

Original Article

*These authors share joint first-authorship.

Cite this article: Choi KW *et al* (2024). Genetic architecture and socio-environmental risk factors for major depressive disorder in Nepal. *Psychological Medicine* 1–9. <https://doi.org/10.1017/S0033291724001284>

Received: 7 December 2023

Revised: 27 March 2024

Accepted: 9 May 2024

Keywords:

depression; gene-environment; genetics; global mental health; heritability; Nepal; polygenic scores; trauma

Corresponding author:

Karmel W. Choi;

Email: kwchoi@mgh.harvard.edu;



Justin D. Tubbs;

Email: jtubbs@mgh.harvard.edu;

Jordan W. Smoller;

Email: jsmoller@mgh.harvard.edu

Genetic architecture and socio-environmental risk factors for major depressive disorder in Nepal

Karmel W. Choi^{1,2,3,*} , Justin D. Tubbs^{1,2,3,*}, Younga H. Lee^{1,2,3}, Yixuan He^{3,4,5}, Kristin Tsuo^{3,4,5}, Mary T. Yohannes^{3,4}, Lethukuthula L. Nkambule^{3,4}, Emily Madsen², Dirgha J. Ghimire⁶, Sabrina Hermosilla^{6,7}, Tian Ge^{1,2,3}, Alicia R. Martin^{3,4}, William G. Axinn⁶  and Jordan W. Smoller^{1,2,3}

¹Department of Psychiatry, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA;

²Psychiatric and Neurodevelopmental Genetics Unit, Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA, USA; ³Stanley Center for Psychiatric Research, Broad Institute, Boston, MA, USA; ⁴Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA, USA; ⁵Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA, USA; ⁶Population Studies Center, Institute for Social Research, University of Michigan, Ann Arbor, MI, USA and ⁷Department of Population and Family Health, Mailman School of Public Health, Columbia University Irving Medical Center, New York, NY, USA

Abstract

Background. Major depressive disorder (MDD) is the leading cause of disability globally, with moderate heritability and well-established socio-environmental risk factors. Genetic studies have been mostly restricted to European settings, with polygenic scores (PGS) demonstrating low portability across diverse global populations.

Methods. This study examines genetic architecture, polygenic prediction, and socio-environmental correlates of MDD in a family-based sample of 10 032 individuals from Nepal with array genotyping data. We used genome-based restricted maximum likelihood to estimate heritability, applied S-LDXX to estimate the cross-ancestry genetic correlation between Nepalese and European samples, and modeled PGS trained on a GWAS meta-analysis of European and East Asian ancestry samples.

Results. We estimated the narrow-sense heritability of lifetime MDD in Nepal to be 0.26 (95% CI 0.18–0.34, $p = 8.5 \times 10^{-6}$). Our analysis was underpowered to estimate the cross-ancestry genetic correlation ($rg = 0.26$, 95% CI -0.29 to 0.81). MDD risk was associated with higher age ($\beta = 0.071$, 95% CI 0.06 – 0.08), female sex ($\beta = 0.160$, 95% CI 0.15 – 0.17), and childhood exposure to potentially traumatic events ($\beta = 0.050$, 95% CI 0.03 – 0.07), while neither the depression PGS ($\beta = 0.004$, 95% CI -0.004 to 0.01) or its interaction with childhood trauma ($\beta = 0.007$, 95% CI -0.01 to 0.03) were strongly associated with MDD.

Conclusions. Estimates of lifetime MDD heritability in this Nepalese sample were similar to previous European ancestry samples, but PGS trained on European data did not predict MDD in this sample. This may be due to differences in ancestry-linked causal variants, differences in depression phenotyping between the training and target data, or setting-specific environmental factors that modulate genetic effects. Additional research among under-represented global populations will ensure equitable translation of genomic findings.

Introduction

Major depressive disorder (MDD) is the leading cause of disability globally (Global Burden of Disease Collaborative Network, 2020). While it is known to be influenced by both environmental and genetic factors (Kendall *et al.*, 2021) more work is needed to understand the relationship between genetic factors and MDD in a wider range of human populations. Polygenic scores (PGS), which aggregate the small effects of many common variants into a single index, have been widely used to capture variation in genetic liability (Wray *et al.*, 2021). PGS derived from large-scale genome-wide association studies (GWAS) of depression have been found to be associated with depression outcomes in a variety of clinical and population-based cohorts (Fang, Scott, Song, Burmeister, & Sen, 2020; Halldorsdottir *et al.*, 2019), but the vast majority of studies have been concentrated in samples of European ancestry (Peterson *et al.*, 2019). A growing body of research has demonstrated that Eurocentric PGS show poorer predictive performance in other populations, and the expansion of PGS research to more diverse samples has been recognized as an urgent priority (Kachuri *et al.*, 2023; Martin *et al.*, 2019; Wang, Tsuo, Kanai, Neale, & Martin, 2022).

Likewise, investigation of genetic factors from different populations with diverse environmental circumstances is a high priority. It is well-established that exposure to early adversity can increase the odds of lifetime depression (McKay *et al.*, 2022), but it remains unclear how

genetic factors combine with environmental factors to influence depression risk across diverse social and economic settings. We address these fundamental questions with newly available environmental and genetic data from the Chitwan Valley Family Study (CVFS). The CVFS is a large, population-based cohort with genomic and carefully ascertained phenotypic data from over 10 000 individuals in Nepal. Though some studies have examined depression PGS in East Asian samples (Amare *et al.*, 2021; Avinun, Nevo, Radtke, Brigidi, & Hariri, 2020; Bigdeli *et al.*, 2017; Pearson-Fuhrhop *et al.*, 2014), the performance of depression PGS in South Asian settings has been largely unexplored apart from one study of South Asians living in the UK (Truong *et al.*, 2024). The CVFS also includes a wide range of community and family environmental measures (Axinn & Pearce, 2006), and we have previously demonstrated that environmental exposures, including childhood trauma and social support, are associated with risk of MDD in the CVFS (Axinn *et al.*, 2022; Benjet *et al.*, 2020; Hermosilla *et al.*, 2022). Leveraging the CVFS, this study examines the heritability and cross-ancestry genetic correlations for lifetime MDD in Nepal. It also assesses the associations between polygenic risk for depression and lifetime MDD, on its own and alongside a potent environmental factor – namely childhood exposure to potentially traumatic events – which may also influence MDD risk.

Methods

Sample

The CVFS features a general population sample of 151 neighborhoods, fully representative of Western Chitwan in Nepal (Axinn, Ghimire, & Williams, 2012). Years of previous research in Nepal by this team, combined with years of living in Chitwan among the study population, drove the innovative combinations of ethnography and survey research that characterize the CVFS (Axinn *et al.*, 2012; Axinn & Pearce, 2006). The panel study was launched 1995, following whole families across time, with the sample being refreshed periodically to maintain full representation of the general population. Across more than 20 years, multiple rounds of interviews have generated very high-quality data with high response rates, low panel attrition, low levels of item missing rates, and high reliability in measures (Axinn *et al.*, 2012; Thornton, Ghimire, & Mitchell, 2012). From 2016–2018 selected modules from the World Mental Health Composite International Diagnostic Interview (WMH-CIDI 3.0) were administered to the full CVFS sample aged 15–59. The current sample was selected to be representative of Western Chitwan in 2016. The response rate for this survey was 93%, generating 10 714 completed interviews (Scott *et al.*, 2021), with 10 308 also providing saliva-based DNA samples. All procedures were approved by the University of Michigan Institutional Review Board (HUM00104171) and by the Nepal Health Research Council. Written or verbal informed consent was obtained from all participants.

