiatry. 2013;70(10):1076–1083.

42. Kendler KS, Prescott CA, Myers J, Neale MC. The structure of genetic and environmental risk factors for common psychiatric and substance use disorders in men and women. Archives of general psychiatry. 2003;60(9):929–937.

43. Krueger RF, Hicks BM, Patrick CJ, Carlson SR, Iacono WG, McGue M. Etiologic connections among substance dependence, antisocial behavior, and personality: modeling the externalizing spectrum. Published online 2009.

44. Khera, A. V. et al. Genetic risk, adherence to a healthy lifestyle, and coronary disease. N. Engl. J. Med. 2016; 375, 2349–2358.

45. Lambert SA, Abraham G, Inouye M. Towards clinical utility of polygenic risk scores. Human molecular genetics. 2019;28(R2):R133–R142.

46. Torkamani A, Wineinger NE, Topol EJ. The personal and clinical utility of polygenic risk scores. Nature Reviews Genetics. 2018;19(9):581.

47. Driver, M. N., Kuo, S. I.-C., & Dick, D. M. Genetic feedback for psychiatric conditions: Where are we now and where are we going. American Journal of Medical Genetics Part B: Neuropsychiatric Genetics; 2020.

48. Mostafavi H, Harpak A, Agarwal I, Conley D, Pritchard JK, Przeworski M. Variable prediction accuracy of polygenic scores within an ancestry group. Elife. 2020;9:e48376.

49. Martin AR, Kanai M, Kamatani Y, Okada Y, Neale BM, Daly MJ. Clinical use of current polygenic risk scores may exacerbate health disparities. Nature genetics. 2019;51(4):584.

50. Rose RJ, Dick DM, Viken RJ, Kaprio J. Gene-environment interaction in patterns of adolescent drinking: regional residency moderates longitudinal influences on alcohol use. Alcoholism: Clinical and Experimental Research. 2001;25(5):637–6

Table 1
Descriptive Statistics for Substance Use Variables Across Time

Age	14	17	20	24	29	34
Alcohol Use Disorder						
M	.08	.50	.90	.97	.60	.45
SD	.59	1.26	1.57	1.59	1.31	1.21
N	1952	2908	2456	2943	2316	802
Autocorrelation	.38	.45	.52	.49	.46	-
Alcohol Quantity						
M	.57	2.08	3.66	3.31	2.65	2.31
SD	1.21	2.76	2.55	2.19	1.90	1.69
N	1787	3033	2442	2926	2314	798
Autocorrelation	.33	.40	.52	.54	.52	-
Nicotine Use Disorder						
M	.27	.76	1.11	1.04	1.00	.82
SD	1.03	1.64	1.77	1.69	1.67	1.47
N	1952	2907	2457	2946	2316	801
Autocorrelation	.46	.56	.69	.74	.68	-
Nicotine Quantity						
M	.45	1.17	1.73	1.55	2.03	1.18
SD	1.09	1.57	1.78	1.77	1.58	1.70
N	1787	3036	2447	2922	3225	796
Autocorrelation	.54	.70	.79	.49	.45	-

M = mean; SD = standard deviation; N = number of respondents; Autocorrelation = correlation between scores at one time point and the immediately subsequent time point (e.g., between scores at age 14 and 17).

Table 2.
Correlations Between PGSs Within MTFS Sample

	1	2	3	4
1. Drinks Per Week				
2. Problematic Alcohol Use	.33			
3. Cigarettes Per Day	.05	.10		
4. Regular Smoking	.23	.24	.16	

Table 3.
Standardized Coefficients to Intercept and Slope Factors From One, Two, and Four PGS Predictor Models

