REVIEW ARTICLE

Precision and Future Medicine 2025;9(4):181-196
https://doi.org/10.23838/pfm.2025.00310
pISSN: 2508-7940 · eISSN: 2508-7959

PRECISION AND FUTURE MEDICINE

Multivariate and multi-trait analyses in psychiatric genetics

A narrative review of tools, challenges, and future directions

Shin-Won Lim[1], Hyun Seok Do[1,2], Soyeon Kim[3], Woojae Myung[1,2]

[1]Department of Neuropsychiatry, Seoul National University Bundang Hospital, Seongnam, Korea
[2]Department of Psychiatry, Seoul National University College of Medicine, Seoul, Korea
[3]Department of Medicine, College of Medicine, Kyung Hee University, Seoul, Korea

Received: November 6, 2025
Revised: November 27, 2025
Accepted: December 1, 2025

Corresponding authors:
Woojae Myung
Department of Neuropsychiatry, Seoul National University Bundang Hospital, Seoul National University College of Medicine, 82 Gumi-ro 173beon-gil, Bundang-gu, Seongnam 13620, Korea
Tel: +82-31-787-7430
E-mail: wmyung@snu.ac.kr

Soyeon Kim
Department of Medicine, College of Medicine, Kyung Hee University, 26 Kyungheedae-ro, Dongdaemun-gu, Seoul 02447, Korea
Tel: +82-2-961-0908
E-mail: kimsoyeon@khu.ac.kr

ABSTRACT

Psychiatric disorders exhibit complex genetic characteristics such as substantial polygenicity, pleiotropy, and genetic overlap, making them difficult to fully understand through studies focused solely on single genes or individual diseases. This review underscores the importance of multivariate and multi-trait analyses in psychiatric genetics and provides a comprehensive overview of major analytical tools, including their concepts, strengths, limitations, and applications. By addressing the current methodological challenges and proposing future directions, we aim to advance our understanding of the genetic architecture underlying psychiatric disorders and support progress towards precision medicine.

Keywords: Mental disorders; Multivariate analysis; Pleiotropy and genetic overlaps; Polygenicity

INTRODUCTION

Psychiatric disorders are major illnesses that place a significant burden on millions of people worldwide [1]. Their etiology is diverse and primarily results from the interaction between genetic predisposition and various environmental factors [2], which include childhood stress, social abuse, neurobiological vulnerabilities, and so on [3]. Understanding these complex mechanisms is essential for effective prevention and intervention of psychiatric illnesses [4].

Genome-wide association studies (GWAS) have been used to identify common genetic variants associated with a wide range of psychiatric disorders [5]. These studies have significantly advanced our understanding of the genetic components underlying disorders by uncovering numerous single-nucleotide polymorphisms (SNPs) associated with disease susceptibility. However, the genetic architecture of psychiatric disorders is highly complex and involves thousands of individual variants with small effect sizes [6]—a phenomenon known as polygenicity [7]. Moreover, pleiotropy, where a single variant influences multiple traits, is frequently ob-

served [8-10], along with a high degree of shared genetic architecture across diagnostic categories [11,12]. These challenges hinder the traditional univariate GWAS framework, which focuses on only one phenotype and may fail to capture broader biological patterns [13].

To overcome these limitations and comprehensively understand the genetic complexity of psychiatric disorders, multivariate analysis methodologies that simultaneously analyze multiple phenotypes or disorders have been emphasized. This approach not only improves statistical power but also helps identify shared genetic loci and unravel overlapping biological mechanisms across disorders [14]. As these methods have become increasingly central to psychiatric genomics, it is necessary to evaluate their utility, assumptions, and limitations.

This narrative review examines the basis of multivariate analysis in psychiatric genetics, systematically compares major multivariate tools, and highlights recent empirical findings using these methods. Furthermore, we discuss current methodological challenges and suggest directions for further research.

NECESSITY OF MULTIVARIATE ANALYSIS

Each psychiatric disorder is characterized by high polygenicity and pleiotropy. Moreover, the extensive genetic overlap across disorders challenges the traditional categorical diagnostic framework. The following section explains why multivariate analysis, which models multiple phenotypes simultaneously, in contrast to conventional univariate approaches focusing on single traits, is essential in psychiatric genomics from three perspectives: polygenicity, pleiotropy, and implications for psychiatric nosology.

Addressing polygenicity: enhancing statistical power for subtle genetic effects

Most psychiatric disorders are highly polygenic, meaning that their etiology involves the cumulative influence of thousands of common variants, each exerting only a small effect on disease risk [7]. Because univariate GWAS independently test each SNP-phenotype association, they often lack the statistical power to detect these subtle effects, leading to many true associations being missed [11,15,16]. This limitation contributes to the phenomenon of 'missing heritability,' in which a larger portion of genetic influences remains unexplained by identified variants [17].

Multivariate analyses directly unravel this issue by simultaneously analyzing multiple genetically linked phenotypes [14]. By integrating information across related traits, these approaches effectively increase the sample size and aggregate the small effects of individual variants, thereby improving the power to detect associations that would be below significance thresholds in single-trait analyses [15]. This enhanced discovery power is crucial for identifying a more complete set of genetic risk factors underlying complex psychiatric disorders, and moving beyond the most penetrant variants to capture the broader genetic landscape.

Unraveling pleiotropy: dissecting shared and distinct genetic influences

Pleiotropy, a phenomenon in which a single genetic variant influences multiple distinct phenotypes, is a pervasive feature of psychiatric genetics [8,18]. It manifests as high rates of comorbidities and duplicated symptom profiles across seemingly distinct psychiatric diagnoses, such as schizophrenia (SCZ) and bipolar disorder (BD), or the clinical and genetic overlap observed between major depressive disorder (MDD) and anxiety disorders [18-21].

Univariate GWAS treats each disorder as an isolated entity, overlooking the complex web of cross-disorder genetic influences that contribute to the observed comorbidities. However, multivariate analysis tools are specifically designed to explicitly model pleiotropic effects, allowing the simultaneous analysis of multiple traits to disentangle the genetic architecture into components shared across disorders and those unique to specific disorders [11,22]. For example, tools like conjunctional false discovery rate (conjFDR) can detect genetic variants with pleiotropic effects even if their associations with a single disorder had been too weak to be identified by traditional univariate analyses [23,24].

Importantly, recent approaches such as Causal mixture model for GWAS summary statistics (MiXeR) [25] and genomic structural equation modeling (genomic SEM) [14] have demonstrated that even among disorders with strong overall genetic correlations (e.g., SCZ and BD [11]), a substantial proportion of variants exert disorder-specific effects, highlighting distinct neurobiological pathways that cannot be explained by shared genetic liability alone [26]. Disentangling shared and unique genetic influences not only deepens our understanding of common biological mechanisms across disorders but also facilitates the identification of disorder-specific pathways, guiding both transdiagnostic and precision-targeted therapeutic development [24,26,27].

Advancing psychiatric nosology through multivariate approaches

The recognition of extensive genetic sharing among psychiatric disorders fundamentally challenges traditional categorical diagnostic systems that often fail to reflect the biological reality of these conditions [6,11,22,28]. To address this limitation, multivariate frameworks, such as genomic SEM, enable the decomposition of genetic covariance structures across multiple disorders into shared latent factors and disorder-specific components [14]. For instance, genomic SEM can partition the genetic architecture of psychiatric traits into general and domain-specific factors, thereby offering a data-driven framework for revising psychiatric classification systems [29]. Simultaneously, tools such as association analysis based on subsets (ASSET) detect genetic variants shared across subsets of disorders [30]. This approach captures the complex continuum of pleiotropy in psychiatric genetics, highlighting that genetic risk factors may transcend diagnostic boundaries while retaining disorder-specific effects [29].

Together, these multivariate approaches uncover both shared etiological dimensions and distinct disorder mechanisms, providing a biologically grounded foundation for refining psychiatric nosology beyond symptom-based categorizations [15,26,27].

OVERVIEW OF MULTIVARIATE ANALYTIC TOOLS

GWAS-based multivariate analysis tools applied in psychiatric genetics serve distinct purposes. They can be broadly classified into three categories depending on whether they rely on individual-level or summary-level GWAS data and the extent to which they jointly model multiple traits. First, individual-level GWAS-based methods directly incorporate genotypic and phenotypic information from each participant into multivariate regression or mixed models. Representative tools include Genome-wide Efficient Mixed Model Association (GEMMA) [31], as well as Genome-wide Complex Trait Analysis-Multivariate Genomic-relatedness-based Restricted Maximum Likelihood (GCTA-mtGREML) [32], Multi-Trait Mixed Model (MTMM) [33], Multivariate Population-based LINKage analysis (mv-PLINK) [34], and joint model of Multiple Phenotypes (MultiPhen) [35]. However, the application of these individual-level approaches is relatively limited because of restricted data accessibility.