Phenotypic measures

Outcome: major depressive disorder (MDD)

As explained in detail elsewhere (Kessler & Üstün, 2004), the WMH-CIDI is a fully structured diagnostic instrument administered by lay interviewers using computer-assisted methods. To create a Nepal-specific version of the WMH-CIDI measures, a multiethnic team of CVFS researchers applied state-of-the-art survey methodology grounded in a mixed-method informed

understanding of the local setting (Scott *et al.*, 2021). The initial process took more than three years to produce clinically valid measures (Ghimire, Chardoul, Kessler, Axinn, & Adhikari, 2013). Building on this accomplishment, the CVFS team then created a setting-specific life history calendar (LHC) to pair with the WMH-CIDI. This LHC approach adds measures of highly memorable personal and family events to the memory cues given to respondents, creating personalized memory ‘anchors’ to facilitate recall of the occurrence of psychiatric symptoms. This LHC-CIDI method significantly improved measurement of lifetime experience with mental disorder (Axinn *et al.*, 2020). The LHC-CIDI was validated in the largest clinical validation ever conducted in Nepal, demonstrating concordance equaling or exceeding the performance of the WMH-CIDI in Europe and the United States (Axinn *et al.*, 2020). Here we use the Nepal-specific LHC-CIDI measure of lifetime occurrence of major depressive disorder (MDD).

Exposure: Childhood exposure to potentially traumatic events (PTE)

Two binary exposure variables, being beaten badly by a parent or caregiver and witnessing serious physical fights at home as a child, collected by trained interviewers each through a single question (e.g. ‘As a child, were you ever badly beaten up by your parents or the people who raised you?’) as part of the WMH-CIDI were combined into a single dichotomous variable of childhood exposure to PTE (coded 1 for endorsement of either experience and 0 for no endorsement of these experiences).

Covariates: sociodemographic variables

Sociodemographic covariates included age (in years), sex, and ethnicity. While race/ethnicity designations are complex in Nepal, the CVFS study population comprises six categories that capture key aspects of variation: Brahmin/Chhetri; Dalits; Hill Janjati (multiple ethnic groups of Tibetan origin, primarily Buddhist); Terai Janjati (multiple plains ethnic groups, primarily of Burmese descent); Newar; and other (primarily recent and temporary migrants from India).

Genomic quality control and imputation

10 294 DNA samples from the CVFS cohort were genotyped using the Illumina Infinium Global Screening Array and aligned to the human genome reference build GRCh38. Genomic quality control (QC) was performed using a high-performance cloud-based pipeline called GWASpy (details available here: <https://github.com/atgu/GWASpy>). A total of 76 samples were excluded due to low genotyping call rate, inbreeding coefficient greater than 0.2, mismatch between self-reported and genetic sex, or more than 10 000 Mendelian errors observed. Additionally, 32 samples were excluded from further analysis which appeared to be duplicates or members of a monozygotic twin pair. After genomic quality control, a total of 10 032 unique samples were available for subsequent analysis. Genetic principal components (PCs) were calculated in an unrelated subset of the data, and the remaining samples were projected onto this PC space. Visual inspection of PC plots revealed no outliers. Of note, phasing and imputation were performed using a new jointly called dataset of harmonized Human Genome Diversity Project (HGDP) and 1000 Genomes (1KG) Project samples (Koenig *et al.*, 2023). In other words, this reference panel is much more diverse and deeply sequenced than existing imputation panels, making it more appropriate for use in this diverse Nepalese cohort.

Heritability estimation

Given the household-based sampling methodology, this study contains related individuals. Therefore, we adopted a genome-based restricted maximum likelihood framework for heritability estimation using the Genome-wide Complex Trait Analysis (GCTA) software version 1.94.1 (Lee, Wray, Goddard, & Visscher, 2011; Yang, Lee, Goddard, & Visscher, 2011). Following the method introduced by Zaitlen et al. (2013) GCTA was used to calculate a dense genetic relatedness matrix (GRM_{all}) from all autosomal SNPs with minor allele frequency greater than 0.01. A sparse GRM (GRM_{close}) was also constructed, where all elements in the GRM less than 0.05 were set to zero, thereby only capturing relatedness among close family members. Subsequently, GCTA was used to simultaneously model the binary lifetime depression outcome as a function of GRM_{all} or GRM_{close} treated as random effects, along with age, sex, ethnicity, and the first 20 genetic PCs fitted as fixed effects. This model thus produces estimates of SNP-based heritability (h^2_{snp}) and narrow-sense heritability (h^2), an estimator which is comparable to that from traditional pedigree-based study designs. h^2_{snp} is estimated as the proportion of phenotypic variance explained by GRM_{all} , while h^2 is the sum of h^2_{snp} and h^2_{kin} , the variance explained by GRM_{close} . We subsequently fitted two additional models for males and females independently. Since this is a population-based sample and the prevalence of depression in Nepal is not well-characterized, the sample prevalence was used when transforming heritability estimates to the liability scale.

Cross-ancestry genetic correlation and GWAS

We used S-LDXR (Shi et al., 2021) to estimate the SNP-based cross-ancestry genetic correlation between summary statistics from GWAS of depression performed in a sample of European ancestry and a GWAS of lifetime MDD conducted in the current Nepalese sample. A generalized linear mixed model GWAS was performed using GCTA (Jiang, Zheng, Fang, & Yang, 2021) in this sample with depression case-control status as the outcome, fitting the GRM_{close} as a random genetic effect along with sex, age, ethnicity, and the first 20 PCs as fixed effect covariates. Summary statistics from a GWAS of strictly defined lifetime MDD in a sample of 67 171 European ancestry participants from the UK Biobank was used (Cai et al., 2020), as this phenotype more closely resembles the definition used in the current Nepalese sample. S-LDXR was run using the default parameters and European and South Asian sub-samples from the 1000 Genomes Project Phase 3 release were used as reference LD panels.

Polygenic scoring

PGS for depression were generated based on a meta-analysis of two sets of publicly available GWAS summary statistics for major depression, from Howard *et al.* (2019, 2018) in $n = 500$ 199 individuals of European ancestry and Giannakopoulou *et al.* (2021) in 98 502 individuals of East Asian ancestry, which were combined using the inverse variance weighted approach implemented in METAL (Willer, Li, & Abecasis, 2010) prior to score generation. PRS-CS (Ge, Chen, Ni, Feng, & Smoller, 2019), a Bayesian polygenic scoring method, was used to generate weighted SNP effect estimates for major depression, which were then scored using PLINK v1.9. Although our primary analysis was based on the continuous PGS, we also divided the continuous PGS into categorical tertiles reflecting relatively low, intermediate, and high polygenic risk, for secondary analyses.

