Supplementary Materials for

Cross-trait assortative mating is widespread and inflates genetic correlation estimates*

Richard Border†, Georgios Athanasiadis, Alfonso Buil, Andrew J. Schork, Na Cai, Alexander I. Young, Thomas Werge, Jonathan Flint, Kenneth S. Kendler, Sriram Sankararaman, Andy W. Dahl, Noah A. Zaitlen

†Correspondence to: border.richard@gmail.com

This PDF file includes:

Materials and Methods
Supplementary Text
Figs. S1 to S20
Captions for Tables S1 to S7

Other Supplementary Materials for this manuscript include the following:

Tables S1 to S7

*This manuscript has been accepted for publication in *Science*. This version has not undergone final editing. Please refer to the complete version of record at http://www.sciencemag.org/.

Materials and Methods

Phenotyping

We examined a broad array of phenotypes across two large cohorts for which cross-mate correlation estimation was feasible. This included 341,997 unrelated European ancestry UK Biobank participants (*14*) and 373,283 spousal pairs drawn from Danish registry data (*30–33*).

With respect to the UKB phenotypes, we a priori selected a variety of previously-studied phenotypes discussed in previous influential publications (*6*, *7*), discarding measures redundant with available measures of higher quality. For example, we excluded college completion (`f.6138`) which factors into the higher-resolution measure of years of education, while retaining the latter. Similarly, we excluded total cholesterol (`f.30690`) while retaining the component phenotypes low-density lipoprotein (LDL) cholesterol (`f.23405`), high-density lipoprotein (HDL) cholesterol (`f.23405`), and triglycerides (`f.30870`). Blood biochemistry phenotypes were further adjusted for statin usage.

With respect to the psychiatric phenotypes, we examined all disorders that were available in the Danish registry data and for which LDSC genetic correlation estimates were available in a recent well-powered investigation of the multivariate genetic architecture of psychiatric disorders (*23*). The Danish Civil Registration System has been registering all people legally residing in Denmark since 1968, and it includes information about sex, date of birth, parental links, and life events (for example, migration or death). The System is linked via anonymized identification numbers to the Danish National Patient Register and the Danish Psychiatric Central Research Register that include all diagnostic information regarding general medical conditions and specific psychiatric conditions, respectively, including all inpatient and outpatient contacts. We estimated estimated cross-mate tetrachoric correlations within and across six psychiatric phenotypes considering both ICD-8 and ICD-10 definitions (see Table S7). The definitions are based on those used by the iPSYCH initiative (*34*).

Identification of mating pairs in the UK Biobank

We identified putative mate pairs in the UK Biobank using a procedure broadly similar to that of Howe and colleagues (*35*) with two distinctions: 1. we incorporated geographical information to impute cohabitation status, and 2. we relaxed the condition that both potential mates must report having lived in the same location for precisely the same number of years. We considered sex-discordant pairs of European ancestry participants from the same assessment centers (`f.54`) who reported living with a spouse (`f.699`), excluding pairs discordant on any of the following measures:

- latitude / longitude of home location rounded to the nearest kilometer (`f.20074`, `f.20075`)
- inverse distance between home and nearest road / major road (`f.24010`, `f.24012`)
- coastal proximity (`f.24508`)
- household size (`f.709`)

- number of vehicles (`f.728`)
- accommodation type (`f.670`)
- rental status (`f.680`)

To minimize the possibility of identifying cohabitating relatives, we required pairs to be discordant on the age of at least one parent (`f.1807`, `f.1846`, `f.2946`, `f.3526`) and removed third degree or closer relatives via estimated kinship coefficients. Finally, we discarded all participant groupings meeting these criteria including more than two individuals. This resulted in a total of 40,697 putative mate pairs (81,394 individuals).

Identification of mating pairs in the Danish population cohort

We obtained empirical estimates of spousal correlations for six psychiatric disorders in the Danish Population using the Danish Civil Registration System (*30*, *31*), the Danish National Patient Register (*32*), and the Danish Psychiatric Central Research Register (*33*). We first randomly selected 500,000 individuals born between 1981 and 2005 from the Danish Civil Registration System. The parents of these individuals served as our sample of mates (373,283 spousal pairs total).

Cross-mate phenotypic correlation estimation

We estimated cross-mate correlations using a structural equation modeling approach via the R package *lavaan* v0.6-8 (*36*). Specifically, for the UK Biobank phenotypes, we estimated the sex-constrained structural model corresponding to Eq. (2), with Y and Z representing a pair of phenotypes after regressing out the effects of sex, age, sex$\times$age and age^2 (with the exception of metabolic phenotypes, which were also adjusted for statin use). For the psychiatric phenotypes, which were dichotomous, we proceded similarly, using a sex-constrained liability-threshold model to estimate cross-mate tetrachoric correlations for all disorders simultaneously. All models were estimated via maximum likelihood using asymptotic standard errors and pairwise-complete observations.

Simulation framework

We present results derived from two forward-time simulation frameworks: a highly realistic but computationally intensive scheme (hereafter referred to as the **large-scale framework**) and a simplified, computationally efficient scheme (hereafter referred to as the **simplified framework**), both of which we describe in detail in the following sections. We justify employing the simplified framework by validating its output against that of the realistic framework, concluding that the simplified framework is sufficient for modeling the effects of xAM. Further, comparing these methods allows us to isolate the impacts (or lack thereof) of several phenomena of interest, including pleiotropy and local linkage disequilibrium (LD). We provide the code necessary to repeat or extend these analyses online (*28*). All analyses were conducted using R v4.0.2 or R v4.1.0 (*37*) unless otherwise stated.

Large-scale framework

With the exception of extending the mating scheme to the bivariate case, we proceeded as in (*10*). Briefly, we used the genotypes of unrelated European ancestry UK Biobank participants at one million phased, imputed HapMap3 SNPs with minor allele frequency (MAF) ≥ 0.01 as the founder population. Meoisis was simulated by dividing the genome into 10 kB blocks and deriving recombination probabilities from a linear interpolation of the 1000 Genomes Project Phase 3 recombination map (*38*), thus preserving the existing local LD structure.

For each simulation, m=1e4 loci were selected as causal variants for the two phenotypes Y, Z such that their standardized effects $\boldsymbol{\beta_y}, \boldsymbol{\beta_z}$ were jointly bivariate Gaussian with respective variances $m^{-1}h^2_{y;\text{pan}}, m^{-1}h^2_{z;\text{pan}}$ and correlation ρ_β. Phenotypes were then generated according to the additive model in Eq. (6). Individuals (represented by pairs of phenotype values) were mated according to the exchangeable cross-mate correlation regime described below.

In order to avoid any potential confounds due to within-sample relatedness, each coupling produced a single offspring, thereby halving the sample size with each successive generation and ensuring no pair of individuals share any indentical-by-descent segments with respect to the founder population. At each generation, we obtained heritability and genetic correlation estimates from HE regression as implemented in GCTA v1.93.2b (*18*), LDSC as implemented in LDSC v1.0.1 (*6*), and REML as implemented in BOLT-LMM v2.3.4 (*19*). For LDSC, we estimated LD scores within sample using input paramater `ld-wind-cm=1.0` and obtained GWAS summary statistics using plink2 v2.0 (*39*). To improve computational efficiency, we analyzed subsamples of n=4e5 individuals, with the exception of the fourth and fifth generation results presented in Fig. 2A, which were respectively limited to n=2e5, 1e5 by the diminishing sample size at each generation.

Simplified framework

Given the computational resources required by the large-scale framework in light of the large number of trait pairs we wanted to base simulations on (each run of the large-scale framework generates several terrabytes of output), we developed a simplified framework using entirely synthetic data and only including causal loci. We constructed founder genotypes by randomly drawing $2m$ haploid SNPs with m allele frequencies distributed uniformly on $[0.01, 0.99]$ independently for n individuals, with meiosis proceeding as described below.

Phenotype simulation with and without pleiotropy. We again generated phenotype pairs according to the additive model in Eq. (6), restricting our focus to the case of genetically orthogonal traits. Given a specified effect correlation (in this case $\rho_\beta = 0$), pleiotropy is irrelevant. We demonstrated this by simulating phenotypes under either complete pleiotropy (all causal variants have non-zero effects on both phenotypes) and zero pleiotropy (all causal variants have non-zero effects on only one of pair of phenotypes). We independently sampled i.i.d. Gaussian genetic effects $\boldsymbol{\beta_y}, \boldsymbol{\beta_z}$, using a single set of m causal variants to simulate pleiotropy at every locus and two independent sets of $m/2$ causal variants to simulate the absence of pleiotropy. Fig. S6 demonstrates that, as expected, the fraction of causal variants shared between phenotypes has no bearing on the impact of xAM. However, as the zero pleiotropy simulation procedure reduced sampling variance across simulations (in a finite sample, ρ_β has greater variance when causal variants overlap), we utilized this

approach for the projections of $\hat{\rho}_{\mathrm{xAM}}$ presented in the primary text. The number of causal variants had no apparent impact on quantities of interest (Fig. S3).

Manipulating local LD. Denoting the first and second copies of an individual's haploid genotypes at the jth diploid locus by $X_j^{[1]}, X_j^{[2]}$, respectively, we manipulated local LD during meiosis I by enforcing recombination events at each single line with probability $p_{\mathrm{recomb.}}$ and each double line with probability 0.5 below:

$$\begin{array}{c|c|c|c||c|c|c|c||c||c} X_1^{[1]} & X_2^{[1]} & \cdots & X_c^{[1]} & X_{c+1}^{[1]} & X_{c+2}^{[1]} & \cdots & X_{2c}^{[1]} & \cdots & X_m^{[1]} \\ X_1^{[2]} & X_2^{[2]} & \cdots & X_c^{[2]} & X_{c+1}^{[2]} & X_{c+2}^{[2]} & \cdots & X_{2c}^{[2]} & \cdots & X_m^{[2]} \end{array} \tag{1}$$

Here $c = m/20$ divides the genome into 20 independently inherited pseudo-chromosomes. Within each chromosome, recombination events occur between contiguous loci with probability $p_{\mathrm{recomb.}} \in (0.0, 0.5]$, such that $p_{\mathrm{recomb.}} = 0.5$ corresponds to unlinked loci, with the strength of local LD increasing as $p_{\mathrm{recomb.}} \to 0$. Fig. S5 demonstrates the irrelevance of $p_{\mathrm{recomb.}}$ with respect to score correlation and estimated effect correlation; simulations with strong local LD ($p_{\mathrm{recomb.}} = 0.01$) and weak local LD ($p_{\mathrm{recomb.}} = 0.5$) yielded results consistent with the realistic framework simulation results.

Mating regimes

For two phenotypes Y, Z, denote their joint distribution across mates by

$$\begin{pmatrix} Y^* \\ Z^* \\ Y^{**} \\ Z^{**} \end{pmatrix} \sim \mathcal{MVN}\left(\mathbf{0}, \begin{pmatrix} 1 & & & \\ s_{yz} & 1 & & \\ r_{yy} & r_{yz} & 1 & \\ r_{yz} & r_{zz} & s_{yz} & 1 \end{pmatrix}\right), \tag{2}$$

where the number of asterisks distinguishes the two mates. Whereas the within-mate single-trait correlation (the conventional phenotype correlation) s_{yz} is unaffected by mating patterns within a given generation, the cross-mate single-trait correlations r_{yy}, r_{zz} and the cross-mate cross-trait correlation r_{yz} are free parameters determined by the mating regime.

For the large-scale simulations presented in Fig. 2, individuals were mated by randomly splitting the sample in half and pairing individuals across the two subsamples, ordering each subsample on a linear combination of their phenotypic values and Gaussian noise. This corresponds to an exchangeable cross-mate correlation structure with all cross-mate correlations r_{yy}, r_{yz}, r_{zz} equal to a single input parameter r_{mate}. Though this method is computationally efficient and thus well-suited to this particular use-case, it is incapable of achieving arbitrary cross-mate correlation structures. Thus, a more flexible approach was needed for the estimation of $\hat{\rho}_{\mathrm{xAM}}$, which employs empirical estimates of the cross-mate correlation parameters.

We achieved arbitrary cross-mate correlation structures, a problem for which we were unable to find a previously published solution, by using an ad hoc approach based on propensity-score matching methods as implemented in the R package MatchIt v4.2.0 (*40*). Specifically, we randomly split the sample in half and used the nearest Mahalanobis distance matching algorithm on a scalar

multiple of each of the two phenotypes, their product, and their difference. The multiplier hyperparameter was chosen numerically by finding the rational function of the correlation parameters in Eq. (2) that minimized the average ℓ_∞ distance between the desired and realized cross-mate correlations across replicate samples. Discrepancies between target correlations and achieved correlations were modest: across all generations of UKB simulations, the average median discrepancy was 0.0007 and 90% of all cross-mate correlations were within 0.02 of their target values. We further validated this approach by directly comparing it to the linear combination ordering approach described above, specifying a single value for each of the parameters $r_{yy} = r_{zz} = r_{yz} = r_{\text{mate}}$. Both methods yielded equivalent results (Fig. S4).

Annotation- and locus-level simulation studies

Partitioned genetic correlation estimates under xAM

We sought to understand how xAM affects genetic correlation estimates across genomic regions with distinct bivariate architectures. To this end, we simulated xAM across two phenotypes, Y and Z, with loci divided across six distinct, equally-sized annotations (Fig. S12): (1) enriched for large, overlapping, correlated effects for both traits ($\rho_\beta = 0.25$, $h^2_{y;\text{pan}} = h^2_{z;\text{pan}} = .122$); (2) enriched for large, overlapping, uncorrelated effects for both traits ($\rho_\beta = 0.0$, $h^2_{y;\text{pan}} = h^2_{z;\text{pan}} = .122$); (3) enriched for large effects for Y only ($\rho_\beta = 0.0$, $h^2_{y;\text{pan}} = .122$, $h^2_{z;\text{pan}} = .002$); (4) enriched for large effects for Z only ($\rho_\beta = 0.0$, $h^2_{y;\text{pan}} = .002$, $h^2_{z;\text{pan}} = .122$); (5) enriched for non-overlapping large effects for Z only (equal parts SNPs as described by the previous two annotations); and (6) not enriched for large effects on either trait ($\rho_\beta = 0.0$, $h^2_{y;\text{pan}} = h^2_{z;\text{pan}} = .002$). We then compared partitioned HE regression estimates:

$$\begin{pmatrix} \text{tr } K_1K_1 & \cdots & \text{tr } K_1K_p & \text{tr } K_1 \\ \vdots & \ddots & \vdots & \vdots \\ \text{tr } K_pK_1 & \cdots & \text{tr } K_pK_p & \text{tr } K_p \\ \text{tr } K_1 & \cdots & \text{tr } K_p & n \end{pmatrix} \begin{pmatrix} \hat{\sigma}^2_{g,1;y} & \hat{\sigma}^2_{g,1;z} & \widehat{\text{cov}}_{g,1} \\ \vdots & \vdots & \vdots \\ \hat{\sigma}^2_{g,p;y} & \hat{\sigma}^2_{g,p;z} & \widehat{\text{cov}}_{g,p} \\ \hat{\sigma}^2_{e;y} & \hat{\sigma}^2_{e;z} & \widehat{\text{cov}}_{e} \end{pmatrix} = \begin{pmatrix} y^\intercal K_1 y & z^\intercal K_1 z & z^\intercal K_1 y \\ \vdots & \vdots & \vdots \\ y^\intercal K_p y & z^\intercal K_p z & z^\intercal K_p y \\ y^\intercal y & z^\intercal z & z^\intercal y \end{pmatrix}$$

after up to five generations of xAM, where K_i denotes the genomic relatedness matrix for annotation $i \in \{1, \ldots, p\}$ (*41*, *42*).