Second, multi-trait analysis methods using GWAS summary statistics have been developed. These tools either jointly model multiple traits or utilize cross-trait information to increase the statistical power of genetic discoveries. Notable examples include the Multi-Trait Analysis of GWAS (MTAG) [15], genomic SEM [14], ASSET [30], Cross-Phenotype Association Test (CPASSOC) [36], Multivariate Meta-Analysis based on a Canonical Correlation Analysis (metaCCA) [37], and Heritability informed power optimization (HIPO) [38]. Additional summary-level methods include metaUSAT [39] and MultiMeta [40].

Third, in a broader sense, several summary-statistics-based approaches are not direct multi-trait GWAS tools, but rather complementary methods for the meta-analysis and characterization of genetic architecture in psychiatric traits. It includes Genome-Wide Association Meta-Analysis (GWAMA) [41], MiXeR [25], pleiotropy-informed false discovery rate (pleioFDR) [23], and Local Analysis of [co]Variant Association (LAVA) [42], which model local genetic correlations, and Sum of Single Effects (SuSiE) [43], which performs fine mapping based on a multiple regression framework.

RECENT RESEARCH APPLICATIONS

This section highlights the recent applications of multi-trait GWAS and major multivariate analysis tools in the field of psychiatric genetics, illustrating their utility in uncovering the complex genetic architecture of disorders. We focus on 10 widely applied tools: MTAG, genomic SEM, ASSET, CPASSOC, metaCCA, HIPO, GWAMA, MiXeR, pleioFDR, and GEMMA, providing their concepts, analytical strategies, strengths and limitations, and suitable applications in psychiatric genetic studies. These also are summarized in Table 1 [23,38,44-68].

MTAG

MTAG jointly analyzes GWAS summary statistics from multiple traits, leveraging shared genetic information to produce trait-specific effect estimates and association statistics for each SNP while appropriately accounting for sample overlap [15,44,45]. MTAG estimates trait-specific SNP effects by combining the homogeneous variance-covariance matrix of SNP effect sizes across correlated traits with an estimation error variance-covariance matrix that accounts for sample overlap obtained through linkage disequilibrium score regression (LDSC) [44,46,47]. The major advantages of MTAG include substantially improved statistical power for genetic discovery, leading to the identification of more genome-wide significant loci, enhanced polygenic risk score (PRS) prediction, and more informative downstream bioinformatic analyses while maintaining computational efficiency and robustness to sam-

Table 1. Summary of major multivariate analysis tools frequently used in psychiatric disorders

Tool	Overview	Input/Output	Advantages	Disadvantages	Suitable application scenarios	Actual paper examples
MTAG (Multi-Trait Analysis of GWAS)	Analyzes jointly GWAS SumStats across multiple traits, leveraging genetic correlations to enhance the discovery power for each trait.	Input: GWAS SumStats, (Z-statistics), genetic covariance matrices via LDSC Output: Trait-specific association statistics, SNP effect estimates, P-values	Increased locus discovery, improved polygenic score accuracy, and accounting for sample overlap	Assumes shared SNP effect covariance across traits; violation can inflate false positives. Limited when genetic correlations across traits are low.	Multi-trait GWAS to increase yield for individual traits such as psychiatric or metabolic disorders.	SCZ, BD, ASD, ADHD and DEP [47] OUD with AUD & CanUD [46] Alcohol EEG coherence [45] Max habitual alcohol intake [44]
Genomic SEM (Genomic structural equation modeling, gSEM)	Extends structural equation modeling to the genomic scale, synthesizing GWAS SumStats and LD information to model multivariate genetic architecture.	Input: GWAS SumStats (genetic & sampling covariance matrices via LDSC) Output: Factor models statistics, latent genetic factors, shared and specific genetic components, and heterogeneity index	Powerful for visually modeling complex genetic relationships between multiple phenotypes and identifying underlying genetic factors. Offers potential for causal inference.	A deep understanding of statistical modeling is required, and model refinement is crucial. Computationally demanding for complex models. Large-scale GWAS samples are required, and the study is limited by a European-centric dataset.	Studying shared etiological factors and causal pathways across correlated traits.	SCZ, BD, MDD, PTSD, anxiety [48] 11 Major psychiatric disorders [48] GSI test to gSEM in 8 psychiatric disorders stratified sex [49]
ASSET (Association analysis based on SubSETs)	Subset-based meta-analysis method to detect SNPs associated with only some traits, no effect on some traits, or with opposite effect directions across traits.	Input: GWAS SumStats, genetic correlation matrix Output: P-value, and best-fit trait subset (concordant/discordant)	Detects heterogeneous and pleiotropic effects. Robust to sample overlap. Improves locus discovery. Control multiple testing.	Computationally intensive with large numbers of traits; interpretation can be complex.	Multi-trait GWAS with heterogeneity or pleiotropy across subset (e.g., psychiatric, immune, metabolic traits)	SCZ, cognition, and education [50] Epilepsy [65]
CPASSOC (Cross-phenotype association test)	Combines GWAS results across phenotypes to test for cross-traits SNP associations, using both homogeneous and heterogeneous effect models.	Input: GWAS SumStats (SNP effect sizes, SE, correlation matrix) Output: SNP-level P-values vector (SHom, SHet), cross-phenotype association	Enhances power across multiple traits, works with correlated or independent phenotypes, and detects opposite risk directions.	SHet can be computationally heavy with large correlation matrices, and its gamma approximation may be inaccurate.	Joint analysis of genetically correlated traits with potential effect heterogeneity.	Psychiatric disorders and hemorrhoidal disease [52] AD and 3 stress-related psychiatric disorders [51] Psycho-metabolic nexus [53]
metaCCA (Multivariate meta-analysis of GWAS using canonical correlation analysis)	Performs multivariate GWAS meta-analysis, modeling linear relationships between multiple SNPs and traits using CCA to increase power and identify shared genetic associations.	Input: GWAS SumStats (SNP effect sizes, SEs), reference correlation matrices Output: Canonical correlation (r), P-value, multivariate SNP-multivariate phenotype analysis with shrinkage adjustment. Pleiotropic SNP & genes	Increases power by aggregating signals across SNPs and traits. Enables cross-study meta-analysis and identifies novel associations.	Cannot determine effect direction without individual data. Noisy correlation estimates may reduce accuracy; shrinkage (metaCCA+) may lower power.	Pleiotropy and shared genetic architecture studies. Cross-disease GWAS meta-analysis. Multi-trait association detection when only summary data are available.	Five major psychiatric disorders [54] ASD and 8 co-occurring traits [55]

(Continued to the next page)