Polygenic score analyses

Building on our heritability analyses, we fitted a series of sequential mixed models using GCTA with lifetime MDD as the outcome, where the baseline model (M0) included GRM_{all} and GRM_{close} as random effects and the first 20 PCs as fixed effects. Models M1 to M3 sequentially updated the preceding model to include additional demographic variables: age, sex, and ethnicity. Model M4 built on model M3 by including the scaled continuous PGS as a fixed effect. Model M5 subsequently included childhood PTE, while model M6 also included a term for the multiplicative interaction between PGS and childhood PTE. As supplementary analyses, Models M4–M6 were refitted using categorical PGS tertiles. The unique phenotypic variance explained (R^2) on the observed scale, controlling for variables included in prior models, was estimated for each fixed effect as the relative change in phenotypic residual variance (Ve) after inclusion of a given fixed effect (e.g. the R^2 for age is calculated as $(Ve_{M0} - Ve_{M1})/Ve_{M0}$). These estimates were transformed to the liability scale using the method proposed by Lee, Goddard, Wray, and Visscher (2012) assuming the sample prevalence (15%) as the population prevalence.

Results

Sample characteristics

Following genomic QC filtering, a total analytic sample of 10 032 individuals was available for modeling with complete data on MDD status, demographic covariates, and childhood PTE. Sample characteristics are reported in Table 1. Of the 10 032 participants, 55% were female and 43% were of Brahmin ethnicity, with a mean age of 35.4 years (s.d. = 12.4). Fifteen percent of

Table 1. Analytic sample characteristics

	Analytic sample ($N = 10\,032$)	
	<i>N</i>	%
Median age in 2019 (median [IQR])	33.0 [25.0–44.0]	
Sex assigned at birth		
Male	4517	45.0
Female	5515	55.0
Ethnicity		
Brahmin/Chhetri (Hindu)	4334	43.2
Hill Janajati (Tibetoburmise)	1988	19.8
Dalit (Hindu)	1242	12.4
Newar	603	6.0
Terai Janajati (Tibetoburmise)	1865	18.6
DSM-IV Diagnosis of MDD (with hierarchy)		
Non-case	8489	84.6
Case	1543	15.4
Childhood exposure to PTE (%)		
Unexposed	8696	86.7
Exposed	1336	13.3

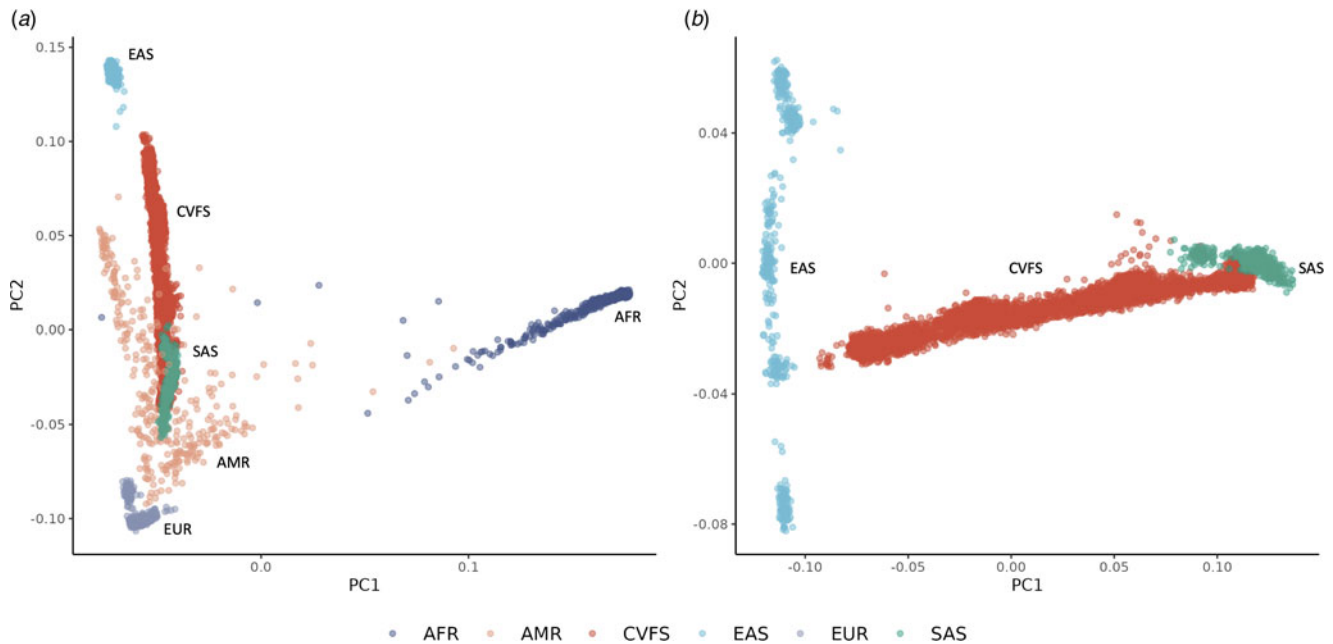


Figure 1. Genetic ancestry of the current CVFS sample.

Note. This figure situates the current Nepalese sample from the Chitwan Valley Family Study (CVFS) in the context of continental superpopulations using samples from the 1000 Genomes Project Phase 3 release. We performed principal components analysis across all 1000 Genomes samples (Panel a) and within the subset of samples with East and South Asian ancestry (Panel b). All samples (including those from CVFS) were subsequently projected onto the principal component space and plotted. The 1000 Genomes project clusters samples into the following superpopulations: African (AFR), Admixed American (AMR), East Asian (EAS), European (EUR) and South Asian (SAS).

the sample met lifetime diagnostic criteria for MDD. The lifetime prevalence of MDD among women in this sample was 22% and among men was 7%. MDD cases were on average 5.9 years older than non-MDD cases.

To further situate the genetic ancestry of this sample among global populations, we performed principal component analysis within the unrelated set of 1000 Genomes Phase 3 samples and projected samples from the current study onto this PC space (Fig. 1a). We also estimated genetic PCs only among East and South Asian ancestry samples from 1000 genomes, and subsequently projected the CVFS samples onto this space (Fig. 1b). As shown in Fig. 1, the CVFS Nepalese samples lie on a continuum between East Asian samples and South Asian samples across the first two principal components, with some samples clustering more closely to samples of East Asian ancestry and some more closely resembling South Asian ancestry samples. This is consistent with findings from previous studies (Arciero et al., 2018; Xing et al., 2010).

