

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

Richard Border<sup>1,2,3</sup>, Georgios Athanasiadis<sup>4,5</sup>, Alfonso Buil<sup>4</sup>, Andrew Schork<sup>4,6</sup>, Na Cai<sup>7</sup>, Alexander Young<sup>1,8</sup>, Thomas Werge<sup>4,9</sup>, Jonathan Flint<sup>10</sup>, Kenneth Kendler<sup>11</sup>, Sriram Sankararaman<sup>2,12,13</sup>, Andy Dahl<sup>14</sup>, and Noah Zaitlen<sup>1,12</sup>

<sup>1</sup>David Geffen School of Medicine, University of California, Los Angeles, California, United States

<sup>2</sup>Department of Computer Science, University of California, Los Angeles, California, United States

<sup>3</sup>Department of Epidemiology, Harvard T.H. Chan School of Public Health, Massachusetts, United States

<sup>4</sup>Institute of Biological Psychiatry, Mental Health Center Sct Hans, Copenhagen University Hospital - Mental Health Services CPH, Copenhagen, Denmark

<sup>5</sup>Department of Evolutionary Biology, Ecology, and Environmental Sciences, University of Barcelona, Catalonia, Spain

<sup>6</sup>Neurogenomics Division, The Translational Genomics Research Institute, Arizona, United States

<sup>7</sup>Helmholtz Pioneer Campus, Helmholtz Zentrum München, Neuherberg, Germany

<sup>8</sup>Anderson School of Management, University of California, Los Angeles, California, United States

<sup>9</sup>Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark

<sup>10</sup>Center for Neurobehavioral Genetics, University of California, Los Angeles, California, United States

<sup>11</sup>Department of Psychiatry, Virginia Institute for Psychiatric Genetics, Virginia Commonwealth University

<sup>12</sup>Department of Computational Medicine, University of California, Los Angeles, California, United States

<sup>13</sup>Department of Human Genetics, University of California, Los Angeles, California, United States

<sup>14</sup>Section of Genetic Medicine, Department of Medicine, University of Chicago, Illinois, United States

The observation of genetic correlations between disparate traits has been interpreted as evidence of widespread pleiotropy, altered theories of human genetic architecture, and spurred considerable research activity across the natural and social sciences. Here, we introduce cross-trait assortative mating (xAM) as an alternative explanation for observed genetic correlations. We observe that xAM is common across a broad array of phenotypes and that phenotypic cross-mate correlation estimates are strongly associated with genetic correlation estimates ( $R^2 = 76\%$ ). Then, we present theoretical and simulation-based results demonstrating that, under xAM, genetic correlation estimators yield significant estimates even for traits with entirely distinct genetic bases. We demonstrate that existing xAM plausibly accounts for substantial fractions of genetic correlation estimates in two large samples ( $N = 827,960$ ). For example, previously reported genetic correlation estimates between many pairs of psychiatric disorders are fully consistent with xAM alone. Finally, we provide evidence for a history of xAM at the genetic level using a novel approach based on cross-trait even/odd chromosome polygenic score correlations. Together, our results demonstrate that previous reports have likely overestimated the true genetic similarity between many phenotypes.

## 1 Introduction

Over the previous decade, considerable research activity has focused on using genetic data to infer the extent to which pairs of complex traits share overlapping genetic bases [1–7]. These efforts have been propelled by the development of genetic correlation estimators designed to use summary statistics from genome-wide association studies (GWAS) [3, 8, 9], which have become a fundamental statistical tool across many domains of human complex trait genetics, including psychiatry [2, 10], economics [11, 12], and medicine [13, 14]. The results of these analyses have been striking: many trait pairs, even those with limited phenotypic similarity, display nontrivial genetic correlations (e.g., 0.209 [ $se = 0.042$ ] for Attention-deficit Hyperactivity Disorder [ADHD] and body mass index [BMI] in [2]). These findings have been broadly interpreted as evidence for widespread pleiotropy across the phenome [3–5, 15], and, in the case of psychiatric disorders, have raised concerns about the suitability of the existing nosology in the face of shared pathogenesis at the genetic level across ostensibly distinct phenotypes [2, 7].

However, such interpretations assume that genetic correlation estimators index shared genetic architectures, and various alternative explanations have been explored in considerable detail (e.g., vertical pleiotropy and diagnostic errors [5, 16]). Here, we consider a previously unexamined source of potential bias: cross-trait assortative mating (xAM), the phenomenon whereby mates are correlated across distinct traits (i.e., individuals' values of phenotype  $Y$  are correlated with their partners' values of phenotype  $Z$ ). There are at least four reasons to be concerned with this potential oversight: First, the single trait random effects model, which genetic correlation estimators generalize, is misspecified under single-trait AM and as a result overestimates SNP heritability [17]. Second, single-trait AM is widespread across multiple domains for which substantial genetic correlations have been observed, including anthropometric, psychosocial, and disease traits [2, 4, 5]. Third, recent work has provided genetic-level evidence for a history of single-trait AM with respect to some of these same phenotypes [18]. Fourth, xAM is known to generate spurious results for other marker-based inference procedures, including Mendelian randomization [19] and GWAS [20].

The present investigation comprises the first systematic investigation of the impact of xAM on genetic correlation estimates. We first compile the largest to-date atlas of cross-mate cross-trait phenotypic correlations, examining a broad array of previously-studied metabolic, anthropometric, psychosocial, and psychiatric phenotypes across two large population based samples in the United Kingdom ( $n=81,394$ ) and Denmark ( $n=746,566$ ). We find that these cross-mate correlations, which are derived from phenotypic measurements, explain a major portion of empirical marker-based genetic correlation estimates for the same trait pairs ( $R^2=76\%$  across datasets). We next present theoretical and

simulation-based results demonstrating that xAM biases genetic correlation estimates and yields significant estimates even among traits with uncorrelated genetic effects. We then use a simulation-based approach to evaluate the extent to which empirical level of xAM alone might plausibly explain empirical genetic correlation estimates among previously studied traits, finding that substantial fractions of empirical genetic correlation estimates are congruent with expectations for biologically independent traits subject to xAM. Lastly, we present a multivariate extension of the work of Yengo and colleagues [18], who utilized correlations between even versus odd chromosome-specific polygenic scores to detect genetic signatures of single trait AM. We find that cross-trait even/odd polygenic score correlations mirror cross-mate phenotypic correlation patterns and, through this association, can explain substantial variation in empirical genetic correlation estimates.

Together, our findings imply that xAM-induced artifacts comprise a common, previously unaddressed systematic source of bias in the genetic correlation literature. We discuss these results in the context of the widespread application of the genetic correlation as a measure of shared etiology, suggesting that many previously reported genetic correlations are overestimated.

## 2 Results

### 2.1 Genetic correlation estimates mirror cross-mate phenotypic correlations

We begin by quantifying the extent to which empirical genetic correlation estimates align with cross-trait spousal correlations across a broad array of previously studied phenotypes: a set of 20 anthropometric, metabolic, and psychosocial traits measured in the UK Biobank (UKB) [21] and a collection of five psychiatric disorder diagnoses ascertained from Danish civil registry data [22–26]. We first identified 40,697 spousal pairs within the UKB sample and randomly selected 373,283 mate pairs within the Danish sample and then estimated cross-mate correlations. For a pair of phenotypes  $Y$ ,  $Z$ , there are three cross-mate correlation parameters:  $r_{yy}$  (resp.  $r_{zz}$ ) the correlation between mates on phenotype  $Y$  (resp. phenotype  $Z$ ) and  $r_{yz}$ , the cross-mate cross-trait correlation; we generically denote these quantities  $r_{\text{mate}}$ , and present these estimates in the diagonal and sub-diagonal entries of Figures 1a and 1b. We also compiled LDSC genetic correlation estimates, denoted  $\hat{\rho}_{\beta;\text{LDSC}}$ , for each pair of phenotypes, which we present in the super-diagonal entries of Figures 1a and 1b. In the context of an ordinary least squares (OLS) regression model,  $\hat{\rho}_{\beta;\text{LDSC}}$  were strongly associated with  $r_{\text{mate}}$  estimates across both samples (Figure 1c; meta-analytic  $R^2=75.99\%$ , 95% CI: 68.60% - 83.39%). We also fit a Bayesian linear model accounting for heteroskedasticity and sampling error in  $r_{\text{mate}}$  es-

estimates, which yielded comparable results ( $R^2=76.95\%$ , 95% CI: 74.19% - 79.72%). All pairwise estimates are provided in Table S1.

## 2.2 Defining genetic correlation

Having established that a large degree of the variance in genetic correlation estimates can be predicted from phenotypic mating correlations, we now provide theoretical intuition as to why this might occur. We start by defining two distinct notions of genetic similarity between phenotypes.

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  $\beta_y, \beta_z$ :

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

Each phenotype is composed of a heritable liability component  $\ell$ , the true polygenic score, and a non-heritable component  $\varepsilon$ . 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  $h_{y;\text{pan}}^2 = \beta_y^T \beta_y$  and  $h_{z;\text{pan}}^2 = \beta_z^T \beta_z$ .

Classically, genetic correlation is defined as the correlation between the heritable components of two traits [5, 27]. We will refer to this quantity as the **score correlation**:

$$\rho_\ell = \text{cor}(\ell_y, \ell_z),$$

as it reflects the correlation between the true polygenic scores. This is distinct from the correlation between effects, which we refer to as the **effect correlation**  $\rho_\beta$ , and which indexes the similarity of variant effects on two phenotypes:

$$\rho_\beta = \text{cor}(\beta_y, \beta_z).$$

We will refer to a pair of traits as *genetically orthogonal* when their effects are uncorrelated; i.e., when  $\rho_\beta = 0$ . Variance components models aim to estimate  $\rho_\beta$  [28–30]. Within the standard random effects model framework, which assumes genotypes are independent of effect sizes,  $\rho_\ell$  and  $\rho_\beta$  are equivalent and hence seldom discussed as separate quantities.

We note that both of these measures are distinct from *pleiotropy*, which refers to the extent to which two traits are affected by the same causal variants. Pleiotropy is a necessary precondition for effect correlation, but it is possible to have independent effects with or without pleiotropy (Section 5.1.2).



Less intuitively, score correlations ( $\rho_\ell$ ) can be nonzero when the effects themselves are independent (i.e.,  $\rho_\beta=0$ ). That is, traits with uncorrelated effects can have collinear polygenic scores. Under xAM, all causal variants affecting trait  $Y$  become correlated with all causal variants affecting trait  $Z$  and these correlations are directionally consistent with their respective effects (see Supplementary Materials S1.1 for details). For example, conditional on having inherited a trait-increasing allele for phenotype  $Y$  from one parent, an individual is more (resp. less) likely to have inherited a trait-increasing (resp. -decreasing) allele for phenotype  $Z$  from the other parent when parents have positively assorted across  $Y$  and  $Z$ . As we will demonstrate in the following section, this results in a non-zero genetic correlation  $\rho_\ell$  in the direction of the cross-mate cross-trait phenotypic correlation, even for genetically orthogonal traits.

### 2.3 The impact of xAM in simulations

We ran a series of forward-time simulations using realistic genotype data (see Online Methods) to investigate the impact of xAM on multiple measures of genetic correlation. We began with a founder population of 335,550 of unrelated European ancestry UK Biobank participants' phased haploid genotypes at 1,000,000 common single nucleotide polymorphisms (SNPs). We then simulated two phenotypes,  $Y$  and  $Z$ , by randomly selecting 10,000 causal variants and drawing two corresponding sets of standardized effects from the bivariate Gaussian distribution with variances  $h_{y;\text{pan}}^2$ ,  $h_{z;\text{pan}}^2$  and correlation  $\rho_\beta$ . Non-heritable components  $\varepsilon_y$ ,  $\varepsilon_z$  were drawn independently from univariate Gaussian distributions with variances  $1 - h_{y;\text{pan}}^2$ ,  $1 - h_{z;\text{pan}}^2$ , such that both phenotypes had unit variance at panmixis.

At each generation, individuals (consisting of a set of genotypes together with two phenotypes) were matched to achieve target cross-mate correlation parameters. Throughout this section, we assume the cross-mate phenotypic correlations are non-negative and exchangeable for simplicity; i.e., that  $r_{yy}$ ,  $r_{zz}$ , and  $r_{yz}$  are all equal to a single parameter, which we denote  $r_{\text{mate}} \geq 0$ . Extension to disassortative mating, wherein  $r_{\text{mate}} < 0$ , is trivially achieved by reversing the sign of either phenotype. We also assume the two phenotypes have equal panmictic heritabilities, which we denote  $h_{\text{pan}}^2 = h_{y;\text{pan}}^2 = h_{z;\text{pan}}^2$ . We later consider general cross-mate correlation structures and differing panmictic heritabilities.

We constructed offspring genotypes by simulating meiosis according to empirically derived recombination frequencies in order to preserve local LD. Each generation's haploid genotypes were thus permutations of the founder haplotypes. We performed sensitivity analyses (see Online Methods) to confirm that our results did not depend on simulation parameters, including the number of causal variants (Figure S1), mate selection algorithm

(Figure S2), recombination scheme (Figure S3), and whether causal variants with orthogonal genetic effects were located on overlapping loci (Figure S4). Having constructed new genotypes, we generated phenotypes as above and estimated effect correlation ( $\hat{\rho}_\beta$ ) using LD score regression (LDSC, denoted  $\hat{\rho}_{\beta;\text{LDSC}}$ ; [3]), Haseman-Elston regression (HE, denoted  $\hat{\rho}_{\beta;\text{HE}}$ ; [31]), and residual maximum likelihood (REML;  $\hat{\rho}_{\beta;\text{REML}}$ ; [28, 32]). We also computed true score correlations ( $\rho_\ell$ ), which is possible when the true genetic effects are known. Additionally, we conducted supplementary simulation studies investigating the impact of xAM on GWAS effect size estimates (Section 5.6, Figure S5) and the impact of xAM on effect correlation estimates in the presence of misdiagnosis (Supplementary Materials S1.2.2, Figure S6).

### 2.3.1 xAM induces nonzero score correlations among genetically orthogonal traits

We confirmed that xAM induces substantial score correlations across a broad array of simulation parameters. This was perhaps most striking for traits with orthogonal effects: Figure 2a demonstrates the increase in the true score correlation across multiple generations of xAM for a pair of traits with  $\rho_\beta = 0$ ,  $r_{\text{mate}} = 0.5$ , and  $h_{\text{pan}}^2 = 0$ . Averaging across simulation replicates, after a single generation of xAM, the score correlation was 0.11, and after three generations of xAM, the score correlation was 0.24. The fraction of shared causal variants had no discernible impact on this behavior (Figure S4); that is, when  $\rho_\beta = 0$ , the extent to which the causal variants affecting two traits overlapped was unimportant.