GWAS effect estimates under xAM

We next examined whether the distorting effects of xAM on estimates at the genome-wide or annotation-specific levels also effected locus-specific analyses. To this end, we simulated equally-heritable phenotypes Y and Z subject to xAM as described to above, this time with fully independent genetic architectures: half of variants were causal for Y but not Z and half were causal for Z but not Y. We then performed standard GWAS on phenotype Y after up to five generations of xAM and examined effect estimate bias, power, and false positive rates (Figs. S7 to S10).

Estimation of GWAS summary statistics in the UK Biobank

From among 488,363 UK Biobank participants, we retained putative "White British" individuals using field `f.22006` (n=409,692). We then filtered out 199 individuals with excess genotype

missingness (>0.05), 312 individuals with a mismatch between self-reported and genetic sex, 999 inviduals with excess heterozygosity (≥5 standard deviations above the mean), and 90 individuals who requested their data be redacted. We then removed 629 individuals related to ten or more individuals (KING coefficient $\geq 2^{-9/2}$) as a preprocessing step to the application of the `maximal_independent_set` algorithm implemented in the NetworkX Python package (*43*). This resulted in 342,257 unrelated individuals. In contrast to (*14*), who estimated kinships using ≈92,000 common SNPs with small loadings onto the first few PCs in the full sample (including multiple ancestries; see S3.7 of (*14*)), we estimated kinships using 561,780 common SNPS in a sample of European ancestry individuals. The close relatives the UKB identified in `f.22021` are a subset of our more conservative approach: we identified all 81,218 related individuals in this subsample identified by the UKB plus an additional 3,261 not identitified by (*14*).

For each phenotype of interest, we estimated association regression weights at 1,157,133 imputed HapMap3 SNPs with missingness < 0.01, Hardy-Weinberg equilibrium $p >$ 1e-5, INFO imputation quality score > 0.9, and MAF > 0.01 using plink v2.0 (*39*). Covariates included sex, age, sex×age, age^2, 21 genomic principal components, assessment center, and genotyping batch (though we note that metabolic phenotypes were further adjusted for statin use as detailed above). Analyses were restricted to unrelated European ancestry individuals.

Empirical heritability and genetic correlation estimates

UK Biobank cohort

We estimated marker-based heritabilities and genetic correlations using LDSC v1.0.1 (*4*) with internal summary statistics (see previous section) and LD scores (computed for regression SNPs with input parameter `ld-wind-cm=1.0`). We obtained pedigree-based heritability estimates from the empirical literature, using estimates derived in UK-based samples whenever available. In the case of classical twin studies, we used results from "ACE" models, which are expected to yield downward-biased heritability estimates under assortative mating (*44*) and thus constitute the more conservative option, when available (Table S2).

Literature-derived estimates for psychiatric disorders

We used a subset of the LDSC genetic correlation estimates reported by Grotzinger and colleagues (*23*) corresponding to the set of psychiatric disorders for which we estimated cross-mate cross-trait correlations in the Danish population cohort. As with the UK Biobank phenotypes, we extracted pedigree-based heritability estimates from demographically comparable samples preferring "ACE" estimates when available (Table S4).

Projected genetic correlation estimates under xAM alone

We implemented a simulation based approach to quantifying the magnitudes of empirical $\hat{\rho}_\beta$ estimates with respect to expectations for genetically orthogonal traits under xAM alone. For each pair of traits Y, Z we ran 400 replicate simulations with a founder population size of $n =$ 64,000, setting the panmictic heritability parameters to empirical estimates and allocating $m = 1000$ distinct causal variants for each phenotype, thus ensuring $\rho_\beta = 0$ (though we note that neither the number

of causal variants nor the extent to which causal variants overlap impacts simulation results; see Figs. S3 and S6, respectively). We then simulated multiple generations of xAM using empirical estimates of the cross-mate correlations $\hat{r}_{yy}$, $\hat{r}_{zz}$, $\hat{r}_{yz}$, at each generation obtaining method-of-moment estimates of ρ_β, which we denote $\hat{\rho}_{\mathrm{xAM}}$, using HE regression, which, as demonstrated in Fig. 2, produces results equivalent to LDSC regression in this setting.

In order to propagate uncertainty in the empirical estimates through the simulations, simulation inputs were randomly sampled from the asymptotic sampling distributions of the corresponding empirical estimates. We first randomly sampled the empirical genetic correlation estimate $\hat{\rho}_{\mathrm{emp}}$ and the input parameters $h^2_{y;\mathrm{pan}}$ and $h^2_{z;\mathrm{pan}}$ from their sampling distributions and r_{yy}, r_{zz}, r_{yz} from their joint sampling distribution:

$$\left(\hat{\rho}_{\beta;\mathrm{emp}}, h^2_{y;0}, h^2_{z;0}, r_{yy}, r_{zz}, r_{yz}\right) \sim \mathcal{MVN}\left(\left(\hat{\rho}_\beta, \hat{h}^2_y, \hat{h}^2_z, \hat{r}_{yy}, \hat{r}_{zz}, \hat{r}_{yz}\right), \widehat{\Omega}\right), \tag{3}$$

where $\widehat{\Omega}$ denotes the empirical variance-covariance matrix. We then compared $\hat{\rho}_{\mathrm{emp}}$ estimates to those projected by the simulation, which we denote $\hat{\rho}_{\mathrm{xAM}}$, to compute the ratio statistic $\hat{\gamma} = \hat{\rho}_{\mathrm{xAM}}/\hat{\rho}_{\mathrm{emp}}$. The empirical quantiles of $\hat{\gamma}$, which approximate its posterior distribution, were then used to compute credible intervals. For primary analyses, which used LDSC genetic correlation estimates and pedigree-based heritability estimates derived from independent samples, only the off-diagonal elements of $\widehat{\Omega}$ corresponding to the cross-mate correlation parameters, which we estimated via structural equation modeling, were nonzero. For the supplementary analyses of the UKB phenotypes using LDSC heritability estimates presented in Fig. S14 and Figs. S15A and S15B, we proceeded analogously as the sample variance-covariances of $\hat{h}^2_y$, $\hat{h}^2_z$, $\hat{\rho}_\beta$ are not reported by the LDSC software used for heritability estimation (though we note that the Genomic SEM implementation of LDSC does provide these quantities (*24*)). We later present sensitivity analyses demonstrating that treating $\hat{h}^2$ and $\hat{\rho}_\beta$ estimates as independent had a limited impact on our results. Additionally, we compared the above definition of $\hat{\gamma}$ to an alternative definition based on genetic covariances, $\hat{\gamma}_{\mathrm{cov}} = \widehat{\mathrm{COV}}_{g;\mathrm{xAM}}/\widehat{\mathrm{COV}}_{g;\mathrm{emp}}$, which yielded similar results (Fig. S19).

Impact of xAM and misclassification errors

We ran an additional set of simulations again using the simplified forward-time simulation framework to characterize the impact of xAM on binary traits subject to misclassification. We simulated two continuous phenotypes with no pleiotropy, this time fixing all the cross-mate correlation parameters to 0.5 while varying trait prevalence as well as the rate and directionality of misclassification errors. Prior to effect correlation estimation, we dichotomized phenotypes using the cutoffs derived from the standard normal quantile function. We introduced misclassification errors by randomly selecting a fixed number of true cases for the first phenotype and relabeling them as controls for the first and cases for the second and preceding analogously for the second phenotype (Fig. S11).

Projected genetic correlation estimates under xAM and misdiagnostic errors

We next investigated the combined impact of xAM and diagnostic errors on genetic correlations among psychiatric disorders. Along with the previously mentioned parameters, we additionally

sampled diagnostic thresholds for each disorder from their joint sampling distribution with the three cross-mate cross-correlation parameters as estimated in our cross-mate tetrachoric correlation structural models. These thresholds were then used to dichotomize latent continuous phenotypes prior to genetic correlation estimation at each generation. Given that the low prevalences for some disorders lead to excessively noisy genetic correlation estimates in the context of a non-ascertained sample (balancing case-control ratios while retaining statistical efficiency would necessitate running much larger, more computationally expensive simulations) we proceeded as follows:

- For each of 250 Monte Carlo samples of the input parameters:

    For each of 50 replicate simulations, proceed through the simplified framework as previously described, except:

        At each generation, dichotomize continuous phenotypes prior to genetic correlation estimation

        For each misdiagnosis rate (0%, 5%, 10%, 15%), for each of 5 replicates:

            Achieve prescribed misdiagnosis rate by randomly permuting cases for the two disorders

            Estimate genetic correlation (that is, compute $\hat{\rho}_{\mathrm{xAM}}$)

        Record the median estimates across the 5 replicates for each misdiagnosis rate

- Record the median estimates across the 50 replicates for each misdiagnosis rate

- Combine with 500 random draws from the sampling distribution of $\hat{\rho}_{\mathrm{emp}}$ to compute 250 $\times$ 500 estimates of $\hat{\gamma}$

- Derive point estimates and credible intervals using the empirical quantiles of $\hat{\gamma}$

This procedure was designed to propagate all uncertainty in the input parameters and empirical genetic correlation estimates while reducing noise related to the simulation procedure. Outside of analyses specifically examining misdiagnostic error, we report $\hat{\gamma}$ values for psychiatric disorders based on simulations of continuous phenotypes. These estimates were virtually identical to those of the corresponding binary phenotypes without diagnostic error (Pearson r=0.999; Welch t-test of difference in means p=.981), but had smaller standard errors (mean difference of 0.078).

Sensitivity analysis of the effects of parameter dependence on $\hat{\gamma}$ estimates

The estimates of $\hat{\gamma}$ presented in the primary manuscript are based on empirical heritability (respectively [resp] genetic correlation) estimates derived from previously published pedigree-based (resp. marker-based) methods. Since they are estimated in independent samples, the sampling distributions of the estimates are necessarily independent. However, when estimated jointly, heritability estimates and genetic correlation estimates tend to be positively correlated. To approximate the extent to which positive dependence among the input parameters might alter our $\hat{\gamma}$ estimates,

we proceeded as follows, using the sampling covariance matrices of the LDSC genetic variance and covariance estimates for the psychiatric disorders reported by Grotzinger and colleagues (*23*), together with pedigree-based heritability estimates.

Across all of the psychiatric disorder pairs studied in the current manuscript, parameter estimates for ANX and MDD had the strongest dependence as measured by the largest single off-diagonal element of the sampling correlation matrix, which we reproduce below:

	$\widehat{\mathrm{var}}_{g;\mathrm{ANX}}$	$\widehat{\mathrm{cov}}_{g}$	$\widehat{\mathrm{var}}_{g;\mathrm{MDD}}$
$\widehat{\mathrm{var}}_{g;\mathrm{ANX}}$	1.00	0.47	0.29
$\widehat{\mathrm{cov}}_{g}$	0.47	1.00	0.79
$\widehat{\mathrm{var}}_{g;\mathrm{MDD}}$	0.29	0.79	1.00

. (4)

For the sake of comparison, the smallest off-diagonal element across all trait pairs was (-0.086).

We then used the multivariate delta method to obtain the corresponding sample covariance matrix of the heritability and genetic correlation estimates. We used this covariance matrix to sample panmictic heritabilities and genetic correlation estimates from the multivariate Gaussian distribution centered on the pedigree-based heritability estimates for ANX and MDD and the LDSC estimate of their genetic correlation. Explicitly, denoting ANX and MDD by y, z, respectively, denoting the covariance matrix corresponding to correlation matrix above by $\mathbf{\Sigma}$, and defining $\boldsymbol{f} : (a, b, c) \mapsto (a, b, c/\sqrt{ab})$, $u = (\widehat{\mathrm{var}}_{g;y;\mathrm{LDSC}}, \widehat{\mathrm{var}}_{g;z;\mathrm{LDSC}}, \widehat{\mathrm{cov}}_{g;\mathrm{LDSC}})$, and $\mu = (\hat{h}^2_{y;\mathrm{ped}}, \hat{h}^2_{z;\mathrm{ped}}, \hat{\rho}_{\mathrm{LDSC}})$, we randomly sampled panmictic heritabilities and genetic correlation estimates from the distribution

$$\begin{pmatrix} h^2_{y;\mathrm{pan}} \\ h^2_{z;\mathrm{pan}} \\ \rho_{\mathrm{emp}} \end{pmatrix} \sim \mathcal{MVN}\left(\mu, \nabla \boldsymbol{f}(u)^{\intercal} \mathbf{\Sigma} \nabla \boldsymbol{f}(u)\right).$$

We then used the sampled panmictic heritability estimates as inputs to our forward-time simulator, which we used to obtain $\hat{\rho}_{\beta}$ estimates, which we in turn compared to the jointly sampled LDSC genetic correlation estimate to obtain $\hat{\gamma}$:

$$\left.\begin{matrix} h^2_{y;\mathrm{pan.}} \\ h^2_{z;\mathrm{pan.}} \end{matrix}\right\} \xrightarrow{\text{simulator}} \rho_{\mathrm{xAM}}, \qquad \hat{\gamma} = \frac{\rho_{\mathrm{xAM}}}{\rho_{\mathrm{emp}}}.$$

This procedure yields a reasonable, if artificial, approximation of the potential qualitative changes to our results caused by strong dependence between the empirical heritability estimates used to seed our simulations and the empirical genetic correlation estimates to which we compare our results. Figure S20 compares the results of this procedure to those of the procedure implemented in the primary manuscript, where panmictic heritabilities and genetic correlations estimates are drawn independently. The dependent sampling procedure yielded larger, less variable $\hat{\gamma}$ estimates.

Cross-chromosome correlation of polygenic scores in the UK Biobank

We first randomly split the UK Biobank into two disjoint subsamples, with 80% of participants comprising the training set and the remaining 20% comprising the test data. Next, we ran a GWAS for each phenotype in the training sample using covariates and procedures identical to those described previously and clumped results using plink v1.90b, employing a 250kB sliding window

and an R^2 threshold of 0.05. We then used the training sample summary statistics to compute polygenic scores for separately for even and odd chromosomes using up to 20 p-value thresholds spaced logarithmically on the interval [5e-8, 1.0], with the resulting number of scores depending on the maximally significant SNP for a given trait (Fig. S18).

For any given pair of phenotypes Y, Z, there are two potential even/odd chromosome polygenic score correlations, $\mathrm{cor}(\hat{\ell}_{y;\mathrm{even}}, \hat{\ell}_{z;\mathrm{odd}})$ and $\mathrm{cor}(\hat{\ell}_{y;\mathrm{odd}}, \hat{\ell}_{z;\mathrm{even}})$, neither of which is of greater interest than the other. Thus, we estimated both quantities via a single parameter, $\rho_{\ell;\mathrm{eo}}$, in the constrained structural model

$$\begin{pmatrix} \hat{\ell}_{y;\mathrm{even}} \\ \hat{\ell}_{y;\mathrm{odd}} \\ \hat{\ell}_{z;\mathrm{even}} \\ \hat{\ell}_{z;\mathrm{odd}} \end{pmatrix} \sim \mathcal{MVN} \left(\mathbf{0}, \begin{pmatrix} 1 & & & \\ \phi_{yy} & 1 & & \\ \psi_{yz} & \rho_{\ell;\mathrm{eo}} & 1 & \\ \rho_{\ell;\mathrm{eo}} & \psi_{zy} & \phi_{zz} & 1 \end{pmatrix} \right), \tag{5}$$

where $\rho_{\ell;\mathrm{eo}}$ is the only free parameter of interest. For the simpler case of single-trait even/odd correlations, we used the correspondingly simpler unconstrained bivariate Gaussian structural model. Again, we estimated structural models using the R package *lavaan* v0.6-8 (*36*) via maximum likelihood with asymptotic standard errors.

Finally, we evaluated the relationships between even/odd chromosome polygenic score correlation estimates and both cross-mate cross-trait correlation estimates and genetic correlation estimates across all pairs of UKB sample phenotypes. We examined these associations in the context of naïve linear models, which don't account for heteroskedasticity and sampling error in the predictor, but also using a Bayesian measurement error model as implemented in the R package *brms* v1.8 (*45*), which does.