Heritability

GCTA analyses indicated that lifetime MDD in this Nepal-based cohort was significantly heritable. Figure 2 plots heritability estimates from the current Nepalese sample with those from previous studies in majority European ancestry samples (Fernandez-Pujals et al., 2015; Polderman et al., 2015). On the liability scale, the narrow-sense h^2 of lifetime MDD was estimated at 0.26 (95% CI 0.18–0.34, $p = 8.5 \times 10^{-6}$). We also report h^2 estimates on the liability scale for several alternative possible assumed prevalences ranging from 5% to 25% (online Supplementary Table S1). Point estimates for narrow-sense h^2 on the liability scale were higher in both the female-only (0.35, 95% CI 0.21–0.49, $p = 2.0 \times 10^{-4}$) and male-only (0.45, 95% CI 0.20 to 0.70, $p = 1.2 \times 10^{-2}$) sub-sample.

Although the method of Zaitlen et al. (2013) produces unbiased estimates of h^2 , the estimates of individual variance components attributable to h^2_{snp} and h^2_{kin} have been found to be severely misestimated (Evans et al., 2018). Thus, we only report estimates of the overall h^2 .

GWAS and cross-ancestry genetic correlation

S-LD XR estimated a genetic correlation of 0.26 (95% CI –0.29 to 0.81) between the Nepalese lifetime MDD GWAS and the European lifetime MDD GWAS. Our current sample was underpowered to identify a statistically significant cross-ancestry genetic correlation with lifetime MDD using GWAS summary statistics.

Associations between polygenic risk, childhood PTE, and lifetime MDD

Regression coefficients for fixed effects estimated from the final saturated model (M6) are shown in Table 2, while the sequential variance explained by each added predictor is provided in Table 3. Sex explained the largest proportion of variance in liability to depression (12.0%, 95% CI 10.8–13.2), followed by age (7.0%, 95% CI 5.9–7.8%) and childhood PTE (0.6, 95% CI 0.3–0.9%). The remaining fixed effect predictors, including ethnicity, the continuous PGS, and the interaction between PGS and childhood PTE, either resulted in a slightly worse model fit after inclusion or had a 95% confidence interval that overlapped with zero h^2_{snp} . Results were consistent when examining the PGS categorically by tertiles, with PGS tertiles accounting for only 0.04% (95% CI –0.04 to 0.12%) of variance in MDD liability, and its interaction with childhood PTE explaining 0.64% (95% CI 0.33–0.95%). As a sensitivity analysis, we also tested the performance of a model where the PGS was modeled with only genetic PCs as covariates.

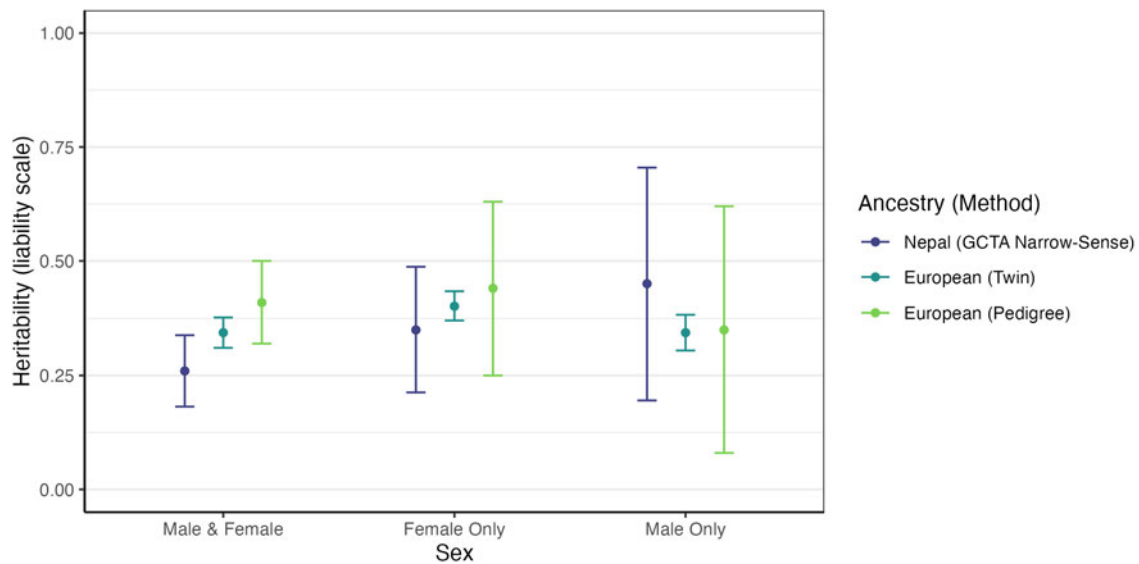


Figure 2. Estimated heritability of MDD in Nepal compared to previous European studies. Note. The current study estimated pedigree-based heritability using the GCTA software in 10 032 individuals from Nepal. Twin-based heritability estimates were obtained from the largest meta-analysis of depression (Polderman et al., 2015) including 178 789 twin pairs in the combined male-female analysis, with the majority of contributing studies performed in samples of European ancestry. European ancestry pedigree-based heritability estimates were obtained from the Scottish Family Health Study, which included 20 198 individuals (Fernandez-Pujals et al., 2015).

Table 2. Estimated effects of demographics, polygenic scores, and childhood trauma on lifetime MDD risk

Coefficient	Beta	s.e.	95% Confidence Interval
Intercept	0.056	0.011	0.034–0.078
PGS	0.004	0.004	–0.004 to 0.012
cPTE	0.050	0.010	0.029–0.070
PGS × cPTE	0.007	0.010	–0.013 to 0.027
Age	0.071	0.003	0.064–0.077
Sex (Female)	0.160	0.007	0.146–0.173
Ethnicity (Dalit)	0.040	0.029	–0.017 to 0.098
Ethnicity (Hill Janajati)	0.001	0.020	–0.039 to 0.040
Ethnicity (Newar)	–0.026	0.030	–0.085 to 0.033
Ethnicity (Terai Janajait)	0.001	0.029	–0.055 to 0.057

Note. This table shows regression coefficient estimates from a linear mixed model fitted using the GCTA software with binary lifetime MDD status as the outcome and all predictors as covariates. Two random genetic effect structures were fitted simultaneously: the full genetic relatedness matrix (GRM) and a sparse GRM where pairwise entries less than 0.05 were set to zero. This model also included the first 20 genetic principal components as fixed effect covariates. All continuous predictors were mean-centered and standardized to have unit variance prior to model fitting. cPTE, childhood exposure to potentially traumatic event; PGS, polygenic score.

This yielded an effect size very similar to, and statistically indistinguishable from, the estimate when additionally controlling for age, sex, and ethnicity.