Intuitively, this behavior occurs because xAM induces correlations between causal loci consistent in sign with the directions of their effects (see S1.1.1 for details). That is, pairs of trait increasing or trait decreasing loci become positively correlated whereas pairs of loci discordant with respect to the direction of their effects become negatively correlated. We can decompose the liability correlation into two disjoint sums:

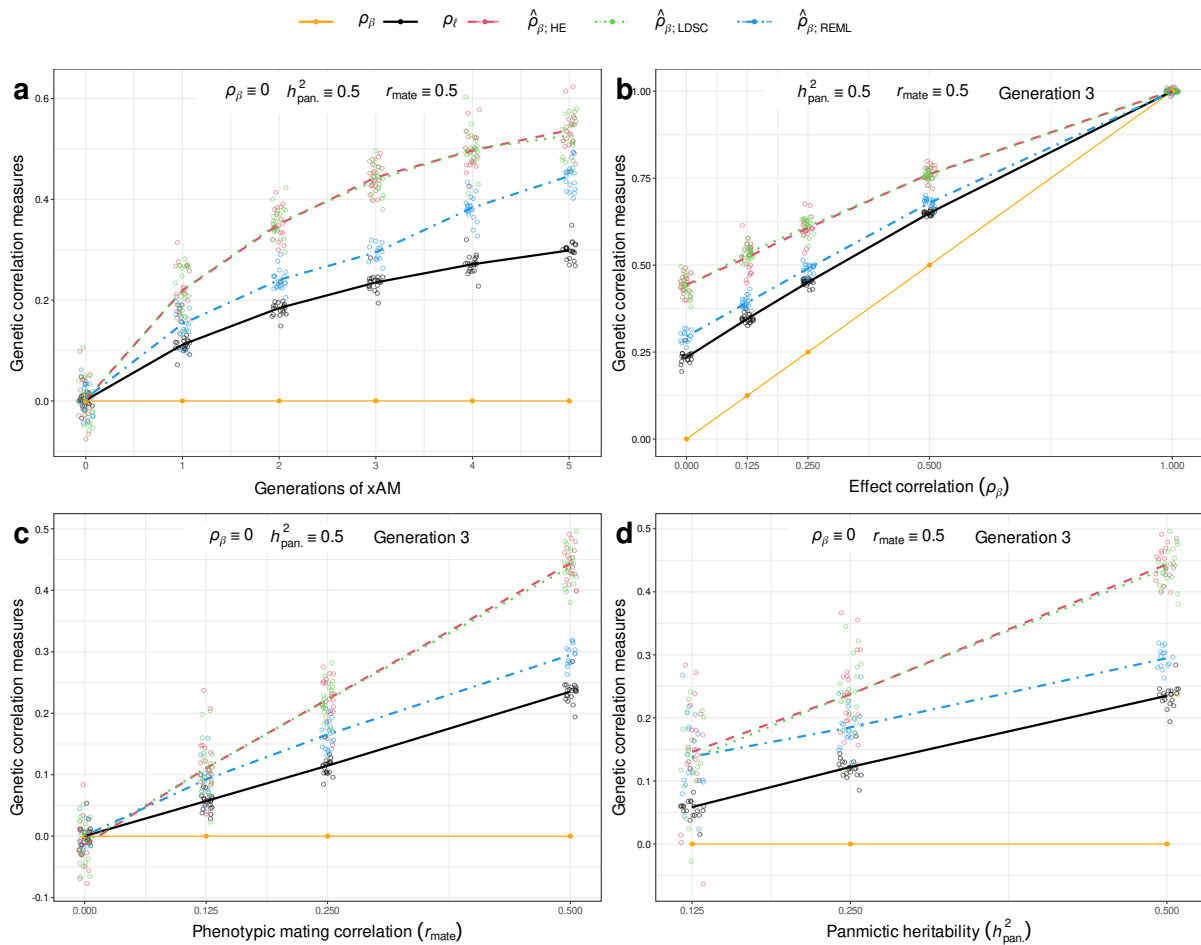
$$\text{cor}(\ell_y, \ell_z) = \sum_{i=j} \beta_{y;i} \beta_{z;j} + \sum_{i \neq j} \text{cor}(x_i, x_j) \beta_{y;i} \beta_{z;j}. \quad (2)$$

The first sum is proportional to  $\rho_\beta$  and will only be nonzero when  $\rho_\beta \neq 0$ . On the other hand, the second sum does not depend on  $\rho_\beta$ , but, due to sign-consistent global LD induced by xAM, will be greater than zero at all pairs of causal loci for  $Y$  and/or  $Z$ .

### 2.3.2 Effect correlation estimates are biased upwards

Although xAM induces nonzero score correlations between genetically orthogonal traits, these correlations are not artifactual in the sense that individuals' latent liabilities truly do become correlated. On the other hand, effect correlation estimators, which assume that

**Figure 2:** Forward-time simulation results demonstrating the impact of xAM on score correlation ( $\rho_\ell$ ) and effect correlation estimates ( $\hat{\rho}_\beta$ ) for two synthetic phenotypes with true effect correlation  $\rho_\beta$  and panmictic heritabilities  $h_{\text{pan}}^2$ . The single mating correlation parameter  $r_{\text{mate}}$  denotes the value of the two cross-mate single-trait correlations and the cross-mate cross-trait correlation. (a) xAM increases the true score correlation among two genetically orthogonal phenotypes. The HE, LDSC, and REML estimators all overestimate  $\rho_\beta$ , and the magnitude of this bias increases over subsequent generations. (b) After three generations of xAM,  $\hat{\rho}_\beta$  estimates are upwardly biased for trait pairs with true effect correlations less than one (i.e., genetically distinct phenotypes). (c) The impact of three generations of xAM increases with the phenotypic correlation among mates. (d) The impact of three generations of xAM increases monotonically with the panmictic heritabilities.



the contributions of causal haplotypes are pairwise independent (i.e.,  $\text{cov}(X_i\beta_{y;i}, X_j\beta_{z;i}) = 0$  for distinct loci  $i, j$ ) [17], are misspecified under xAM and yield upwardly biased estimates of the true effect correlation  $\rho_\beta$ . Returning to the case of genetically orthogonal traits presented in Figure 2a, after a single generation of assortative mating, the REML estimator yielded  $\hat{\rho}_\beta=0.15$  and the method-of-moments estimators, HE and LDSC, which are mathematically equivalent to each other [33, 34], yielded average estimates of  $\hat{\rho}_\beta=0.21$  and  $\hat{\rho}_\beta=0.21$ , respectively, all of which are substantially higher than the true effect correlation of zero. After three generations of xAM, this upward bias becomes more pronounced,



with REML and LDSC yielding effect correlation estimates of  $\hat{\rho}_\beta=0.30$  and  $\hat{\rho}_\beta=0.44$ , respectively.

### 2.3.3 The true score correlation and estimated effect correlation are monotonically related to $\rho_\beta$ , $r_{\text{mate}}$ , and $h_{\text{pan}}^2$

So far we have restricted our focus to the instructive case of genetically orthogonal traits, though many trait pairs subject to cross-trait xAM will truly have correlated genetic effects. Figure 2b illustrates the relationship between  $\rho_\beta$ ,  $\rho_\ell$ , and  $\hat{\rho}_\beta$  for two traits with  $h_{\text{pan}}^2 = 0.5$  and subject to three generations of xAM with  $r_{\text{mate}} = 0.5$ . Excepting the case of  $\rho_\beta = 1.0$  (a pair of genetically identical phenotypes), the pattern of results remains consistent with the genetically orthogonal case, where the correlation between effects is lower than the true score correlation, which is in turn lower than the upwardly biased  $\hat{\rho}_\beta$  estimates provided by REML, HE, and LDSC. For example, when  $\rho_\beta = 0.25$ , LDSC yields  $\hat{\rho}_\beta=0.62$  after three generations of xAM. The impact of cross-trait xAM on both  $\rho_\ell$  and  $\hat{\rho}_\beta$  is greater for traits subject to higher cross-mate correlations (i.e., stronger assortment induces more substantial bias; Figure 2c) and for traits with greater panmictic heritabilities (i.e., the genetic consequences of xAM are more substantial for more heritable phenotypes; Figure 2d).

## 2.4 xAM alone can plausibly explain substantial variance in empirical genetic correlation estimates

Having established in simulations that xAM can induce substantial  $\hat{\rho}_\beta$  estimates regardless of shared biological basis, we sought to quantify the extent to which empirical  $\hat{\rho}_\beta$  estimates for previously-studied trait pairs could be explained by xAM alone, assuming genetic orthogonality. As there are currently no unbiased estimators of  $\rho_\beta$  under xAM, we proceeded with a simulation-based approach, restricting our focus to method-of-moments estimators in light of the widespread usage of LDSC regression. We present two self-contained analyses focused on the UKB (Figure 1a) and Danish sample phenotypes (Figure 1b), respectively.

Within each sample, we first identified mate pairs and estimated the phenotypic cross-mate correlations  $r_{yy}$ ,  $r_{zz}$ , and  $r_{yz}$  for each pair of traits as described above. We then used these estimates, together with empirical heritability estimates, as inputs to a forward-time simulation where separate, non-overlapping collections of causal variants were assigned to the two phenotypes, thereby constraining the true effect correlation to zero. At each generation, we estimated the effect correlation,  $\rho_\beta$ , using method-of-moments. These projected effect correlation estimates, which we denote  $\hat{\rho}_{\text{xAM}}$ , can be interpreted as the

expected LDSC effect correlation estimate for the two traits in the complete absence of pleiotropy after a given number of generations of xAM with mating proceeding in accordance with empirical spousal correlation estimates.

Given that both marker-based and pedigree-based heritability estimators (which we denote  $\hat{h}_{\text{SNP}}^2$  and  $\hat{h}_{\text{PED}}^2$ , respectively) are biased under AM (albeit in different directions; [17, 35]) and that neither approach directly estimates the panmictic heritability, we conducted two sets of simulations, one assuming  $h_{\text{pan}}^2 = \hat{h}_{\text{SNP}}^2$  and the other assuming  $h_{\text{pan}}^2 = \hat{h}_{\text{PED}}^2$ . We focus discussion in the primary manuscript on the simulations which assumed  $h_{\text{pan}}^2 = \hat{h}_{\text{PED}}^2$  as the impact of xAM depends on the total heritability of the traits involved, not just that tagged by common variants, though results assuming  $h_{\text{pan}}^2 = \hat{h}_{\text{SNP}}^2$  are presented in the supplement.

We next compared the projected correlation estimates under xAM alone,  $\hat{\rho}_{\text{xAM}}$ , to empirical LDSC effect correlation estimates using real (rather than synthetic) phenotype data, which we denote  $\hat{\rho}_{\text{emp}}$ . To simplify discussion, we define the ratio

$$\hat{\gamma} = \hat{\rho}_{\text{xAM}} / \hat{\rho}_{\text{emp}}, \quad (3)$$

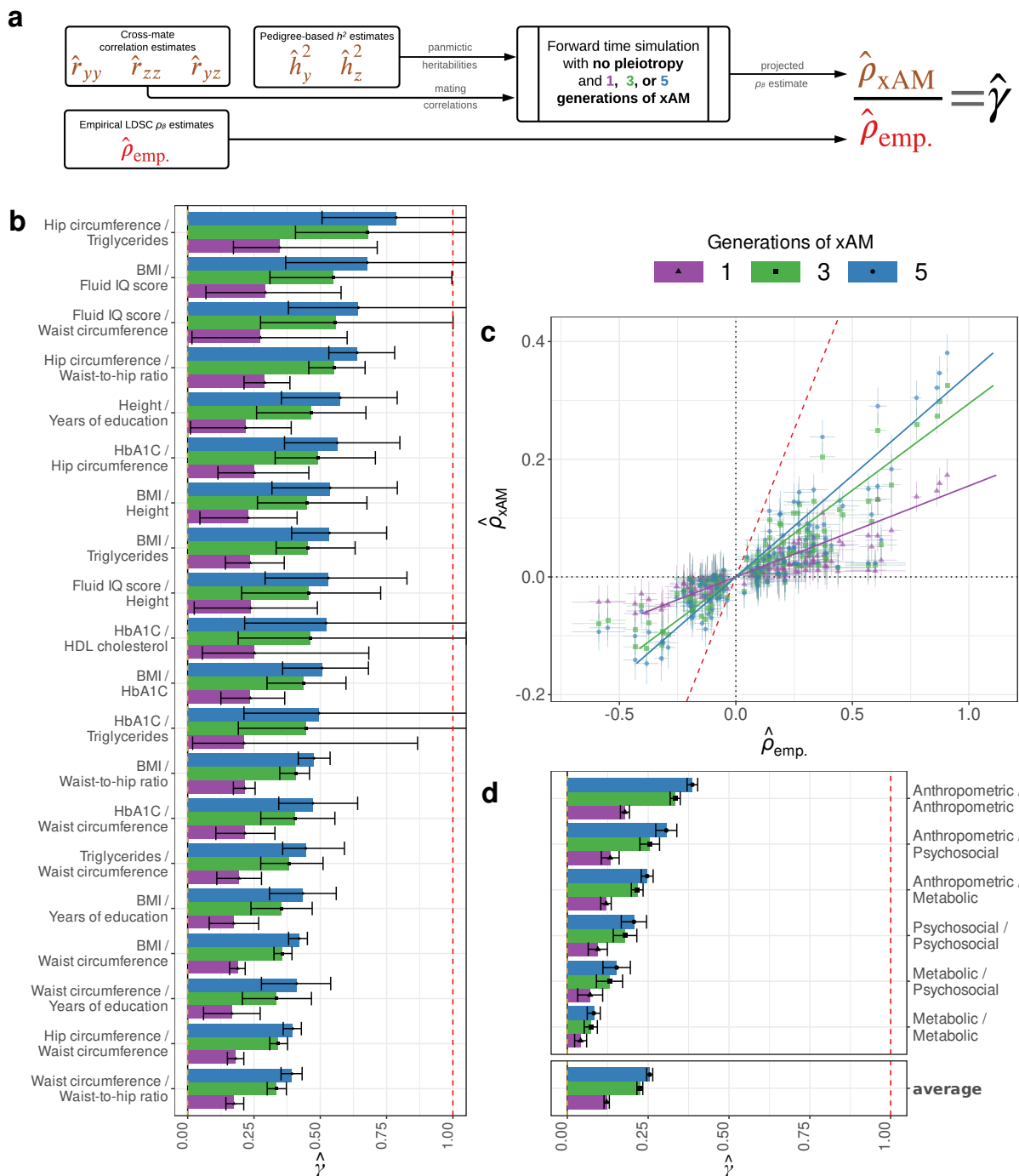
which measures the strength of the projected LDSC effect correlation estimate due to xAM-induced artifact relative to the empirical LDSC effect correlation estimate for a given pair of phenotypes. When  $\hat{\gamma}$  is close to zero, empirical effect correlation estimates are far greater than what might be attributable to xAM alone. On the other hand, when  $\hat{\gamma}$  is close to one, empirical estimates are entirely consistent with xAM-induced artifact. We caution that  $\hat{\gamma}$  does not constitute a direct estimate of the fraction of the empirical genetic correlation estimate attributable to xAM.

#### 2.4.1 Expected effect correlation estimates for UK Biobank phenotypes in the absence of pleiotropy

As we were concerned with the misinterpretation of xAM-induced artifacts for evidence of pleiotropy, we restricted our attention to 132 (of 190 possible) pairs of phenotypes with nominally significant LDSC genetic correlation estimates (Table S1). We first obtained pedigree-based heritability estimates for each of the traits of interest from the existing literature, using estimates derived in UK-based samples when available (Table S2). Together with the phenotypic mating correlations (Figure 1a), these comprised inputs for the forward time simulations used to compute  $\hat{\rho}_{\text{xAM}}$ , which we in turn compared to empirical LDSC effect correlation estimates ( $\hat{\rho}_{\text{emp}}$ ) to obtain  $\hat{\gamma}$  as illustrated in Figure 3a.