	Alcohol Use Disorder		Alcohol Quantity		Nicotine Use Disorder		Nicotine Quantity	
	Intercept	Slope	Intercept	Slope	Intercept	Slope	Intercept	Slope
Drinks Per Week PGS								
1 PGS Model	**.12 [.07, .22]**	**.06 [.01, .11]**	.10 [-.04, .42]	**.07 [.01, .11]**	**.10 [.03, .18]**	.04 [-.02, .10]	**.12 [.02, .29]**	.02 [-.04, .08]
2 PGS Model	**.09 [.03, .19]**	.03 [-.03, .09]	.01 [-.17, .23]	.06 [-.01, .14]	.06 [-.02, .14]	.03 [-.04, .09]	.07 [-.03, .22]	.02 [-.04, .09]
4 PGS Model	**.08 [.02, .16]**	.01 [-.04, .07]	-.05 [-.32, .13]	.06 [-.01, .14]	.03 [-.04, .11]	.00 [-.06, .06]	.02 [-.09, .15]	.01 [-.05, .08]
Problematic Alcohol Use PGS								
1 PGS Model	**.12 [.05, .26]**	**.10 [.05, .16]**	**.31 [.12, 1.12]**	**.10 [.05, .16]**	**.14 [.07, .22]**	.04 [-.01, .10]	**.18 [.07, .42]**	-.02 [-.09, .04]
2 PGS Model	**.08 [.02, .22]**	**.09 [.04, .15]**	**.30 [.11, 1.10]**	.03 [-.04, .11]	**.12 [.05, .20]**	.03 [-.03, .10]	**.15 [.05, .38]**	-.03 [-.10, .04]
4 PGS Model	.07 [-.01, .20]	**.08 [.02, .13]**	.24 [-.06, .92]	.03 [-.05, .10]	**.08 [.01, .16]**	.00 [-.06, .06]	.10 [-.01, .30]	-.05 [-.12, .02]
Cigarettes Per Day PGS								
1 PGS Model	.06 [-.03, .19]	.04 [-.01, .09]	.12 [-.02, .53]	.04 [-.01, .09]	**.11 [.03, .20]**	**.09 [.03, .15]**	**.14 [.02, .33]**	.06 [-.01, .13]
2 PGS Model	.04 [-.05, .16]	.02 [-.03, .07]	.07 [-.09, .36]	.03 [-.04, .10]	.08 [.00, .16]	.06 [-.01, .12]	.09 [-.03, .25]	.04 [-.03, .12]
4 PGS Model	.04 [-.06, .15]	.02 [-.04, .07]	.06 [-.11, .33]	.03 [-.04, .09]	.07 [-.01, .15]	.06 [-.01, .12]	.08 [-.04, .24]	.05 [-.02, .12]
Regular Smoking PGS								
1 PGS Model	**.12 [.06, .26]**	**.12 [.07, .17]**	**.36 [.18, 1.26]**	.05 [-.02, .12]	**.19 [.13, .28]**	**.18 [.13, .23]**	**.30 [.20, .64]**	**.09 [.02, .16]**
2 PGS Model	**.11 [.06, .25]**	**.12 [.06, .16]**	**.35 [.18, 1.20]**	.04 [-.03, .11]	**.18 [.12, .27]**	**.17 [.12, .23]**	**.29 [.19, .64]**	**.08 [.01, .15]**
4 PGS Model	**.08 [.02, .20]**	**.10 [.04, .15]**	**.33 [.14, 1.16]**	.02 [-.05, .10]	**.15 [.09, .24]**	**.17 [.11, .23]**	**.27 [.16, .59]**	**.09 [.02, .16]**

Bold = 95% confidence interval does not include 0. In the 1 PGS Models only a single PGS was entered as a predictor of the intercept and slope factors; in the 2 PGS Models both PGS' for a specific substance (alcohol or nicotine) were entered as predictors of the intercept and slope factors; in the 4 PGS models all four PGS' were entered as predictors of the intercept and slope factors. Eigenvalues 1 through 5 and sex were entered into each model along with the PGS' as control variables; coefficients for control variables not presented. Confidence intervals derived via clustered (to account for nesting by family) non-parametric percentile bootstrap with 10,000 draws.