Table 1. Continued

Tool	Overview	Input/Output	Advantages	Disadvantages	Suitable application scenarios	Actual paper examples
HIPO (Heritability informed power optimization)	Summary-statistics-based multi-trait GWAS method that derives optimal linear combinations of association coefficients across traits, informed by heritability and sample size, to maximize power.	Input: GWAS SumStats (SNP effect sizes, SE); estimates of heritability and genetic/phenotypic variance-covariance matrix via LDSC Output: HIPO components with association Z-scores (P-value) and genome-wide significant loci	Increases power by leveraging pleiotropy; accounts for heritability, sample size variation, and sample overlap; robust to stratification; maintains type I error	Interpretation less direct than single-trait results; may sacrifice power for strongest single-trait hits; complicating interpretation.	Cross-trait GWAS with correlated, heritable traits where pleiotropy is expected (e.g., psychiatric disorders). Discovering new susceptibility loci. Reduction of high-dimensional phenotypic data.	Five psychiatric disorders [38]
GWAMA (Genome-Wide Association Meta-Analysis)	A software tool for fixed/random-effects meta-analysis based on heterogeneity in GWAS SumStats of multiple studies on a single-trait	Input: GWAS SumStats (e.g., marker identifier, allelic effect estimates, SE) Output: Meta-analyzed association results, heterogeneity metrics, fixed/random-effects analysis result	User-friendly scripts for data alignment and pre-processing; performs error trapping; calculates heterogeneity metrics (Q, I^2); supports fixed- and random-effects	Restricted to single-trait analyses; cannot perform multivariate modeling directly; requires downstream tools for multi-trait applications.	Large-scale single-trait GWAS meta-analyses across diverse cohorts.	AD [56] ADHD [57] Cortical structure [66]
MiXeR (Causal mixture model for GWAS summary statistics)	Quantifies polygenicity, discoverability, and genetic overlap between traits beyond infinitesimal models; Models causal mixed effect directions.	Input: GWAS SumStats (z-scores, effective N). LD structure estimates, MAF, information of sample overlap Output: Polygenic overlap, shared/specific causal variant number estimates	Captures genetic overlap beyond genetic correlation; detects mixed effect directions; identifies shared vs. specific variants	Does not localize individual SNPs; assumes additive/uniform model; needs large GWAS; sensitive to LD/ancestry; mainly bivariate	Exploring genetic architecture of complex/Psychiatric traits. Quantifying polygenic overlaps even without genetic correlation	SCZ & brain structures [58,60] Psychiatric disorders (SCZ, BD, MDD, ADHD) [59] AUD-BMI [61] GSA-MiXeR [62]
pleioFDR (Pleiotropy-informed false discovery rate)	A model-free statistical approach. Pleiotropy-informed false discovery rate approaches: condFDR improves discovery in one trait conditional on another; conjFDR identifies shared loci.	Input: GWAS SumStats (P-values) for multiple phenotypes from non-overlapping samples Output: Trait-specific loci (condFDR) and shared loci (conjFDR)	Increases statistical power to discover SNPs that might be missed by univariate analysis. Effective in identifying common genetic risk factors across multiple related phenotypes.	Limited to pairwise or small sets of traits; requires strong cross-traits enrichment to be effective.	To identify 'missing heritability' and increase statistical power, when genetic correlation between two traits is low or absent, and when the direction of genetic variation on two traits is mixed.	SCZ and BD [23] BMI and major psychiatric disorder [67] Cannabis use, SCZ, BD [68]
GEMMA (Genome-wide Efficient Mixed Model Association)	A tool implementing LMM and multivariate LMM for GWAS individual data, accounting for population stratification and relatedness.	Input: Individual-level genotype data, phenotype(s), covariates Output: SNP-level association statistics (P-values, effect sizes), heritability estimates	Controls for population stratification and relatedness, supports univariate and multivariate traits, efficient computation, rapid and exact statistics	Efficiency gains limited when multiple variance components are included; approximate methods may be needed in complex models.	GWAS in related individuals (e.g., twins, family data), small-to-moderate sample size studies, multivariate trait analysis	Depression, cognition, and memory [63] Suicide death [64]

GWAS, genome-wide association study; SumStats, summary statistics; LDSC, linkage disequilibrium score regression; SNP, single-nucleotide polymorphism; SCZ, schizophrenia; BD, bipolar disorder; ASD, autism spectrum disorder; ADHD, attention-deficit/hyperactivity disorder; DEP, depression; OUD, opioid use disorder; AUD, alcohol use disorder; CanUD, cannabis use disorder; EEG, electroencephalographical; SEM, structural equation modeling; LD, linkage disequilibrium; MDD, major depressive disorder; PTSD, post-traumatic stress disorder; GSI, genomic structural invariance; SE, standard error; SHom, summary statistic homogeneous effect test; SHet, summary statistic heterogeneous effect test; AD, Alzheimer's disease; CCA, canonical correlation analysis; MAF, minor allele frequency; BMI, body mass index; GSA, gene set analysis; condFDR, conditional FDR; conjFDR, conjunctional FDR; LMM, linear mixed model.

ple overlap [15,45]. However, a key limitation is the strong assumption of a homogeneous SNP effect size covariance. If this assumption is violated, MTAG can yield biased estimates and inflated false discovery rates for individual SNPs, thereby reducing trait specificity [15,46].

The empirical application of the MTAG illustrates its utility. For example, it has been applied to three traits: increasing the number of genome-wide significant loci for depressive symptoms (from 32 to 64 loci), neuroticism (from 9 to 37 loci), and subjective well-being (SWB) (from 13 to 49 loci), as well as improving PRS prediction and bioinformatics analyses [15]. In another study, MTAG was used to combine GWAS summary statistics for post-traumatic stress disorder (PTSD) and MDD (together with cardiovascular disease traits), significantly enhancing the genetic predictive performance of the PTSD diagnostic status [69]. More recently, MTAG was applied to opioid use disorder (OUD) and genetically correlated substance use disorders such as alcohol use disorder and cannabis use disorder (CanUD), increasing the equivalent sample size and identifying 19 independent genome-wide significant OUD risk loci while explaining a larger proportion of OUD variance in the PRS [46].

Genomic SEM

Genomic SEM models the multivariate genetic architecture of complex traits using GWAS summary statistics, enabling the identification of genetic variants contributing to both general liability (latent factors) and trait-specific divergence (heterogeneity) [14,70]. The genetic framework extends LDSC to estimate genetic covariance matrices, capture pleiotropy, and model higher-order latent factors, thereby offering the potential for causal inference by clarifying the pathways through which shared and trait-specific genetic influences affect multiple disorders [14,29,48]. The advantages of genomic SEM include a flexible and efficient framework that enhances discoverability, improves polygenic prediction, and provides unbiased estimation while robustly addressing sample overlap, heterogeneity, and population stratification [14,19,71,72]. Limitations include reliance on well-powered GWAS of common variants, restricted generalizability across ancestries, and potential oversimplification of complex trait interactions [19,70,73-75].

Genomic SEM was used to dissect the shared and heterogeneous genetic architectures of psychiatric disorders. Multivariate GWAS analysis using genomic SEM has highlighted loci contributing to multiple disorders through shared factors, as well as loci acting heterogeneously within individual traits, providing insights at the biobehavioral, functional genomic, and molecular genetic levels for 11 major mental disorders [48]. A recent study added the genomic structural invariance test to genomic SEM, allowing for the comparison of autosomal multivariate genetic architecture of shared variants in eight major psychiatric disorders according to biologically stratified sex [49]. Finally, genomic SEM has been applied in a 'GWAS-by-subtraction' approach to isolate non-cognitive genetic variation in educational attainment, which was later associated with mental illness [75].

ASSET

ASSET is a pleiotropy-oriented meta-analysis framework that systematically identifies and characterizes genetic associations that may be shared across multiple traits or specific to subsets of traits [30]. Unlike conventional fixed-effects meta-analyses, which assume homogeneous SNP effects across studies, ASSET performs an exhaustive search across all possible subsets of traits to detect associations that are present in only some traits, absent in others, or operating in directions across phenotypes. Methodologically, it integrates weighted Z-statistics within a multivariate normal framework, accounting for correlations among test statistics, adjusting for multiple testing, and correcting for sample overlap to maintain rigorous control of type I errors [30]. The key advantages of ASSET include enhanced locus discovery, the ability to capture heterogeneous or even discordant cross-trait genetic effects, and robustness in the presence of overlapping samples. Nonetheless, limitations remain, such as a high computational burden, reduced statistical power in highly homogeneous study settings owing to the subset search procedure, and limited sensitivity in detecting associations with rare variants.

Using ASSET, Lam et al. [50] showed that concordant loci (high cognition/education and lower SCZ risk) are linked to early neurodevelopment, whereas discordant loci (higher education and higher SCZ risk) are linked to synaptic pruning. Johnson et al. [76] identified 327 pleiotropic loci across SCZ, CanUD, and smoking, including 150 novel loci, using ASSET, highlighting shared brain tissue enrichment.

CPASSOC

CPASSOC is a statistical framework designed to detect genetic variants associated with multiple traits by leveraging GWAS summary statistics [36,51]. This is particularly relevant in the study of pleiotropy, including clinically distinct conditions. CPASSOC tests the null hypothesis that a genetic marker is not associated with any of the studied phenotypes, against the al-

ternative that it is associated with at least one. By jointly analyzing multiple traits, it enhances locus discovery [36]. Methodologically, it accommodates heterogeneity across traits, including opposite effect directions, correlated and independent phenotypes, and continuous and binary outcomes. The framework also corrects for sample overlap and relies only on summary-level data, making it suitable for large-scale consortia, where individual-level data sharing is restricted.

CPASSOC offers two test statistics: summary statistic homogeneous effect test (Shom), which is more powerful when genetic effects are homogeneous across traits, and summary statistic heterogeneous effect test (SHet), which is an extension of SHom that increases power when effect sizes vary. Its main advantages include high power, flexibility in modeling diverse phenotypic structures, and robustness against overlapping samples. Limitations include the computational burden of estimating the SHet statistic in high-dimensional correlation matrices and potential inaccuracy of gamma distribution approximation under such condition [36,51].