Across 132 trait pairs, 42 evidenced  $\hat{\gamma}$  values significantly greater than zero after a single generation of xAM, which increased to 74 trait pairs after three generations. Figure 3b

**Figure 3:** Comparison of empirical genetic correlations ( $\hat{\rho}_{\text{emp}}$ ) to expected estimates in the absence of pleiotropy ( $\hat{\rho}_{\text{xAM}}$ ) for assorted UK Biobank phenotypes. Error bars throughout represent 95% credible intervals. (a) For each trait pair, we ran 400 replicate forward time simulations without pleiotropy with panmictic heritabilities obtained from pedigree-based estimates and selecting mates to match empirical cross-mate correlation estimates. After each generation, we computed the method-of-moments genetic correlation estimate ( $\hat{\rho}_{\text{xAM}}$ ), which we interpret as the expected LDSC genetic correlation estimate in the absence of pleiotropy after a given number of generations of xAM. We compared this quantity to empirical LDSC genetic correlation estimates in the UK Biobank ( $\hat{\rho}_{\text{emp}}$ ) to obtain the ratio  $\hat{\gamma} := \hat{\rho}_{\text{xAM}}/\hat{\rho}_{\text{emp}}$ . (b) The top 20  $\hat{\gamma}$  estimates after up to five generations of simulated xAM across previously-studied UK Biobank phenotype pairs with nominally significant LDSC genetic correlations. (c) Projected versus empirical LDSC estimates for all UK Biobank phenotypes with nominally significant genetic correlation estimates. (b) Inverse variance weighted average  $\hat{\gamma}$  estimates within and between UK Biobank phenotypes in qualitatively similar domains.



presents the first 20 pairs in descending order of  $\hat{\gamma}$  and Figure 3c presents the raw projected and empirical effect correlation estimates across all 132 pairs (see Table S3, Figures S7 to S9 for detailed results spanning one to five generations of xAM). Averaging across all trait pairs (including those not significantly different from zero), the inverse variance weighted average  $\hat{\gamma}$  estimate was 0.25 ( $se=0.005$ ). The relative strengths of within- and cross-domain projections varied substantially when aggregating across qualitative clusters of phenotypes (see Figure 1a and Table S2 for domain definitions). For example, relatively small fractions of empirical  $\hat{\rho}_\beta$  estimates among metabolic phenotypes were congruent with expectations under five generations of xAM alone ( $\hat{\gamma}=0.08$ ,  $se=0.01$ ) in comparison to those between psychosocial phenotypes ( $\hat{\gamma}=0.21$ ,  $se=0.01$ ). We note that these domain-averages are highly dependent on the set of traits considered and should be interpreted with care.

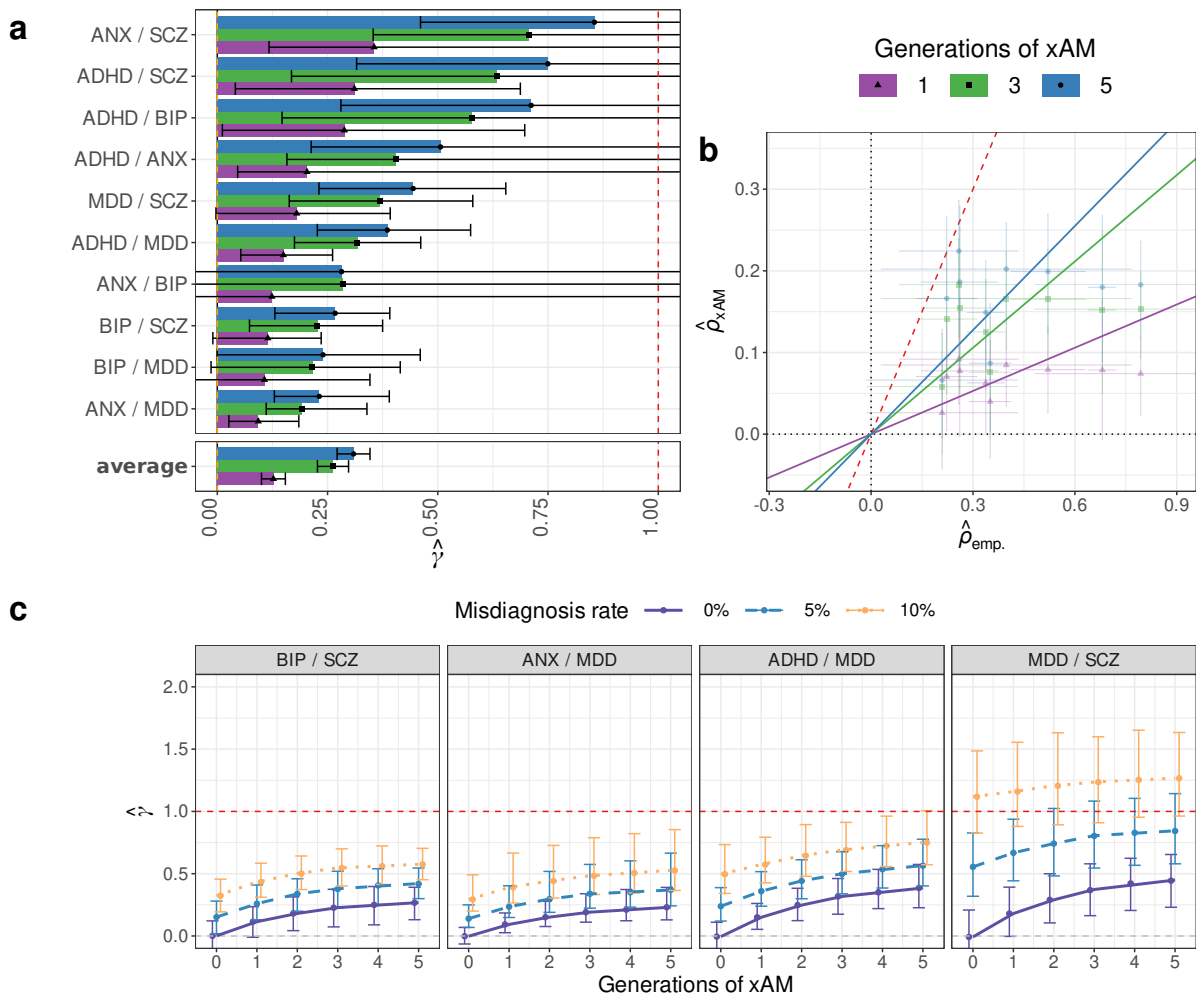
#### 2.4.2 Expected effect correlation estimates among psychiatric disorders in the absence of pleiotropy

We next computed projected effect correlation estimates assuming up to five generations of xAM under empirically derived cross-mate correlation structures for a collection of five psychiatric disorders, as estimated in sample of 373,283 spousal pairs randomly selected from the Danish population (see S1.3 for details). Again, we obtained pedigree-based heritabilities for each disorder, preferring studies using samples from Scandinavian countries when available (Table S4), and used these estimates, together with the cross-mate-correlations, to compute  $\hat{\rho}_{xAM}$ . We then compared these projections to the LDSC genetic correlation estimates presented by Anttila and colleagues [2]. We focused on Anttila et al.'s manuscript as their interpretation of the genetic correlation as a direct measure of biological overlap is broadly representative of the views of many complex trait geneticists [3–5, 7].

Across all pairwise combinations of disorders, we observed an average ratio of  $\hat{\gamma} = 0.309$  ( $se=0.037$ ; Figures 4a and 4b) after five generations of xAM. Some trait pairs evidenced considerably greater empirical effect correlation estimates than might be explained by xAM alone (e.g., for anxiety disorders and major depression,  $\hat{\gamma}=0.231$ , 95% CI: 0.129 - 0.390), whereas for other pairs this discrepancy was modest (e.g., for anxiety disorders and Schizophrenia,  $\hat{\gamma}=0.855$ , 95% CI: 0.461 - 2.583). We present detailed results across one to five generations of xAM in Table S5 and Figures S10 and S11.

**xAM exacerbates the impact of misdiagnosis on effect correlation estimates among psychiatric disorders** Given the combined effects of xAM and measurement error in terms of inflating effect correlation estimates (see S1.2.2), we next extended our method

**Figure 4:** Comparison of empirical effect correlation estimates among psychiatric phenotypes to expectations in the absence of pleiotropy. (a) Ratios ( $\hat{\gamma} := \hat{\rho}_{xAM} / \hat{\rho}_{emp}$ ) of projected LDSC effect correlation estimate ( $\hat{\rho}_{xAM}$ ) assuming xAM and no pleiotropy relative to the empirical effect correlation estimates ( $\hat{\rho}_{emp}$ ) of [2] for five psychiatric disorders with 95% credible intervals. Projections assume up to five generations of xAM matching empirical cross-mate correlations in the Danish sample and assume panmictic heritabilities equal to pedigree-based estimates from demographically-comparable studies (Table S4). (b) Projected versus empirical LDSC estimates across psychiatric phenotype pairs. (c) The potential joint impacts of bidirectional errors in diagnosis and xAM on effect correlation estimates for selected psychiatric disorder pairs. The red dashed line corresponds to  $\hat{\gamma} = 1$  across all panes.



for deriving  $\hat{\rho}_{xAM}$  to incorporate errors in diagnosis and applied this extension to each pair of psychiatric disorders. At each generation, we dichotomized continuous phenotypes using empirical prevalence estimates (Table S4) and introducing misdiagnosis errors prior to estimating the effect correlation. We considered bidirectional misclassification errors (one disorder might be mistaken for another and vice-versa) ranging from 0% to 10% and up to three generations of xAM in accordance with empirical spousal-correlation estimates (Figure S11). Results demonstrated that some of the empirical effect correlation estimates presented by [2] were entirely consistent with artifacts induced by xAM coupled with moderate diagnostic error rates for some trait pairs but not others. Whereas high

(i.e.,  $\geq 10\%$ ) rates of misdiagnostic errors are necessary to fully account observed effect correlations between ADHD and anxiety disorders in the absence of xAM, five generations of xAM coupled with a 5% misdiagnosis rate would account for nearly all of the empirical genetic correlation estimate ( $\hat{\gamma}=0.808$ , 95% CI: 0.363 – 5.106). On the other hand, the same extent of xAM and misdiagnosis would account for a considerably smaller fraction of the empirical genetic correlation estimate between anxiety disorders and major depression ( $\hat{\gamma}=0.370$ , 95% CI: 0.242 – 0.665). Figure 4c highlights the potential impacts of xAM and diagnostic errors on four selected trait pairs.

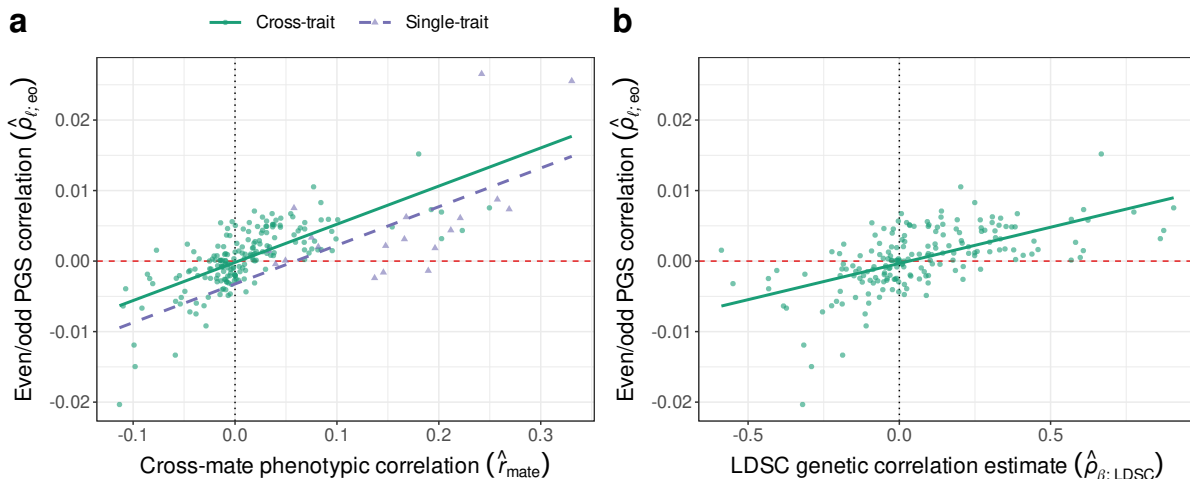
## 2.5 Genetic evidence for xAM recapitulates empirical cross-mate correlations

We have thus far established that one or more generations of xAM consistent with empirical cross-mate correlation estimates could plausibly induce substantial bias in genetic correlation estimates for many previously studied traits. We now assess evidence for a history of phenotypically-mediated xAM at the genetic level. Previous work has used even/odd chromosome polygenic score correlations to examine evidence for AM with respect to a single trait [18]. We adapted this approach to the bivariate case by computing *cross-trait* even/odd polygenic score correlations ( $\hat{\rho}_{\ell;eo}$ ; Table S6). In the absence of xAM or other sources of population structure, polygenic scores computed on even chromosomes for one trait versus those computed on odd chromosomes for a second trait are expected to be orthogonal. On the other hand, the long-range sign-consistent LD induced by xAM will induce correlations between polygenic scores on separate chromosomes (see Supplemental Materials S1.1.3). Further, these correlations should positively associate with cross-mate phenotypic correlations.

In order to compute cross-chromosome polygenic score correlations, we first computed GWAS effect estimates and clumped results in a training subsample of 80% of unrelated European ancestry UK Biobank participants and then computed polygenic scores specific to even and odd chromosomes for the remaining 20% using multiple  $p$ -value thresholds. We then regressed the estimated even/odd polygenic score correlations on the corresponding cross-mate cross-trait correlations (or cross-mate same-trait correlations in the case of univariate even/odd correlations; Section 5.9). Finally, we regressed the estimated cross-trait even/odd polygenic score correlations on the corresponding LDSC effect-correlation estimates.

In an OLS model using scores built with all clumped variants, cross-mate phenotypic correlation estimates ( $r_{\text{mate}}$ ) explained substantial variance in the cross-chromosome even/odd polygenic score correlations ( $\hat{\rho}_{\ell;eo}$ ;  $R^2=47.66\%$ ,  $se=0.060$ ; Figure 5a). As before, account-

**Figure 5:** Genetic level evidence consistent with xAM in the UK Biobank. (a) Correlation between even and odd chromosome-specific polygenic scores as a function of the cross-mate phenotypic correlation. For a single trait, the vertical axis reflects the correlation between even and odd chromosome polygenic scores  $\hat{\ell}_{\text{even}}$ ,  $\hat{\ell}_{\text{odd}}$  and the horizontal axis reflects the cross-mate correlation. For a pair of traits  $Y$ ,  $Z$ , the vertical axis reflects a single parameter to which the correlations between  $\hat{\ell}_{y;\text{even}}$ ,  $\hat{\ell}_{z;\text{odd}}$  and between  $\hat{\ell}_{y;\text{odd}}$ ,  $\hat{\ell}_{z;\text{even}}$  are both constrained, and the horizontal axis reflects the cross-mate cross-trait correlation. (b) Cross-trait even/odd polygenic score correlations as a function of empirical LDSC effect correlation estimate. The positive association of these quantities is consistent with our hypothesis that empirical effect correlations capture genetic structure in addition to signals of biological overlap.



ing for measurement error and heteroskedasticity in a Bayesian linear model yielded nearly identical results. This association persisted across polygenic score  $p$ -value thresholds ranging from  $3.00e-3$  to 1.00; Figure S12).

Additionally, cross-trait even/odd chromosome polygenic score correlations were positively associated with empirical LDSC effect correlation estimates ( $\hat{\rho}_{\beta;LDSC}$ ;  $R^2=34.81\%$ ,  $se=0.045$ ; Figure 5b). This is consistent with the hypothesis that empirical effect correlation estimates are capturing additional structure beyond the signatures of biological overlap. Further, regressing  $\hat{\rho}_{\beta;LDSC}$  on  $\hat{\rho}_{\ell;eo}$  and  $\hat{r}_{yz}$  simultaneously revealed that the association between  $\hat{\rho}_{\beta;LDSC}$  and  $\hat{\rho}_{\ell;eo}$  is mediated via  $\hat{r}_{yz}$  ( $\Delta R^2 < 0.001$ ; partial effect  $p=0.48$  for  $\hat{\rho}_{\ell;eo}$  versus  $p < 5e-8$  for  $\hat{r}_{yz}$ ). Thus, any alternative sources of population structure beyond xAM captured by  $\hat{\rho}_{\ell;eo}$  do not appear to explain the positive association between  $\hat{\rho}_{\ell;eo}$  and  $\hat{\rho}_{\beta;LDSC}$ .