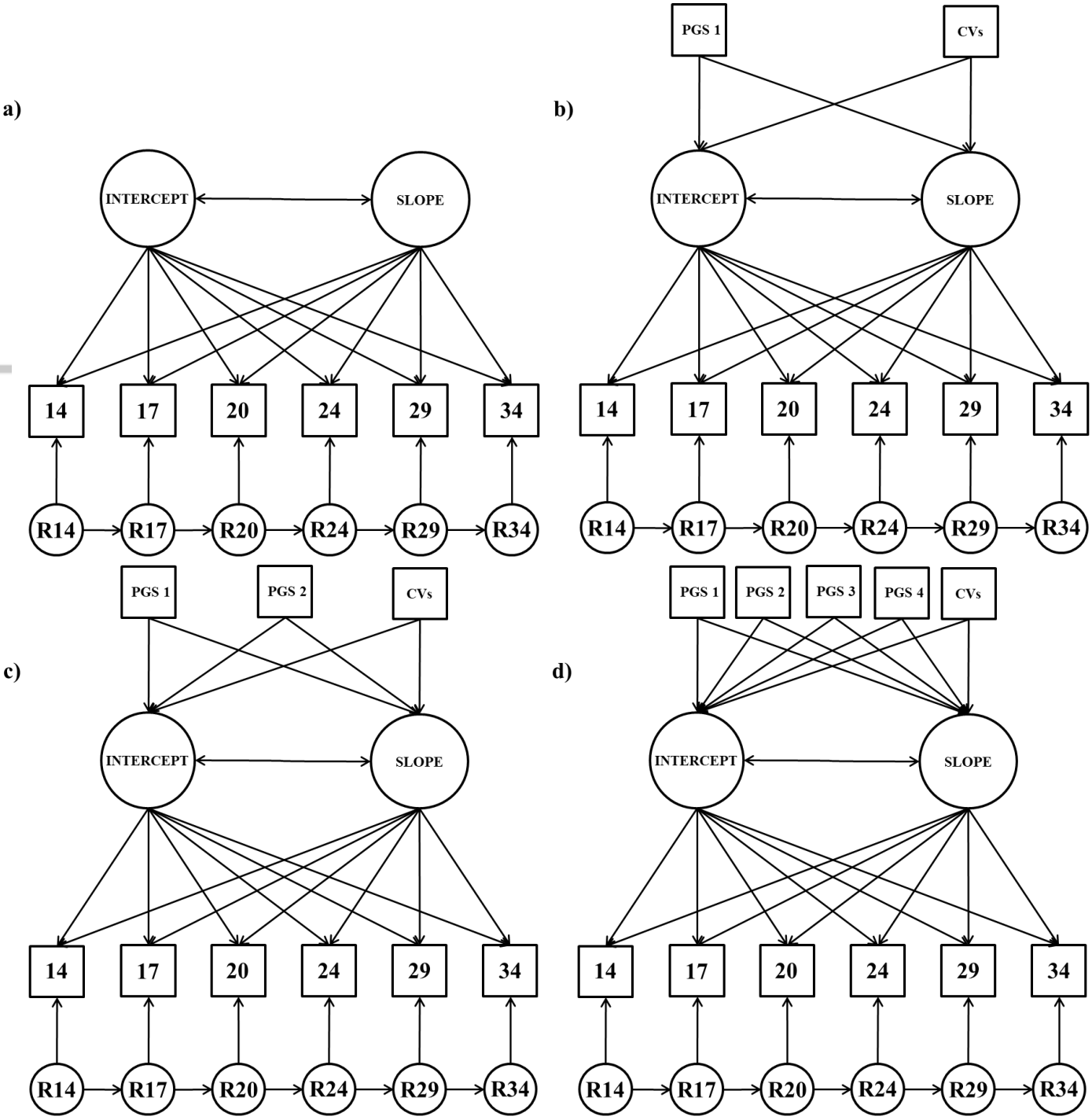

Figure 1. Unconditional and Conditional Latent Growth Models with Structured Residuals. Panel a depicts the unconditional latent growth model; Panel b depicts the 1 PGS conditional latent growth model; Panel c depicts the 2 PGS conditional latent growth model; Panel d depicts the 4 PGS conditional latent growth model. R = residual factor; PGS = polygenetic risk score; CVs = covariates (first 5 eigenvalues and sex). Variances and mean structure omitted from figure for clarity of presentation.