Supplementary Text

Here we present theoretical explanations for the impact of xAM on genetic architecture and statistical estimators. We consider a pair of phenotypes Y, Z composed of the additive effects of m standardized haploid variants $X_1, \dots X_m$ with phenotype-specific effects $\boldsymbol{\beta}_y$, $\boldsymbol{\beta}_z$:

$$Y = \underbrace{\sum_{i=1}^{m} X_i \beta_{y;i}}_{:=\ell_y} + \varepsilon_y, \qquad Z = \underbrace{\sum_{i=1}^{m} X_i \beta_{z;i}}_{:=\ell_z} + \varepsilon_z. \tag{6}$$

Each phenotype is composed of a heritable liability component ℓ, the true polygenic score, and a non-heritable component ε. For convenience, we assume that causal variants are initially unlinked and both Y and Z have unit variance under random mating (panmixis), such that the panmictic heritabilities are $h^2_{y;\text{pan}} = \boldsymbol{\beta}_y^\intercal \boldsymbol{\beta}_y$ and $h^2_{z;\text{pan}} = \boldsymbol{\beta}_z^\intercal \boldsymbol{\beta}_z$.

Long range sign-consistent LD

Denote quantities relating to the members of a parent-parent-offspring trio by $[\cdot]^*$, $[\cdot]^{**}$, $\tilde{[\cdot]}$, respectively. Let $\mathbb{E}_t[\cdot]$ denote the expectation of a quantity after t generations of positive xAM. Let X_i, X_j denote mean-deviated trait-increasing-allele counts at two causal loci for mean-deviated phenotypes Y, Z, respectively. We assume that X_i, X_j are unlinked at panmixis (when $t = 0$). For simplicity, we assume each variant is causal for one and only one of the two phenotypes, that is,

$$\beta_{y,i} \neq 0 = \beta_{y,j}, \qquad \text{and } \beta_{z,i} = 0 \neq \beta_{z,j},$$

and we assume the cross-mate covariances are positive and symmetric such that for all $t \geq 0$

$$\mathbb{E}_t[Y^* Y^{**}] \geq 0, \qquad \mathbb{E}_t[Z^* Z^{**}] \geq 0,$$
$$\mathbb{E}_t[Y^* Z^{**}] = \mathbb{E}_t[Y^{**} Z^*] > 0.$$

We assume **primary phenotypic assortment**: mates' genotypes are conditionally independent given the heritable component of either mate's phenotype. Denoting the heritable components of Y, Z by ℓ_y, ℓ_z, respectively, we assume

$$p[X_i^* X_j^{**} | \ell_y^*, \ell_z^*] = p[X_i^* | \ell_y^*, \ell_z^*] p[X_j^{**} | \ell_y^*, \ell_z^*].$$

Denote the conditional expectation of an individual's causal variant genotype at locus i given the heritable components of their phenotypes by

$$\phi_{y,i,t}(y) := \mathbb{E}_t[X_i^* | \ell_y^* = y] = \mathbb{E}_t[X_i^{**} | \ell_y^{**} = y],$$

and denote the conditional expectation of an individual's causal variant genotype given their mate's heritable phenotype components by

$$\psi_{y,i,t}(z) := \mathbb{E}_t[X_i^* | \ell_z^{**} = z] = \mathbb{E}_t[X_i^{**} | \ell_z^* = z].$$

We assume that $\phi_{i,t}(y), \psi_{i,t}(z)$ are monotone functions in their respective arguments that cross the origin and that agree in sign with the effect of locus i such that

$$\phi_{y,i,t}(y) = \beta_{y,i} \cdot \tilde{\phi}_{y,t}(y),$$
$$\psi_{y,i,t}(z) = \beta_{y,i} \cdot \tilde{\psi}_{y,t}(z),$$

where $\tilde{\phi}, \tilde{\psi}$ are non-trivial monotone increasing functions that cross the origin. That is, we assume that for a *trait-increasing* allele for Y (that is, when $\beta_{y,i} > 0$) $\phi_{y,i,t}(y)$ is an *increasing* function, and further, as Y and Z are positively correlated across mates, $\psi_{y,i,t}(z)$ will also be increasing. Likewise, if X_i is *trait-decreasing*, $\phi_{y,i,t}(y)$ and $\psi_{y,i,t}(z)$ will be *decreasing* in their respective arguments. Though rather technical in specification, these assumptions are intuitive: if I know one mate is "high" with respect to (the heritable component of) phenotype Y, I expect "higher" genotypic values at their own loci and their partner's loci that increase Y and "lower" genotypic values at their own and their partner's loci that decrease Y. Further, this is a weaker assumption than those of previous authors (for instance, in the single-trait case (*46*) assumes linearity such that $\phi_{y,i,t}(y) \propto \beta_{y,i} \cdot y$).

Having introduced our assumptions, we seek to investigate the linkage disequilibrium between X_i, an arbitrary causal variant for phenotype Y and X_j, an arbitrary causal variant for phenotype Z, after $t > 0$ generations of positive xAM across Y and Z. At every generation, we can factor the covariance $\mathbb{E}[\tilde{X}_i\tilde{X}_j]$ using the possible patterns of inheritance:

$$\begin{aligned}\mathbb{E}[\tilde{X}_i\tilde{X}_j] &= \frac{1}{4}\mathbb{E}[X_i^*X_j^*] + \frac{1}{4}\mathbb{E}[X_i^*X_j^{**}] + \frac{1}{4}\mathbb{E}[X_i^{**}X_j^*] + \frac{1}{4}\mathbb{E}[X_i^{**}X_j^{**}] \\ &= \frac{1}{2}\mathbb{E}[X_i^*X_j^*] + \frac{1}{2}\mathbb{E}[X_i^*X_j^{**}].\end{aligned}$$

As by definition

$$\mathbb{E}_{t-1}[\tilde{X}_i\tilde{X}_j] = \mathbb{E}_t[X_i^*X_j^*],$$

this induces the following recurrence relation:

$$\begin{aligned}\mathbb{E}_t[X_iX_j] &= \frac{1}{2}\mathbb{E}_{t-1}[X_i^*X_j^*] + \frac{1}{2}\mathbb{E}_{t-1}[X_i^*X_j^{**}] \\ &= 2^{-t}\mathbb{E}_0[X_iX_j] + \sum_{l=1}^{t-1} 2^{t-l}\mathbb{E}_l[X_i^*X_j^{**}].\end{aligned}$$

By hypothesis, X_i, X_j are unlinked at panmixis and thus $\mathbb{E}_0[X_iX_j]$ is zero. That leaves us with the sequence of cross-mate cross-locus moments $\{\mathbb{E}_l[X_i^*X_j^{**}]\}_{l=0}^{t-1}$ terms.

Applying the assumption of primary phenotypic assortment, we have

$$\begin{aligned}\mathbb{E}_t[X_i^*X_j^{**}] &= \int \mathbb{E}_t[X_i^*|\ell_y^*]\mathbb{E}_t[X_j^{**}|\ell_z^*]dP(\ell_y^*, \ell_z^*) \\ &= \beta_{y,i}\beta_{z,j}\int \tilde{\phi}_{y,0}(\ell_y^*)\tilde{\psi}_{y,0}(\ell_z^*)dP(\ell_y^*, \ell_z^*).\end{aligned}$$

The above integral is strictly positive, thereby yielding

$$\operatorname{sgn}\left[\mathbb{E}_t[X_i^*X_j^{**}]\right] = \operatorname{sgn}\left[\beta_{y,i}\right] \cdot \operatorname{sgn}\left[\beta_{z,j}\right] \neq 0.$$

All together, we have

$$\mathbb{E}_t[X_i X_j] = \sum_{l=1}^{t-1} 2^{t-l} \mathbb{E}_l[X_i^* X_j^{**}],$$

where each of the terms in the above sum has sign $\operatorname{sgn}[\beta_{y,i}] \cdot \operatorname{sgn}[\beta_{z,j}]$, establishing that

$$\operatorname{sgn}\ [\mathbb{E}_t[X_i X_j]] = \operatorname{sgn}[\beta_{y,i}] \cdot \operatorname{sgn}[\beta_{z,j}].$$

Inflation of GWAS statistics

We now show that assortment on a single trait (sAM) leads to moderately inflated GWAS effect estimates. This is a simplification of the bivariate case.

By the same argument used in the previous section, it is easy to see that sAM on phenotype Y induces sign-consistent long-range LD such that

$$\operatorname{sgn}[\mathbb{E}[X_i X_j]] = \operatorname{sgn}[\beta_{y;i}]\operatorname{sgn}[\beta_{y;j}]$$

for all causal variants indexed i, j. Without loss of generality, assume all variants $i, j \in \{1, \ldots, m\}$ are causal and denote the correlation between standardized variants at loci index i, j by

$$\omega_{ij} := \mathbb{E}[X_i X_j].$$

We estimate the GWAS effect of variant X_i on Y as

$$\begin{aligned}
\mathbb{E}[\hat{\beta}_i] &= n^{-1}\mathbb{E}[\beta^\intercal \boldsymbol{X}^\intercal \boldsymbol{X}_{\cdot i}] \\
&= \sum_{j=1}^{m} \omega_{ij}\beta_j \\
&= \beta_i + \sum_{j \neq i} \omega_{ij}\beta_j.
\end{aligned}$$

Each of the above summands has sign

$$\operatorname{sgn}\ \omega_{ij}\beta_j = \operatorname{sgn}\ \beta_i \cdot \operatorname{sgn}\ \beta_j \cdot \operatorname{sgn}\ \beta_j = \operatorname{sgn}\ \beta_i.$$

Thus, when $\beta_i > 0$ we have $\mathbb{E}[\hat{\beta}_i] > \beta_i$ and when $\beta_i < 0$ we have $\mathbb{E}[\hat{\beta}_i] < \beta_i$. That is, the magnitude of the GWAS effect estimate is inflated upwards at causal variants.

Correlation of even/odd chromosome cross-trait polygenic scores

We now consider the correlation between polygenic scores for Y and Z constructed on disjoint sets of loci (as in even/odd chromosome polygenic score correlations). Partition the indices of the genome into two non-empty sets $\mathcal{E}$, $\mathcal{O}$, such that $\{1, \ldots, m\}$ is their disjoint union. Define the estimated polygenic scores

$$\hat{\ell}_{y;\mathcal{E}} := \sum_{i \in \mathcal{E}} X_i \hat{\beta}_{y;i}, \qquad \hat{\ell}_{z;\mathcal{O}} := \sum_{j \in \mathcal{O}} X_j \hat{\beta}_{z;j}.$$

These correspond to a polygenic score for Y built only on even chromsomes and a polygenic score for Z built only on odd chromsomes. The expected cross-trait even/odd correlation polygenic score covariance is then equal to

$$\begin{aligned}\mathbb{E}[\hat{\rho}_{\ell;\text{eo}}] &= \sum_{i\in\mathcal{E}}\sum_{j\in\mathcal{O}}\omega_{ij}\mathbb{E}[\hat{\beta}_{y;i}\hat{\beta}_{z;j}] \\ &= \sum_{i\in\mathcal{E}}\sum_{j\in\mathcal{O}}\sum_{k,l}\omega_{ij}\omega_{ik}\omega_{jl}\beta_{y;k}\beta_{z;l.}\end{aligned}$$

Compare this to the true cross-trait even/odd chromsome polygenic score correlation

$$\rho_{\ell;\text{eo}} = \sum_{i\in\mathcal{E}}\sum_{j\in\mathcal{O}}\omega_{ij}\beta_{y;i}\beta_{z;j}.$$

Pulling this out of the previous expression yields

$$\mathbb{E}[\hat{\rho}_{\ell;\text{eo}}] = \rho_{\ell;\text{eo}} + \sum_{i\in\mathcal{E}}\sum_{j\in\mathcal{O}}\sum_{k\neq i}\sum_{l\neq j}\omega_{ij}\omega_{ik}\omega_{jl}\beta_{y;k}\beta_{z;l.}$$

Dealing with the above summand requires understanding the relationship between $\beta_{y;i}$, $\beta_{z;j}$ and ω_{ij} for all i, j, which is beyond the scope of the current manuscript. For now, we address the simpler, but analogous, case of the correlation between polygenic scores for a single phenotype Y restricted to disjoint sets of chromosomes as analyzed in (*11*). Again, we assume that $i \in \mathcal{E}\cup\mathcal{O} \implies \beta_{y;i} \neq 0$ for simplicity. In this case, we have

$$\mathbb{E}[\hat{\rho}_{\ell;\text{eo}}] = \rho_{\ell;\text{eo}} + \sum_{i\in\mathcal{E}}\sum_{j\in\mathcal{O}}\sum_{k\neq i}\sum_{l\neq j}\omega_{ij}\omega_{ik}\omega_{jl}\beta_{y;k}\beta_{y;l}, \tag{7}$$

where each of the above summands has sign

$$\text{sgn}\ \omega_{ij}\omega_{ik}\omega_{jl}\beta_{y;k}\beta_{z;l} = (\text{sgn}\ \beta_{y;i})^2(\text{sgn}\ \beta_{y;j})^2(\text{sgn}\ \beta_{y;k})^2(\text{sgn}\ \beta_{y;l})^2 = +1.$$

Likewise, the summands comprising the true correlation

$$\rho_{\ell;\text{eo}} = \sum_{i\in\mathcal{E}}\sum_{j\in\mathcal{O}}\omega_{ij}\beta_{y;i}\beta_{y;j}$$

also all have sign $+1$, altogether yielding the chain of strict inequalities

$$\mathbb{E}[\hat{\rho}_{\ell;\text{eo}}] > \rho_{\ell;\text{eo}} > 0.$$

That is, as anticipated by Yengo et al. (*11*), single-trait AM induces a true correlation between genetic liabilities restricted to disjoint collections of loci. Note that we do not require the assumption of equilibrium.

Supplementary Figures

Figure S1: Cross-mate cross-trait correlations versus cross-mate single-trait correlations. Cross-mate cross-trait phenotypic correlation estimates ($\hat{r}_{yz}$) plotted against the average of the two corresponding cross-mate single-trait correlation estimates ($\hat{r}_{yy}$, $\hat{r}_{zz}$). The red dashed line is the unit slope line. Whereas $\hat{r}_{yz}$ is close to $\hat{r}_{yy}$ and $\hat{r}_{zz}$ for highly similar phenotypes (for example, BMI and waist circumference), there are many phenotype pairs for which the cross-mate cross-trait correlation and cross-mate single-trait correlations are of opposing signs (for example, BMI and years of education). Only phenotype pairs with nominally significant ($p < .05$) empirical genetic correlation estimates are shown. Labels for selected trait pairs were removed algorithmically to aid readability.