CPASSOC has been used in psychiatric genetics. SHet statistics have been used with ASSET in a cross-trait GWAS to discover shared risk variants between Alzheimer's disease and three stress-related disorders (PTSD, anxiety disorder, and MDD), improving the power to detect heterogeneous genetic effects [51]. Another study identified shared SNPs and specific loci between psychiatric disorders and hemorrhoidal disease [52]. It has also been used for sensitivity analysis in cross-trait meta-analyses to identify pleiotropic SNPs underlying the shared genetic architecture between psychiatric disorders and metabolic traits [53].

MetaCCA

MetaCCA is a framework for summary statistics-based multivariate meta-analysis of GWAS using CCA [34]. It models linear relationships between multiple SNPs and multiple traits from univariate GWAS inputs while estimating correlation structures from external reference panels, such as the 1000 Genomes Project. To improve the robustness, the method incorporates a covariance shrinkage algorithm [37,54]. Novel pleiotropic associations can be discovered by aggregating weak signals across phenotypes, without requiring individual-level data [77]. The limitations include the inability to infer effect directions in the absence of individual-level data, and reduced accuracy when correlation estimates are noisy, although a shrinkage-based extension, metaCCA+, partly addresses this issue [37].

One notable application is the utility of the metaCCA. A multivariate GWAS identified pleiotropic genes shared across five psychiatric disorders (BD, SCZ, MDD, autism spectrum disorder [ASD], and attention-deficit/hyperactivity disorder [ADHD]) using Psychiatric Genomics Consortium (PGC) summary statistics, yielding 1,147 SNPs and 246 candidate genes. After refinement with VEGAS2 [78], 37 pleiotropic genes remained, including known loci as well as 24 novel candidates, underscoring the ability of metaCCA to uncover shared genetic architecture [54]. More recently, Salenius et al. [55] conducted the largest multivariate GWAS of ASD and eight co-occurring traits using metaPhat/metaCCA. Their analysis identified 637 significant associations and novel ASD-related genes, such as KAT8 regulatory NSL complex subunit 1 (*KANSL1),* N-ethylmaleimide-sensitive factor (*NSF*), and *neurotrimin* (*NTM)*, which are involved in immune response and synaptic transmission [55].

HIPO

HIPO utilizes GWAS summary statistics across multiple traits to uncover shared genetic associations. It works by using the genetic covariance between traits estimated from SNP-level summary statistics and then identifying variants that are likely to influence more than one phenotype. HIPO derives an optimal linear combination of association coefficients that maximizes the non-centrality parameter, thereby boosting its power to detect cross-trait genetic signals. This framework leverages pleiotropy, genome-wide estimates of heritability, and genetic covariance to enhance the discovery. Compared to standard meta-analyses, HIPO offers several advantages. It increases the detection of novel loci, controls type I error rates, is robust to population stratification, and can reduce dimensionality in high-dimensional phenotypic data. However, it has limitations: it may be less efficient at identifying the strongest single-trait associations ('top hits'), and, unlike MTAG, it does not directly assign SNPs to individual traits, which can complicate trait-specific interpretation [38].

Examples of applications in psychiatric diseases: In the PGC cross-disorder data, HIPO increased the number of genome-wide significant loci by 200%, notably discovering a new locus (rs13072940) associated with BD and SCZ that was not found by individual trait analysis or standard meta-analysis [38].

GWAMA

GWAMA is an open-source software framework designed to combine GWAS summary statistics from multiple cohorts, thereby increasing its power to detect genetic variants associ-

ated with complex traits [41]. It harmonizes the study results with a common reference allele, performs comprehensive error checking, and applies genomic controls to correct for population stratification. Meta-analysis is conducted inverse-variance weighted models: fixed-effects or random-effects when heterogeneity is detected. Heterogeneity is quantified using Cochran's Q statistic and the I^2 index [41]. Strengths of GWAMA include its user-friendly compatibility with standard GWAS outputs (e.g., Population-based LINKage analysis Single Nucleotide Polymorphism Test [PLINK SNPTEST]), automated quality control, and flexibility between fixed- and random-effects models. A key limitation is that GWAMA is restricted to single-trait meta-analyses; multi-trait analyses require downstream frameworks, such as MTAG, CPASSOC, and HIPO, that build on univariate results [79].

GWAMA has been widely applied in psychiatric and complex trait genetics. It was used in a meta-analysis of anorexia nervosa (AN) that revealed three risk loci associated with body mass index (BMI) [56]. In African American cohorts, GWAMA identified four novel common loci (endoplasmic reticulum degradation-enhancing alpha-mannosidase-like protein 1 [*EDEM1*], activated leukocyte cell adhesion molecule [*ALCAM*], glypican 6 [*GPC6*], VRK serine/threonine kinase 3 [*VRK3*]) and one rare locus near insulin-like growth factor I receptor (*IGF1R*) associated with AD [80]. A landmark GWAMA of ADHD identified the first significant genome-wide loci associated with this disorder [57]. More recently, GWAMA outputs for the age at onset of walking (AOW) were further leveraged in multi-trait analyses (genomic SEM, MiXeR), demonstrating genetic overlap between AOW, ADHD, BMI, and brain morphology traits [81].

MiXeR

MiXeR is a statistical framework designed to characterize the genetic architecture of complex traits using GWAS summary statistics [25,58,82,83]. Unlike the infinitesimal model, which assumes all variants contribute a small effect [84], MiXeR distinguishes between 'non-null' (causal) and 'null' (non-causal) variants by fitting mixture models to SNP effect size distributions. This approach enables estimation of polygenicity (number of causal variants), discoverability (variance of causal effect sizes), and genetic overlap between traits, even when effect directions differ [25,59]. The strength of MiXeR is its ability to provide a more nuanced view of the shared architecture than standard genetic correlation methods, which may underestimate overlap in the presence of mixed effect directions [25,60,61]. Its extension, gene set analysis (GSA)-MiXeR, facilitates biological interpretation by quantifying heritability and enrichment at the gene-set level [62]. Nevertheless, MiXeR assumes additive effects and normally distributed causal variants independent of allele frequency, linkage disequilibrium (LD), or genomic location [82], assumptions which may not hold across all traits. It also requires well-powered GWASs and accurate LD reference panels, limiting its applicability in underpowered studies or in non-European populations [85]. In practice, MiXeR is often complemented by pleioFDR approaches, which help localize the specific loci underlying the overlap it estimates [60,61,82,83,86].

MiXeR has been widely studied in several psychiatric disorders to elucidate its complex genetic architecture and interrelationships. For SCZ, it has revealed high polygenicity and extensive genetic overlap with conditions like BD, cortical and subcortical brain volumes, and even educational attainment, often identifying shared variants despite low or non-existent genetic correlations [58,60]. In the case of major depression, MiXeR analyses uncovered significant polygenic overlap with intelligence and subcortical brain volumes, often with mixed effect directions influencing the traits [82,83]. For ADHD, MiXeR showed it to be less polygenic compared to some other mental disorders and identified substantial overlap with irritable bowel syndrome [59,87].

pleioFDR: conditioned/conjunctional FDR

Conditioned FDR (condFDR) and its extension, conjFDR, are pleiotropy-informed approaches that enhance locus discovery by leveraging cross-trait enrichment of GWAS summary statistics [23,88]. Implemented within an empirical Bayesian framework [89], pleioFDR re-ranks SNP test statistics according to enrichment patterns across traits, thereby increasing the statistical power while remaining cost-effective and model-free [60,88,90,91]. Unlike traditional methods, pleioFDR can detect both trait-specific and shared loci, irrespective of the overall genetic correlation or mixed effect directions. Effect sizes and directions are typically inferred *post hoc* from the original GWAS summary statistics estimates (e.g., z-scores, regression coefficients, or odds ratios) [83,85]. Despite these advantages, pleioFDR has limitations, including sensitivity to LD structures in complex genomic regions (which are typically excluded), inability to differentiate true from mediated pleiotropy or pinpoint causal variants, dependence on well-powered GWAS, and potential inflation from sample overlap. ConjFDR, although more conservative, mitigates some of these issues at the expense of reduced power [88,90,92].