Although the mixed model framework maximizes the amount of available data by explicitly modeling relatedness, there may be a concern that the random effects modeled by the GRM overlap with the effects captured by the PGS, resulting in larger standard errors and reduced power to detect an effect of the PGS. Thus, we performed a sensitivity analysis, using linear regression on binary case-control status to model the effects of PGS while controlling

Table 3. Sequential variance explained in lifetime MDD

Coefficient	$\Delta R^2_{liability}$	s.e.	95% Confidence Interval
Age	6.8×10^{-2}	4.9×10^{-3}	5.9×10^{-2} – 7.8×10^{-2}
Sex	1.2×10^{-1}	6.1×10^{-3}	1.1×10^{-1} – 1.3×10^{-1}
Ethnicity	-1.6×10^{-4}	NA	NA
PGS	2.0×10^{-4}	2.8×10^{-4}	-3.5×10^{-4} to 7.4×10^{-4}
cPTE	6.4×10^{-3}	1.6×10^{-3}	3.3×10^{-3} – 9.5×10^{-3}
PGS × cPTE	-3.9×10^{-5}	NA	NA

Note. Sequential linear mixed models were fit using GCTA with lifetime major depression status as the outcome variable, building on a baseline model where random effects included the full genetic relatedness matrix (GRM) and a sparse GRM where pairwise entries less than 0.05 were set to zero, while fixed effects included the first 20 genetic principal components. Age, sex, ethnicity, depression polygenic score (PGS), childhood exposure to potentially traumatic event (cPTE), and their multiplicative interaction were included as fixed effects in a stepwise manner. The unique phenotypic variance explained (R^2) on the observed scale, controlling for variables included in prior models, was estimated for each fixed effect as the relative change in phenotypic residual variance after inclusion of a given fixed effect. These were transformed to variance explained on the liability scale ($R^2_{liability}$) assuming a population prevalence of 15%. Since the inclusion of ethnicity and PGS × cPTE terms resulted in a slightly worse fit, it is not meaningful to compute standard errors or confidence intervals.

for age, sex, and PCs in an unrelated (less than 2nd degree) subset of 4606 subjects from the full dataset. Results from this analysis were highly consistent with the mixed model results, with the PGS effect estimated as 0.005 (95% CI –0.003 to 0.012), nominally higher than the mixed model effect estimate, but statistically indistinguishable.

Discussion

In this study, we examined genetic factors contributing to major depression risk in an ancestrally diverse, household-based sample of 10 032 individuals with high relatedness in Western Chitwan,

Nepal. 15% of the sample met diagnostic criteria on a validated interview survey for lifetime MDD, comparable to that seen in US samples (Kessler *et al.*, 2003). Lifetime MDD showed statistically significant narrow-sense heritability; however, polygenic scores for depression trained on large European and East Asian ancestry GWAS did not significantly predict lifetime MDD status in this Nepalese sample. Demographic variables and environmental exposures explained a far greater proportion of variance in liability to lifetime MDD in the current sample.

Genetic contributions to depression are not well characterized outside of European ancestry populations. In this Nepalese sample, lifetime MDD was found to have a statistically significant narrow-sense heritability, explaining 26% of variance in risk for MDD on the liability scale. This estimate is slightly lower but comparable to previously reported heritability estimates of lifetime MDD from European ancestry twin and family-based samples (Fig. 2), which range from 30%–50% (Kendall *et al.*, 2021) indicating a substantial genetic basis. In addition, it is similar to the liability-scale SNP heritability (26%) reported from CONVERGE, a cohort of East Asian ancestry (Giannakopoulou *et al.*, 2021), though the latter was estimated based on severe recurrent depression at a lower prevalence.

In contrast, depression PGS trained on a meta-analysis of two large GWAS from European and East Asian ancestry samples performed poorly in predicting lifetime MDD in this sample. Limited previous studies have yielded mixed evidence on the predictive power of PGS for depression trained on European GWAS for depression outcomes in non-European populations: PGS associations, albeit with lower variance explained, have been observed for depressive symptoms in a cohort of over 3000 pregnant individuals in Peru (Shen *et al.*, 2020) or for combined anxiety/depression status among individuals of African and Hispanic ancestry in US health systems (Coombes *et al.*, 2023), though a large sample of US veterans with African ancestry reported no such association (Bigdeli *et al.*, 2022). While it is known that PGS show reduced effect sizes when applied to samples that differ in ancestry from the training sample (Wang *et al.*, 2022), our observed discrepancy may have a number of possible explanations. First, it could reflect differences in the underlying genetic architecture of MDD across ancestral populations. We attempted to quantify these differences using S-LDXR, a method for estimating cross-ancestry genetic correlation, which suggested only a modest correlation ($r_g = 0.26$) in common genome-wide factors for lifetime depression across European and Nepalese samples, although we were underpowered to make precise estimates. This is consistent with a previous GWAS that revealed only a partial overlap (11%) between the genetic loci associated with major depression in a European sample and those in individuals of East Asian ancestry (Giannakopoulou *et al.*, 2021). Second, low cross-ancestry PGS performance could reflect differences in how lifetime MDD is measured and reported across settings, including the experience and expression of specific symptoms. However, the Nepal-CIDI instrument has been clinically validated (Axinn *et al.*, 2020). Importantly to note, a large proportion of samples used to train our PGS came from a GWAS which used a shallow definition of depression (Howard *et al.*, 2018), which have been shown to have lower heritability and decreased specificity (Cai *et al.*, 2020). Third, different factors may contribute to MDD risk in the Nepalese setting; for example, this population had higher levels of exposure to poverty and violent events relative to other settings assessed with the CIDI. At the same time, recent studies document a lower population-based prevalence of MDD and

other disorders in the CVFS compared to more wealthy European diaspora settings (Scott *et al.*, 2021). The sources of this ‘resilience’ remain to be documented, but rates of psychiatric disorder are steadily rising across birth cohorts in this study sample, possibly due to higher rates of schooling, work for pay, and geographic moves for work (Scott *et al.*, 2021). As social life in Nepal becomes more similar to other settings represented in our GWAS training data, the PGS may also become a stronger predictor of MDD.

Exposure to childhood trauma, a robust predictor of depression (McKay *et al.*, 2022), was statistically significantly associated with lifetime MDD in this sample, although it explained modest variance compared to the demographic factors. Factors that most strongly explained lifetime MDD status included sex and age, with females and older participants reporting higher rates of lifetime MDD. Stratified heritability analyses suggested that the heritability of lifetime MDD liability may be higher in males than females, although the difference between the two estimates was not statistically significant. Previous findings in European samples (Polderman *et al.*, 2015) have found heritability to be higher among females than males. However, our study appears to be underpowered to detect sex differences in heritability, since the 95% confidence interval for the male-only heritability estimate (0.20–0.70) overlapped with those from both the combined sample estimate (0.18–0.34) and the female-only estimate (0.21–0.49). One explanation for higher point estimates in this sample when males and females are modeled separately is that the genetic factors influencing depression liability may be different for males and females, introducing additional heterogeneity in effects and deflating heritability estimates when modeled together. Indeed, there is some evidence from family studies in European samples that genetic risk factors differ between males and females (Kendler, Gardner, Neale, & Prescott, 2001; Kendler, Gatz, Gardner, & Pedersen, 2006). Another explanation may be indirect parental genetic effects that differ between mothers and fathers and can bias GCTA estimates of heritability (Barry *et al.*, 2023). One recent study has provided preliminary evidence for parent-specific genetic nurture effects in samples of European ancestry (Tubbs & Sham, 2023).