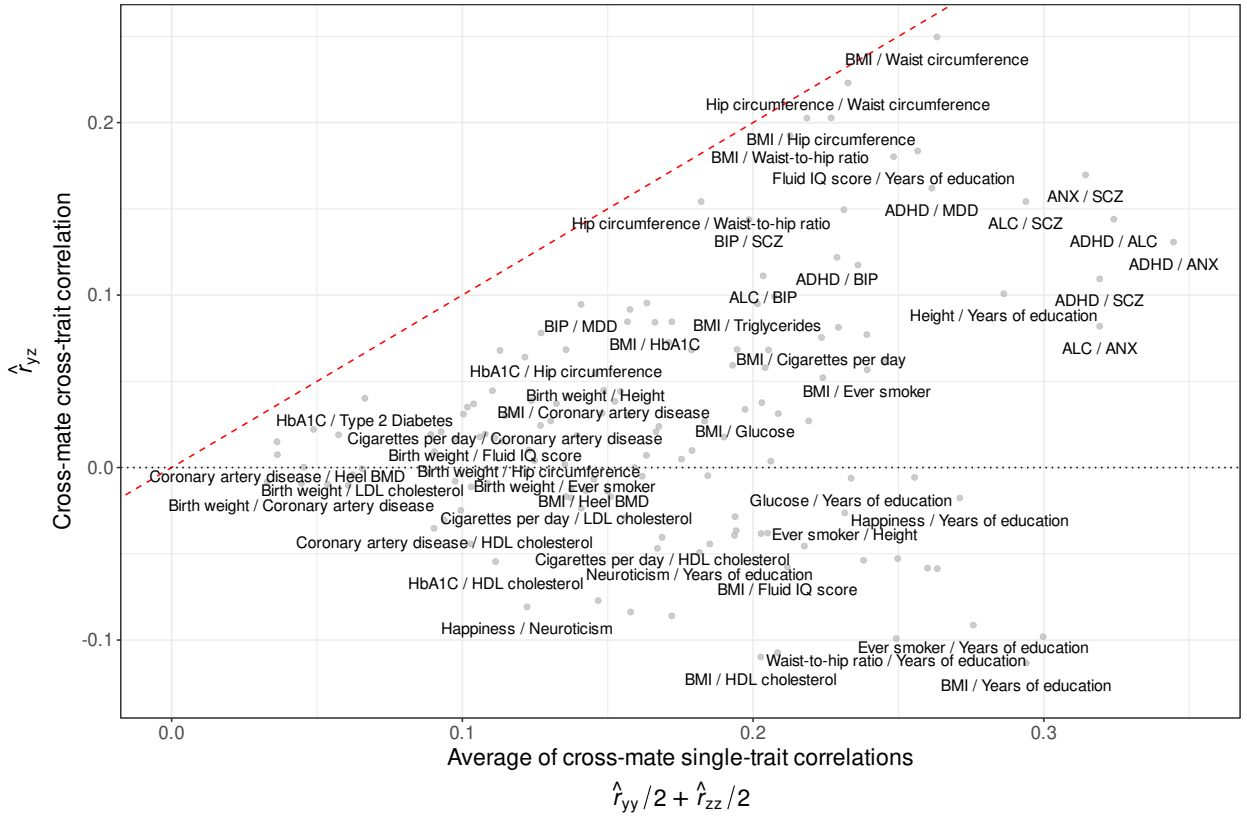


Figure S2: Empirical versus expected mate correlations under single-trait assortative mating. Empirical versus expected single-trait cross-mate correlations for a secondary trait Z assuming assortment only occurs on a related primary trait Y. In this case, the cross-mate correlation for Z is expected to be $r_{zz} = r_{yy} \cdot s_{yz}^2$, where r_{yy} is the cross-mate correlation for Y, and s_{yz} is the within-individual correlation between Y and Z. Whereas cross-mate correlation structures are consistent with single-trait AM for some highly similar phenotypes (for instance, waist-to-hip ratio and waist circumference), this is not case in general. The red dashed line is the unit slope line and the blue solid line is the ordinary least squares line of best fit (estimated slope = 0.18, significantly different from 1.00 [p =1.66e-6]). For each pair of traits, we display the assignment to Y and Z that minimized the distance between the observed and expected values of r_{zz}.Only phenotype pairs with nominally significant ($p < .05$) empirical genetic correlation estimates were included. Labels for selected trait pairs were removed algorithmically to aid readability.

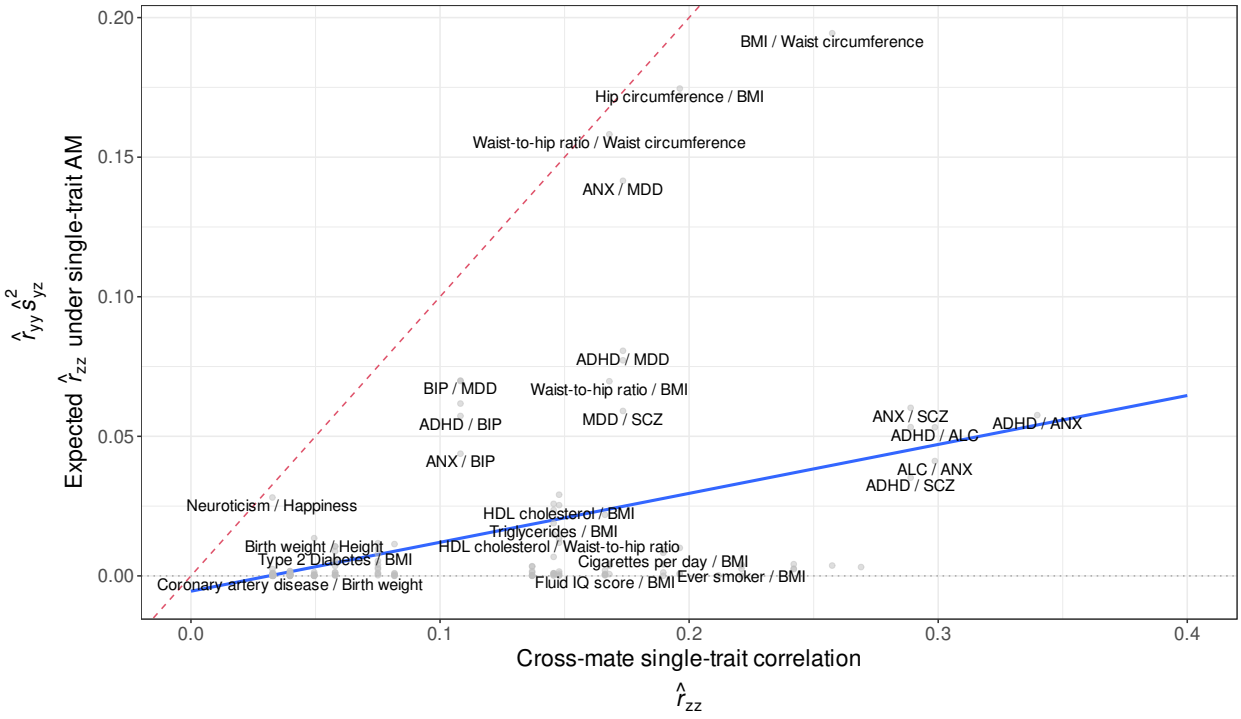


Figure S3: xAM effects and the number of casual variants. Varying the number of causal variants has no impact on either (A) true score correlations or (B) estimated effect correlations across simulations of genetically orthogonal traits with panmictic heritabilities and cross-mate correlations fixed at 0.5.

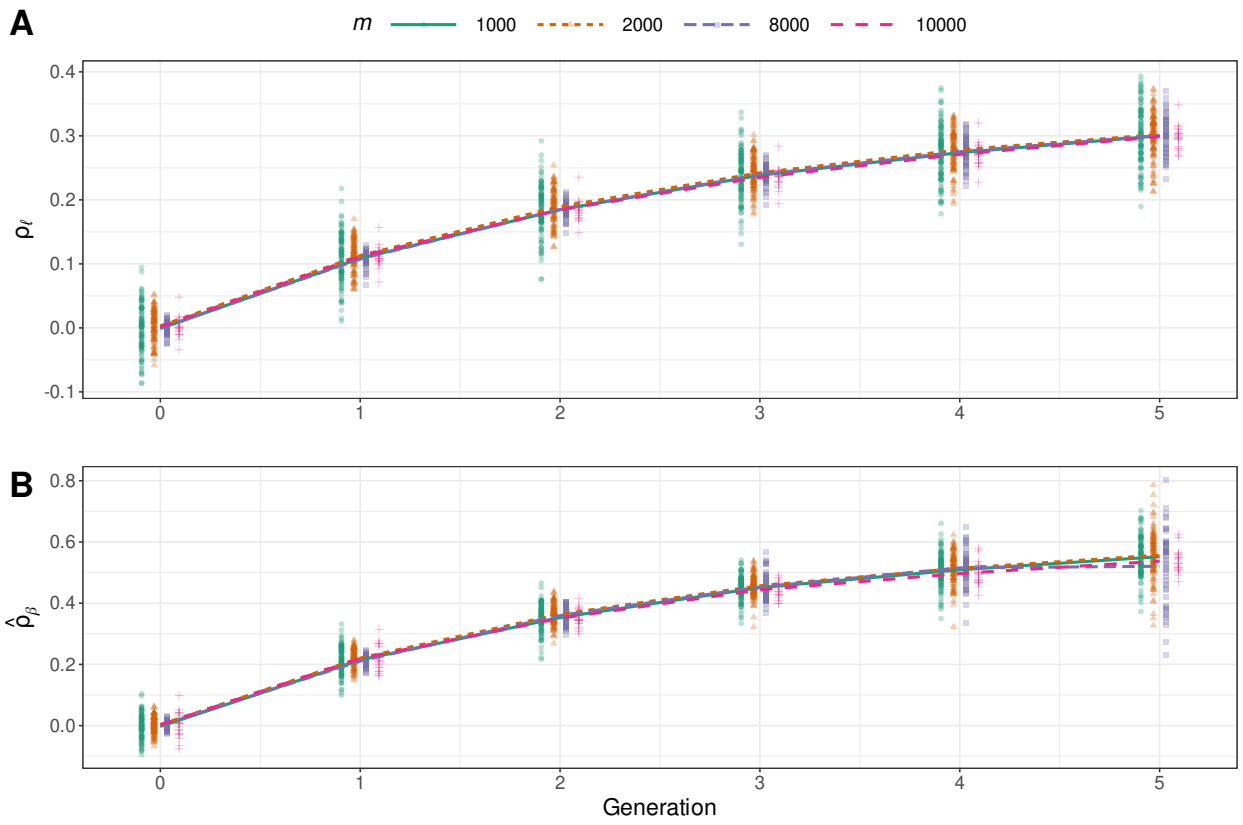


Figure S4: Comparison of synthetic mating regimes. Equivalence of mating regimes under exchangeable correlation structure for genetically orthogonal traits with respect to (A) true score correlation or (B) estimated effect correlation. Panmictic heritabilities and cross-mate correlations were fixed at 0.5 across simulations.

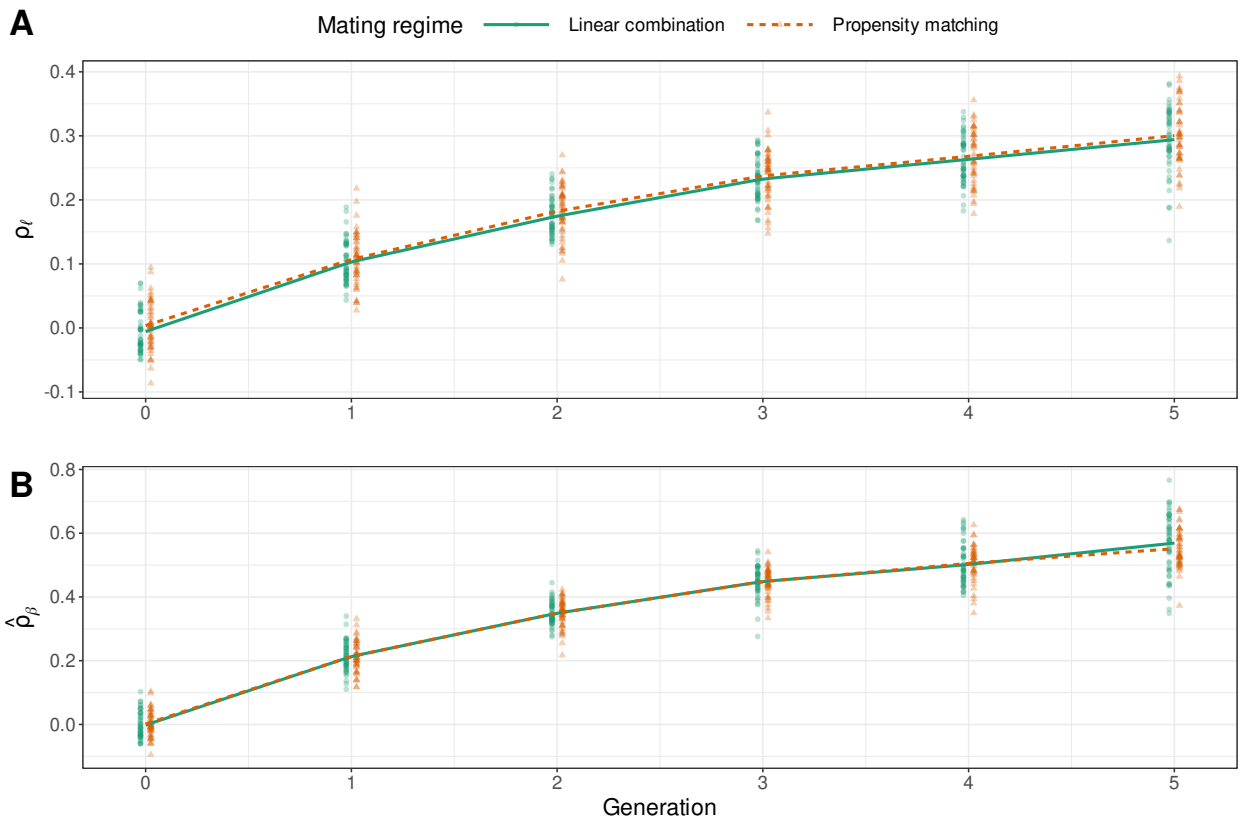


Figure S5: Effects of local LD. Impact of local LD in synthetic data for genetically orthogonal traits with panmictic heritabilities and cross-mate correlations fixed at 0.5. Enforcing recombination probabilities between continguous loci ($\rho_{\text{recom.}}$) at varying fixed values or using an empirical recombination map has no impact on the (A) true score correlation or (B) estimated effect correlation across simulations.

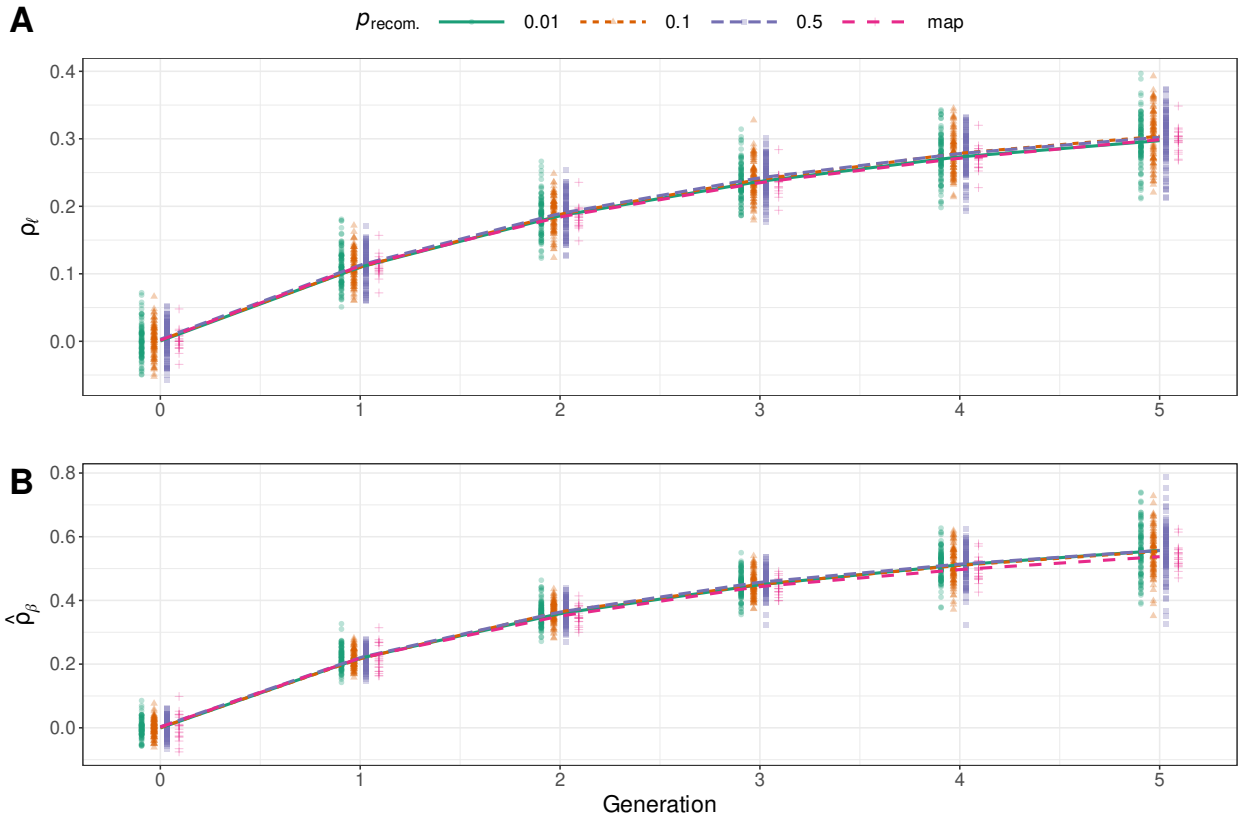


Figure S6: Pleiotropy and xAM. Score correlation and estimated effect correlation for genetically orthogonal traits subject to xAM with and without pleiotropy under orthogonal effects. Across simulations, panmictic heritabilities and cross-mate correlations were fixed at 0.5. Under the separate casual variants regime, $\beta_{y;i} \neq 0$ if and only if $\beta_{z;i} = 0$ for each causal variant indexed i, whereas, under the shared causal variants regime, every variant is causal for both Y and Z but the effects are drawn independently.