Applications of pleioFDR to psychiatric genetics include:

Early work on SCZ and BD identified 58 SCZ loci and 35 BD loci at condFDR of less than 0.05, with conjFDR revealing 14 additional loci jointly associated with both disorders, many of which were novel beyond univariate GWAS [23]. Subsequent studies extended the framework to SCZ and cardiometabolic traits (BMI, smoking initiation, and type 2 diabetes), uncovering extensive genetic overlaps with both concordant and discordant effect directions [85]. Recently, pleioFDR applied to SWB and multiple psychiatric disorders (MDD, BD1, SCZ, AN, ADHD, cannabis use disorder (CNB), and ASD) identified numerous additional loci, including 101 shared between SWB and MDD, and 30 shared between SWB and SCZ, highlighting complex pleiotropic relationships, even in the presence of low genetic correlations [93].

GEMMA

GEMMA applies linear mixed models to individual-level GWAS data to simultaneously test for SNP associations with one or multiple phenotypes. The key idea is to model genetic similarity (genomic relatedness matrix) among individuals as a random effect that corrects for confounding due to population stratification and relatedness (e.g., family or cryptic relatedness) [31]. In practice, GEMMA calculates exact association statistics that are mathematically identical to the well-known Efficient Mixed Model Association (EMMA) method; however, it does so far more efficiently, approximately 'n' times faster, making genome-wide analyses computationally feasible for large samples. This efficiency results from the reformulation of the computations, which avoids the heavy repeated matrix decompositions required by the EMMA [31]. GEMMA is particularly useful for small or moderately sized datasets and provides a robust foundation for multivariate phenotype analyses [63]. Its main strengths are the precise control of type I error rates, effective adjustment for relatedness and stratification, and computational tractability. However, its efficiency gains are limited when more than one random effect (variance component) is included in the model, in which case, approximate methods may still be required.

GEMMA has been applied to a multivariate GWAS of depression, cognition, and memory phenotypes in Chinese individuals, as well as a meta-analysis of a GWAS of suicide mortality [63]. GEMMA was utilized for GWAS analysis of suicide deaths in the University of Utah Cohort 1 samples, specifically to examine the association between variants and suicide deaths while modeling population stratification [64].

LIMITATIONS AND FUTURE DIRECTIONS OF MULTIVARIATE METHODOLOGY

Although multivariate GWAS approaches have made significant contributions to psychiatric genetics, they have several methodological limitations.

Methodological limitations

First, residual confounding from sample overlap and population stratification remains a fundamental challenge [94]. Although statistical approaches are partially correct for sample overlap [15,30,38] and population structure [41,63], their effectiveness is limited when individual-level data are inaccessible or when subtle unmeasured substructures exist. In psychiatric genetics, where consortia integrate data collected under heterogeneous recruitment strategies and diagnostic practices, cryptic relatedness and subtle population stratification can produce inflated effect estimates or spurious associations, thereby complicating interpretation and hindering clinical translation [14,23].

Second, the complexity of biological and clinical interpretations of multivariate findings poses a significant hurdle. When pleiotropic loci or shared genetic factors are identified, the challenge extends beyond statistical significance to understanding their precise biological mechanisms and clinical implications [58]. It remains difficult to ascertain how a genetic variant affecting multiple disorders differentially contributes to each specific phenotype, or through which molecular pathways these diffuse effects are mediated. This inherent complexity can impede the direct translation of multivariate genetic discoveries into actionable insights for personalized diagnosis, prognosis, or treatment selection in psychiatric practice.

Third, limited generalizability across diverse populations is a critical barrier. The vast majority of large-scale GWAS underlying multivariate analyses are based on European ancestry cohorts [75,85]. This Eurocentric bias constrains the transferability of the genetic architecture, PRS, and pleiotropic relationships to non-European populations [95]. Genetic effects and their frequencies can vary significantly across ancestries due to differences in LD patterns, allele frequencies, and environmental exposures, posing a substantial barrier to the equitable global implementation of precision psychiatry.

Fourth, heterogeneity in phenotypic definitions and diagnostic criteria across studies presents a persistent challenge [96,97]. Multivariate analyses rely on the harmonization of phenotypic data from multiple sources. However, psychiatric diagnoses are often based on symptom clusters rather than

on objective biomarkers, resulting in considerable clinical heterogeneity within diagnostic categories and overlap between disorders. Inconsistent diagnostic criteria or varying levels of phenotypic granularity across different cohorts can introduce noise and bias into multivariate models, impacting the validity and comparability of results, and potentially obscuring true genetic signals relevant to specific psychiatric constructs [98].

Finally, inherent difficulty in model specification and validation for complex multivariate tools, particularly for methods like genomic SEM, represents a significant academic and methodological challenge [14,29]. The model requires *a priori* hypotheses about the underlying genetic relationships among psychiatric traits, and the choice of the model structure can profoundly influence the results. Mis-specification of the genetic covariance structure or the relationships between latent factors and observed phenotypes can lead to biased parameter estimates and erroneous conclusions regarding shared etiologies or causal pathways. This highlights the need for continuous improvements, robust validation strategies, and collaboration between geneticists and clinicians to ensure biologically plausible and clinically meaningful models.

Future directions

To overcome above limitations and advance psychiatric genetics, several directions are necessary: Firstly, integration and utilization of multi-ancestry GWAS data are essential. This involves moving beyond predominantly European-centric datasets and integrating genetic data from diverse global populations. From a practical clinical perspective, this is not merely a matter of statistical robustness, but a critical step towards achieving equitable precision psychiatry [99]. The portability and predictive accuracy of PRS across ancestries can be improved by identifying novel population-specific genetic risk factors and refining existing ones [100]. This will enable more accurate risk stratification, earlier intervention, and personalized treatment recommendations for all individuals, directly addressing current health disparities in psychiatric care [101].

Second, integration with functional genomics and multi-omics data must be strengthened. To better understand the biological significance of genetic variants identified through multivariate GWAS, it is crucial to integrate these findings with multi-omics data, including transcriptomics, proteomics, epigenomics, and metabolomics [102]. This multi-layered approach can bridge the critical gap between genetic associations and their underlying molecular mechanisms, moving beyond mere statistical correlations to a functional understanding of disease pathways. Clinically, this can lead to the identification of actionable molecular targets, facilitating the development of novel pharmacotherapies or the repurposing of existing drugs based on specific biological dysfunctions identified in patient subgroups. This provides a pathway for mechanism-based interventions in psychiatric disorders [103].

Third, advancements in causal inference methodologies are required. Multivariate analyses often identify numerous genetic associations. However, distinguishing genuine causal relationships from statistical correlations is suitable for translating these findings into effective clinical interventions. By combining robust causal inference approaches, such as multivariable Mendelian randomization (MR), summary-data-based MR (SMR), and two-sample MR with pleiotropy robust methods, multivariate genetic data can substantially enhance the reliability of causal inference. In particular, SMR integrates GWAS and expression quantitative trait loci (eQTL) data to test whether changes in gene expression causally mediate disease risk, thereby enabling the rigorous prioritization of candidate genes and potential therapeutic targets. Large-scale multi-ancestry omics datasets are indispensable to fully leverage these strategies. Such diversity improves the generalizability of causal discoveries and facilitates the identification of modifiable risk factors (e.g., lifestyle, environmental exposures) associated with psychiatric outcomes [104]. Ultimately, by identifying causal rather than correlative pathways, clinicians can focus on interventions that are most likely to alter disease trajectory, support evidence-based prevention strategies, and develop personalized treatment plans by targeting causal levers rather than correlated markers.

Fourth, the development of high-dimensional multivariate tools is required to analyze dozens or hundreds of phenotypes simultaneously [105]. This enables fine resolution of symptom dimensions, cognitive endophenotypes, longitudinal trajectories, and treatment response profiles. Analyzing such high-dimensional clinical data along with genetic information can lead to the discovery of more homogeneous and clinically relevant genetic subtypes of psychiatric disorders [54,106]. This stratification will enable clinicians to select the most effective treatment for individual patients, moving away from a 'one-size-fits-all' approach to personalized care.

Fifth, user-friendly tools and platforms are essential for translating the multivariate findings into practice. This necessitates intuitive interactive platforms that allow practitioners to easily visualize, interpret, and apply PRS, genetic correlations, and other multivariate insights into routine clinical de-

cision-making [107]. Such platforms would facilitate the integration of genomic information into electronic health records and clinical guidelines, nurturing a data-driven approach to patient management and promoting collaborative research across diverse clinical and academic settings [108].

