

Dimensional Phenotypes in Psychiatric Genetics: Lessons from Genome-Wide Association Studies of Alcohol Use Phenotypes

Travis T. Mallard^a Sandra Sanchez-Roige^{b, c}

^aDepartment of Psychology, University of Texas at Austin, Austin, TX, USA; ^bDepartment of Psychiatry, University of California San Diego, La Jolla, CA, USA; ^cDivision of Genetic Medicine, Department of Medicine, Vanderbilt University Medical Center, Nashville, TN, USA

Alcohol use disorders (AUDs) are complex psychiatric disorders that can be viewed as the progression of a series of transitions (e.g., initiation of use, escalation to hazardous drinking, and transition to compulsive and harmful use despite negative consequences) – each influenced by different sets of risk factors. Despite this, the field of psychiatric genetics has traditionally used case-control designs to compare individuals who meet diagnostic criteria for a particular disorder, such as AUD, to those who do not. While this approach is extremely valuable for understanding pathological states, it can also be slow and expensive. Moreover, it can obfuscate genetic associations that may be relevant to specific symptoms or clinical presentations, as it treats continuous and multifaceted phenomena as categorical and unidimensional. However, several of these concerns can be ameliorated by using dimensional phenotypes, like various aspects of alcohol consumption, which conceptualize and measure psychiatric phenomena as quantitative traits [1, 2].

Dimensional phenotypes, which are more easily collected in population-based cohorts via self-report questionnaires, can complement the study of psychiatric diag-

noses in several ways. In this editorial, we review the opportunities offered by studying dimensional traits together with psychiatric conditions, using alcohol consumption and AUD as exemplars. We go on to discuss relevant challenges and the solutions that we and others have employed to overcome them. In essence, this editorial attempts to address the following questions: *To what extent can we use dimensional phenotypes in nonclinical populations to learn about the etiology of psychiatric disorders? What considerations should be taken when working with dimensional phenotypes? Can our existing statistical methods overcome the challenges associated with ascertainment and related biases?*

Over the past decade, genome-wide association studies (GWASs) have proved to be useful tools for advancing our knowledge of genetic influences on complex traits like AUD. In 2018, the Psychiatric Genomics Consortium (PGC) Substance Use Disorder Workgroup performed the inaugural trans-ancestral GWAS of alcohol dependence. This case-control study of a diagnostic phenotype, which combined data from nearly 53,000 participants, served as the gold standard for modern genetic associa-

tion studies of AUDs [3]. However, despite its methodological rigor and large sample size, the study only identified 2 genome-wide significant loci – both located in the familiar *ADH1B* gene and previously reported in the literature [4]. It is now abundantly clear that to better understand the massively polygenic architecture of AUDs, we need to increase statistical power even further.

Dimensional phenotypes measured in large-scale cohorts provide a means of accomplishing this goal. Dimensional measures of psychopathology can be inexpensively collected in large cohorts with hundreds of thousands to millions of participants [5]. Indeed, numerous GWASs have now used this approach, leveraging the large population-based cohort of UK Biobank (UKB), the mega-biobank of US veterans, the Million Veteran Program (MVP), and other smaller population-based cohorts to identify hundreds of loci associated with alcohol-related phenotypes [6–14]. In this piece, we will focus on 2 of these studies – both of which are GWASs of the Alcohol Use Disorders Identification Test (AUDIT [15]) that we conducted [7, 16] – in order to illustrate how self-reported dimensional phenotypes can serve as cost-effective strategies to complement more traditional, clinically ascertained phenotypes.