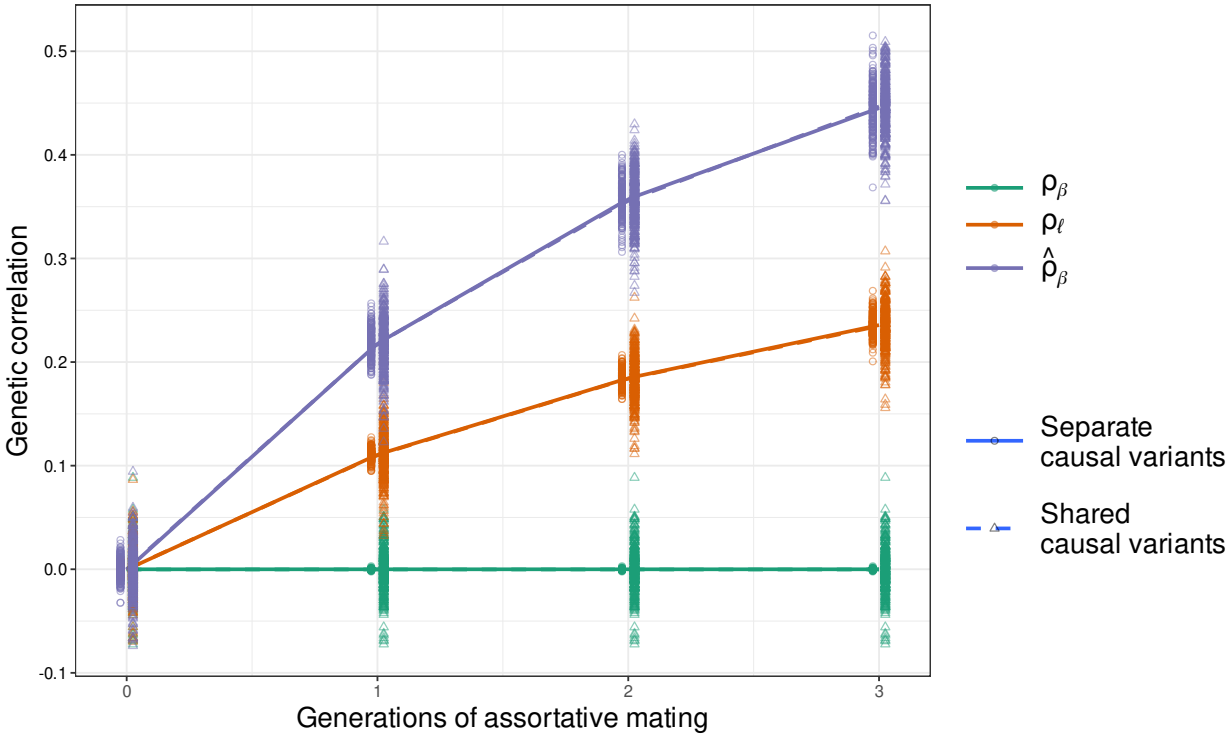


Figure S7: Signed bias of GWAS effect estimates under xAM. In the context of a GWAS of phenotype Y, variants that are causal for Y but not Z are biased upwards in magnitude (left pane); on the other hand, variants that are causal for Z but not Y are biased towards the their true effects on Z (right pane). Estimates are based on fifty simulations of bivariate xAM with exchangeable cross-mate correlations of 0.5, heritabilities of 0.5, and $m = 4000$ mutually-exclusive causal variants per phenotype. GWAS sample size was n=16000. The dashed line at $\sqrt{1/\pi}$ corresponds to the average magnitude of a Gaussian with mean zero and variance $h^2 = 0.5$.

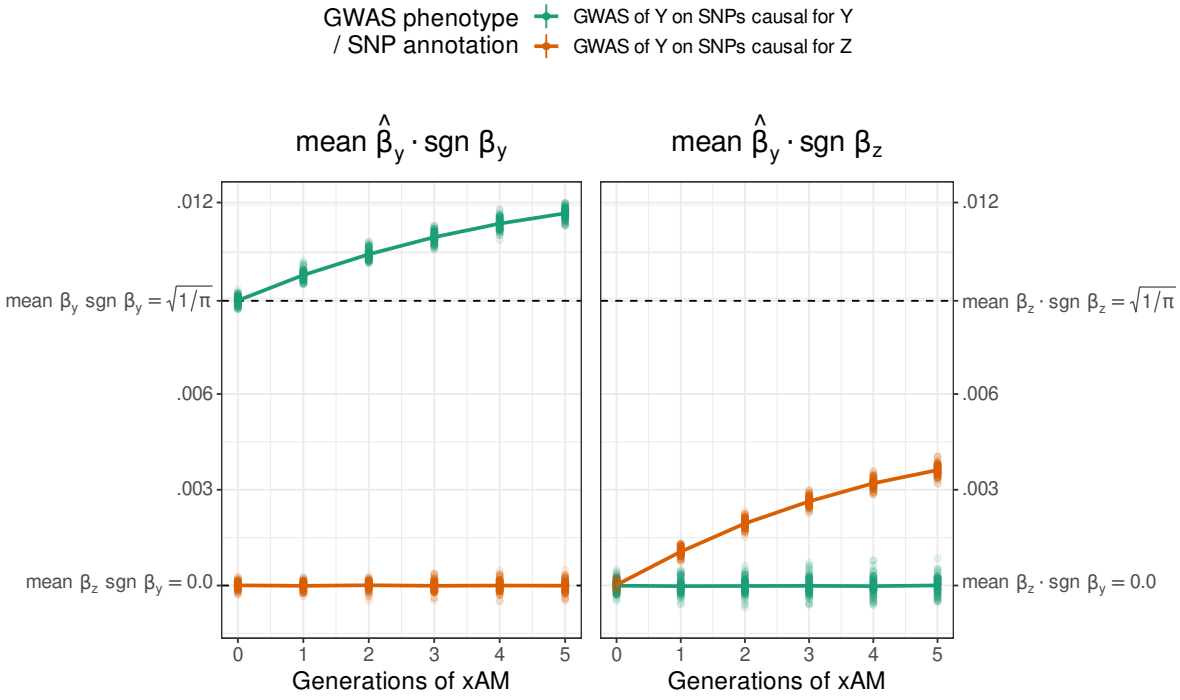


Figure S8: Example of GWAS effect estimate bias under xAM. GWAS effect estimates versus true values for a single instance of the fifty replicate simulations presented in Fig. S7. In the context of a GWAS of phenotype Y, variants that are causal for Y but not Z are biased upwards in magnitude (top row); on the other hand, variants that are causal for Z but not Y are biased towards the their true effects on Z (bottom row).

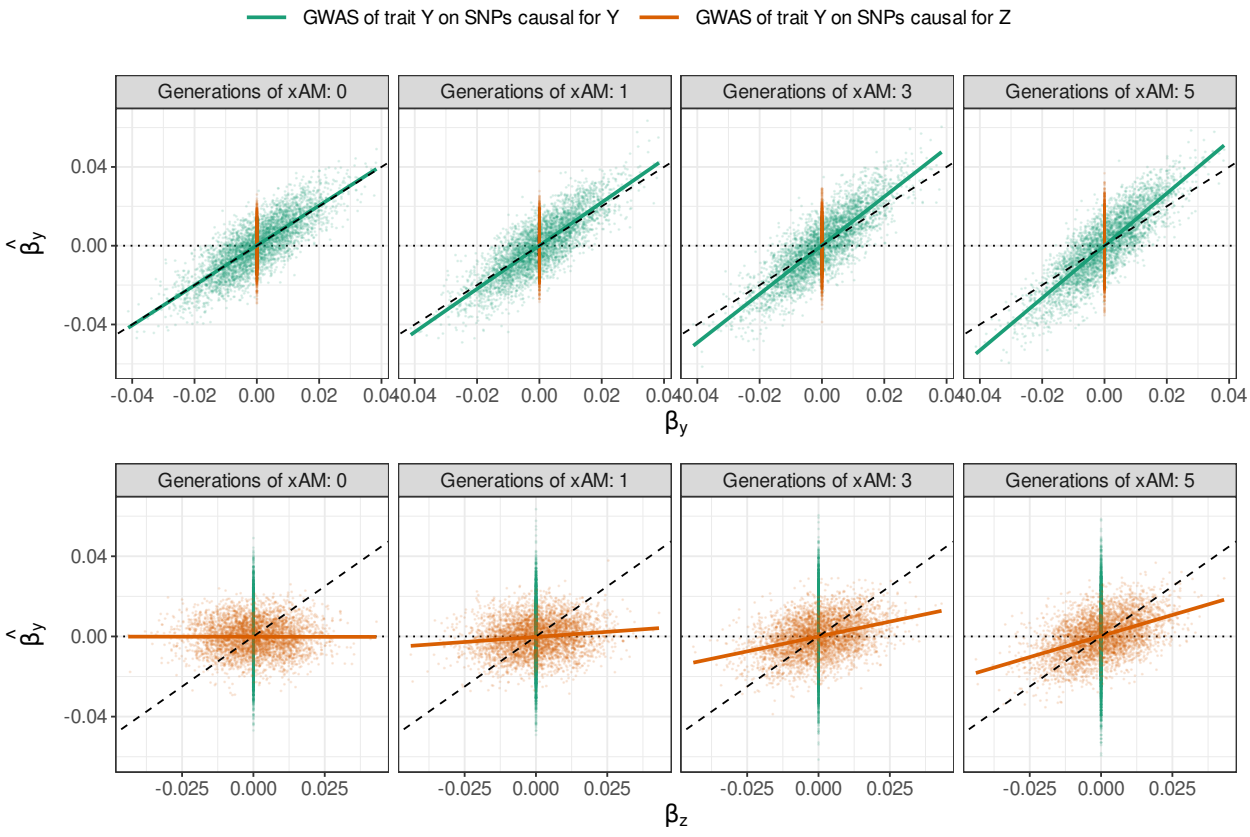


Figure S9: Signed bias of GWAS estimates for varying m, n. Signed bias of GWAS effect estimates under xAM is independent of the number of causal variants per phenotype (m) and sample size n after scaling by $\sqrt{m}$. Estimates are based on fifty simulations of bivariate xAM with exchangeable cross-mate correlations of 0.5 and heritabilities of 0.5 per combination of m and n. The dashed line at $\sqrt{1/\pi}$ corresponds to the average magnitude of a Gaussian with mean zero and variance $h^2 = 0.5$.

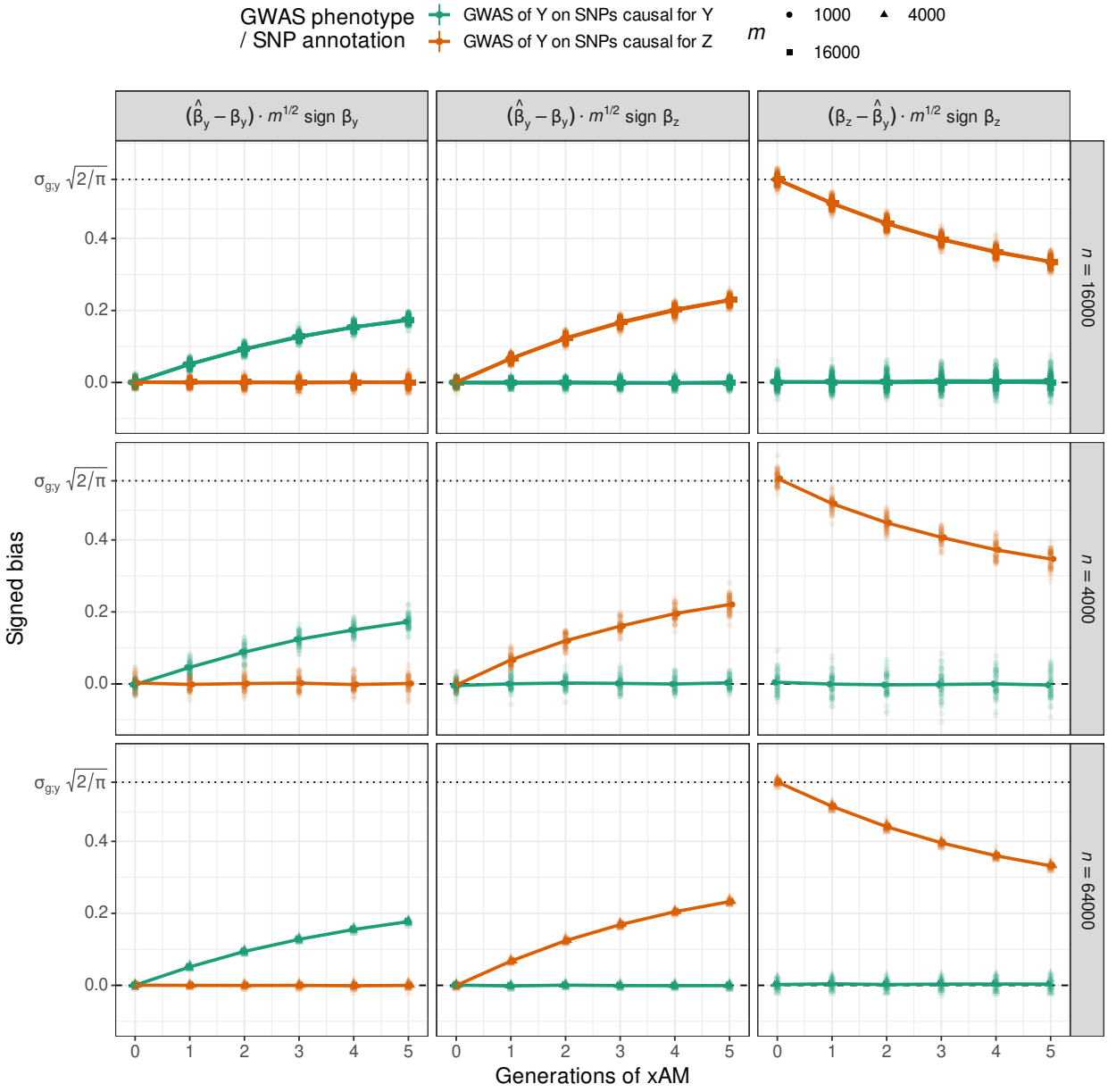


Figure S10: GWAS falsely identifies pleiotropic SNPs under xAM. SNPs affecting trait Z but not trait Y will nevertheless approach genome-wide significance (p = 5e-8) in the context of a GWAS on Y. Additionally, xAM increases power to detect causal variants for the focal phenotype due to the upward bias shown in Figs. S7 to S9. Both power and the false positive rate increase with ratio of sample size to average effect size (as modulated via the number of causal variants m under fixed heritability and sample size; top row), but also as both n and m grow large proportionately (bottom row). Estimates are based on fifty simulations of bivariate xAM with exchangeable cross-mate correlations of 0.5 and heritabilities of 0.5 per combination of m and n. To allow use of a logarithmic scale, values of zero have been replaced with the genome-wide type-I error rate.