Strengths of this study include, first, its focus on a genetically characterized population-based sample in Western Chitwan in Nepal, addressing the gap of limited research in non-European ancestry populations (Martin *et al.*, 2019; Wang *et al.*, 2022). To date, there have only been a handful of psychiatric genetics studies in South Asian populations (Periyasamy *et al.*, 2019) and none focused on MDD. Second, this study leverages a highly unique cohort where depression has been rigorously ascertained, with detailed and culturally validated diagnostic assessment of lifetime MDD using life history calendar methods to ensure recall in the context of a long-term, household-based longitudinal study, which is usually challenging to obtain in genomic studies. Third, because it is from a whole-family longitudinal study, this population-representative sample captured large numbers of related individuals to maximize power for heritability estimation.

This study also had several limitations. Notably, while one of the largest of its kind, this sample was still underpowered for detecting smaller polygenic effects or estimate statistically significant cross-ancestry genetic correlations, or for identifying potential sex differences in the heritability of MDD in this Nepalese population. This points to the need to collect more data to enable more well-powered genetic analyses of psychiatric phenotypes in

South Asia. We also relied on GWAS summary statistics derived from European and East Asian ancestry populations, which are the largest available but less likely to be directly transferable to this population. Additionally, we examined a relatively simple measure of childhood trauma as a candidate exposure, measured retrospectively, and lifetime MDD is likely to be associated with a wider range of environmental factors (Köhler et al., 2018) that may benefit from more comprehensive study in this population in conjunction with genetic factors.

Conclusion

In this unique cohort of 10 032 individuals in Nepal, we find evidence that lifetime MDD has a substantial genetic basis in this population, which partially overlaps with the genetic architecture of MDD in previous European ancestry studies. However, the low genetic correlation for MDD between this sample and European samples ($r_g = 0.26$) and the limited predictive performance of PGS derived from European and East Asian studies highlight the need for expanded genetic studies in South Asian populations. Demographic variables and specific environmental exposures such as childhood trauma emerged as more influential risk factors for lifetime MDD. Future research should focus on characterizing the genetic architecture of MDD in settings like Nepal and identifying population-specific risk factors for depression and other common psychiatric disorders. Findings underscore the importance of considering genetic, environmental, and sociocultural factors to enhance our understanding of MDD in diverse populations.

Supplementary material. The supplementary material for this article can be found at <https://doi.org/10.1017/S0033291724001284>

Acknowledgements. The authors thank the respondents of the CVFS, whose generous contributions made this research possible; the survey staff of the Institute for Social and Environmental Research–Nepal for collecting the data reported here; the staff of the Survey Research Operations unit of the University of Michigan's Survey Research Center for development and support of the technical systems that made the fieldwork in Nepal possible; and the World Mental Health Consortium leadership and staff at Harvard University for their input into the design and all subsequent steps of collecting and analyzing the data reported here. The authors thank the genomics team at the Broad Institute for sample processing and genotyping. The authors also thank Paul Schulz and Heather Gatny for data support and Jennifer Mamer and Alison Shereda for research and editorial assistance. The authors alone remain responsible for any errors or omissions in this manuscript.

Funding statement. This project was supported by the National Institute of Mental Health (JWS, WA, grant number R01MH110872), (KWC, grant number K08MH127413); a NARSAD Brain and Behavior Foundation Young Investigator Award (KWC, no grant number); the National Human Genome Research Institute (JDT, grant number T32HG010464); and a Eunice Kennedy Shriver National Institute of Child Health and Human Development Center Grant (DG, WA, P2CHD041028) to the Population Studies Center at the University of Michigan.

Competing interests. Ghimire is faculty at the University of Michigan and also the Director of the Institute for Social and Environmental Research in Nepal (ISER-N) that collected the data for the research reported here. Ghimire's conflict of interest management plan is approved and monitored by the Regents of the University of Michigan.

Ethical standards. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

References

- Amare, A. T., Schubert, K. O., Hou, L., Clark, S. R., Papiol, S., Cearnas, M., ... Baune, B. T. (2021). Association of polygenic score for major depression with response to lithium in patients with bipolar disorder. *Molecular Psychiatry*, 26(6), 2457–2470. <https://doi.org/10.1038/s41380-020-0689-5>
- Arciero, E., Kraaijenbrink, T., Asan, Haber, M., Mezzavilla, M., Ayub, Q., ... Tyler-Smith, C. (2018). Demographic history and genetic adaptation in the Himalayan region inferred from genome-wide SNP genotypes of 49 populations. *Molecular Biology and Evolution*, 35(8), 1916–1933. <https://doi.org/10.1093/molbev/msy094>
- Avinun, R., Nevo, A., Radtke, S. R., Brigidi, B. D., & Hariri, A. R. (2020). Divergence of an association between depressive symptoms and a dopamine polygenic score in Caucasians and Asians. *European Archives of Psychiatry and Clinical Neuroscience*, 270(2), 229–235. <https://doi.org/10.1007/s00406-019-01040-x>
- Axinn, W. G., & Pearce, L. D. (2006). *Mixed method data collection strategies*. New York, NY: Cambridge University Press.
- Axinn, W. G., Ghimire, D. J., & Williams, N. E. (2012). Collecting survey data during armed conflict. *Journal of Official Statistics*, 28(2), 153–171.
- Axinn, W. G., Chardoul, S., Gatny, H., Ghimire, D. J., Smoller, J. W., Zhang, Y., & Scott, K. M. (2020). Using life history calendars to improve measurement of lifetime experience with mental disorders. *Psychological Medicine*, 50(3), 515–522. Cambridge Core. <https://doi.org/10.1017/S0033291719000394>
- Axinn, W. G., Choi, K. W., Ghimire, D. J., Cole, F., Hermsilla, S., Benjet, C., ... Smoller, J. W. (2022). Community-level social support infrastructure and adult onset of major depressive disorder in a South Asian postconflict setting. *JAMA Psychiatry*, 79(3), 243–249. <https://doi.org/10.1001/jamapsychiatry.2021.4052>
- Barry, C.-J. S., Walker, V. M., Cheesman, R., Davey Smith, G., Morris, T. T., & Davies, N. M. (2023). How to estimate heritability: A guide for genetic epidemiologists. *International Journal of Epidemiology*, 52(2), 624–632. <https://doi.org/10.1093/ije/dyac224>
- Benjet, C., Axinn, W. G., Hermsilla, S., Schulz, P., Cole, F., Sampson, L., & Ghimire, D. (2020). Exposure to armed conflict in childhood vs older ages and subsequent onset of major depressive disorder. *JAMA Network Open*, 3(11), e2019848–e2019848. <https://doi.org/10.1001/jamanetworkopen.2020.19848>
- Bigdeli, T. B., Ripke, S., Peterson, R. E., Trzaskowski, M., Bacanu, S.-A., Abdellaoui, A., ... Kendler, K. S. (2017). Genetic effects influencing risk for major depressive disorder in China and Europe. *Translational Psychiatry*, 7(3), e1074. <https://doi.org/10.1038/tp.2016.292>
- Bigdeli, T. B., Voloudakis, G., Barr, P. B., Gorman, B. R., Genovese, G., Peterson, R. E., ... Cooperative Studies Program (CSP) #572 and Million Veteran Program (MVP). (2022). Penetrance and pleiotropy of polygenic risk scores for schizophrenia, bipolar disorder, and depression among adults in the US Veterans Affairs health care system. *JAMA Psychiatry*, 79(11), 1092–1101. <https://doi.org/10.1001/jamapsychiatry.2022.2742>
- Cai, N., Revez, J. A., Adams, M. J., Andlauer, T. F. M., Breen, G., Byrne, E. M., ... Flint, J. (2020). Minimal phenotyping yields genome-wide association signals of low specificity for major depression. *Nature Genetics*, 52(4), 437–447. <https://doi.org/10.1038/s41588-020-0594-5>
- Coomes, B. J., Landi, I., Choi, K. W., Singh, K., Fennessy, B., Jenkins, G. D., ... Biernacka, J. M. (2023). The genetic contribution to the comorbidity of depression and anxiety: A multi-site electronic health records study of almost 178 000 people. *Psychological Medicine*, 53(15), 7368–7374. <https://doi.org/10.1017/S0033291723000983>
- Evans, L. M., Tahmasbi, R., Vrieze, S. I., Abecasis, G. R., Das, S., Gazal, S., ... Keller, M. C. (2018). Comparison of methods that use whole genome data to estimate the heritability and genetic architecture of complex traits. *Nature Genetics*, 50(5), 737–745. <https://doi.org/10.1038/s41588-018-0108-x>
- Fang, Y., Scott, L., Song, P., Burmeister, M., & Sen, S. (2020). Genomic prediction of depression risk and resilience under stress. *Nature Human Behaviour*, 4(1), 111–118. <https://doi.org/10.1038/s41562-019-0759-3>
- Fernandez-Pujals, A. M., Adams, M. J., Thomson, P., McKechnie, A. G., Blackwood, D. H. R., Smith, B. H., ... McIntosh, A. M. (2015). Epidemiology and heritability of major depressive disorder, stratified by age of onset, sex, and illness course in Generation Scotland: Scottish