The AUDIT is a prime example of such a phenotype. It is a 10-item self-report questionnaire, with well-established psychometric properties, that can assess 2 different dimensions of AUD risk: alcohol consumption (AUDIT-C, items 1–3) and problematic consequences of drinking (AUDIT-P, items 4–10). Using this tool, we showed that genetic influences on consumption and misuse are only partially overlapping [7], emphasizing the value of studying separate dimensions of AUD symptoms. In addition, and unexpectedly, AUDIT-C and AUDIT-P showed a divergent set of genetic correlations that puzzled many of us at first [7]. For example, AUDIT-C (a composite measure of frequency and quantity of drinking – including binge drinking) measured in the UKB was *only modestly genetically correlated* with alcohol dependence ($r_g \sim 0.3$), whereas there was a large degree of genetic overlap between alcohol dependence and AUDIT-P ($r_g \sim 0.6$). While this finding in and of itself might be intuitive (i.e., most people drink alcohol without developing physical or mental health problems), it was unexpected given the clear genetic (and conceptual) link between alcohol consumption and AUDs (e.g., genetic variation in alcohol metabolizing genes directly influences alcohol consumption, which, in turn, directly influences the individual's likelihood of developing an AUD) [17, 18]. We were further surprised by the positive genetic correlations between AUDIT-C and

well-being and socioeconomic status and negative genetic correlations with several psychiatric conditions [7]. Notably, these findings were not observed for AUDIT-P [7, 19]. As these relationships were consistent with those reported by other studies of alcohol consumption phenotypes in independent cohorts (AUDIT-C in MVP [10] and in a hospital-based cohort [19]; drinks per week in UKB [13]; and alcohol consumption in GSCAN, 23andMe, UKB, and other smaller cohorts [11]), we initially concluded that measures of alcohol consumption were not a viable proxy for studying AUD.

Indeed, it is not difficult to imagine that alcohol consumption is an example of a phenotype for which the population matters greatly. Drinking customs differ from country to country, as well as within a country, and they are related to socioeconomic status, age, and gender. However, it is worth noting that the aforementioned studies have used large population-based cohorts (e.g., 23andMe and UKB), which tend to have sampling strategies that inadvertently recruit individuals who are healthier or of higher socioeconomic status than the broader population. Therefore, although the apparent divergence might reflect true differences in the biology associated with alcohol consumption versus problems, we suspected that the unexpected genetic correlations may be confounded by other factors, such as sources of ascertainment or selection bias [20].

We also hypothesized that the unusual relationships might be a consequence of the sum score approach used to construct the AUDIT-C and AUDIT-P phenotypes. As sum scores assume each survey item is equally informative of the construct that is being measured, it is possible that this approach may be confounded by genetic heterogeneity among alcohol phenotypes (frequency, quantity, and binge drinking) in a way that yielded incongruous associations. In this sense, using “incorrect” item weights might have introduced error in downstream genetic analyses, such as those reported above. We anticipated that we might be able to resolve some of the previous paradoxical associations by implementing an approach that allowed survey items to vary in their contribution to the target construct (e.g., giving more weight to quantity and binge drinking and less weight to frequency of drinking). Furthermore, we hypothesized that this approach might refine the polygenic signal of interest, strengthening the relationship between alcohol consumption and AUD.

To test these hypotheses, we used Genomic Structural Equation Modeling (Genomic SEM) [21] to analyze each individual AUDIT item in a novel multivariate framework that modeled 2 latent genetic factors, *Consumption*

and *Problems*. These 2 factors conceptually paralleled AUDIT-C and AUDIT-P but, critically, allowed for more nuanced and empirically derived weights to be applied to each of the AUDIT items. This was a stark departure from our previous sum score approach that assumed that all items were equally informative of alcohol consumption and problematic use. By giving appropriate weights to each phenotype, we resolved the paradoxical associations and identified strong positive genetic correlations between both *Consumption* and *Problems* and alcohol dependence (both $r_g \sim 0.7$). Moreover, although *Problems* was more strongly genetically correlated with psychiatric disorders, *Consumption* was no longer negatively correlated with major depressive disorder or ADHD, as we previously reported [7].