### 3 Discussion

Nonzero effect correlation estimates have been widely interpreted as evidence that two traits share overlapping genetic bases. It is therefore surprising that substantial variation in genetic correlation estimates can be explained by cross-mate phenotypic correlations

(Figure 1c). Further, given the strength of this association, the consequences of the random mating assumption implicit in all commonly used genetic correlation estimators [3, 8, 28, 31, 36] warrant critical attention.

We have demonstrated that xAM increases the score correlation ( $\rho_\ell$ ) between a pair of traits by inducing long-range sign-consistent LD across all pairs of causal variants, even when the causal effects themselves are uncorrelated ( $\rho_\beta = 0$ ). Marker-based genetic correlation estimators, however, implicitly assume that  $\rho_\ell = \rho_\beta$ , and under this misspecification substantially overestimate  $\rho_\beta$  after even a single generation of xAM (Figure 2a). In this context, we sought to quantify this potential source of bias across a broad array of phenotypes by compiling the largest to-date collection of empirical cross-mate cross-trait correlation estimates and using these to anticipate  $\hat{\rho}_{\text{xAM}}$ , the projected effect correlation estimate purely due to xAM-induced artifact. Across numerous pairs of nominally genetically correlated traits, xAM alone could plausibly produce substantial  $\hat{\rho}_\beta$  values relative to empirical LDSC estimates, though the relative magnitude of such effects varied considerably across phenotype pairs. Additionally, we provided evidence congruent with a history of xAM at the genetic level by demonstrating that cross-trait even/odd chromosome polygenic score correlations mirror phenotypic cross-mate correlations. Moreover, we showed that cross-trait even/odd chromosome polygenic score correlations explain substantial variation in genetic correlation estimates through their association with cross-mate phenotypic correlations, consistent with the hypothesis that changes to global LD structure induced by previous generations of xAM have resulted in biased genetic correlation estimates.

Taken together, our results show that assortment across phenotypes will bias effect correlation estimates, that cross-mate cross-trait phenotypic correlations among many pairs of phenotypes are strong enough that one or more generations of assortment would substantially inflate effect correlation estimates, and that the joint distribution of causal variants and their estimated effects coincides with what we would expect after one or more generations of xAM. We therefore conclude that xAM provides a previously ignored source of systematic bias in the study of genetic similarity across complex traits, one that precludes the interpretation of the effect correlation estimate as a direct index of biological similarity.

This is perhaps most salient in scientific domains where effect correlation estimates may be interpreted through a clinical lens. In order to better understand the potential consequences of xAM-naïve approaches, we reexamined the results Anttila and colleagues [2], who, in chorus with others [7], interpret substantial effect correlation estimates among psychiatric disorders as evidence of shared pathogenesis that contradicts existing disease classifications. Though our results are agnostic with respect to the adequacy of



psychiatric nosology, they illustrate that effect correlation estimates alone do not imply shared pathogenesis and suggest that a substantial fraction of the effect correlation estimates among many pairs of psychiatric traits are largely consistent with xAM alone. Further, modest levels of diagnostic error, when coupled with xAM, can fully account for some previously published estimates (Figures 4c and S11). Reassuringly, our results also demonstrate that particular trait pairs with high rates of comorbidity or characterized by similar symptomatology (i.e., anxiety disorders/major depression and bipolar disorders/Schizophrenia; [26, 37]) exhibit effect correlation estimates greater than what we could plausibly attribute to xAM and diagnostic error. Though it is still possible that such estimates are biased upwards to some extent by either or both factors, the general qualitative interpretation that these disorders share overlapping genetic bases appears robust.

Our results also complicate the interpretation of a number of multivariate analytic frameworks beyond effect correlation estimators. For example, genomic structural equation modeling [6], which takes marker-based genetic correlation estimates as inputs, will propagate xAM induced biases. The consequences for methods such as MTAG [38] are not obvious as xAM will not necessarily decrease prediction  $R^2$  (e.g., see Figure S5b). That is, the apparent gain in predictive performance of multi-phenotype polygenic score aggregates could arise from the global patterns sign-consistent LD induced by xAM. Broadly speaking, our findings mirror recent results regarding the potential impacts of assortative mating across other areas of statistical genetics, including marker-based heritability estimation [17], and Mendelian randomization [19, 39].

There are a number of limitations to the current investigation. Foremost among these are the constellation of assumptions required to predict the expected effect correlation estimate under xAM alone ( $\hat{\rho}_{\text{xAM}}$ ), which here reflect a particular model of assortment (see Supplemental Materials S1.1 for further details). Even in the simplified context of two additive traits with uncorrelated effects, anticipating  $\hat{\rho}_{\beta}$  values under xAM alone necessitates numerous unconfirmed assumptions about the dynamics of a particular population, including but not limited to: the panmictic heritabilities, the stability of cross-mate phenotypic correlation structures over succeeding generations, and the stability and transmissibility of environmental factors, each of which comprises an active research area [17, 20, 40]. We proceeded under a relatively tractable dynamical framework assuming a fixed number of generations of assortment on two additive phenotypes with constant cross-mate correlations, stable non-heritable sources of variation, and no vertical transmission. Though each of these assumptions is likely to prove untenable for particular trait pairs and thereby compromise the accuracy of our projections, we hypothesize that the general qualitative phenomenon whereby xAM inflates effect correlation estimates is likely to persist for a number of traits. Nonetheless, we caution that these projections

are contingent upon a number of consequential decisions, noting that all existing methods are only able to sidestep these decisions by making the strong (and likely incorrect) assumption that mating is random.

Separate from this collection of necessary modeling decisions is the more elemental assumption of *primary phenotypic assortment*—that mate selection is mediated through both the heritable and non-heritable factors underlying a pair of phenotypes. This assumption represents a middle ground between the two extreme scenarios of *genetic homogamy*, in which mates assort exclusively on genotype, and *social homogamy*, where mates assort exclusively on environmental factors [41]. Critically, xAM under pure social homogamy, which is conceptually equivalent to primary phenotypic assortment across 0% heritable phenotypes, would have no impact on either the score correlation or effect correlation estimator (this can be intuited by extrapolating toward  $h_{\text{pan}}^2 = 0$  in Figure 2d). However, if pure social homogamy were widespread, we would not expect to observe any association between cross-trait even/odd chromosome polygenic score correlations and cross-mate cross-trait correlations, which was not the case empirically (Figure 5a). To the extent that the assumption of primary phenotypic assortment is violated, it is most likely to be a matter of degree (i.e., heritable and non-heritable phenotype components may be weighted disproportionately but both impact mate choice to some extent), and can be conceived of as another modeling procedure unlikely to dramatically alter our qualitative conclusions. Still, constructing a generative model that reconciles the association between empirical mating patterns and genetic correlation estimates is a ill-posed inverse problem for which there are multiple solutions, of which we have only explored one.

Finally, though we demonstrate that empirical levels of xAM are substantial enough to have induced systemic bias in the genetic correlation literature, we have not explicitly estimated the extent to which empirical genetic correlation estimates reflect biological versus structural factors. Thus, the development of quantitative methods that jointly model causal loci and their effects (i.e., do not assume  $\rho_\ell$  and  $\rho_\beta$  are equivalent) comprises a prime target for future research. We further remark that xAM is, in essence, a form of population structure, and suggest that the development of such methods may provide insight into residual sources of variation not captured by conventional principal component or mixed-model based correction. Given the increasing evidence that existing methods fail to completely address structural factors, even in ostensibly ancestrally homogenous groups [42, 43], a broader characterization of population structure and new methods for addressing such structure will like be necessary to generate results that are maximally clinically relevant and can be applied equitably.

## 4 References

- [1] S Hong Lee et al. “Genetic Relationship between Five Psychiatric Disorders Estimated from Genome-Wide SNPs”. In: *Nat Genet* 45.9 (9 Sept. 2013), pp. 984–994. ISSN: 1546-1718. DOI: [10.1038/ng.2711](https://doi.org/10.1038/ng.2711). URL: <https://www.nature.com/articles/ng.2711> (visited on 10/29/2021).
- [2] Verner Anttila et al. “Analysis of Shared Heritability in Common Disorders of the Brain”. In: *Science* 360.6395 (2018).
- [3] Brendan Bulik-Sullivan et al. “An Atlas of Genetic Correlations across Human Diseases and Traits”. In: *Nature Genetics* 47.11 (Nov. 2015), pp. 1236–1241. ISSN: 1546-1718. DOI: [10.1038/ng.3406](https://doi.org/10.1038/ng.3406). URL: <https://www.nature.com/articles/ng.3406> (visited on 07/25/2018).
- [4] Jie Zheng et al. “LD Hub: A Centralized Database and Web Interface to Perform LD Score Regression That Maximizes the Potential of Summary Level GWAS Data for SNP Heritability and Genetic Correlation Analysis”. In: *Bioinformatics* 33.2 (Jan. 15, 2017), pp. 272–279. ISSN: 1367-4803. DOI: [10.1093/bioinformatics/btw613](https://doi.org/10.1093/bioinformatics/btw613). pmid: 27663502. URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5542030/> (visited on 06/10/2018).
- [5] Wouter van Rheenen et al. “Genetic Correlations of Polygenic Disease Traits: From Theory to Practice”. In: *Nat Rev Genet* 20.10 (10 Oct. 2019), pp. 567–581. ISSN: 1471-0064. DOI: [10.1038/s41576-019-0137-z](https://doi.org/10.1038/s41576-019-0137-z). URL: <https://www.nature.com/articles/s41576-019-0137-z> (visited on 08/17/2021).
- [6] Andrew D. Grotzinger et al. “Genomic Structural Equation Modelling Provides Insights into the Multivariate Genetic Architecture of Complex Traits”. In: *Nat Hum Behav* 3.5 (5 May 2019), pp. 513–525. ISSN: 2397-3374. DOI: [10.1038/s41562-019-0566-x](https://doi.org/10.1038/s41562-019-0566-x). URL: <https://www.nature.com/articles/s41562-019-0566-x> (visited on 08/26/2021).
- [7] Phil H. Lee et al. “Genomic Relationships, Novel Loci, and Pleiotropic Mechanisms across Eight Psychiatric Disorders”. In: *Cell* 179.7 (Dec. 12, 2019), 1469–1482.e11. ISSN: 0092-8674. DOI: [10.1016/j.cell.2019.11.020](https://doi.org/10.1016/j.cell.2019.11.020). URL: <https://www.sciencedirect.com/science/article/pii/S0092867419312760> (visited on 08/26/2021).
- [8] Omer Weissbrod, Jonathan Flint, and Saharon Rosset. “Estimating SNP-Based Heritability and Genetic Correlation in Case-Control Studies Directly and with Summary Statistics”. In: *The American Journal of Human Genetics* 103.1 (July 5, 2018), pp. 89–99. ISSN: 0002-9297, 1537-6605. DOI: [10.1016/j.ajhg.2018.06.002](https://doi.org/10.1016/j.ajhg.2018.06.002). pmid: 29979983. URL: [https://www.cell.com/ajhg/abstract/S0002-9297\(18\)30195-2](https://www.cell.com/ajhg/abstract/S0002-9297(18)30195-2) (visited on 05/26/2019).

- [9] Doug Speed and David J. Balding. “SumHer Better Estimates the SNP Heritability of Complex Traits from Summary Statistics”. In: *Nat Genet* 51.2 (2 Feb. 2019), pp. 277–284. ISSN: 1546-1718. DOI: [10.1038/s41588-018-0279-5](https://doi.org/10.1038/s41588-018-0279-5). URL: <https://www.nature.com/articles/s41588-018-0279-5> (visited on 08/02/2021).
- [10] Phil H. Lee, Yen-Chen A. Feng, and Jordan W. Smoller. “Pleiotropy and Cross-Disorder Genetics Among Psychiatric Disorders”. In: *Biological Psychiatry*. Disentangling Psychiatric Polygenicity 89.1 (Jan. 1, 2021), pp. 20–31. ISSN: 0006-3223. DOI: [10.1016/j.biopsych.2020.09.026](https://doi.org/10.1016/j.biopsych.2020.09.026). URL: <https://www.sciencedirect.com/science/article/pii/S0006322320319879> (visited on 11/22/2021).
- [11] Riccardo E. Marioni et al. “Molecular Genetic Contributions to Socioeconomic Status and Intelligence”. In: *Intelligence* 44 (May 1, 2014), pp. 26–32. ISSN: 0160-2896. DOI: [10.1016/j.intell.2014.02.006](https://doi.org/10.1016/j.intell.2014.02.006). URL: <https://www.sciencedirect.com/science/article/pii/S0160289614000178> (visited on 11/22/2021).
- [12] Andries T. Marees et al. “Genetic Correlates of Socio-Economic Status Influence the Pattern of Shared Heritability across Mental Health Traits”. In: *Nat Hum Behav* 5.8 (8 Aug. 2021), pp. 1065–1073. ISSN: 2397-3374. DOI: [10.1038/s41562-021-01053-4](https://doi.org/10.1038/s41562-021-01053-4). URL: <https://www.nature.com/articles/s41562-021-01053-4> (visited on 11/22/2021).
- [13] Niek Verweij et al. “Identification of 15 Novel Risk Loci for Coronary Artery Disease and Genetic Risk of Recurrent Events, Atrial Fibrillation and Heart Failure”. In: *Sci Rep* 7.1 (1 June 5, 2017), p. 2761. ISSN: 2045-2322. DOI: [10.1038/s41598-017-03062-8](https://doi.org/10.1038/s41598-017-03062-8). URL: <https://www.nature.com/articles/s41598-017-03062-8> (visited on 11/22/2021).
- [14] Yeda Wu et al. “Genome-Wide Association Study of Medication-Use and Associated Disease in the UK Biobank”. In: *Nat Commun* 10.1 (1 Apr. 23, 2019), p. 1891. ISSN: 2041-1723. DOI: [10.1038/s41467-019-09572-5](https://doi.org/10.1038/s41467-019-09572-5). URL: <https://www.nature.com/articles/s41467-019-09572-5> (visited on 11/22/2021).
- [15] Peter M. Visscher and Jian Yang. “A Plethora of Pleiotropy across Complex Traits”. In: *Nat Genet* 48.7 (7 July 2016), pp. 707–708. ISSN: 1546-1718. DOI: [10.1038/ng.3604](https://doi.org/10.1038/ng.3604). URL: <https://www.nature.com/articles/ng.3604> (visited on 11/23/2021).
- [16] Naomi R Wray, Sang Hong Lee, and Kenneth S Kendler. “Impact of Diagnostic Misclassification on Estimation of Genetic Correlations Using Genome-Wide Genotypes”. In: *Eur J Hum Genet* 20.6 (June 2012), pp. 668–674. ISSN: 1018-4813. DOI: [10.1038/ejhg.2011.257](https://doi.org/10.1038/ejhg.2011.257). pmid: [22258521](https://pubmed.ncbi.nlm.nih.gov/22258521/). URL: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3355255/> (visited on 08/04/2017).