- Family Health Study (GS:SFHS). *PLoS ONE*, 10(11), e0142197. <https://doi.org/10.1371/journal.pone.0142197>
- Ge, T., Chen, C.-Y., Ni, Y., Feng, Y.-C. A., & Smoller, J. W. (2019). Polygenic prediction via Bayesian regression and continuous shrinkage priors. *Nature Communications*, 10(1), 1776. <https://doi.org/10.1038/s41467-019-09718-5>
- Ghimire, D. J., Chardoul, S., Kessler, R. C., Axinn, W. G., & Adhikari, B. P. (2013). Modifying and validating the Composite International Diagnostic Interview (CIDI) for use in Nepal. *International Journal of Methods in Psychiatric Research*, 22(1), 71–81. (23494758). <https://doi.org/10.1002/mpr.1375>
- Giannakopoulou, O., Lin, K., Meng, X., Su, M.-H., Kuo, P.-H., Peterson, R. E., ... 23andMe Research Team, China Kadoorie Biobank Collaborative Group, and Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium. (2021). The genetic architecture of depression in individuals of east Asian ancestry: A genome-wide association study. *JAMA Psychiatry*, 78(11), 1258–1269. <https://doi.org/10.1001/jamapsychiatry.2021.2099>
- Global Burden of Disease Collaborative Network. (2020). Global Burden of Disease Study 2019 (GBD 2019) Results. Retrieved August 25, 2022, from Institute for Health Metrics and Evaluation website. <https://vizhub.healthdata.org/gbd-results/>. Accessed August 25, 2022.
- Halldorsdottir, T., Piechaczek, C., Soares de Matos, A. P., Czamara, D., Pehl, V., Wagenbuechler, P., ... Binder, E. B. (2019). Polygenic risk: Predicting depression outcomes in clinical and epidemiological cohorts of youths. *American Journal of Psychiatry*, 176(8), 615–625. <https://doi.org/10.1176/appi.ajp.2019.18091014>
- Hermosilla, S., Choi, K. W., Askari, M. S., Marks, T., Denckla, C., Axinn, W., ... Benjet, C. (2022). What can we learn about polytrauma typologies by comparing population-representative to trauma-exposed samples: A Nepali example. *Journal of Affective Disorders*, 314, 201–210. <https://doi.org/10.1016/j.jad.2022.07.006>
- Howard, D. M., Adams, M. J., Shirali, M., Clarke, T.-K., Marioni, R. E., Davies, G., ... McIntosh, A. M. (2018). Genome-wide association study of depression phenotypes in UK Biobank identifies variants in excitatory synaptic pathways. *Nature Communications*, 9(1), 1470. <https://doi.org/10.1038/s41467-018-03819-3>
- Howard, D. M., Adams, M. J., Clarke, T.-K., Hafferty, J. D., Gibson, J., Shirali, M., ... McIntosh, A. M. (2019). Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nature Neuroscience*, 22(3), 343–352. <https://doi.org/10.1038/s41593-018-0326-7>
- Jiang, L., Zheng, Z., Fang, H., & Yang, J. (2021). A generalized linear mixed model association tool for biobank-scale data. *Nature Genetics*, 53, 1616–1621. <https://doi.org/doi.org/10.1038/s41588-021-00954-4>
- Kachuri, L., Chatterjee, N., Hirbo, J., Schaid, D. J., Martin, I., Kullo, I. J., ... Ge, T. (2023). Principles and methods for transferring polygenic risk scores across global populations. *Nature Reviews Genetics*, 25(1), 8–25. <https://doi.org/10.1038/s41576-023-00637-2>
- Kendall, K. M., Assche, E. V., Andlauer, T. F. M., Choi, K. W., Luykx, J. J., Schulte, E. C., & Lu, Y. (2021). The genetic basis of major depression. *Psychological Medicine*, 51(13), 2217–2230. <https://doi.org/10.1017/S0033291721000441>
- Kendler, K. S., Gardner, C. O., Neale, M. C., & Prescott, C. A. (2001). Genetic risk factors for major depression in men and women: Similar or different heritabilities and same or partly distinct genes? *Psychological Medicine*, 31(4), 605–616. <https://doi.org/10.1017/S0033291701003907>
- Kendler, K. S., Gatz, M., Gardner, C. O., & Pedersen, N. L. (2006). A Swedish national twin study of lifetime major depression. *American Journal of Psychiatry*, 163(1), 109–114. <https://doi.org/10.1176/appi.ajp.163.1.109>
- Kessler, R. C., & Üstün, T. B. (2004). The World Mental Health (WMH) survey initiative version of the World Health Organization (WHO) Composite International Diagnostic Interview (CIDI). *International Journal of Methods in Psychiatric Research*, 13(2), 93–121. <https://doi.org/10.1002/mpr.168>
- Kessler, R. C., Berglund, P., Demler, O., Jin, R., Koretz, D., Merikangas, K. R., ... Wang, P. S. (2003). The epidemiology of major depressive disorder: Results from the national comorbidity survey replication (NCS-R). *JAMA*, 289(23), 3095–3105. <https://doi.org/10.1001/jama.289.23.3095>
- Koenig, Z., Yohannes, M. T., Nkambule, L. L., Goodrich, J. K., Kim, H. A., Zhao, X., ... Martin, A. R. (2023). A harmonized public resource of deeply sequenced diverse human genomes. *BioRxiv: The Preprint Server for Biology*, 01.23.525248. <https://doi.org/10.1101/2023.01.23.525248>
- Köhler, C. A., Evangelou, E., Stubbs, B., Solmi, M., Veronese, N., Belbasis, L., ... Carvalho, A. F. (2018). Mapping risk factors for depression across the lifespan: An umbrella review of evidence from meta-analyses and Mendelian randomization studies. *Journal of Psychiatric Research*, 103, 189–207. <https://doi.org/10.1016/j.jpsychires.2018.05.020>
- Lee, S. H., Wray, N. R., Goddard, M. E., & Visscher, P. M. (2011). Estimating missing heritability for disease from genome-wide association studies. *American Journal of Human Genetics*, 88(3), 294–305. <https://doi.org/10.1016/j.ajhg.2011.02.002>
- Lee, S. H., Goddard, M. E., Wray, N. R., & Visscher, P. M. (2012). A better coefficient of determination for genetic profile analysis. *Genetic Epidemiology*, 36(3), 214–224. <https://doi.org/10.1002/gepi.21614>
- Martin, A. R., Kanai, M., Kamatani, Y., Okada, Y., Neale, B. M., & Daly, M. J. (2019). Clinical use of current polygenic risk scores may exacerbate health disparities. *Nature Genetics*, 51(4), 584–591. <https://doi.org/10.1038/s41588-019-0379-x>
- McKay, M. T., Kilmartin, L., Meagher, A., Cannon, M., Healy, C., & Clarke, M. C. (2022). A revised and extended systematic review and meta-analysis of the relationship between childhood adversity and adult psychiatric disorder. *Journal of Psychiatric Research*, 156, 268–283. <https://doi.org/10.1016/j.jpsychires.2022.10.015>
- Pearson-Fuhrhop, K. M., Dunn, E. C., Mortero, S., Devan, W. J., Falcone, G. J., Lee, P., ... Cramer, S. C. (2014). Dopamine genetic risk score predicts depressive symptoms in healthy adults and adults with depression. *PLoS One*, 9(5), e93772. (24834916). <https://doi.org/10.1371/journal.pone.0093772>
- Periyasamy, S., John, S., Padmavati, R., Rajendren, P., Thirunavukkarasu, P., Gratten, J., ... Mowry, B. J. (2019). Association of schizophrenia risk with disordered niacin metabolism in an Indian genome-wide association study. *JAMA Psychiatry*, 76(10), 1026–1034. <https://doi.org/10.1001/jamapsychiatry.2019.1335>
- Peterson, R. E., Kuchenbaecker, K., Walters, R. K., Chen, C.-Y., Popejoy, A. B., Periyasamy, S., ... Duncan, L. E. (2019). Genome-wide association studies in ancestrally diverse populations: Opportunities, methods, pitfalls, and recommendations. *Cell*, 179(3), 589–603. <https://doi.org/10.1016/j.cell.2019.08.051>
- Polderman, T. J. C., Benyamin, B., de Leeuw, C. A., Sullivan, P. F., van Bochoven, A., Visscher, P. M., & Posthuma, D. (2015). Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nature Genetics*, 47(7), 702–709. <https://doi.org/10.1038/ng.3285>
- Scott, K. M., Zhang, Y., Chardoul, S., Ghimire, D. J., Smoller, J. W., & Axinn, W. G. (2021). Resilience to mental disorders in a low-income, non-Westernized setting. *Psychological Medicine*, 51(16), 2825–2834. <https://doi.org/10.1017/S0033291720001464>
- Shen, H., Gelaye, B., Huang, H., Rondon, M. B., Sanchez, S., & Duncan, L. E. (2020). Polygenic prediction and GWAS of depression, PTSD, and suicidal ideation/self-harm in a Peruvian cohort. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, 45(10), 1595–1602. <https://doi.org/10.1038/s41386-020-0603-5>
- Shi, H., Gazal, S., Kanai, M., Koch, E. M., Schoech, A. P., Siewert, K. M., ... Price, A. L. (2021). Population-specific causal disease effect sizes in functionally important regions impacted by selection. *Nature Communications*, 12, 1098. <https://doi.org/10.1038/s41467-021-21286-1>
- Thornton, A., Ghimire, D. J., & Mitchell, C. (2012). The measurement and prevalence of an ideational model of family and economic development in Nepal. *Population Studies*, 66(3), 329–345.
- Truong, B., Hull, L. E., Ruan, Y., Huang, Q. Q., Hornsby, W., Martin, H., ... Natarajan, P. (2024). Integrative polygenic risk score improves the prediction accuracy of complex traits and diseases. *Cell Genomics*, 4(4), 100523. <https://doi.org/10.1016/j.xgen.2024.100523>
- Tubbs, J. D., & Sham, P. C. (2023). Preliminary evidence for genetic nurture in depression and neuroticism through polygenic scores. *JAMA Psychiatry*, 80(8), 832–841. <https://doi.org/10.1001/jamapsychiatry.2023.1544>