This multivariate approach also allowed us to partition and characterize the genetic variance that was unique to each AUDIT item (i.e., the residual variance that was not related to the broader latent construct). For example, inspection of the latent genetic factor model revealed that the residual genetic variance in item 1 (i.e., frequency of consumption) was negatively related to alcohol dependence and other psychiatric disorders and positively related to socioeconomic outcomes like educational attainment, household income, and intelligence. This would suggest that many of the puzzling genetic correlations that we and others previously reported for alcohol consumption were partially driven by variance related to socially stratified differences in behavior rather than variance related to the alcohol phenotypes of clinical interest. These findings emphasize how some phenotypes may be more strongly influenced by known (and unknown) sample characteristics than others (e.g., the effect of socioeconomic status on frequency vs. quantity of consumption [22]). As these relationships have primarily been studied in low-risk populations, it will be critical to determine whether they are consistent across the full spectrum of alcohol consumption (i.e., normative consumption to high-intensity drinking) in higher-risk populations.

References

- 1 Kotov R, Krueger RF, Watson D, Achenbach TM, Althoff RR, Bagby RM, et al. The Hierarchical taxonomy of psychopathology (HiTOP): a dimensional alternative to traditional nosologies. *J Abnorm Psychol*. 2017 May; 126(4):454–77.
- 2 Insel T, Cuthbert B, Garvey M, Heinssen R, Pine DS, Quinn K, et al. Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. *Am J Psychiatry*. 2010 Jul;167(7):748–51.
- 3 Walters RK, Polimanti R, Johnson EC, McClintick JN, Adams MJ, Adkins AE, et al. Transancestral GWAS of alcohol dependence reveals common genetic underpinnings with psychiatric disorders. *Nat Neurosci*. 2018 Dec;21(12):1656–69.
- 4 Gelernter J, Kranzler HR, Sherva R, Almasy L, Koesterer R, Smith AH, et al. Genome-wide association study of alcohol dependence: significant findings in African- and European-Americans including novel risk loci. *Mol Psychiatry*. 2014 Jan;19(1):41–9.

Collectively, these findings exemplify how dimensional phenotypes in population-based cohorts can be used to advance genetic discovery for psychiatric disorders without solely relying on clinical diagnostic phenotypes. We advocate for the use of multivariate methods, like Genomic SEM, which can be used to (i) increase GWAS sample size by studying the joint genetic architecture of related phenotypes [23] and (ii) partition genetic variance that is specific to particular phenotypes [24]. With regard to alcohol genetics, we may be able to use these methods to leverage the enormous samples with alcohol consumption phenotypes in future gene discovery efforts for AUDs. By disentangling genetic effects that are specific to a given phenotype from those that influence a broader psychological construct, we may also be able to generate novel hypotheses that advance translational research.

In closing, this editorial has used the AUDIT to exemplify how dimensional phenotyping and multivariate methods can be combined to increase sample size and refine the polygenic signal of interest. This work, we believe, sets the stage for discoveries related to many other psychiatric conditions in the years to come.

Conflict of Interest Statement

S.S.R. is an editorial board member for *Complex Psychiatry* but has no other conflicts of interest to declare. T.T.M. has no conflicts of interest to declare.

Funding Sources

S.S.R. was supported by funds from the California Tobacco-Related Disease Research Program (TRDRP; Grant No. T29KT0526).

Author Contributions

S.S.R. conceived the idea; T.T.M. and S.S.R. wrote the article.