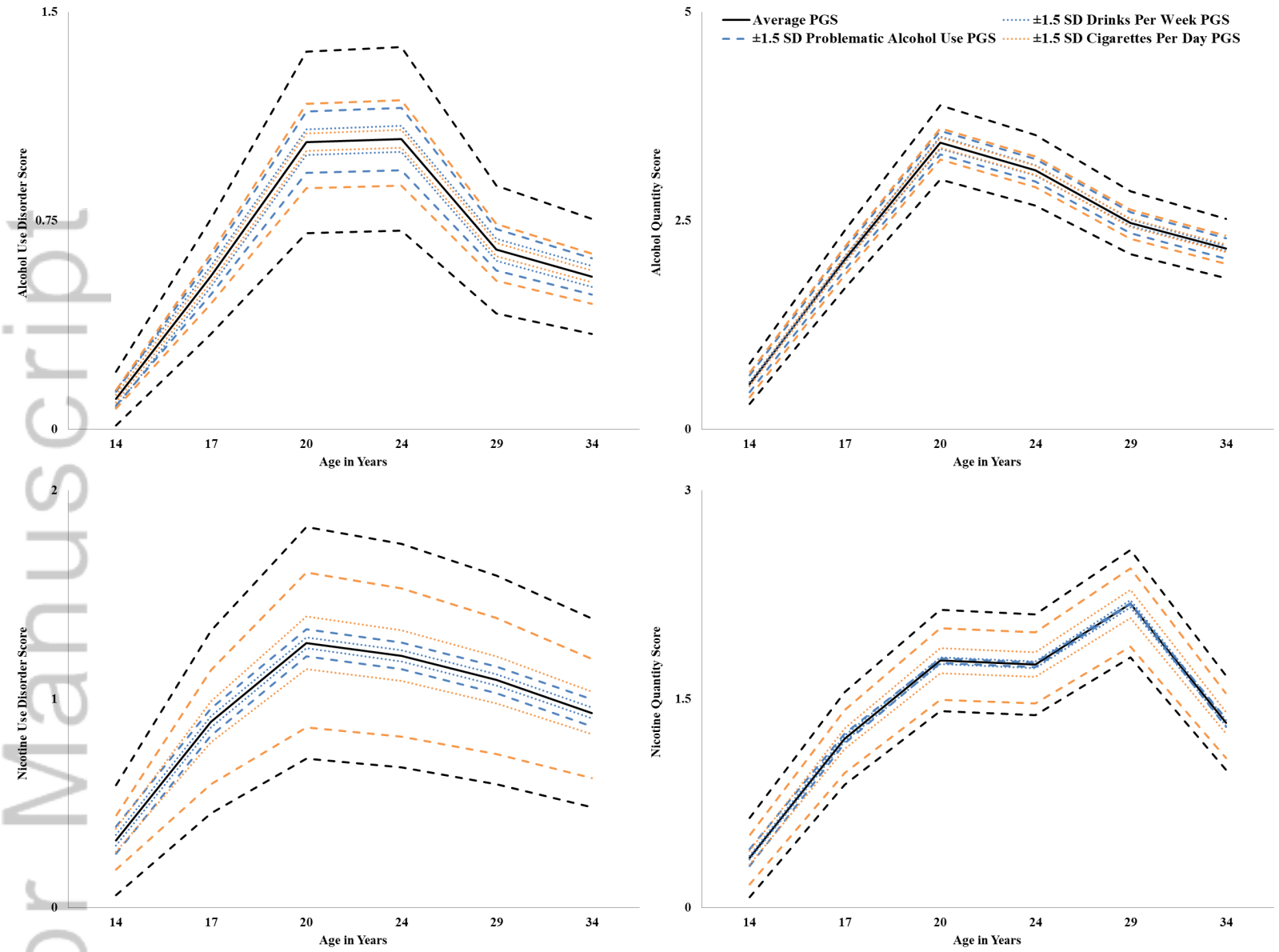

Figure 2. Growth Trajectories from 4 PGS Conditional Growth Models. Age in years presented on X axis, substance use scores presented on Y axis. Trajectories are based on the parameter estimates from the full 4 PGS models. The lines depict trajectories for those with average scores on all four PGSs (solid black line), and either high or low scores (±1.5 standard deviation around the mean) on either the drinks per week PGS (dotted blue line), problematic alcohol use PGS (dashed blue line), cigarettes per day PGS (dotted orange line), regular smoking PGS (dashed orange line), or all four PGS' (dashed black line).