- [17] Richard Border et al. “Assortative Mating Biases Marker-based Heritability Estimators”. In: (Mar. 20, 2021), p. 2021.03.18.436091. DOI: [10.1101/2021.03.18.436091](https://doi.org/10.1101/2021.03.18.436091). URL: <https://www.biorxiv.org/content/10.1101/2021.03.18.436091v1> (visited on 08/17/2021).
- [18] Loic Yengo et al. “Imprint of Assortative Mating on the Human Genome”. In: *Nature human behaviour* 2.12 (2018), pp. 948–954.
- [19] Fernando Pires Hartwig, Neil Martin Davies, and George Davey Smith. “Bias in Mendelian Randomization Due to Assortative Mating”. In: *Genetic Epidemiology* 42.7 (2018), pp. 608–620. ISSN: 1098-2272. DOI: [10.1002/gepi.22138](https://doi.org/10.1002/gepi.22138). URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/gepi.22138> (visited on 08/04/2021).
- [20] Augustine Kong et al. “The Nature of Nurture: Effects of Parental Genotypes”. In: *Science* 359.6374 (Jan. 26, 2018), pp. 424–428. ISSN: 0036-8075, 1095-9203. DOI: [10.1126/science.aan6877](https://doi.org/10.1126/science.aan6877). pmid: [29371463](https://pubmed.ncbi.nlm.nih.gov/29371463/). URL: <https://science.sciencemag.org/content/359/6374/424> (visited on 05/16/2019).
- [21] Clare Bycroft et al. “The UK Biobank Resource with Deep Phenotyping and Genomic Data”. In: *Nature* 562.7726 (2018), pp. 203–209.
- [22] Ole Mors, Gurli P. Perto, and Preben Bo Mortensen. “The Danish Psychiatric Central Research Register”. In: *Scandinavian journal of public health* 39 (7\_suppl 2011), pp. 54–57.
- [23] Carsten B. Pedersen et al. “The Danish Civil Registration System”. In: *Dan Med Bull* 53.4 (2006), pp. 441–449.
- [24] Carsten Bøcker Pedersen. “The Danish Civil Registration System”. In: *Scandinavian journal of public health* 39 (7\_suppl 2011), pp. 22–25.
- [25] Morten Schmidt et al. “The Danish National Patient Registry: A Review of Content, Data Quality, and Research Potential”. In: *Clinical epidemiology* 7 (2015), p. 449.
- [26] Oleguer Plana-Ripoll et al. “Exploring Comorbidity within Mental Disorders among a Danish National Population”. In: *JAMA psychiatry* 76.3 (2019), pp. 259–270.
- [27] Douglas Scott Falconer. *Introduction to Quantitative Genetics*. Oliver And Boyd; Edinburgh; London, 1960.
- [28] Jian Yang et al. “GCTA: A Tool for Genome-wide Complex Trait Analysis”. In: *Am J Hum Genet* 88.1 (Jan. 7, 2011), pp. 76–82. ISSN: 0002-9297. DOI: [10.1016/j.ajhg.2010.11.011](https://doi.org/10.1016/j.ajhg.2010.11.011). pmid: [21167468](https://pubmed.ncbi.nlm.nih.gov/21167468/). URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3014363/> (visited on 01/17/2019).
- [29] Po-Ru Loh et al. “Mixed-Model Association for Biobank-Scale Datasets”. In: *Nature genetics* (2018), p. 1.

- [30] Brendan K. Bulik-Sullivan et al. “LD Score Regression Distinguishes Confounding from Polygenicity in Genome-Wide Association Studies”. In: *Nature Genetics* 47.3 (Mar. 2015), pp. 291–295. ISSN: 1546-1718. DOI: [10.1038/ng.3211](https://doi.org/10.1038/ng.3211). URL: <https://www.nature.com/articles/ng.3211> (visited on 07/25/2018).
- [31] J. K. Haseman and R. C. Elston. “The Investigation of Linkage between a Quantitative Trait and a Marker Locus”. In: *Behav. Genet.* 2.1 (Mar. 1972), pp. 3–19. ISSN: 0001-8244. pmid: [4157472](https://pubmed.ncbi.nlm.nih.gov/4157472/).
- [32] H. D. Patterson and R. Thompson. “Recovery of Inter-Block Information When Block Sizes Are Unequal”. In: *Biometrika* 58.3 (Dec. 1971), p. 545. ISSN: 00063444. DOI: [10.2307/2334389](https://doi.org/10.2307/2334389). JSTOR: [2334389](https://www.jstor.org/stable/2334389).
- [33] Brendan Bulik-Sullivan. *Relationship between LD Score and Haseman-Elston Regression*. preprint. Genetics, Apr. 20, 2015. DOI: [10.1101/018283](https://doi.org/10.1101/018283). URL: <http://biorxiv.org/lookup/doi/10.1101/018283> (visited on 05/16/2019).
- [34] Xiang Zhou. “A Unified Framework for Variance Component Estimation with Summary Statistics in Genome-Wide Association Studies”. In: *The annals of applied statistics* 11.4 (2017), p. 2027.
- [35] Lindon Eaves. “The Use of Twins in the Analysis of Assortative Mating”. In: *Heredity* 43.3 (3 Dec. 1979), pp. 399–409. ISSN: 1365-2540. DOI: [10.1038/hdy.1979.90](https://doi.org/10.1038/hdy.1979.90). URL: <https://www.nature.com/articles/hdy197990> (visited on 08/07/2020).
- [36] Doug Speed et al. “Reevaluation of SNP Heritability in Complex Human Traits”. In: *Nature Genetics* 49.7 (7 July 2017), pp. 986–992. ISSN: 1546-1718. DOI: [10.1038/ng.3865](https://doi.org/10.1038/ng.3865). URL: <https://www.nature.com/articles/ng.3865> (visited on 04/05/2021).
- [37] Ronald C. Kessler et al. “Lifetime Prevalence and Age-of-Onset Distributions of DSM-IV Disorders in the National Comorbidity Survey Replication”. In: *Arch Gen Psychiatry* 62.6 (June 1, 2005), pp. 593–602. ISSN: 0003-990X. DOI: [10.1001/archpsyc.62.6.593](https://doi.org/10.1001/archpsyc.62.6.593). URL: <https://jamanetwork.com/journals/jamapsychiatry/fullarticle/208678> (visited on 07/25/2018).
- [38] Patrick Turley et al. “Multi-Trait Analysis of Genome-Wide Association Summary Statistics Using MTAG”. In: *Nature genetics* 50.2 (2018), pp. 229–237.
- [39] Ben Brumpton et al. “Avoiding Dynastic, Assortative Mating, and Population Stratification Biases in Mendelian Randomization through within-Family Analyses”. In: *Nature communications* 11.1 (1 2020), pp. 1–13.
- [40] Alexander I. Young et al. “Mendelian Imputation of Parental Genotypes for Genome-Wide Estimation of Direct and Indirect Genetic Effects”. In: (July 3, 2020). DOI: [10.1101/2020.07.02.185199](https://doi.org/10.1101/2020.07.02.185199). URL: <https://www.biorxiv.org/content/10.1101/2020.07.02.185199v1> (visited on 08/17/2021).

- [41] Matthew C. Keller, Sarah E. Medland, and Laramie E. Duncan. “Are Extended Twin Family Designs Worth the Trouble? A Comparison of the Bias, Precision, and Accuracy of Parameters Estimated in Four Twin Family Models”. In: *Behav Genet* 40.3 (May 1, 2010), pp. 377–393. ISSN: 1573-3297. DOI: [10.1007/s10519-009-9320-x](https://doi.org/10.1007/s10519-009-9320-x). URL: <https://doi.org/10.1007/s10519-009-9320-x> (visited on 08/27/2021).
- [42] Jeremy J Berg et al. “Reduced Signal for Polygenic Adaptation of Height in UK Biobank”. In: *eLife* 8 (Mar. 21, 2019). Ed. by Magnus Nordborg et al., e39725. ISSN: 2050-084X. DOI: [10.7554/eLife.39725](https://doi.org/10.7554/eLife.39725). URL: <https://doi.org/10.7554/eLife.39725> (visited on 08/27/2021).
- [43] Simon Haworth et al. “Apparent Latent Structure within the UK Biobank Sample Has Implications for Epidemiological Analysis”. In: *Nat Commun* 10.1 (1 Jan. 18, 2019), p. 333. ISSN: 2041-1723. DOI: [10.1038/s41467-018-08219-1](https://doi.org/10.1038/s41467-018-08219-1). URL: <https://www.nature.com/articles/s41467-018-08219-1> (visited on 08/27/2021).
- [44] R Core Team. *R: A Language and Environment for Statistical Computing*. manual. R Foundation for Statistical Computing. Vienna, Austria, 2021. URL: <https://www.R-project.org/>.
- [45] The 1000 Genomes Project Consortium. “A Global Reference for Human Genetic Variation”. In: *Nature* 526.7571 (Oct. 2015), pp. 68–74. ISSN: 1476-4687. DOI: [10.1038/nature15393](https://doi.org/10.1038/nature15393). URL: <https://www.nature.com/articles/nature15393> (visited on 10/28/2019).
- [46] Christopher C. Chang et al. “Second-Generation PLINK: Rising to the Challenge of Larger and Richer Datasets”. In: *GigaScience* 4 (2015), p. 7. ISSN: 2047-217X. DOI: [10.1186/s13742-015-0047-8](https://doi.org/10.1186/s13742-015-0047-8). pmid: [25722852](https://pubmed.ncbi.nlm.nih.gov/25722852/).
- [47] Daniel E. Ho et al. “MatchIt: Nonparametric Preprocessing for Parametric Causal Inference”. In: *Journal of Statistical Software* 42.8 (2011), pp. 1–28. URL: <https://www.jstatsoft.org/v42/i08/>.
- [48] Laurence J. Howe et al. “Genetic Evidence for Assortative Mating on Alcohol Consumption in the UK Biobank”. In: *Nat Commun* 10.1 (Nov. 19, 2019), pp. 1–10. ISSN: 2041-1723. DOI: [10.1038/s41467-019-12424-x](https://doi.org/10.1038/s41467-019-12424-x). URL: <https://www.nature.com/articles/s41467-019-12424-x> (visited on 11/27/2019).
- [49] Yves Rosseel. “Lavaan: An R Package for Structural Equation Modeling”. In: *Journal of Statistical Software* 48.2 (2012), pp. 1–36. URL: <https://www.jstatsoft.org/v48/i02/>.

- [50] Paul-Christian Bürkner. “Brms: An R Package for Bayesian Multilevel Models Using Stan”. In: *Journal of Statistical Software* 80 (Aug. 29, 2017), pp. 1–28. ISSN: 1548-7660. DOI: [10.18637/jss.v080.i01](https://doi.org/10.18637/jss.v080.i01). URL: <https://doi.org/10.18637/jss.v080.i01> (visited on 11/30/2021).
- [51] Thomas Nagylaki. “Assortative Mating for a Quantitative Character”. In: *J. Math. Biology* 16.1 (Dec. 1, 1982), pp. 57–74. ISSN: 1432-1416. DOI: [10.1007/BF00275161](https://doi.org/10.1007/BF00275161). URL: <https://doi.org/10.1007/BF00275161> (visited on 04/05/2019).
- [52] D. Gianola. “Assortative Mating and the Genetic Correlation”. In: *Theor Appl Genet* 62.3 (Sept. 1982), pp. 225–231. ISSN: 0040-5752. DOI: [10.1007/BF00276244](https://doi.org/10.1007/BF00276244). pmid: [24270615](https://pubmed.ncbi.nlm.nih.gov/24270615/).
- [53] Michael J. Gandal et al. “Shared Molecular Neuropathology across Major Psychiatric Disorders Parallels Polygenic Overlap”. In: *Science* 359.6376 (2018), pp. 693–697.
- [54] *NetworkX*. NetworkX, Dec. 1, 2021. URL: <https://github.com/networkx/networkx> (visited on 12/01/2021).

## 5 Online methods

### 5.1 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 section. 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 at [https://github.com/rborder/xAM\\_and\\_gen\\_corr](https://github.com/rborder/xAM_and_gen_corr). All analyses were conducted using R v4.0.2 or R v4.1.0 [44] unless otherwise stated.

#### 5.1.1 Large-scale framework

With the exception of extending the mating scheme to the bivariate case, we proceeded as in [17]. Briefly, we used the genotypes of unrelated European ancestry UK Biobank participants at one million phased, imputed HapMap3 SNPs with minor allele frequency (MAF)  $\geq 0.01$  as founder population. Meiosis was simulated by dividing the genome into



10 kB blocks and deriving recombination probabilities from a linear interpolation of 1000 Genomes Project Phase 3 recombination map [45], 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  $\beta_y, \beta_z$  were jointly bivariate Gaussian with respective variances  $m^{-1}h_{y;\text{pan}}^2, m^{-1}h_{z;\text{pan}}^2$  and correlation  $\rho_\beta$ . Phenotypes were then generated according to the additive model in (1). Individuals (represented by pairs of phenotype values) were mated according to the exchangeable cross-mate correlation regime described in Section 5.1.3.

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 pairs of individuals share any identical-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 [28], LDSC regression as implemented in LDSC v1.0.1 [3], and REML as implemented in BOLT-LMM v2.3.4 [29]. For LDSC, we estimated LD scores within sample using input parameter `ld-window-cm=1.0` and obtained GWAS summary statistics using plink2 v2.0 [46]. To improve computational efficiency, we analyzed subsamples of  $n=4e5$  individuals, with the exception of the fourth and fifth generation results presented in Figure 2a, which were respectively limited to  $n=2e5, 1e5$  by the diminishing sample size at each generation.

### 5.1.2 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 terabytes 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 in Section 5.1.2. Any potential confounds due to relatedness were avoided as in the large-scale framework described above.

**Phenotype simulation with and without pleiotropy** We again generated phenotype pairs according to the additive model (1), 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 genetic effects  $\beta_y, \beta_z$  from the univariate Gaussian distributions with respective variances  $m^{-1}h_{y;\text{pan}}^2, m^{-1}h_{z;\text{pan}}^2$ , using a single set of  $m$  causal variants to simulate complete pleiotropy and two independent sets of  $m/2$  causal variants for to simulate zero pleiotropy. Figure S4 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,  $\rho_\beta$  has greater variance when causal variants overlap), we utilized this approach for the results presented in Section 2.4. The number of causal variants had no apparent impact on quantities of interest (Figure S1).

**Manipulating local LD** Denoting the first and second copies of an individual’s haploid genotypes at the  $j$ th 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  $p_{\text{recomb}}$  and each double line with probability 0.5 below:

$$\begin{array}{c} X_1^{[1]} \\ X_1^{[2]} \end{array} \left| \begin{array}{c} X_2^{[1]} \\ X_2^{[2]} \end{array} \right| \cdots \left| \begin{array}{c} X_c^{[1]} \\ X_c^{[2]} \end{array} \right| \left\| \begin{array}{c} X_{c+1}^{[1]} \\ X_{c+1}^{[2]} \end{array} \right\| \left| \begin{array}{c} X_{c+2}^{[1]} \\ X_{c+2}^{[2]} \end{array} \right| \cdots \left| \begin{array}{c} X_{2c}^{[1]} \\ X_{2c}^{[2]} \end{array} \right\| \cdots \left\| \cdots \right\| \cdots \left| \begin{array}{c} X_m^{[1]} \\ X_m^{[2]} \end{array} \right| \quad (4)$$

Here  $c = m/20$  divides the genome into 20 independently inherited “chromosomes”. Within each chromosome, recombination events occurs between contiguous loci with probability  $p_{\text{recomb}} \in (0.0, 0.5]$ , such that  $p_{\text{recomb}} = 0.5$  corresponds to unlinked loci, with the strength of local LD in the subsequent generation increasing as  $p_{\text{recomb}} \rightarrow 0$ . Figure S3 demonstrates the irrelevance of  $p_{\text{recomb}}$  with respect to liability correlation and estimated effect correlation; simulations with strong local LD ( $p_{\text{recomb}} = 0.01$ ) and weak local LD ( $p_{\text{recomb}} = 0.5$ ) yielded results consistent with the realistic framework simulation results.

### 5.1.3 Mating regimes

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

$$\begin{pmatrix} Y^* \\ Z^* \\ Y^{**} \\ Z^{**} \end{pmatrix} \sim \mathcal{MVN} \left( \mathbf{0}, \begin{pmatrix} 1 & & & \\ s_{yz} & 1 & & \\ r_{yy} & r_{yz} & 1 & \\ r_{yz} & r_{zz} & s_{yz} & 1 \end{pmatrix} \right), \quad (5)$$

where the number of asterisks distinguish the two mates. Whereas the within-mate single-trait correlation (i.e., 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 Section 2.3, 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  $r_{\text{mate}}$ . 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 required for the simulations presented in Section 2.4, which were based on 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 [47]. 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 (5) that minimized the average  $\ell_\infty$  distance between the desired and achieved 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 (see Tables S3 and S5). 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 (Figure S2).