- 5 Sanchez-Roige S, Palmer AA. Emerging phenotyping strategies will advance our understanding of psychiatric genetics. *Nat Neurosci*. 2020 Apr;23(4):475–80.
- 6 Schumann G, Liu C, O'Reilly P, Gao H, Song P, Xu B, et al. KLB is associated with alcohol drinking, and its gene product β -Klotho is necessary for FGF21 regulation of alcohol preference. *Proc Natl Acad Sci U S A*. 2016 13; 113(50):14372–7.
- 7 Sanchez-Roige S, Palmer AA, Fontanillas P, Elson SL; 23andMe Research Team, the Substance Use Disorder Working Group of the Psychiatric Genomics Consortium, Adams MJ, 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 Feb 1; 176(2):107–18.
- 8 Sanchez-Roige S, Fontanillas P, Elson SL; 23andMe Research Team, Gray JC, Gray JC, de Wit H, et al. Genome-wide association study of alcohol use disorder identification test (AUDIT) scores in 20 328 research participants of European ancestry. *Addict Biol*. 2019 Jan;24(1):121–31.
- 9 Sanchez-Roige S, Palmer AA, Clarke TK. Recent efforts to dissect the genetic basis of alcohol use and abuse. *Biol Psychiatry*. 2020 Apr 1;87(7):609–18.
- 10 Kranzler HR, Zhou H, Kember RL, Vickers Smith R, Justice AC, Damrauer S, et al. Genome-wide association study of alcohol consumption and use disorder in 274,424 individuals from multiple populations. *Nat Commun*. 2019 02;10(1):1499.
- 11 Liu M, Jiang Y, Wedow R, Li Y, Brazel DM, Chen F, et al. Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. *Nat Genet*. 2019;51(2):237–44.
- 12 Evangelou E, Gao H, Chu C, Ntritsos G, Blakeley P, Butts AR, et al. New alcohol-related genes suggest shared genetic mechanisms with neuropsychiatric disorders. *Nat Hum Behav*. 2019;3(9):950–61.
- 13 Clarke TK, Adams MJ, Davies G, Howard DM, Hall LS, Padmanabhan S, et al. Genome-wide association study of alcohol consumption and genetic overlap with other health-related traits in UK Biobank (N=112117). *Mol Psychiatry*. 2017;22(10):1376–84.
- 14 Zhou H, Sealock JM, Sanchez-Roige S, Clarke T-K, Levey D, Cheng Z, et al. Meta-analysis of problematic alcohol use in 435,563 individuals identifies 29 risk variants and yields insights into biology, pleiotropy and causality. *Nature Neurosci*. 2019 Aug 16:738088.
- 15 Saunders JB, Aasland OG, Babor TF, de la Fuente JR, Grant M. Development of the alcohol use disorders identification test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption--II. *Addiction*. 1993 Jun;88(6): 791–804.
- 16 Mallard TT, Savage JE, Johnson EC, Huang Y, Edwards AC, Hottenga JJ, et al. Item-level genome-wide association study of the alcohol use disorders identification test in three population-based cohorts. *AJP*. 2021 May 14: appi.ajp.2020.2.
- 17 Edenberg HJ, McClintick JN. Alcohol dehydrogenases, aldehyde dehydrogenases, and alcohol use disorders: a critical review. *Alcohol Clin Exp Res*. 2018 Dec;42(12):2281–97.
- 18 Deak JD, Johnson EC. Genetics of substance use disorders: a review. *Psychol Med*. 2021 Apr 21:1–12.
- 19 Sanchez-Roige S, Cox NJ, Johnson EO, Hancock DB, Davis LK. Alcohol and cigarette smoking consumption as genetic proxies for alcohol misuse and nicotine dependence. *Drug Alcohol Depend*. 2021 Apr 1;221: 108612.
- 20 Xue A, Jiang L, Zhu Z, Wray NR, Visscher PM, Zeng J, et al. Genome-wide analyses of behavioural traits biased by misreports and longitudinal changes. *Genet Genomic Med*. 2020 Jun.
- 21 Grotzinger AD, Rhemtulla M, de Vlaming R, Ritchie SJ, Mallard TT, Hill WD, et al. Genomic SEM Provides Insights into the Multivariate Genetic Architecture of Complex Traits [Internet]. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6520146/>.
- 22 Marees AT, Smit DJA, Ong JS, MacGregor S, An J, Denys D, et al. Potential influence of socioeconomic status on genetic correlations between alcohol consumption measures and mental health. *Psychol Med*. 2020 Feb;50(3): 484–98.
- 23 Karlsson Linnér R, Mallard TT, Barr PB, Sanchez-Roige S, et al. Multivariate analysis of 1.5 million people identifies genetic associations with traits related to self-regulation and addiction. *Nat Neurosci*. Forthcoming 2021.
- 24 Barr PB, Mallard TT, Sanchez-Roige S, Poore HE, Karlsson Linnér R, Collaborators COGA, et al. Parsing genetically influenced risk pathways: Genetic loci impact problematic alcohol use via externalizing and specific risk. *medRxiv*. 2021. <https://doi.org/10.1101/2021.07.20.21260861>.