- Wang, Y., Tsuo, K., Kanai, M., Neale, B. M., & Martin, A. R. (2022). Challenges and opportunities for developing more generalizable polygenic risk scores. *Annual Review of Biomedical Data Science*, 5, 293–320. <https://doi.org/10.1146/annurev-biodatasci-111721-074830>
- Willer, C. J., Li, Y., & Abecasis, G. R. (2010). METAL: Fast and efficient meta-analysis of genomewide association scans. *Bioinformatics (Oxford, England)*, 26(17), 2190–2191. <https://doi.org/10.1093/bioinformatics/btq340>
- Wray, N. R., Lin, T., Austin, J., McGrath, J. J., Hickie, I. B., Murray, G. K., & Visscher, P. M. (2021). From basic science to clinical application of polygenic risk scores: A primer. *JAMA Psychiatry*, 78(1), 101–109. <https://doi.org/10.1001/jamapsychiatry.2020.3049>
- Xing, J., Watkins, W. S., Shlien, A., Walker, E., Huff, C. D., Witherspoon, D. J., ... Jorde, L. B. (2010). Toward a more uniform sampling of human genetic diversity: A survey of worldwide populations by high-density genotyping. *Genomics*, 96(4), 199–210. <https://doi.org/10.1016/j.ygeno.2010.07.004>
- Yang, J., Lee, S. H., Goddard, M. E., & Visscher, P. M. (2011). GCTA: A tool for genome-wide complex trait analysis. *American Journal of Human Genetics*, 88(1), 76–82. <https://doi.org/10.1016/j.ajhg.2010.11.011>
- Zaitlen, N., Kraft, P., Patterson, N., Pasaniuc, B., Bhatia, G., Pollack, S., & Price, A. L. (2013). Using extended genealogy to estimate components of heritability for 23 quantitative and dichotomous traits. *PLOS Genetics*, 9(5), e1003520. <https://doi.org/10.1371/journal.pgen.1003520>