## 5.2 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 [21] and 373,283 spousal pairs drawn from Danish registry data [22–25].

We *a priori* selected a variety of previously-studied phenotypes discussed in two influential publications [2, 3], discarding measures redundant with available measures of higher quality. For example, in the UK Biobank, 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 LDL cholesterol (f.23405), HDL cholesterol (f.23405), and triglycerides (f.30870). Blood biochemistry phenotypes were further adjusted for statin usage.

### 5.3 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 [48] 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 as via estimated kinship coefficients (see Supplemental Materials S1.3.1 for kinship estimation details). Finally, we removed all participant groupings meeting these criteria including more than two individuals. This resulted in a total of 40,697 putative mate pairs (i.e., 81,394 individuals).

### 5.4 Identification of mating pairs in the Danish population cohort

We obtained empirical estimates of spousal correlations for five psychiatric disorders in the Danish Population using the Danish Civil Registration System [23, 24], the Danish National Patient Register [25], and the Danish Psychiatric Central Research Register [22]. 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).

## 5.5 Cross-mate phenotypic correlation estimation

We estimated cross-mate correlations using a structural equation modeling approach via the R package *lavaan* v0.6-8 [49]. Specifically, for the UK Biobank phenotypes, we estimated the sex-constrained structural model corresponding to Eq. (5), with  $Y$  and  $Z$  representing a pair of phenotypes after regressing out the effects of sex, age, sex $\times$ age and age<sup>2</sup> (with the exception of metabolic phenotypes, which were also adjusted for statin use). For the psychiatric phenotypes, which were dichotomous, we proceeded similarly, using the corresponding liability-threshold models to estimate cross-mate tetrachoric correlations. All models were estimated via maximum likelihood with asymptotic standard errors and using pairwise complete observations.

## 5.6 Estimation of GWAS summary statistics in the UK Biobank

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 [46]. Covariates included sex, age, sex $\times$ age, age<sup>2</sup>, 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 a subset of 341,997 unrelated European ancestry individuals (see Supplemental Materials S1.3.1 for details of relatedness estimation).

## 5.7 Empirical heritability and genetic correlation estimates

### 5.7.1 UK Biobank cohort

We estimated marker-based heritabilities and genetic correlations using LDSC v1.0.1 [30] with internal summary statistics (see previous section) and LD scores (computed for regression SNPs with input parameter `ld-window-cm 1.0`). We obtained pedigree-based heritability estimates from the empirical literature, preferring 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 AM [35] and thus constitute the more conservative option, when available (Table S2).

### 5.7.2 Literature-derived estimates for psychiatric disorders

We used a subset of the LDSC heritability and effect correlation estimates reported by Anttila and colleagues [2] 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).

## 5.8 Projected genetic correlation estimates under xAM alone

We implemented a simulation based approach to quantify 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 marker or pedigree-based heritability estimates and allocating  $m = 1000$  distinct causal variants for each phenotype (thus ensuring  $\rho_\beta = 0$ ). We then simulated three 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  $\rho_\beta$ , which we denote  $\hat{\rho}_{\text{xAM}}$ , using HE regression, which, as demonstrated in Figures 2a to 2d, produces results equivalent to LDSC regression.

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  $\hat{\rho}_{\text{emp}}$  and the input parameters  $h_{y;\text{pan}}^2, h_{z;\text{pan}}^2, r_{yy}, r_{zz}, r_{yz}$  from their joint sampling distribution

$$\left(\hat{\rho}_{\beta;\text{emp}}, h_{y;0}^2, h_{z;0}^2, r_{yy}, r_{zz}, r_{yz}\right) \sim \mathcal{MVN}\left(\left(\hat{\rho}_\beta, \hat{h}_y^2, \hat{h}_z^2, \hat{r}_{yy}, \hat{r}_{zz}, \hat{r}_{yz}\right), \hat{\Omega}\right), \quad (6)$$

where  $\hat{\Omega}$  denotes the empirical variance-covariance matrix. We then compared the  $\hat{\rho}_{\text{emp}}$  to the estimate projected by the simulation, which we denote  $\hat{\rho}_{\text{xAM}}$ , to compute the ratio statistic  $\hat{\gamma} = \hat{\rho}_{\text{xAM}}/\hat{\rho}_{\text{emp}}$ . The empirical quantiles of  $\hat{\gamma}$ , which approximate its posterior distribution, were then used to compute credible intervals. For each trait pair, we produced two sets of results, once substituting pedigree-based heritability estimates  $\hat{h}_{\text{PED}}^2$  for the panmictic heritability inputs and again substituting marker-based heritability estimates  $\hat{h}_{\text{SNP}}^2$ . Unfortunately, the sample variance-covariances of  $\hat{h}_y^2, \hat{h}_z^2, \hat{\rho}_\beta$  are not reported by the LDSC software and were not available for any of the literature-derived estimates, and were treated as independent by necessity. That is, only the off-diagonal elements of  $\hat{\Omega}$  corresponding to the cross-mate correlation parameters, which we estimated via structural equation modeling, were nonzero.

## 5.9 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 in Section 5.6 and clumped results using plink v1.90b6.21 using 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.

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

$$\begin{pmatrix} \hat{\ell}_{y;\text{even}} \\ \hat{\ell}_{y;\text{odd}} \\ \hat{\ell}_{z;\text{even}} \\ \hat{\ell}_{z;\text{odd}} \end{pmatrix} \sim \mathcal{MVN} \left( \mathbf{0}, \begin{pmatrix} 1 & & & \\ \phi_{yy} & 1 & & \\ \psi_{yz} & \rho_{\ell;\text{eo}} & 1 & \\ \rho_{\ell;\text{eo}} & \psi_{zy} & \phi_{zz} & 1 \end{pmatrix} \right), \quad (7)$$

where  $\rho_{\ell;\text{eo}}$  is the only free parameter of interest. For the simpler case of single-trait even/odd correlations, we used the correspondingly simple bivariate Gaussian structural model. Again, we estimated structural models using the R package *lavaan* v0.6-8 [49] 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 [50], which does.

## S1 Supplemental material

### Contents

---

S1.1 Theory of assortative mating . . . . .	1
S1.2 Supplementary simulation study results . . . . .	6
S1.3 Psychiatric phenotype definitions . . . . .	7

---

### Supplementary Figures

---

S1 xAM effects and the number of casual variants . . . . .	9
S2 Mating regimes . . . . .	10
S3 Effects of local LD . . . . .	11
S4 Pleiotropy and xAM . . . . .	12
S5 Effects of xAM on GWAS and polygenic scores . . . . .	13
S6 Effects of xAM and misdiagnosis . . . . .	14
S7 Projected versus empirical $\hat{\rho}_\beta$ for UKB traits . . . . .	15
S8 $\hat{\gamma}$ for UKB traits (1 of 2) . . . . .	16
S9 $\hat{\gamma}$ for UKB traits (2 of 2) . . . . .	17
S10 Projected versus empirical $\hat{\rho}_\beta$ for psychiatric trait pairs . . . . .	18
S11 $\hat{\gamma}$ for psychiatric trait pairs subject to misdiagnosis . . . . .	19
S12 Even/odd score correlations for varying $p$ -value thresholds . . . . .	20

---

## S1.1 Theory of assortative mating

### S1.1.1 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, i.e.,

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



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

$$\begin{aligned}\mathbb{E}_t[Y^*Y^{**}] &\geq 0, & \mathbb{E}_t[Z^*Z^{**}] &\geq 0, \\ \mathbb{E}_t[Y^*Z^{**}] &= \mathbb{E}_t[Y^{**}Z^*] > 0.\end{aligned}$$

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  $\ell_y, \ell_z$ , 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

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

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  $Y$  (i.e., 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 (e.g. in the single-trait case [51] 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:

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

By hypothesis,  $X_i, X_j$  are unlinked at panmixis and thus  $\mathbb{E}_0[X_i X_j]$  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

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

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  $\text{sgn} [\beta_{y,i}] \cdot \text{sgn} [\beta_{z,j}]$ , establishing that

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

### S1.1.2 Inflation of GWAS statistics

We now show that assortment on a single trait (sAM) leads to 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

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

for all causal variants indexed  $i, j$ . Without loss of generality, assume all variants  $i, j \in \{1, \dots, m\}$  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

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

Each of the above summands has sign

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

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.

### S1.1.3 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  $\{1, \dots, m\}$  is their disjoint union. Define the estimated polygenic scores

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

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

$$\mathbb{E}[\hat{\rho}_{\ell;\text{eo}}] = \sum_{i \in \mathcal{E}} \sum_{j \in \mathcal{O}} \omega_{ij} \mathbb{E}[\hat{\beta}_{y;i} \hat{\beta}_{z;j}]$$

$$= \sum_{i \in \mathcal{E}} \sum_{j \in \mathcal{O}} \sum_{k,l} \omega_{ij} \omega_{ik} \omega_{jl} \beta_{y;k} \beta_{z;l}.$$

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

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

Pulling this out of the previous expression yields

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

Dealing with the above summand requires fulling understanding the relationship  $\beta_{y;i}$ ,  $\beta_{z;j}$  and  $\omega_{ij}$  for all  $i, j$ , which is beyond the scope of the current manuscript. For now, we address the simpler case of the correlation between polygenic scores for a single phenotype  $Y$  restricted to disjoint sets of chromosomes as analyzed in [18]. Again, we assume that  $i \in \mathcal{E} \cup \mathcal{O} \implies \beta_{y;i} \neq 0$  for simplicity. In this case, we have

$$\mathbb{E}[\hat{\rho}_{\ell;eo}] = \rho_{\ell;eo} + \sum_{i \in \mathcal{E}} \sum_{j \in \mathcal{O}} \sum_{k \neq i} \sum_{l \neq j} \omega_{ij} \omega_{ik} \omega_{jl} \beta_{y;k} \beta_{y;l}, \quad (8)$$

where each of the above summands has sign

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

Likewise, the summands comprising the true correlation

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

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

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

That is, as anticipated by Yengo et al. [18], sAM induces a true correlation between genetic liabilities restricted to disjoint collections of loci. Note that we do not require the assumption of equilibrium.

## S1.2 Supplementary simulation study results

### S1.2.1 GWAS effect estimates are biased upwards under xAM

We performed additional simulation studies examining the impact of xAM on GWAS effect estimates by comparing estimated regression slopes to their true values at each generation in the context of the simplified forward-time simulation framework described in the Online Methods. In contrast to effect correlation estimators, on which single-trait AM (sAM) has limited effects [5, 52], association statistics are biased upwards in magnitude under sAM as each causal variant weakly will tag every other causal variant (see S1.1.1 and S1.1.2 ). We therefore sought to distinguish the effects of sAM and xAM by varying the cross-mate cross-trait correlation  $r_{yz}$  independently of the cross-mate single-trait correlations  $r_{yy}$  and  $r_{zz}$ . Specifically, we consider two phenotypes  $Y, Z$ , each with  $m = 2,000$  non-overlapping causal variants and panmictic heritabilities equal to 0.5. Letting asterisks distinguish mates, we varied the cross-mate correlation parameters governing assortment

$$\text{cor} \left( \begin{pmatrix} Y^* \\ Z^* \\ Y^{**} \\ Z^{**} \end{pmatrix} \right) = \begin{pmatrix} 1 & & & \\ s_{yz} & 1 & & \\ r_{yy} & r_{yz} & 1 & \\ r_{yz} & r_{zz} & s_{yz} & 1 \end{pmatrix}, \quad (9)$$

fixing  $r_{zz} \equiv 0.5$  and varying  $(r_{yy}, r_{yz}) \in \{0.0, 0.25, 0.5\}^2$ , running 100 simulation replicates per experimental condition. Focusing on  $Y$ , we have sAM when  $r_{yy} > 0$  and  $r_{yz} = 0$  and xAM when  $r_{yz} > 0$

At each generation, we emulate a GWAS by regressing  $Y$  on each of the  $2m$  standardized genotypes to estimate  $\beta_j$ ,  $j = 1, \dots, 2m$  in a sample of size  $n = 32,000$ . We then report two metrics: the average magnitude of GWAS slope error relative to the average effect size ( $\text{mean}((\beta - \hat{\beta}) \cdot \text{sgn } \beta) / \sqrt{(h_{y;\text{pan}}^2/2m)}$ ; hereafter termed *slope inflation*), which we present in Figure S5a, and the correlation between the true liability  $\ell_y = \sum_i X_i \beta_i$  and estimated polygenic score ( $\hat{\ell}_y = \sum_i X_i \hat{\beta}_i$ ; hereafter termed *score correlation* ), which we present in Figure S5b.

As expected, we observed that the strength of single-trait correlation biases GWAS effects upwards in magnitude regardless of the cross-trait correlation (e.g., after three generations, the average slope inflation per 0.05 increase in  $r_{yy}$  was 0.143 [se=0.0374%]). We also saw that cross-mate cross-trait correlation alone causes upwards inflation and exacerbates the inflation caused by single trait correlations (e.g., after three generations, the average slope inflation per 0.05 increases in  $r_{yz}$  were 0.157% [SE 0.0097%] and 0.143% [SE 0.0310%] for  $r_{yz}=0$  and  $r_{yz}>0$ , respectively). To summarize, both assortment on a

single trait and assortment across multiple traits upwardly bias GWAS effect estimates in subsequent generations, and these biases are exacerbated when both forms of assortment are present.

Polygenic score results are less straightforward as single- versus cross-trait assortment exert countervailing influences on the correlation between the true and estimated polygenic scores. Regressing the difference between the score correlations for phenotype  $Y$  after three generations of assortment and those at generation zero (panmixis), the difference score increases by  $3.46\text{e-}3$  (SE:  $2.59\text{e-}5$ ) per 0.05 increase in  $r_{yy}$  and decreases by  $1.10\text{e-}3$  (SE:  $2.59\text{e-}5$ ) per 0.05 increase in  $r_{yz}$ . We interpret this to mean that single-trait assortment, even though it increases the average effect estimate error, leads to a stronger correlation between the true and estimated polygenic scores, presumably because all causal variants tag one another. On the other hand, cross-trait assortment, which induces correlations between causal variants for one trait and causal variants for another, inflates GWAS slope estimates while decreasing the score correlation, at least when the two traits are genetically orthogonal.

### **S1.2.2 xAM exacerbates the impacts of misclassification errors on effect correlation estimates**

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 again 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, an preceding analogous for the second phenotype. As anticipated, misclassification errors induced artifactual  $\hat{\rho}_\beta$  estimates among genetically orthogonal traits and this inflation was in turn exacerbated by the extent of xAM across the two traits (Figure S6). Further, in contrast to xAM, the impacts of misclassification errors depended on trait prevalence such that classification-induced  $\hat{\rho}_\beta$  inflation was more pronounced for less common binary traits (Figure S6).

## **S1.3 Psychiatric phenotype definitions**

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 (e.g., 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 five psychiatric phenotypes considering both ICD-8 and ICD-10 definitions (Table S1). The definitions are based on those by the iPSYCH initiative [53].