CONCLUSION

Multivariate analysis is an essential tool for deciphering the complex genetic basis of psychiatric disorders and understanding their shared and distinct genetic mechanisms. Tools such as MTAG, genomic SEM, MiXeR, and pleioFDR each offer unique strengths and have been successfully applied to identify genetic risk factors and elucidate inter-disease relationships. Despite these advances, challenges such as sample overlap, interpretability, and limited generalizability remain. Future progress depends on the integration of multi-ancestral data, linking findings with multi-omics, advancing causal inference methodologies, and harnessing high-dimensional approaches. The continued evolution of multivariate analyses will be pivotal in transforming psychiatric genetics into a foundation for precision medicine, ultimately innovating prevention, diagnosis, and treatment strategies. Building on this foundation, multivariate analytical approaches may also enable the identification of more robust biomarkers for psychiatric disorders. Incorporating such biomarkers into improved diagnostic frameworks could create a virtuous cycle in which diagnostic accuracy is progressively refined and more precise interventions become possible, ultimately advancing the vision of precision psychiatry [109].

CONFLICTS OF INTEREST

This work was supported by the NAVER Digital Bio Innovation Research Fund funded by NAVER Corporation (Grant No. 37-2023-0140).

ACKNOWLEDGMENTS

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT, RS-2024-00335261).

ORCID

Shin-Won Lim https://orcid.org/0000-0002-5309-2143
Hyun Seok Do https://orcid.org/0009-0004-3035-3787
Soyeon Kim https://orcid.org/0000-0002-2798-2257
Woojae Myung https://orcid.org/0000-0001-9985-2032

AUTHOR CONTRIBUTIONS

Conception or design: WM.
Acquisition, analysis, or interpretation of data: SWL, WM.
Drafting the work or revising: SWL, HSD, SK, WM.
Final approval of the manuscript: SWL, HSD, SK, WM.

REFERENCES

1. Fan Y, Fan A, Yang Z, Fan D. Global burden of mental disorders in 204 countries and territories, 1990-2021: results from the global burden of disease study 2021. BMC Psychiatry 2025;25:486.
2. Sullivan PF, Daly MJ, O'Donovan M. Genetic architectures of psychiatric disorders: the emerging picture and its implications. Nat Rev Genet 2012;13:537-51.
3. Teicher MH, Samson JA. Annual research review: enduring neurobiological effects of childhood abuse and neglect. J Child Psychol Psychiatry 2016;57:241-66.
4. Insel TR. Disruptive insights in psychiatry: transforming a clinical discipline. J Clin Invest 2009;119:700-5.
5. Collins AL, Sullivan PF. Genome-wide association studies in psychiatry: what have we learned? Br J Psychiatry 2013;202:1-4.
6. Visscher PM, Wray NR, Zhang Q, Sklar P, McCarthy MI, Brown MA, et al. 10 Years of GWAS discovery: biology, function, and translation. Am J Hum Genet 2017;101:5-22.
7. Martin AR, Daly MJ, Robinson EB, Hyman SE, Neale BM. Predicting polygenic risk of psychiatric disorders. Biol Psychiatry 2019;86:97-109.
8. Stearns FW. One hundred years of pleiotropy: a retrospective. Genetics 2010;186:767-73.
9. Sivakumaran S, Agakov F, Theodoratou E, Prendergast JG, Zgaga L, Manolio T, et al. Abundant pleiotropy in human complex diseases and traits. Am J Hum Genet 2011;89:607-18.
10. Solovieff N, Cotsapas C, Lee PH, Purcell SM, Smoller JW. Pleiotropy in complex traits: challenges and strategies. Nat Rev Genet 2013;14:483-95.
11. Lee SH, Ripke S, Neale BM, Faraone SV, Purcell SM, Perlis RH, et al. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. Nat Genet 2013;45:984-94.

12. Brainstorm Consortium; Anttila V, Bulik-Sullivan B, Finucane HK, Walters RK, Bras J, et al. Analysis of shared heritability in common disorders of the brain. Science 2018;360:eaap8757.
13. Watanabe K, Stringer S, Frei O, Umicevic Mirkov M, de Leeuw C, Polderman TJC, et al. A global overview of pleiotropy and genetic architecture in complex traits. Nat Genet 2019;51:1339-48.
14. Grotzinger AD, Rhemtulla M, de Vlaming R, Ritchie SJ, Mallard TT, Hill WD, et al. Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. Nat Hum Behav 2019;3:513-25.
15. Turley P, Walters RK, Maghzian O, Okbay A, Lee JJ, Fontana MA, et al. Multi-trait analysis of genome-wide association summary statistics using MTAG. Nat Genet 2018;50:229-37.
16. Purcell SM, Moran JL, Fromer M, Ruderfer D, Solovieff N, Roussos P, et al. A polygenic burden of rare disruptive mutations in schizophrenia. Nature 2014;506:185-90.
17. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, et al. Finding the missing heritability of complex diseases. Nature 2009;461:747-53.
18. Hettema JM, Chen X, Sun C, Brown TA. Direct, indirect and pleiotropic effects of candidate genes on internalizing disorder psychopathology. Psychol Med 2015;45:2227-36.
19. Richards AL, Cardno A, Harold G, Craddock NJ, Di Florio A, Jones L, et al. Genetic liabilities differentiating bipolar disorder, schizophrenia, and major depressive disorder, and phenotypic heterogeneity in bipolar disorder. JAMA Psychiatry 2022;79:1032-9.
20. Tesfaye M, Shadrin A, Parker N, Jaholkowski P, Parekh P, Kutrolli G. Comorbidity alters the genetic relationship between anxiety disorders and major depression [Preprint]. Posted 2024 Nov 20. Medrxiv 2024.11.19.24317523. https://doi.org/10.1101/2024.11.19.24317523
21. Cross-Disorder Group of the Psychiatric Genomics Consortium. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. Lancet 2013;381:1371-9.
22. Grotzinger AD. Shared genetic architecture across psychiatric disorders. Psychol Med 2021;51:2210-6.
23. Andreassen OA, Thompson WK, Schork AJ, Ripke S, Mattingsdal M, Kelsoe JR, et al. Improved detection of common variants associated with schizophrenia and bipolar disorder using pleiotropy-informed conditional false discovery rate. PLoS Genet 2013;9:e1003455.
24. Zhang C, Yang Z, Li X, Zhao L, Guo W, Deng W, et al. Unraveling NEK4 as a potential drug target in schizophrenia and bipolar I disorder: a proteomic and genomic approach. Schizophr Bull 2024;50:1185-96.
25. Frei O, Holland D, Smeland OB, Shadrin AA, Fan CC, Maeland S, et al. Bivariate causal mixture model quantifies polygenic overlap between complex traits beyond genetic correlation. Nat Commun 2019;10:2417.
26. Mallard TT, Linner RK, Grotzinger AD, Sanchez-Roige S, Seidlitz J, Okbay A, et al. Multivariate GWAS of psychiatric disorders and their cardinal symptoms reveal two dimensions of cross-cutting genetic liabilities. Cell Genom 2022;2:100140.
27. Hatoum AS, Gorelik AJ, Blaydon L, Huggett SB, Chi T, Baranger DA, et al. Psychiatric genome-wide association study enrichment shows promise for future psychopharmaceutical discoveries. Commun Med (Lond) 2025;5:176.
28. Smoller JW, Andreassen OA, Edenberg HJ, Faraone SV, Glatt SJ, Kendler KS. Psychiatric genetics and the structure of psychopathology. Mol Psychiatry 2019;24:409-20.
29. Cross-Disorder Group of the Psychiatric Genomics Consortium. Genomic relationships, novel loci, and pleiotropic mechanisms across eight psychiatric disorders. Cell 2019;179:1469-82.
30. Bhattacharjee S, Rajaraman P, Jacobs KB, Wheeler WA, Melin BS, Hartge P, et al. A subset-based approach improves power and interpretation for the combined analysis of genetic association studies of heterogeneous traits. Am J Hum Genet 2012;90:821-35.
31. Zhou X, Stephens M. Genome-wide efficient mixed-model analysis for association studies. Nat Genet 2012;44:821-4.
32. De Vlaming R, Slob EA, Groenen PJ, Rietveld CA. Multivariate estimation of factor structures of complex traits using SNP-based genomic relationships. BMC Bioinformatics 2022;23:305.
33. Korte A, Vilhjalmsson BJ, Segura V, Platt A, Long Q, Nordborg M. A mixed-model approach for genome-wide association studies of correlated traits in structured populations. Nat Genet 2012;44:1066-71.
34. Ferreira MA, Purcell SM. A multivariate test of association. Bioinformatics 2009;25:132-3.
35. O'Reilly PF, Hoggart CJ, Pomyen Y, Calboli FC, Elliott P, Jarvelin MR, et al. MultiPhen: joint model of multiple phenotypes can increase discovery in GWAS. PLoS One