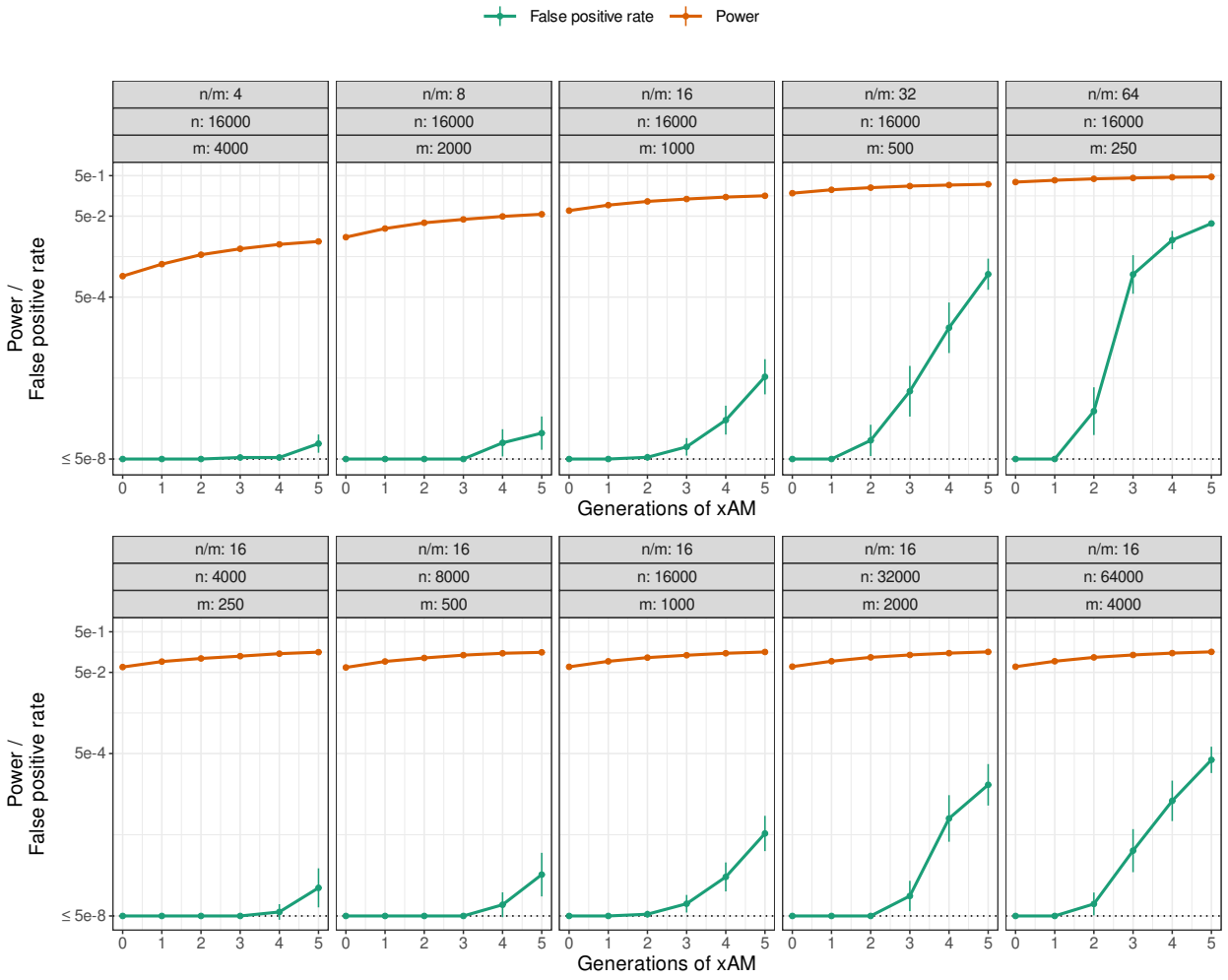


Figure S11: Effects of xAM and misdiagnosis. Impact of xAM and misdiagnosis errors on genetic correlation estimates between genetically orthogonal binary traits. Panmictic heritabilities and cross-mate correlations are all fixed at 0.5. Under the unidirectional scheme, individuals with disorder A are mislabeled as controls for disorder A and cases for disorder B, regardless of their true status for disorder B, at the rate reflected on the x axis. Under the bidirectional scheme, the analogous misdiagnoses are enforced for disorder B as well. The induced bias in both cases is more pronounced for less common disorders.

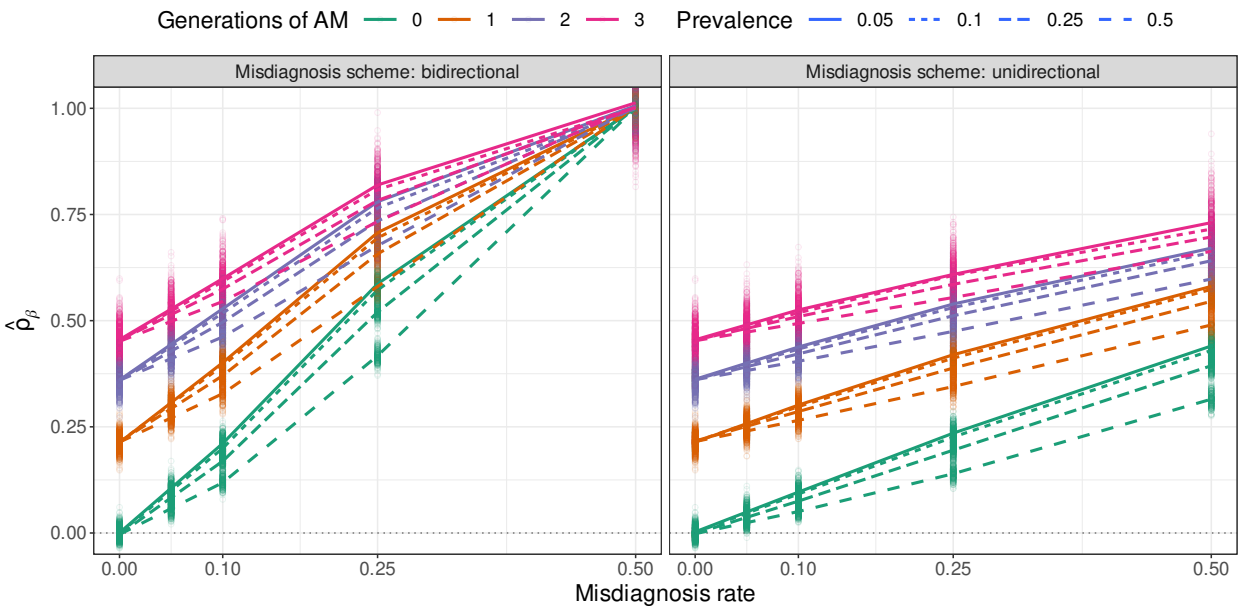


Figure S12: Partitioned genetic correlation under xAM. Partitioned HE regression estimates of genetic correlation under xAM. Plots reflect estimates across 250 replicate simulations of bivariate xAM with exchangeable cross-mate correlations of 0.5, with $m = 2{,}400$ causal variants equally divided into six mutually exclusive annotations (see legend) and a sample size of $n = 16{,}000$. Horizontal lines show the true effect correlation for each annotation. Annotations with orthogonal effects and with little relevance to one or both traits evidence the largest estimates: after three generations of xAM, the average annotation-specific genetic correlation estimates were respectively $\hat{\rho}_\beta^{(1)} = 0.63$, $\hat{\rho}_\beta^{(2)} = 0.46$, $\hat{\rho}_\beta^{(3)} = 0.87$, $\hat{\rho}_\beta^{(4)} = 0.87$, $\hat{\rho}_\beta^{(5)} = 0.46$, and $\hat{\rho}_\beta^{(6)} = 0.70$, all of which were substantially larger than the true annotation-specific effect correlations (0.25 for annotation (1), 0.00 for all others). The largest annotation-specific estimates were attributed to annotations (3) and (4) followed by annotation (6). Annotations explaining less variation than others will generate annotation-specific genetic covariance estimates that are relatively large compared to their annotation-specific heritability estimates, leading to extreme upward bias and distorting the relative ordering of annotation-specific signals. Estimates for which the geometric mean heritability estimates across the two traits was less than 1e-3, which yielded unstable genetic correlation estimates, were dropped for clarity. Error bars reflect 95% credible intervals.

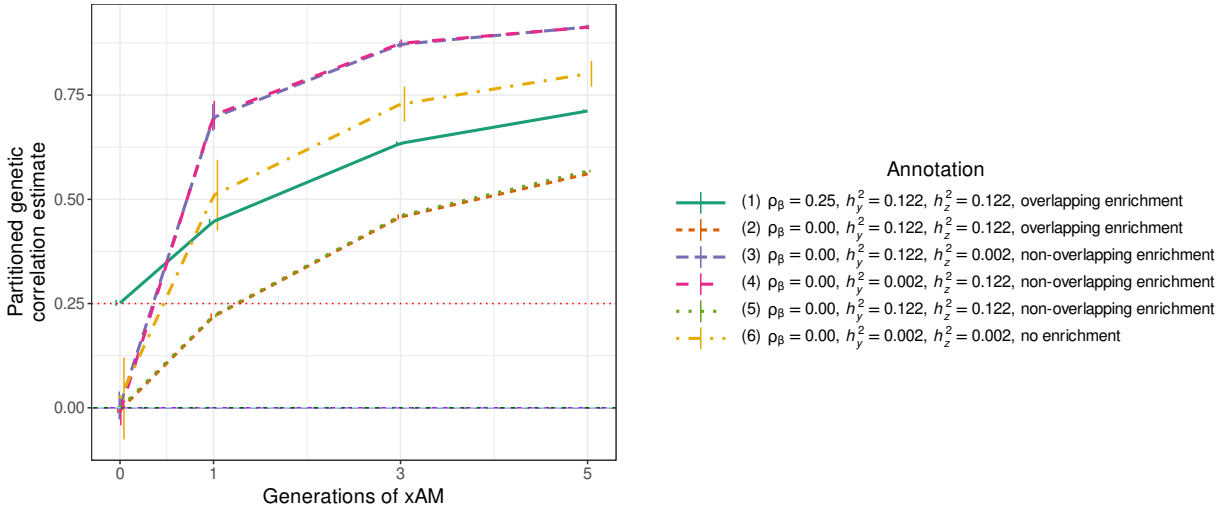


Figure S13: Impact of xAM on genetic correlation measures versus genetic covariance measures. Results are those presented in Figs. 2C and 2D (A), or reformulated in terms of covariances (B). Compared to genetic correlations, genetic covariances are unbounded and thus do not converge to one as $\rho_\beta \to 1$. Additionally, whereas HE and LDSC genetic correlation estimates are identical, LDSC genetic covariance estimates are attenuated as the LDSC genetic variance estimator is more susceptible to missing heritability issues. This is particularly evident in the attenuated LDSC genetic variances we observe here, where all causal variants are directly measured in the simulated marker data; the only source of attenuation here comes from imperfect LD scores, which as is always the case in practice, have been estimated using a sliding window approach.

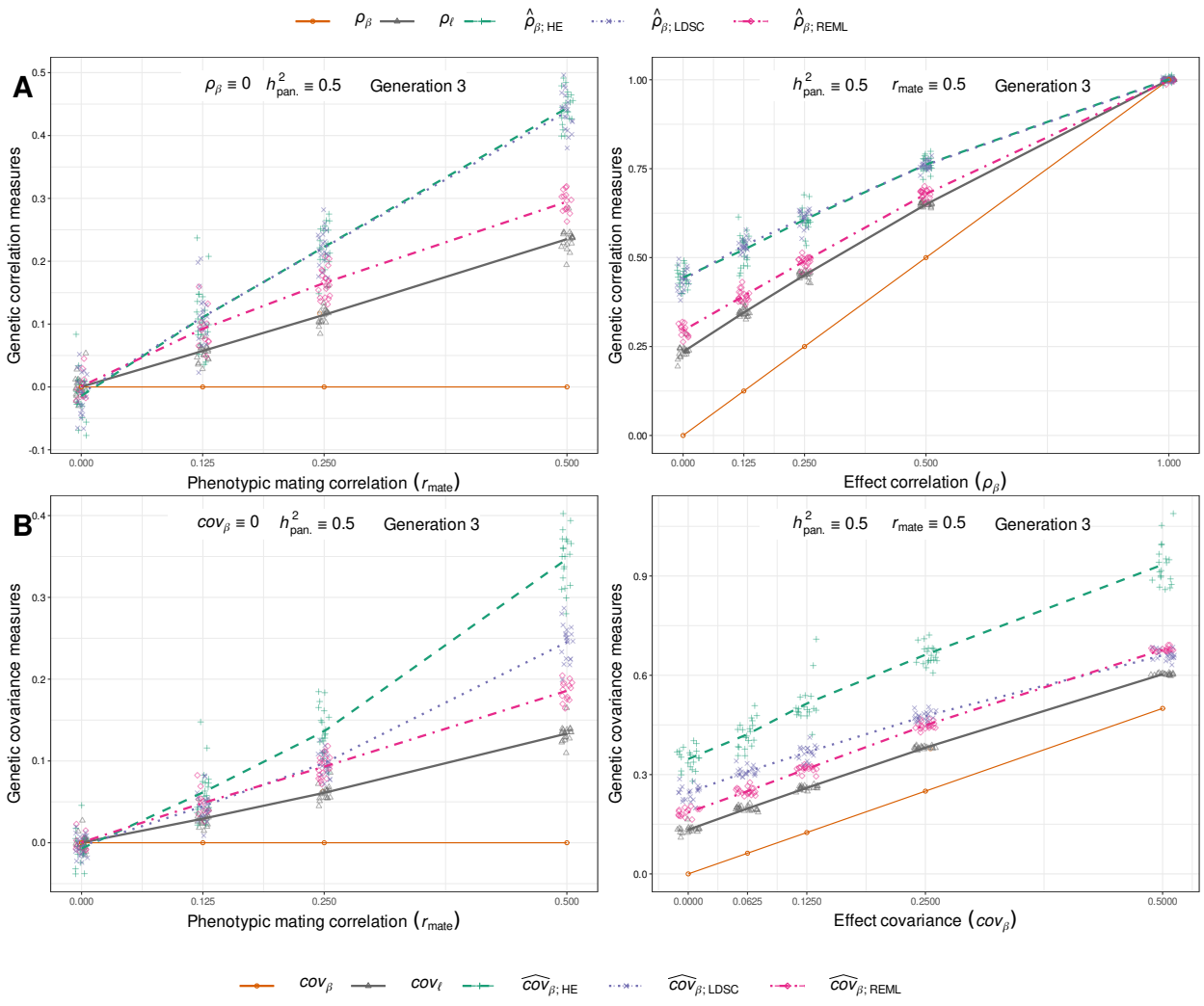


Figure S14: Projected versus empirical $\hat{\rho}_\beta$ for UKB traits. Projections under xAM alone ($\hat{\rho}_{\mathrm{xAM}}$) versus empirical $\hat{\rho}_\beta$ estimates for UK Biobank trait pairs, substituting either pedigree-based ($\hat{h}^2_{\mathrm{PED}}$) or marker-based ($\hat{h}^2_{\mathrm{SNP}}$) heritability estimates for panmictic heritabilities (h^2_{pan}), as a function of number of generations of xAM. Error bars reflect 95% credible intervals.