**Table S1:** Psychiatric phenotype definitions

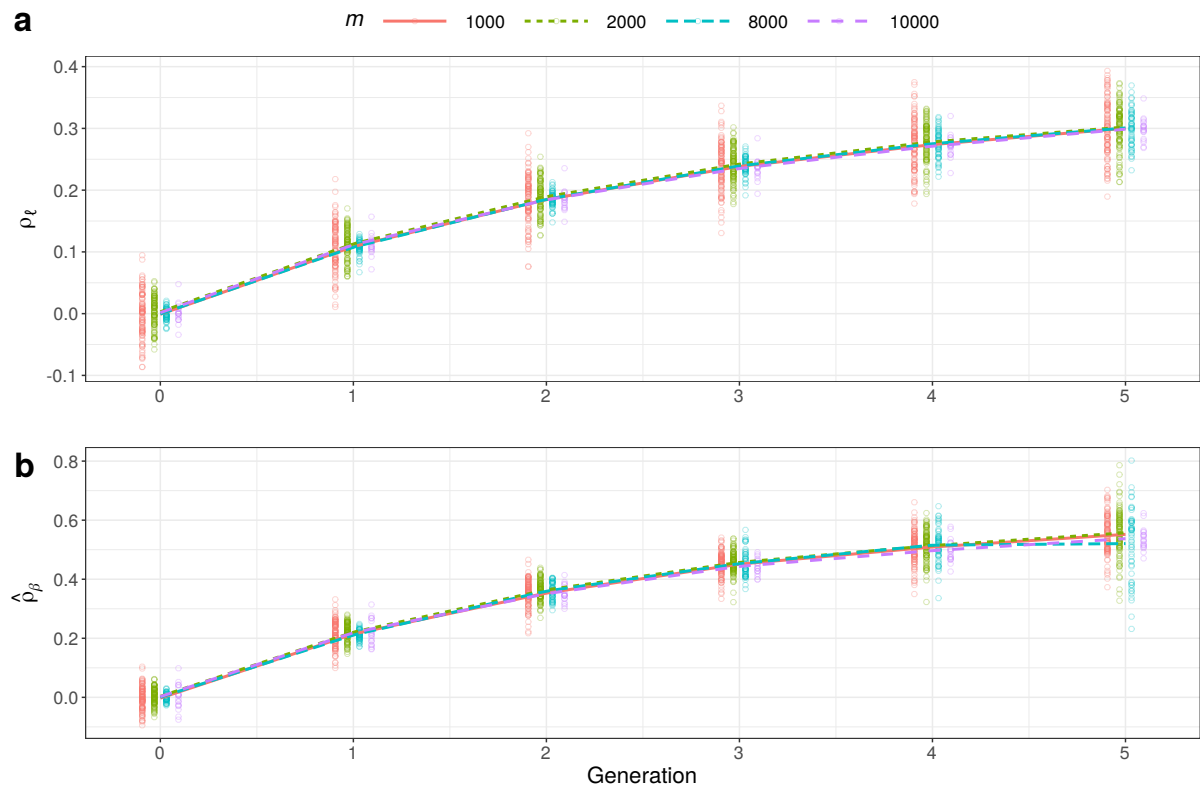
	Phenotype	ICD-8 code	ICD-10 code
<b>ANX</b>	Anxiety disorders	NA	F40.0-F40.2, F41.0-F41.1, F42, F43.0-F43.1
<b>ADHD</b>	Attention deficit/ hyperactivity disorder	NA	F90.0
<b>BIP</b>	Bipolar disorders	296.19, 296.39, 298.19	F30-F31
<b>MDD</b>	Single and recurrent depressive disorder	296.09, 296.29, 298.09, 300.49	F32-F33
<b>SCZ</b>	Schizophrenia	295.x9 (excl 295.79)	F20

### S1.3.1 Identification of unrelated European-ancestry individuals 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 individuals with excess heterozygosity ( $\geq 5$  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 [54]. This resulted in 342,257 unrelated individuals. In contrast to [21], who estimated kinships using  $\approx 92,000$  common SNPs with small loadings onto the first few PCs in the full sample (including multiple ancestries; see S3.7 of [21]), 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 identified by [21].

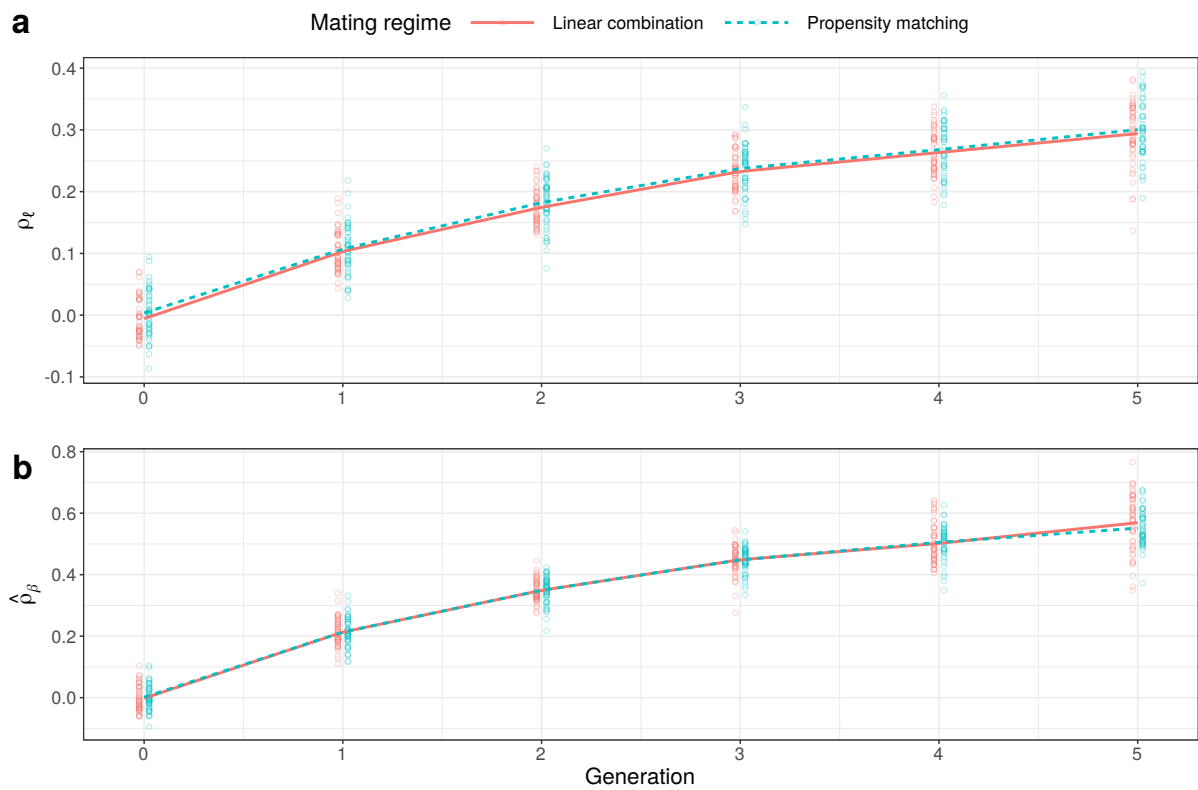
## Supplementary Figures

**Figure S1:** Varying the number of causal variants has no impact on (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.

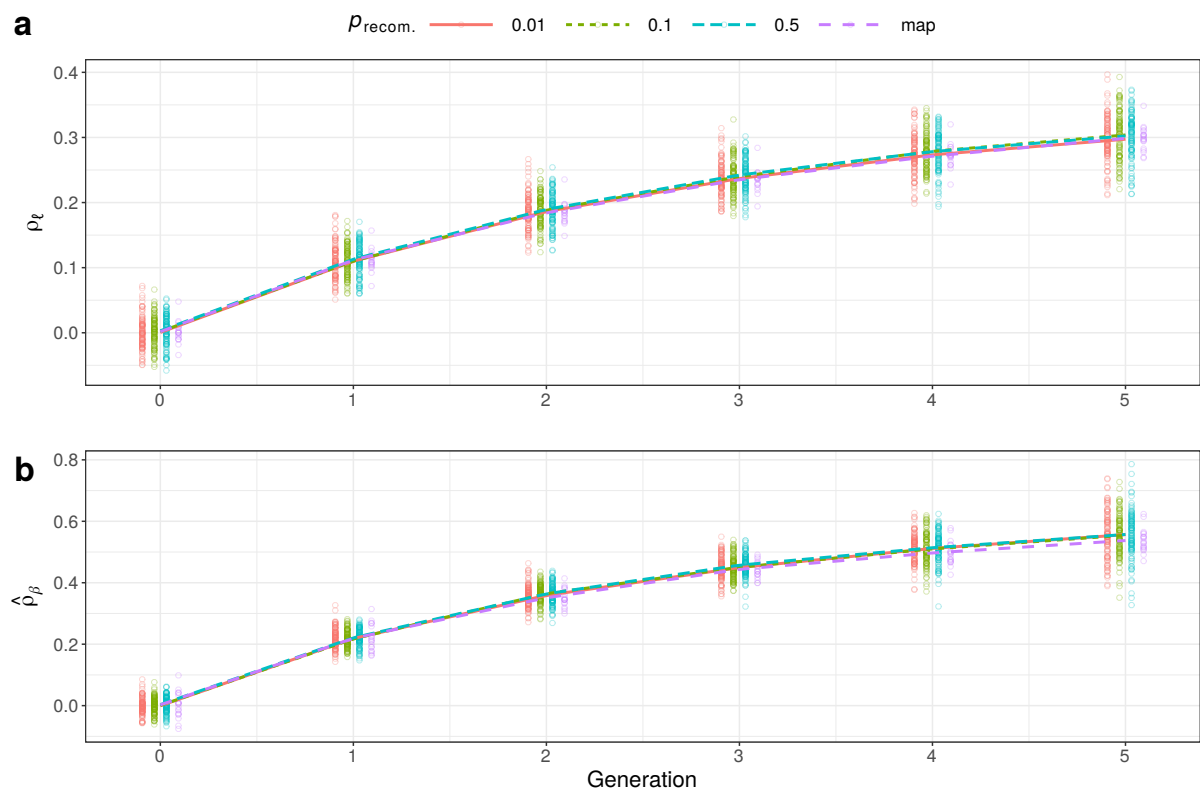




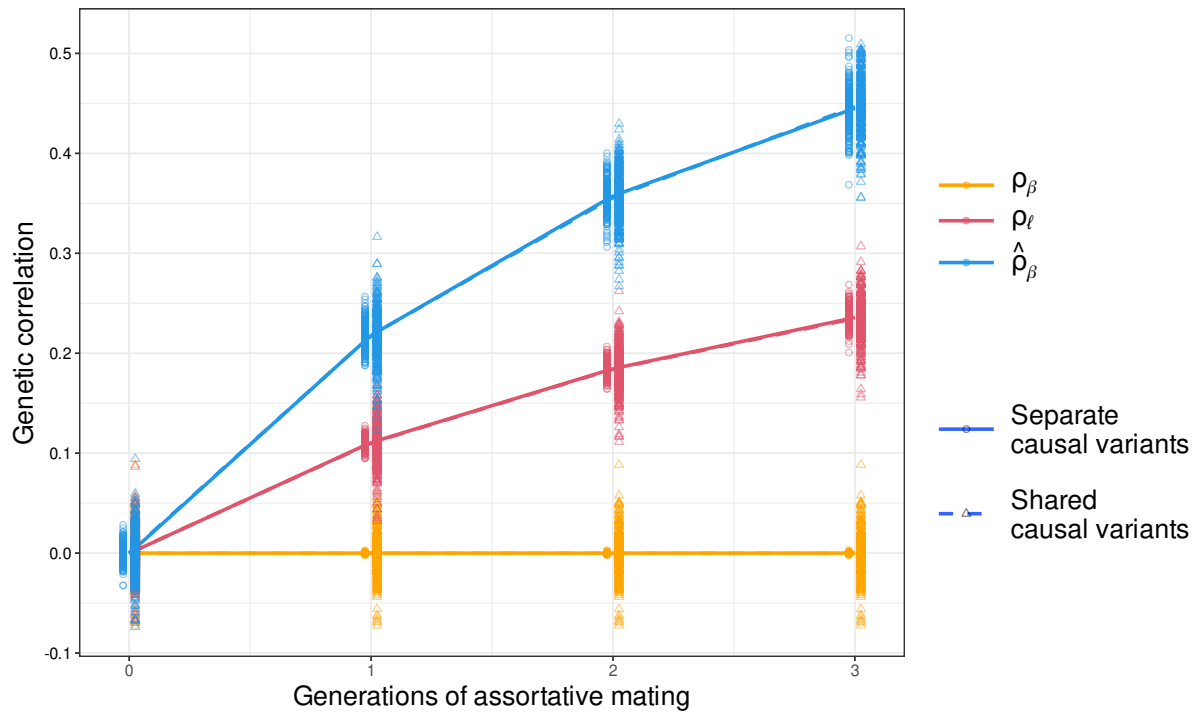
**Figure S2:** Equivalence of mating regimes under exchangeable correlation structure for genetically orthogonal traits with respect to (a) true score correlation or (b) estimated effect correlation. Panmictic heritabilities and cross-mate correlations were fixed at 0.5 across simulations.



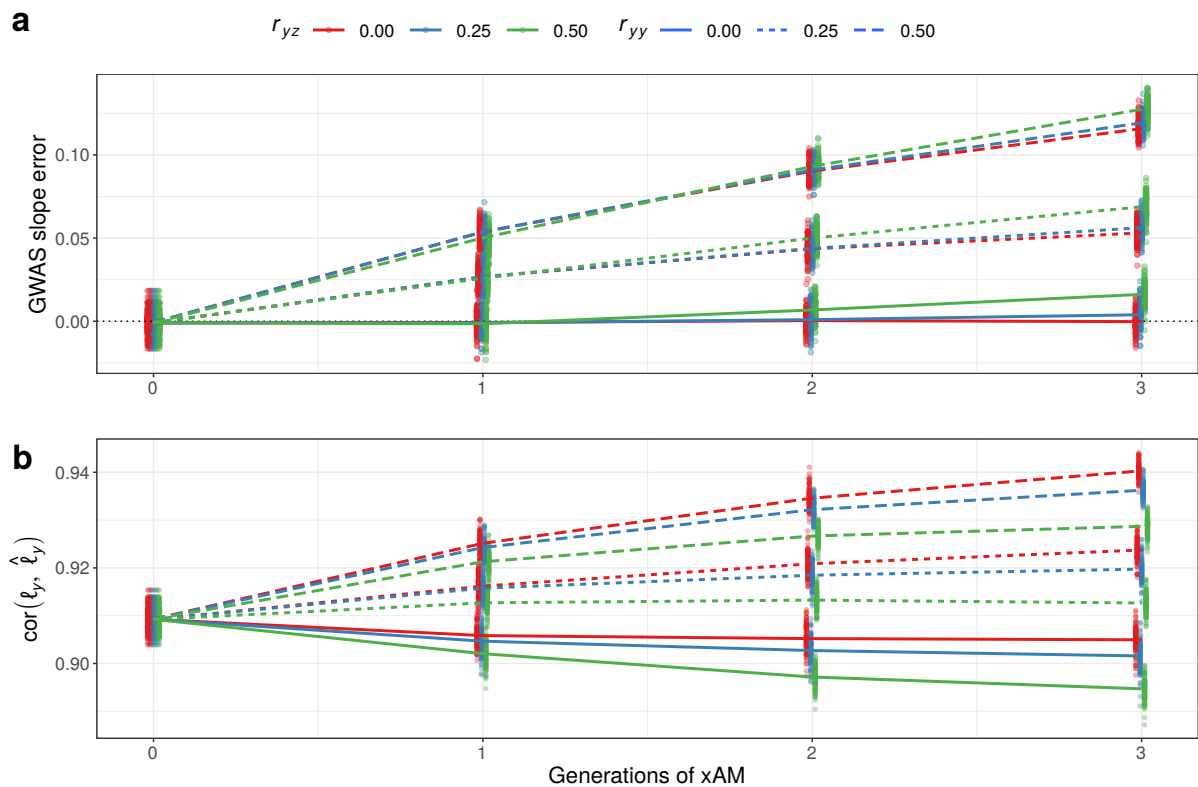
**Figure S3:** Impact of local LD in synthetic data for genetically orthogonal traits with panmictic heritabilities and cross-mate correlations fixed at 0.5. Enforcing recombination probabilities between contiguous loci ( $\rho_{\text{recom.}}$ ) at varying fixed values or using an empirical recombination map has no impact on the (a) true score correlation or (b) estimated effect correlation across simulations.