2012;7:e34861.
36. Li X, Zhu X. Cross-phenotype association analysis using summary statistics from GWAS. Methods Mol Biol 2017; 1666:455-67.
37. Cichonska A, Rousu J, Marttinen P, Kangas AJ, Soininen P, Lehtimaki T, et al. MetaCCA: summary statistics-based multivariate meta-analysis of genome-wide association studies using canonical correlation analysis. Bioinformatics 2016;32:1981-9.
38. Qi G, Chatterjee N. Heritability informed power optimization (HIPO) leads to enhanced detection of genetic associations across multiple traits. PLoS Genet 2018;14: e1007549.
39. Ray D, Boehnke M. Methods for meta-analysis of multiple traits using GWAS summary statistics. Genet Epidemiol 2018;42:134-45.
40. Vuckovic D, Gasparini P, Soranzo N, Iotchkova V. MultiMeta: an R package for meta-analyzing multi-phenotype genome-wide association studies. Bioinformatics 2015; 31:2754-6.
41. Magi R, Morris AP. GWAMA: software for genome-wide association meta-analysis. BMC Bioinformatics 2010;11: 288.
42. Werme J, van der Sluis S, Posthuma D, de Leeuw CA. An integrated framework for local genetic correlation analysis. Nat Genet 2022;54:274-82.
43. Wang G, Sarkar A, Carbonetto P, Stephens M. A simple new approach to variable selection in regression, with application to genetic fine mapping. J R Stat Soc Series B Stat Methodol 2020;82:1273-300.
44. Deak JD, Levey DF, Wendt FR, Zhou H, Galimberti M, Kranzler HR, et al. Genome-wide investigation of maximum habitual alcohol intake in US veterans in relation to alcohol consumption traits and alcohol use disorder. JAMA Netw Open 2022;5:e2238880.
45. Meyers JL, Zhang J, Chorlian DB, Pandey AK, Kamarajan C, Wang JC, et al. A genome-wide association study of interhemispheric theta EEG coherence: implications for neural connectivity and alcohol use behavior. Mol Psychiatry 2021;26:5040-52.
46. Deak JD, Zhou H, Galimberti M, Levey DF, Wendt FR, Sanchez-Roige S, et al. Genome-wide association study in individuals of European and African ancestry and multi-trait analysis of opioid use disorder identifies 19 independent genome-wide significant risk loci. Mol Psychiatry 2022;27:3970-9.
47. Wu Y, Cao H, Baranova A, Huang H, Li S, Cai L, et al. Multi-trait analysis for genome-wide association study of five psychiatric disorders. Transl Psychiatry 2020;10:209.
48. Grotzinger AD, Mallard TT, Akingbuwa WA, Ip HF, Adams MJ, Lewis CM, et al. Genetic architecture of 11 major psychiatric disorders at biobehavioral, functional genomic and molecular genetic levels of analysis. Nat Genet 2022; 54:548-59.
49. Schwaba T, Mallard TT, Maihofer AX, Rhemtulla M, Lee PH, Smoller JW, et al. Comparison of the multivariate genetic architecture of eight major psychiatric disorders across sex. Nat Genet 2025;57:583-90.
50. Lam M, Hill WD, Trampush JW, Yu J, Knowles E, Davies G, et al. Pleiotropic meta-analysis of cognition, education, and schizophrenia differentiates roles of early neurodevelopmental and adult synaptic pathways. Am J Hum Genet 2019;105:334-50.
51. Dang W, Hao T, Li N, Zhang H, Li Z, Yu H, et al. Investigating shared risk variants and genetic etiology between Alzheimer's disease and three stress-related psychiatric disorders: a large-scale genome-wide cross-trait analysis. Front Aging 2025;6:1488528.
52. Chen Z, Hu B, Sun J, Jiang Y, Chen Z, Yang C, et al. Shared genetic architecture of psychiatric disorders and hemorrhoidal disease: a large-scale genome-wide cross-trait analysis. Front Psychiatry 2024;15:1456182.
53. Guo X, Feng Y, Ji X, Jia N, Maimaiti A, Lai J, et al. Shared genetic architecture and bidirectional clinical risks within the psycho-metabolic nexus. EBioMedicine 2025;111: 105530.
54. Jia X, Yang Y, Chen Y, Cheng Z, Du Y, Xia Z, et al. Multivariate analysis of genome-wide data to identify potential pleiotropic genes for five major psychiatric disorders using MetaCCA. J Affect Disord 2019;242:234-43.
55. Salenius K, Valja N, Thusberg S, Iris F, Ladd-Acosta C, Roos C, et al. Exploring autism spectrum disorder and co-occurring trait associations to elucidate multivariate genetic mechanisms and insights. BMC Psychiatry 2024; 24:934.
56. Hinney A, Kesselmeier M, Jall S, Volckmar AL, Focker M, Antel J, et al. Evidence for three genetic loci involved in both anorexia nervosa risk and variation of body mass index. Mol Psychiatry 2017;22:192-201.
57. Demontis D, Walters RK, Martin J, Mattheisen M, Als TD, Agerbo E, et al. Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. Nat Genet 2019;51:63-75.
58. Cheng W, van der Meer D, Parker N, Hindley G, O'Connell

KS, Wang Y, et al. Shared genetic architecture between schizophrenia and subcortical brain volumes implicates early neurodevelopmental processes and brain development in childhood. Mol Psychiatry 2022;27:5167-76.
59. Hindley G, Frei O, Shadrin AA, Cheng W, O'Connell KS, Icick R, et al. Charting the landscape of genetic overlap between mental disorders and related traits beyond genetic correlation. Am J Psychiatry 2022;179:833-43.
60. Cheng W, Frei O, van der Meer D, Wang Y, O'Connell KS, Chu Y, et al. Genetic association between schizophrenia and cortical brain surface area and thickness. JAMA Psychiatry 2021;78:1020-30.
61. Malone SG, Davis CN, Piserchia Z, Setzer MR, Toikumo S, Zhou H, et al. Alcohol use disorder and body mass index show genetic pleiotropy and shared neural associations. Nat Hum Behav 2025;9:1056-66.
62. Frei O, Hindley G, Shadrin AA, van der Meer D, Akdeniz BC, Hagen E, et al. Improved functional mapping of complex trait heritability with GSA-MiXeR implicates biologically specific gene sets. Nat Genet 2024;56:1310-8.
63. Sun J, Wang W, Zhang R, Duan H, Tian X, Xu C, et al. Multivariate genome-wide association study of depression, cognition, and memory phenotypes and validation analysis identify 12 cross-ethnic variants. Transl Psychiatry 2022;12:304.
64. Li QS, Shabalin AA, DiBlasi E, Gopal S, Canuso CM, Palotie A, et al. Genome-wide association study meta-analysis of suicide death and suicidal behavior. Mol Psychiatry 2023;28:891-900.
65. International League Against Epilepsy Consortium on Complex Epilepsies. GWAS meta-analysis of over 29,000 people with epilepsy identifies 26 risk loci and subtype-specific genetic architecture. Nat Genet 2023;55:1471-82.
66. Grasby KL, Jahanshad N, Painter JN, Colodro-Conde L, Bralten J, Hibar DP. The genetic architecture of the human cerebral cortex. Science (1979) 2020;367:eaay6690.
67. Bahrami S, Steen NE, Shadrin A, O'Connell K, Frei O, Bettella F, et al. Shared genetic loci between body mass index and major psychiatric disorders: a genome-wide association study. JAMA Psychiatry 2020;77:503-12.
68. Cheng W, Parker N, Karadag N, Koch E, Hindley G, Icick R, et al. The relationship between cannabis use, schizophrenia, and bipolar disorder: a genetically informed study. Lancet Psychiatry 2023;10:441-51.
69. Seligowski AV, Misganaw B, Duffy LA, Ressler KJ, Guffanti G. Leveraging large-scale genetics of PTSD and cardiovascular disease to demonstrate robust shared risk and improve risk prediction accuracy. Am J Psychiatry 2022;179:814-23.
70. Quintero Reis A, Newton BA, Kessler R, Polimanti R, Wendt FR. Functional and molecular characterization of suicidality factors using phenotypic and genome-wide data. Mol Psychiatry 2023;28:1064-71.
71. 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. Am J Psychiatry 2022;179:58-70.
72. Park S, Kim S, Kim B, Kim DS, Kim J, Ahn Y, et al. Multivariate genomic analysis of 5 million people elucidates the genetic architecture of shared components of the metabolic syndrome. Nat Genet 2024;56:2380-91.
73. Stein MB, Levey DF, Cheng Z, Wendt FR, Harrington K, Pathak GA, et al. Genome-wide association analyses of post-traumatic stress disorder and its symptom subdomains in the Million Veteran Program. Nat Genet 2021;53:174-84.
74. Karlsson Linner R, Mallard TT, Barr PB, Sanchez-Roige S, Madole JW, Driver MN, et al. Multivariate analysis of 1.5 million people identifies genetic associations with traits related to self-regulation and addiction. Nat Neurosci 2021;24:1367-76.
75. Demange PA, Malanchini M, Mallard TT, Biroli P, Cox SR, Grotzinger AD, et al. Investigating the genetic architecture of noncognitive skills using GWAS-by-subtraction. Nat Genet 2021;53:35-44.
76. Johnson EC, Austin-Zimmerman I, Thorpe HH, Levey DF, Baranger DA, Colbert SM, et al. Cross-ancestry genetic investigation of schizophrenia, cannabis use disorder, and tobacco smoking. Neuropsychopharmacology 2024;49:1655-65.
77. Wang Y, Yang Y, Jia X, Zhao C, Yang C, Fan J, et al. Identifying pleiotropic genes for major psychiatric disorders with GWAS summary statistics using multivariate adaptive association tests. J Psychiatr Res 2022;155:471-82.
78. Mishra A, Macgregor S. VEGAS2: software for more flexible gene-based testing. Twin Res Hum Genet 2015;18:86-91.
79. Grove J, Ripke S, Als TD, Mattheisen M, Walters RK, Won H, et al. Identification of common genetic risk variants for autism spectrum disorder. Nat Genet 2019;51:431-44.
80. Kunkle BW, Schmidt M, Klein HU, Naj AC, Hamilton-Nelson KL, Larson EB, et al. Novel Alzheimer disease risk loci