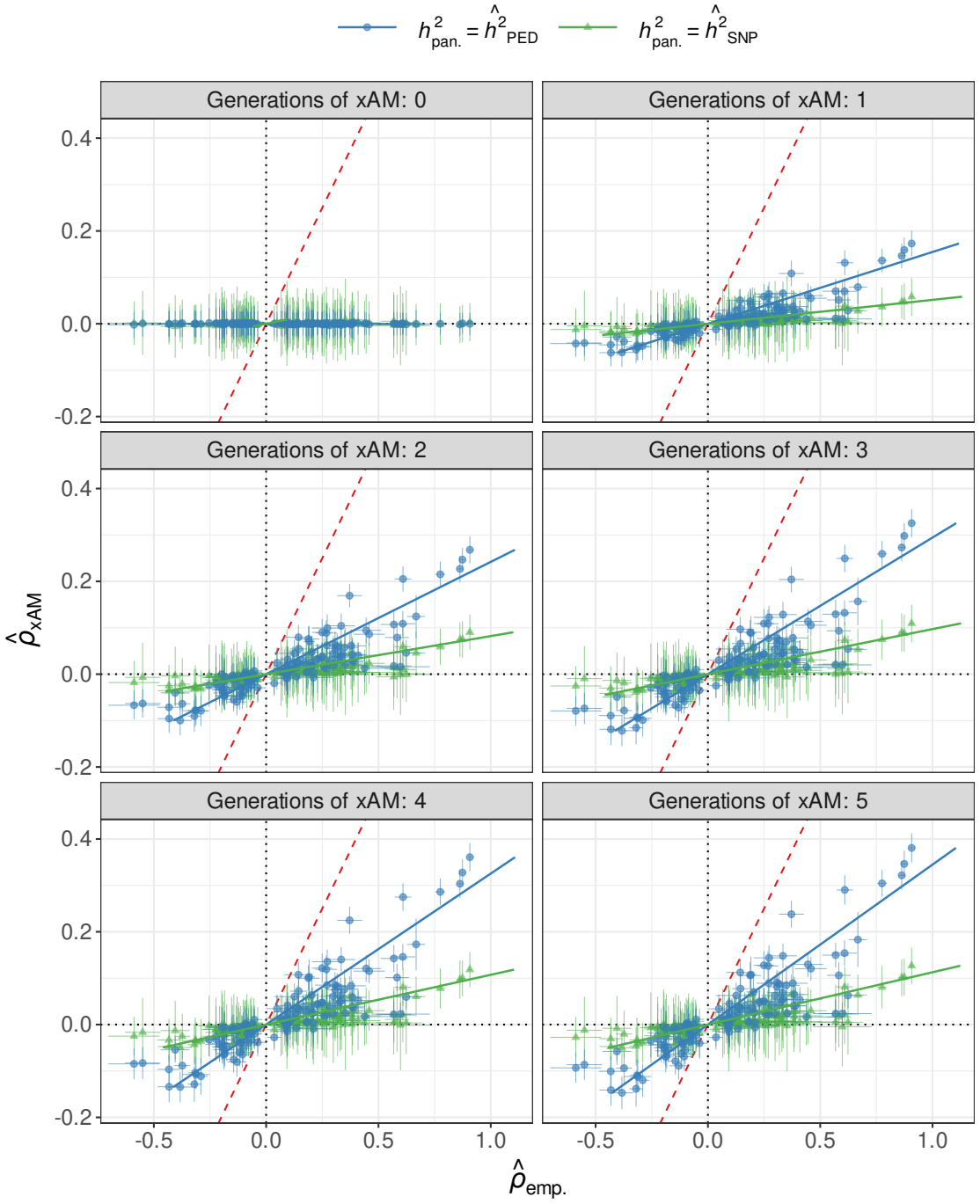


Figure S15A: Projected $\hat{\gamma}$ for UKB traits 1-20. Projected $\hat{\rho}_\beta$ estimates relative to empirical estimates for UK Biobank trait pairs 1-20, substituting either pedigree-based ($\hat{h}^2_{\text{PED}}$) or marker-based ($\hat{h}^2_{\text{SNP}}$) heritability estimates for panmictic heritabilities (h^2_{pan}), as a function of number of generations of xAM. Error bars reflect 95% credible intervals.

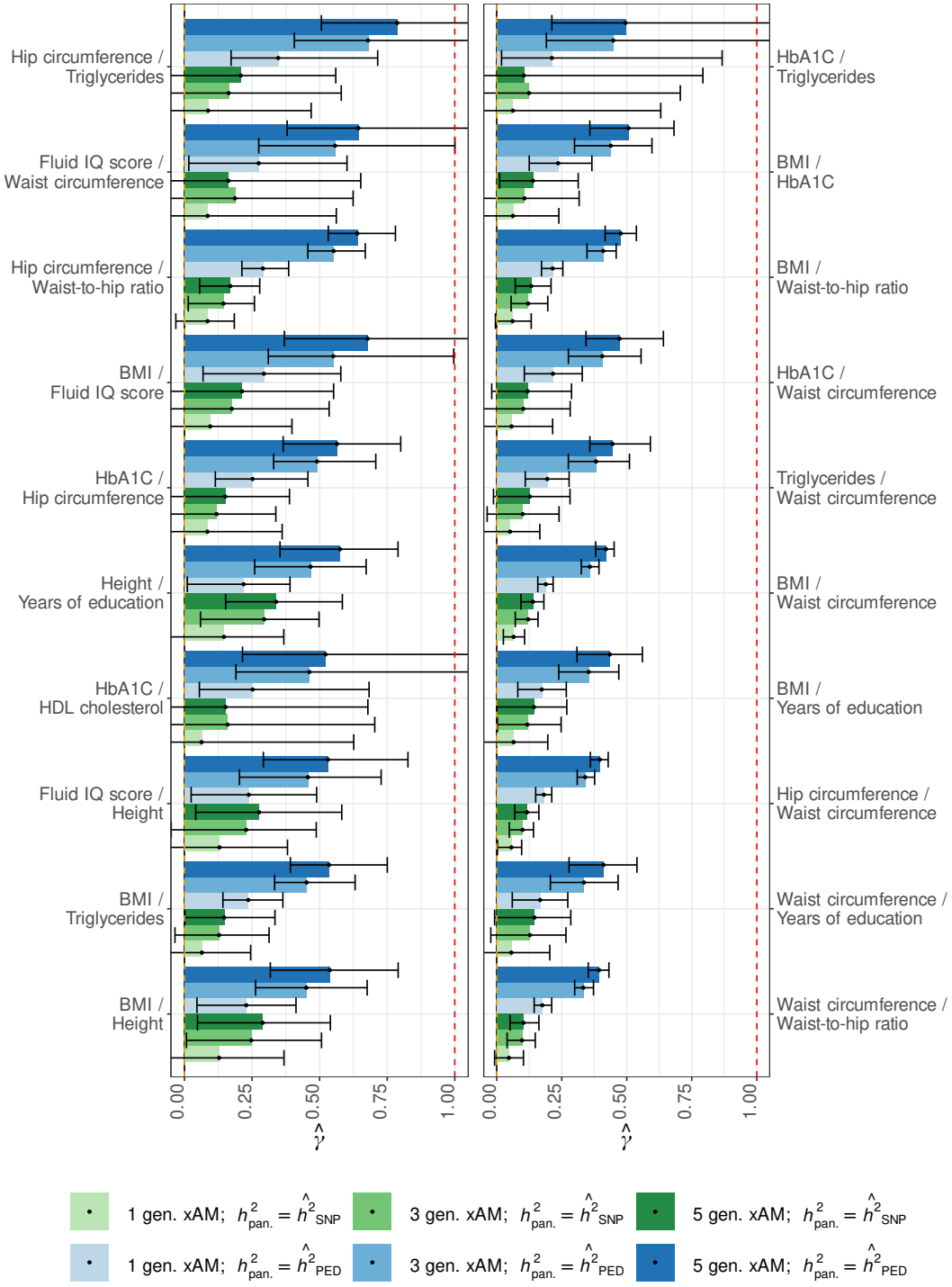


Figure S15B: Projected $\hat{\gamma}$ for UKB traits 21-42. Projected $\hat{\rho}_\beta$ estimates relative to empirical estimates for UK Biobank trait pairs 21-42, substituting either pedigree-based ($\hat{h}^2_{\text{PED}}$) or marker-based ($\hat{h}^2_{\text{SNP}}$) heritability estimates for panmictic heritabilities (h^2_{pan}), as a function of number of generations of xAM. Error bars reflect 95% credible intervals.

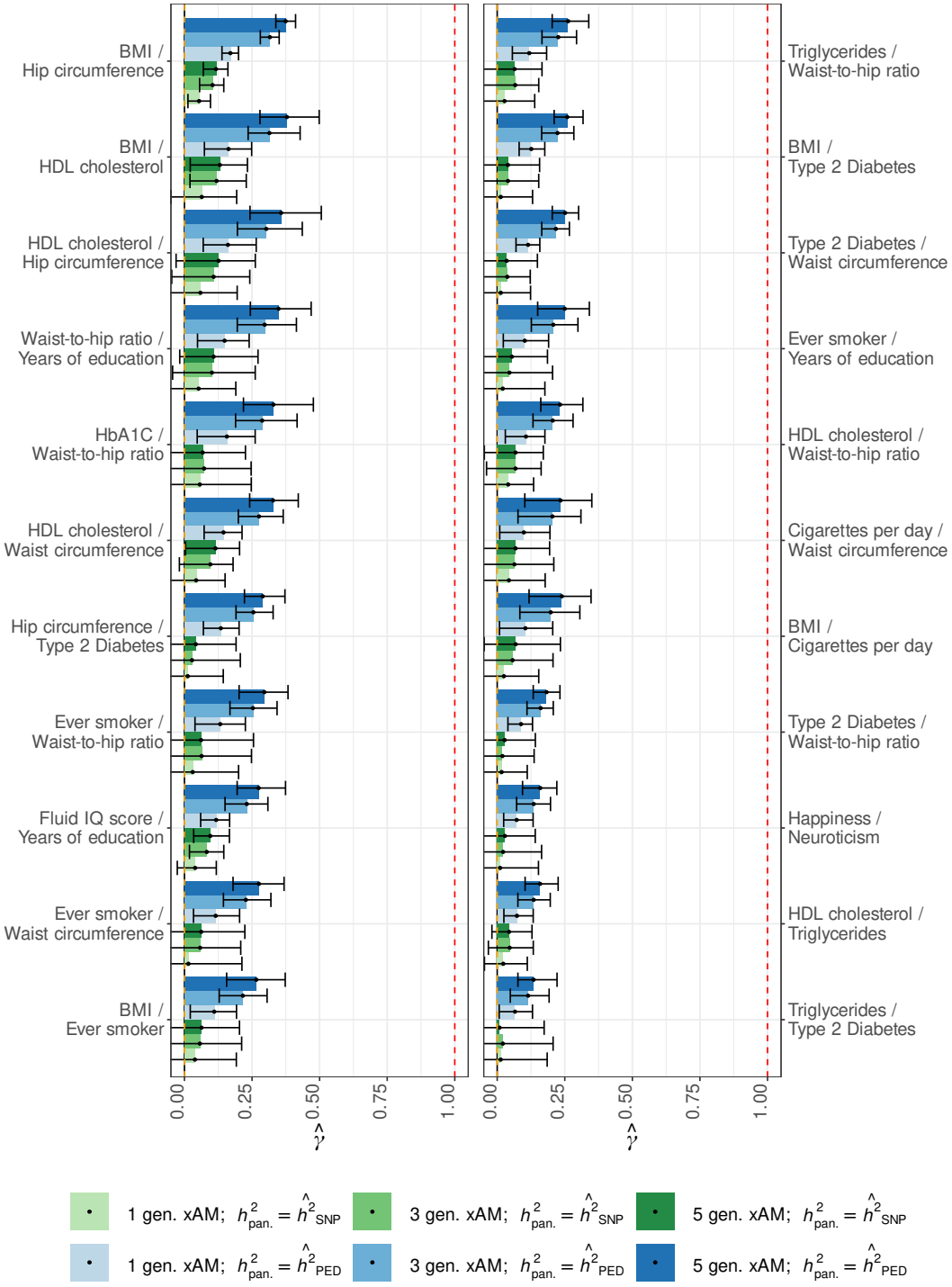


Figure S16: Projected versus empirical $\hat{\rho}_\beta$ for psychiatric trait pairs. Projections reflect up to five generations of xAM congruent with empirical spousal correlation estimates. Selected points labeled for emphasis. Error bars reflect 95% credible intervals.

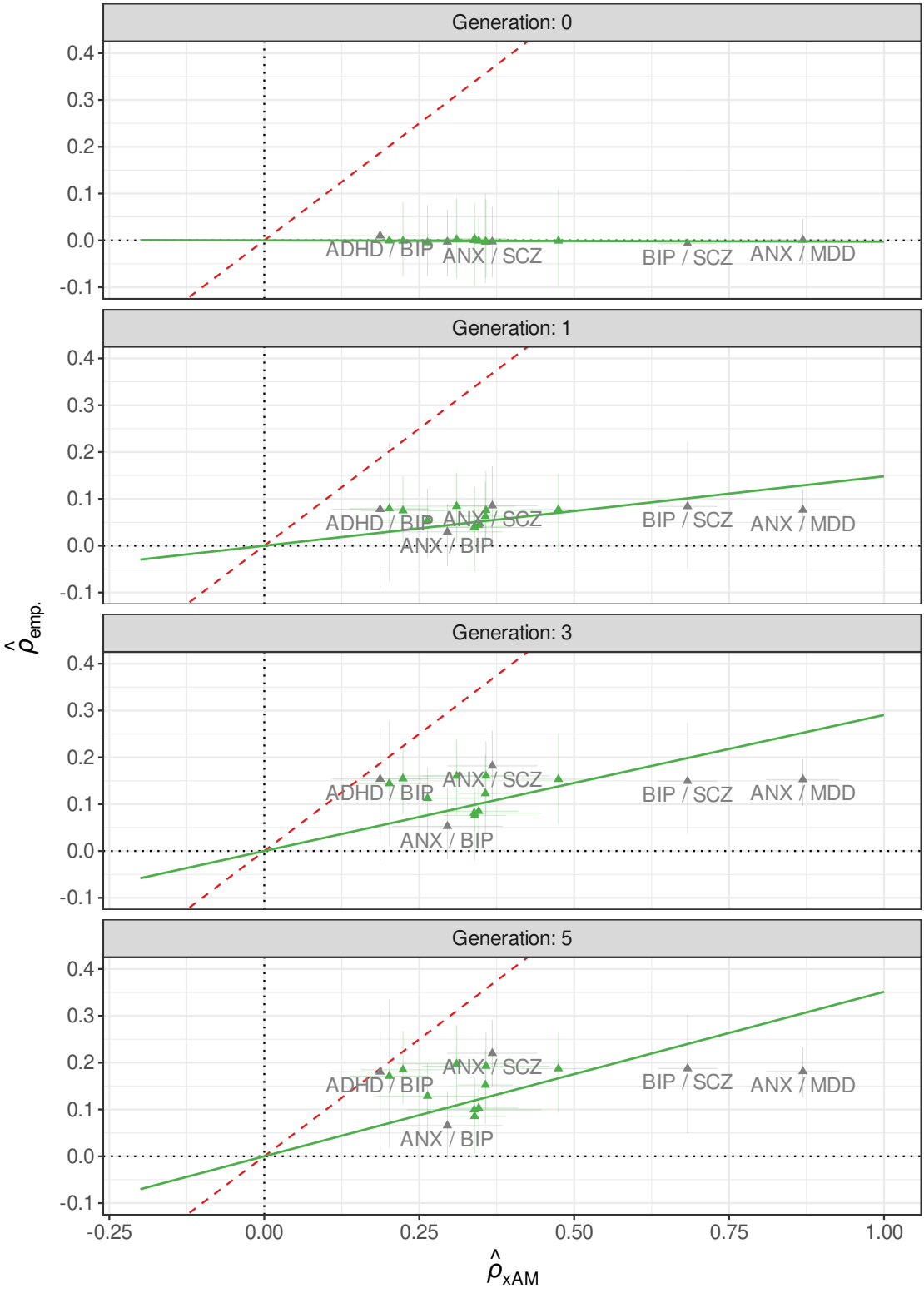


Figure S17: Projected versus empirical $\hat{\rho}_\beta$ for psychiatric trait pairs subject to misdiagnosis. Projections reflect bidirectional misclassification errors and up to five generations of xAM congruent with empirical spousal correlation estimates. Error bars reflect 95% credible intervals.