and pathways in African American individuals using the African genome resources panel: a meta-analysis. JAMA Neurol 2021;78:102-13.
81. Gui A, Hollowell A, Wigdor EM, Morgan MJ, Hannigan LJ, Corfield EC, et al. Genome-wide association meta-analysis of age at onset of walking in over 70,000 infants of European ancestry. Nat Hum Behav 2025;9:1470-87.
82. Liu M, Wang L, Zhang Y, Dong H, Wang C, Chen Y, et al. Investigating the shared genetic architecture between depression and subcortical volumes. Nat Commun 2024; 15:7647.
83. Bahrami S, Shadrin A, Frei O, O'Connell KS, Bettella F, Krull F, et al. Genetic loci shared between major depression and intelligence with mixed directions of effect. Nat Hum Behav 2021;5:795-801.
84. Fisher RA. XV: the correlation between relatives on the supposition of Mendelian inheritance. Earth Environ Sci Trans R Soc Edinb 1919;52:399-433.
85. Rodevand L, Rahman Z, Hindley GF, Smeland OB, Frei O, Tekin TF, et al. Characterizing the shared genetic underpinnings of schizophrenia and cardiovascular disease risk factors. Am J Psychiatry 2023;180:815-26.
86. Ahn Y, Kim J, Jung K, Lee DJ, Jung JY, Eom Y, et al. Relationship between problematic alcohol use and various psychiatric disorders: a genetically informed study. Am J Psychiatry 2025;182:671-82.
87. van der Meer D, Hindley G, Shadrin AA, Smeland OB, Parker N, Dale AM, et al. Mapping the genetic landscape of psychiatric disorders with the MiXeR toolset. Biol Psychiatry 2025;98:455-65.
88. Smeland OB, Frei O, Shadrin A, O'Connell K, Fan CC, Bahrami S, et al. Discovery of shared genomic loci using the conditional false discovery rate approach. Hum Genet 2020;139:85-94.
89. Efron B. Large-scale inference: empirical bayes methods for estimation, testing, and prediction. Cambridge University Press; 2010. p. 360.
90. O'Connell KS, Shadrin A, Bahrami S, Smeland OB, Bettella F, Frei O, et al. Identification of genetic overlap and novel risk loci for attention-deficit/hyperactivity disorder and bipolar disorder. Mol Psychiatry 2021;26:4055-65.
91. Andreassen OA, Thompson WK, Dale AM. Boosting the power of schizophrenia genetics by leveraging new statistical tools. Schizophr Bull 2014;40:13-7.
92. Zhao Q, Wang S, Xiong D, Liu M, Zhang Y, Zhao G, et al. Genome-wide analysis identifies novel shared loci between depression and white matter microstructure. Mol Psychiatry 2025;30:3455-65.
93. Jung JY, Ahn Y, Park JW, Jung K, Kim S, Lim S, et al. Polygenic overlap between subjective well-being and psychiatric disorders and cross-ancestry validation. Nat Hum Behav 2025;9:1272-82.
94. Sillanpaa MJ. Overview of techniques to account for confounding due to population stratification and cryptic relatedness in genomic data association analyses. Heredity (Edinb) 2011;106:511-9.
95. Duncan L, Shen H, Gelaye B, Meijsen J, Ressler K, Feldman M, et al. Analysis of polygenic risk score usage and performance in diverse human populations. Nat Commun 2019;10:3328.
96. Jermy BS, Glanville KP, Coleman JRI, Lewis CM, Vassos E. Exploring the genetic heterogeneity in major depression across diagnostic criteria. Mol Psychiatry 2021;26:7337-45.
97. Waszczuk MA, Eaton NR, Krueger RF, Shackman AJ, Waldman ID, Zald DH, et al. Redefining phenotypes to advance psychiatric genetics: implications from hierarchical taxonomy of psychopathology. J Abnorm Psychol 2020;129: 143-61.
98. Feczko E, Miranda-Dominguez O, Marr M, Graham AM, Nigg JT, Fair DA. The heterogeneity problem: approaches to identify psychiatric subtypes. Trends Cogn Sci 2019; 23:584-601.
99. Kachuri L, Chatterjee N, Hirbo J, Schaid DJ, Martin I, Kullo IJ, et al. Principles and methods for transferring polygenic risk scores across global populations. Nat Rev Genet 2024;25:8-25.
100. Lewis CM, Vassos E. Polygenic scores in psychiatry: on the road from discovery to implementation. Am J Psychiatry 2022;179:800-6.
101. Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium. Trans-ancestry genome-wide study of depression identifies 697 associations implicating cell types and pharmacotherapies. Cell 2025;188:640-52.
102. Guan F, Ni T, Zhu W, Williams LK, Cui LB, Li M, et al. Integrative omics of schizophrenia: from genetic determinants to clinical classification and risk prediction. Mol Psychiatry 2022;27:113-26.
103. Jiang Y, Liu Q, Stridh P, Kockum I, Olsson T, Alfredsson L, et al. Multiomics integration prioritizes potential drug targets for multiple sclerosis. Proc Natl Acad Sci U S A 2025;122:e2425537122.
104. Wootton RE, Jones HJ, Sallis HM. Mendelian randomisa-

tion for psychiatry: how does it work, and what can it tell us? Mol Psychiatry 2022;27:53-7.
105. Guo B, Xie Z, He W, Islam SMS, Gottlieb A, Chen H, et al. Efficient multi-phenotype genome-wide analysis identifies genetic associations for unsupervised deep-learning-derived high-dimensional brain imaging phenotypes [Preprint]. Posted 2024 Dec 8. Medrxiv 2024.12.06.24318618. https://doi.org/10.1101/2024.12.06.24318618
106. Holleman AM, Broadaway KA, Duncan R, Todor A, Almli LM, Bradley B, et al. Powerful and efficient strategies for genetic association testing of symptom and questionnaire data in psychiatric genetic studies. Sci Rep 2019;9:7523.
107. Page ML, Vance EL, Cloward ME, Ringger E, Dayton L, Ebbert MT, et al. The polygenic risk score knowledge base offers a centralized online repository for calculating and contextualizing polygenic risk scores. Commun Biol 2022;5:899.
108. Pain O, Gillett AC, Austin JC, Folkersen L, Lewis CM. A tool for translating polygenic scores onto the absolute scale using summary statistics. Eur J Hum Genet 2022;30:339-48.
109. Carroll BJ. Biomarkers in DSM-5: lost in translation. Aust N Z J Psychiatry 2013;47:676-8.