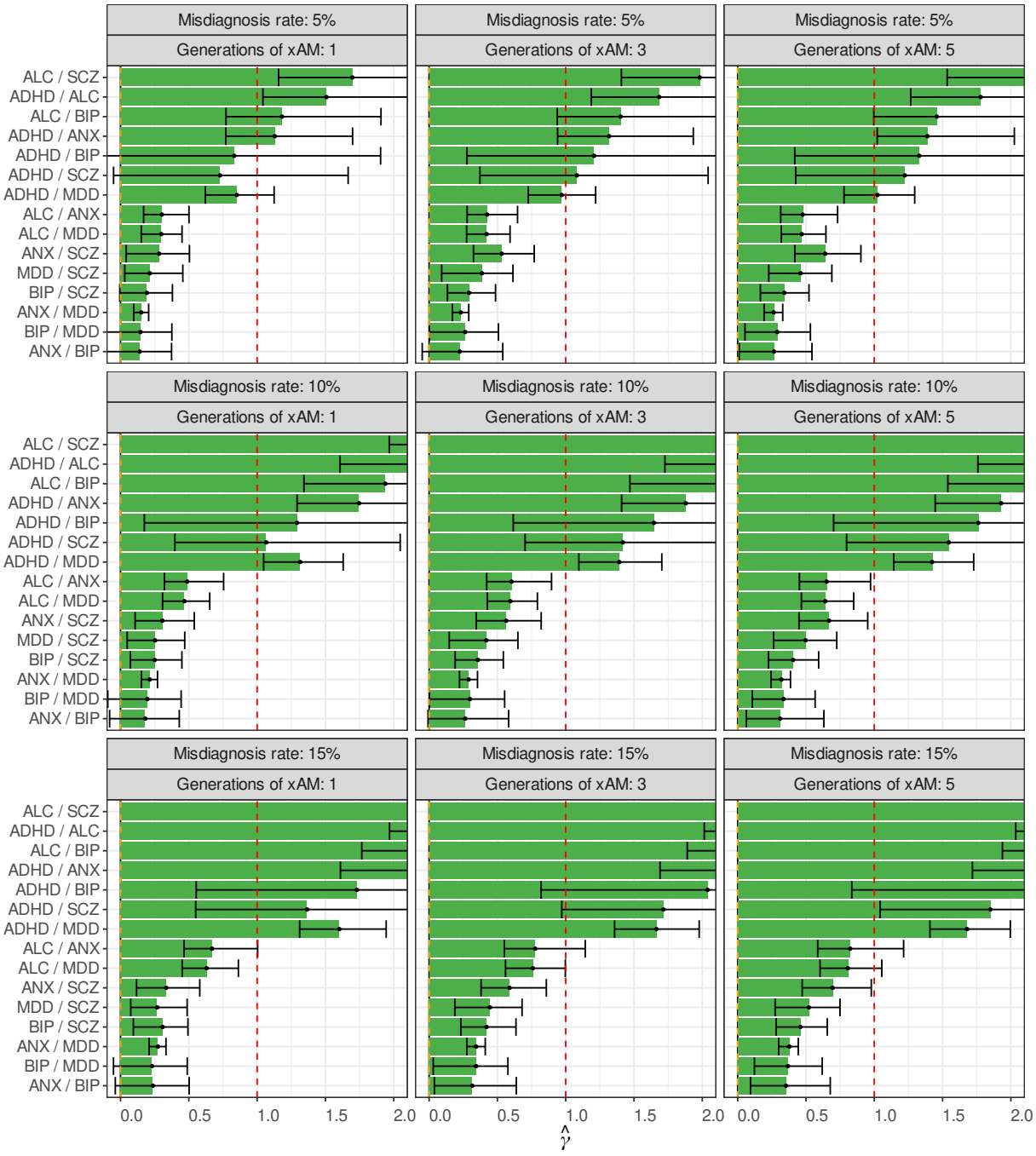


Figure S18: Even/odd PGS correlations for varying thresholds. Associations between cross-chromosome polygenic score correlations (y-axis) and cross-mate phenotypic correlations (x-axis) are consistent across varying p-value thresholds.

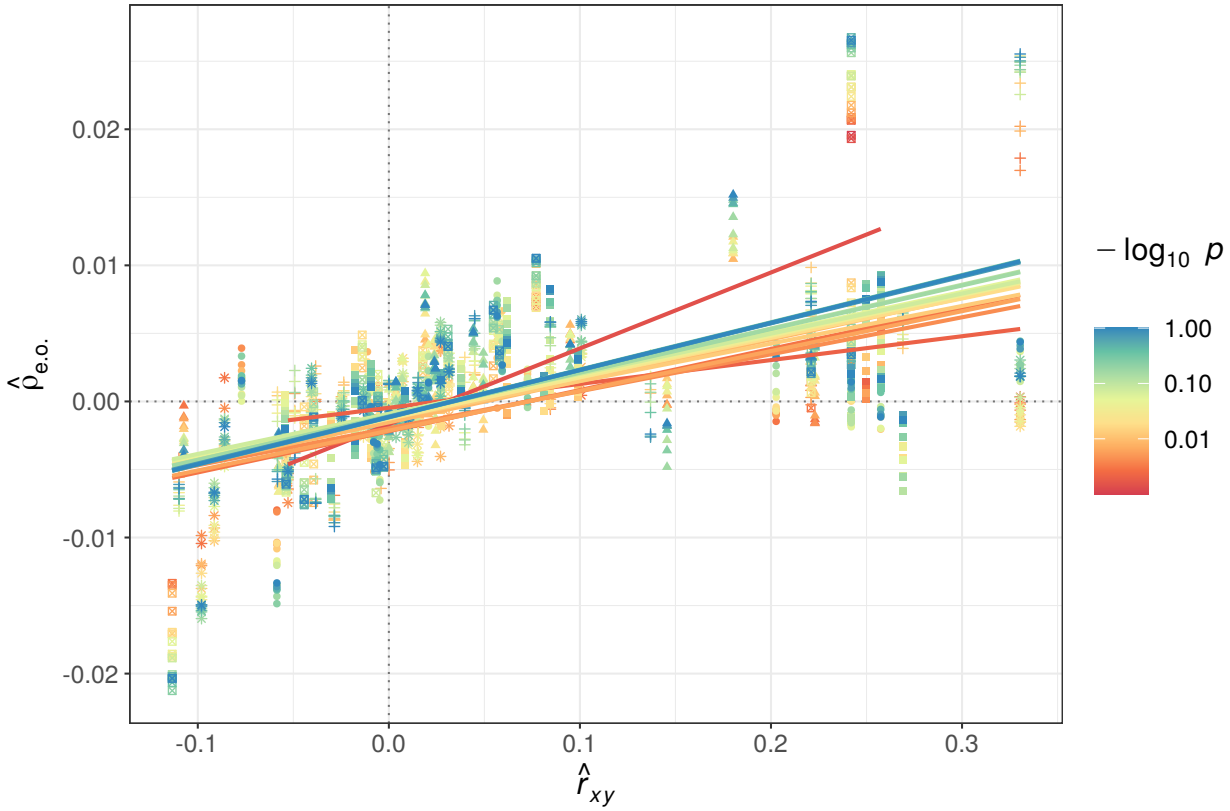


Figure S19: Alternative $\hat{\gamma}$ definition estimates. An alternative definition of $\hat{\gamma}$ based on the ratio of expected and observed genetic covariance estimates yields similar results as the correlation based definition used in the primary manuscript. Across all projections for UKB phenotypes, the covariance ratio was slightly higher as it also reflects xAM-induced inflation of heritability estimates.

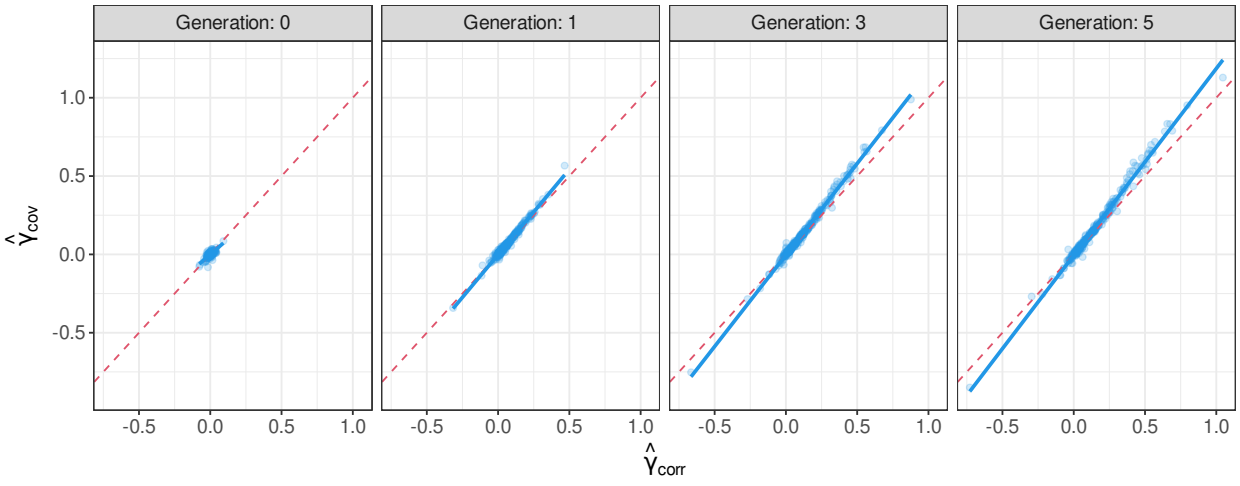
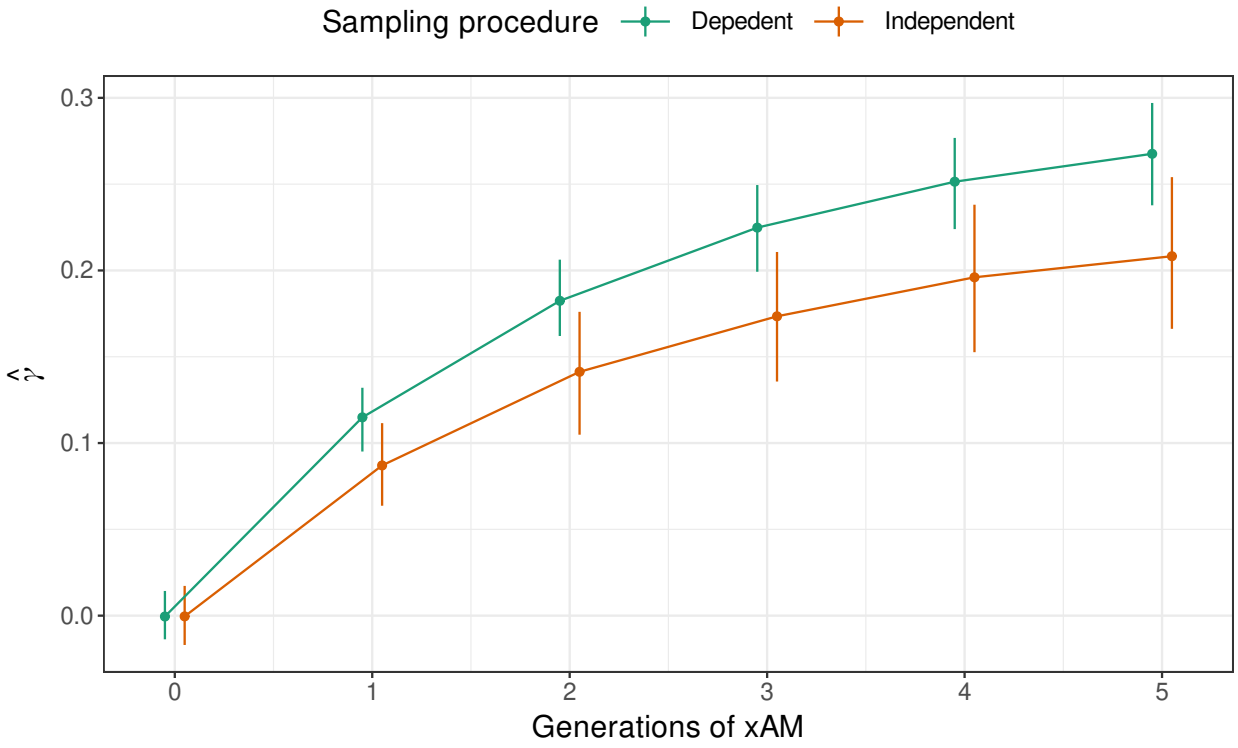


Figure S20: Impact of parameter dependence on $\hat{\gamma}$ estimates. Sensitivity analyses depicting the approximate impact of parameter dependence on $\hat{\gamma}$ estimates for ANX and MDD. Strong parameter dependence leads to larger, less variable $\hat{\gamma}$ estimates.



Captions for Tables S1 to S7

Table S1: Pairwise estimates. Genetic correlation and cross-mate correlation estimates with standard errors. For each trait pair, `rg` denotes the LD score regression genetic correlation, `rx_yz` is the cross-mate cross-trait correlation, `rw_yz` is the within-individual correlation, and `rx_yy`, `rx_zz` are the cross-mate correlations for the first and second phenotypes, respectively.

Table S2: UK Biobank univariate data. Pedigree- (`h2_ped`) and marker-based (`h2_snp`) heritability estimates and standard errors for UK Biobank phenotypes. `h2_ped_source` lists sources for pedigree estimates and `h2_ped_notes` includes additional details when relevant.

Table S3: UK Biobank phenotype simulation results. Simulation results for each trait pair after zero to five generations of xAM consistent with empirical cross-mate correlations. `rhobeta_empirical` denotes empirical LD score regression genetic correlation estimates. `rhoxAM_PED` and `rhoxAM_SNP` denote $\hat{\rho}_{\mathrm{xAM}}$ projections computed by setting panmictic heritabilities to pedigree- versus marker-based estimates respectively, as is the case for `gammahat_PED` and `gammahat_SNP`, which correspond to $\hat{\gamma}$ estimates. `rxmate_discrepancy*` denotes the average discrepancy from the target cross-mate correlation versus that achieved in simulation. Also included are standard errors and empirical quantiles.

Table S4: Psychiatric phenotype univariate data. Pedigree-based heritability estimates (`h2_ped`) and standard errors for psychiatric phenotypes. `h2_ped_country` lists the country in which the source sample was ascertained, `h2_ped_source` lists the source study, and `h2_notes` includes additional details when relevant.

Table S5: Psychiatric phenotype simulation results. Pairwise data and simulation results for each trait pair after zero to five generations of xAM consistent with empirical cross-mate correlations, under varying levels of diagnostic errors (`MC_rate`). `rg` and `rg_HE` correspond to $\hat{\rho}_{\mathrm{emp}}$ and $\hat{\rho}_{\mathrm{xAM}}$, respectively, and `rg_rel` corresponds to $\hat{\gamma}$. Also included are standard errors and upper and lower bounds of 95% credible intervals. `rx_yz` is the cross-mate cross-trait tetrachoric correlation, `rw_yz` is the within-individual tetrachoric correlation, and `rx_yy`, `rx_zz` are the cross-mate tetrachoric correlations for the first and second phenotypes, respectively.

Table S6: Even / odd chromosome PGS correlations. Correlations between polygenic scores (PGS) on odd versus even chromosomes across pairs of UK Biobank phenotypes.

Table S7: Psychiatric phenotype definitions. Definitions of psychiatric disorder phenotypes used in Danish registry data in terms of ICD-8 and ICD-10 codes.