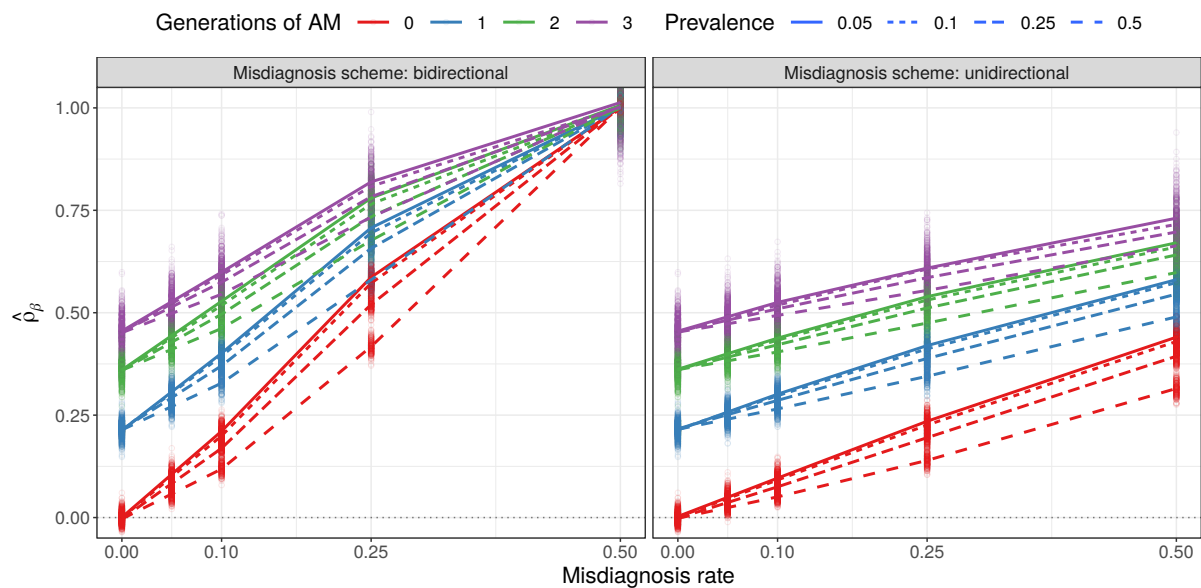
**Figure S4:** Score correlation and estimated effect correlation for genetically orthogonal traits subject to xAM with and without pleiotropy under orthogonal effects. Across simulations, panmictic heritabilities and cross-mate were correlations fixed at 0.5. Under the separate casual variants regime,  $\beta_{y;i} \neq 0$  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.



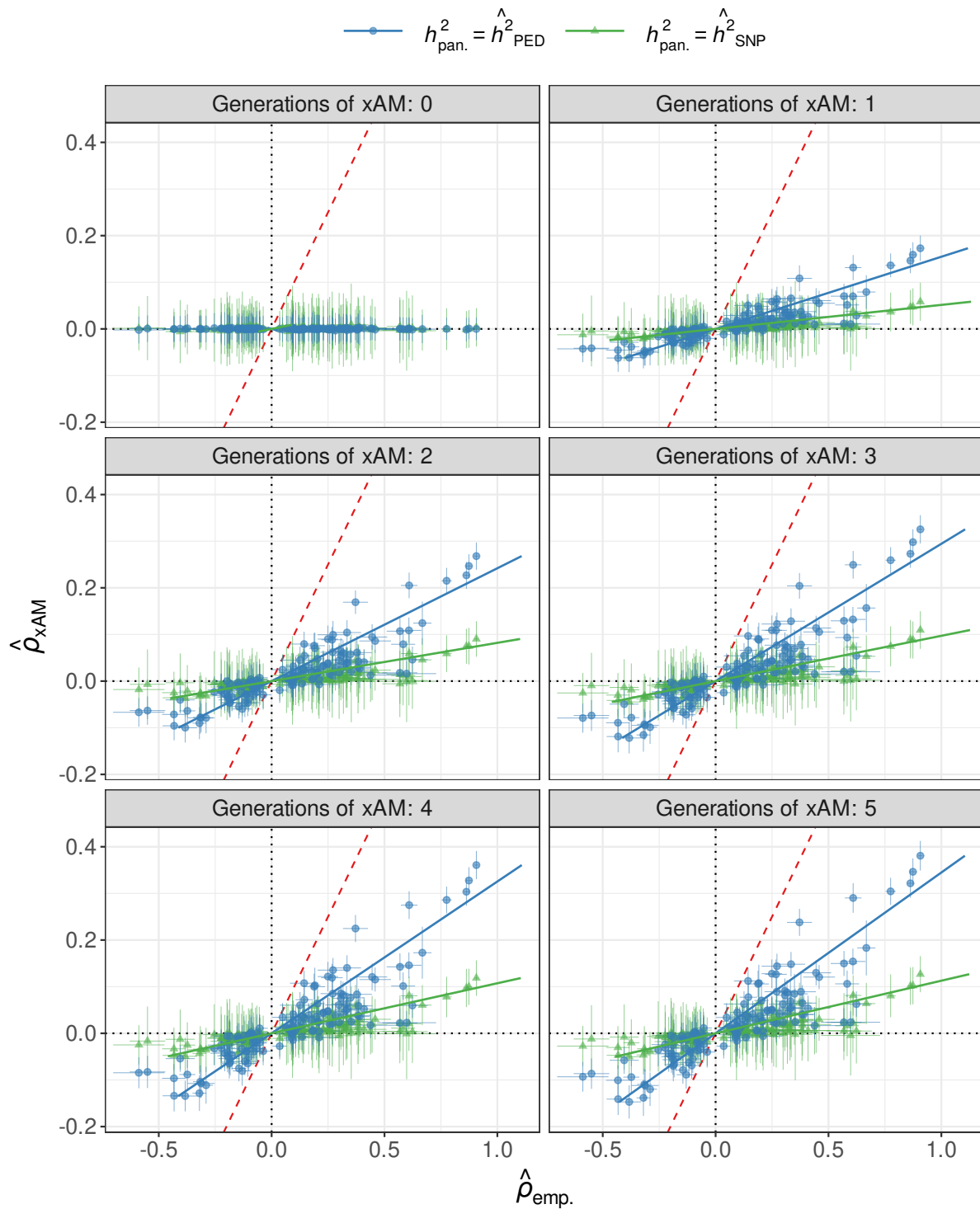
**Figure S5:** Impact of assortative mating on GWAS regression weights and polygenic scores. Across all simulations, the panmictic heritabilities for  $Y$  and  $Z$  are set at 0.5, the true effect correlation is fixed at 0.0, and the cross-mate correlation with respect to  $Z$ ,  $r_{zz}$ , is fixed at 0.5, whereas  $r_{yy}$  and  $r_{yz}$  vary across experimental conditions. (a) Slope inflation of GWAS estimates with respect to  $Y$  under positive xAM depends strongly on  $r_{yy}$  and modestly on  $r_{yz}$ . (b) Correlation between the true and estimated polygenic scores for  $Y$  is inflated by  $r_{yy}$  and deflated by  $r_{yz}$ . Intuitively, increasing  $r_{yy}$  induces sign-consistent LD among causal variants for  $Y$  whereas increasing  $r_{yz}$  induces sign-consistent LD across causal variants for  $Y$  and  $Z$ .



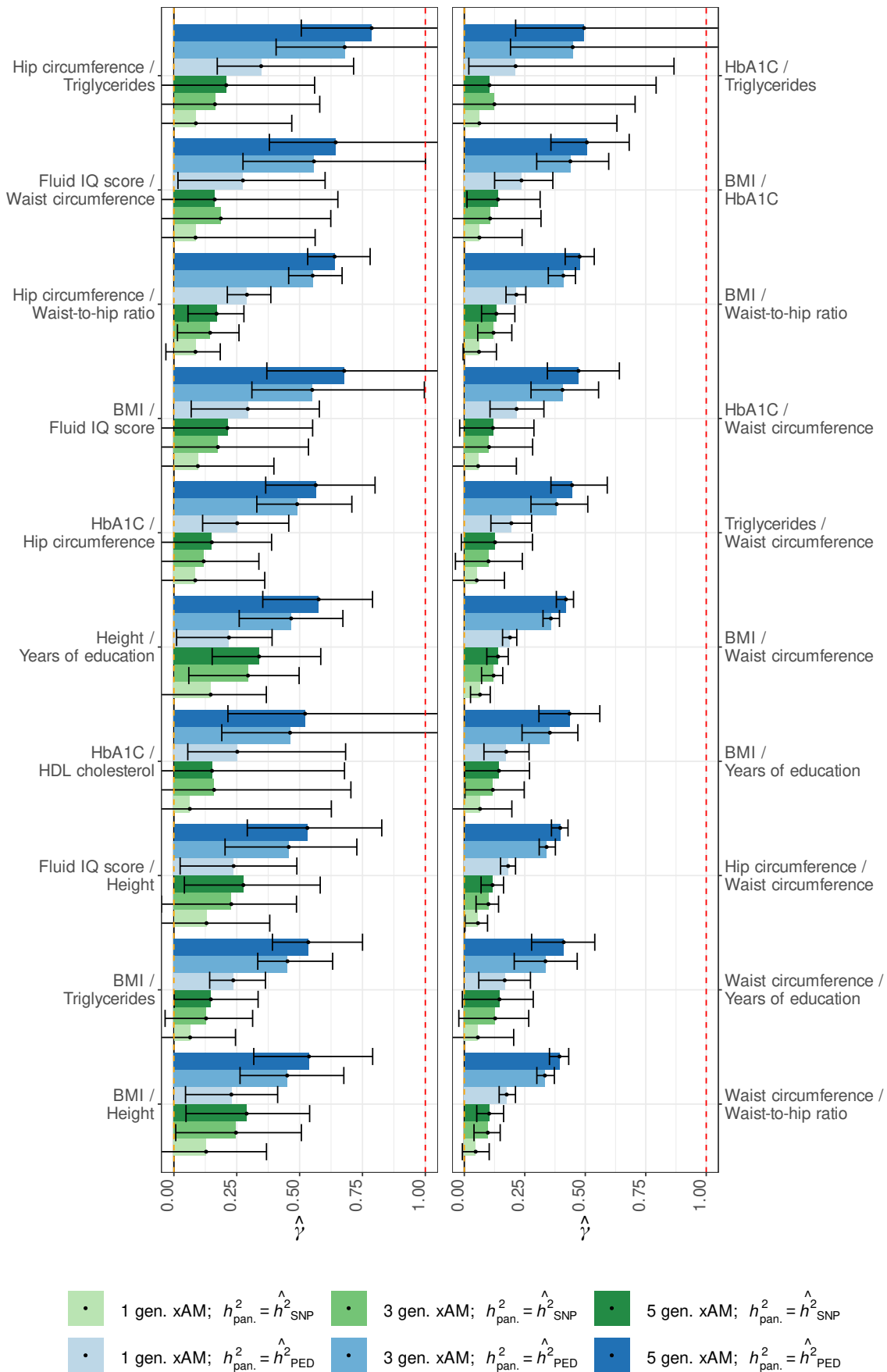
**Figure S6:** Impact of xAM and misdiagnosis errors on effect correlation estimates between genetically orthogonal binary traits. Panmictic heritabilities and cross-mate correlations are all fixed at 0.5. Under the unidirectional scheme, individuals with disorder *A* are mislabeled as controls for disorder *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.



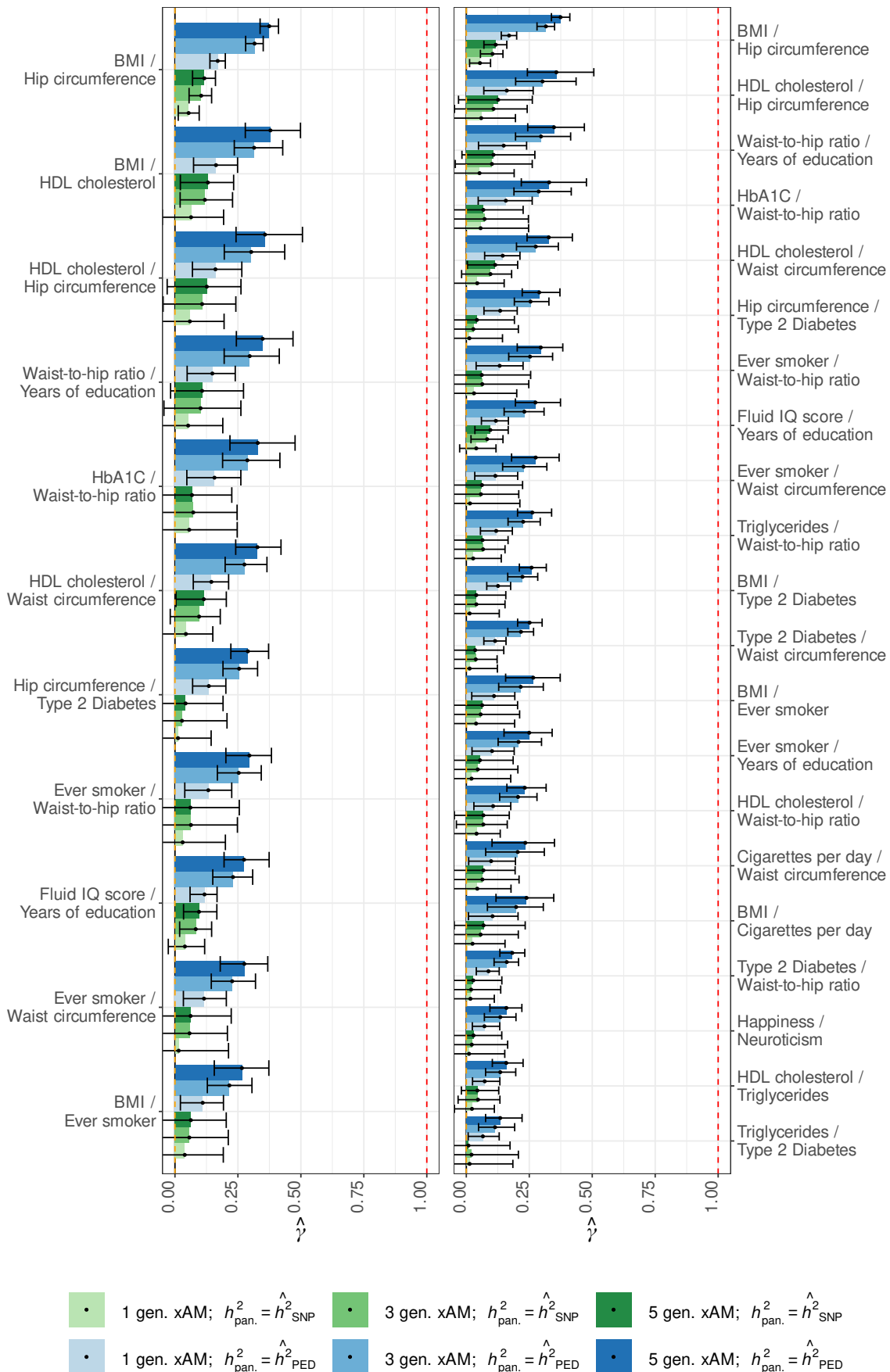
**Figure S7:** Projected versus empirical  $\hat{\rho}_\beta$  for UK Biobank trait pairs as a function of number of generations of xAM.



**Figure S8:** Projected  $\hat{\rho}_\beta$  estimates relative to empirical estimates for UK Biobank trait pairs (1 of 2).

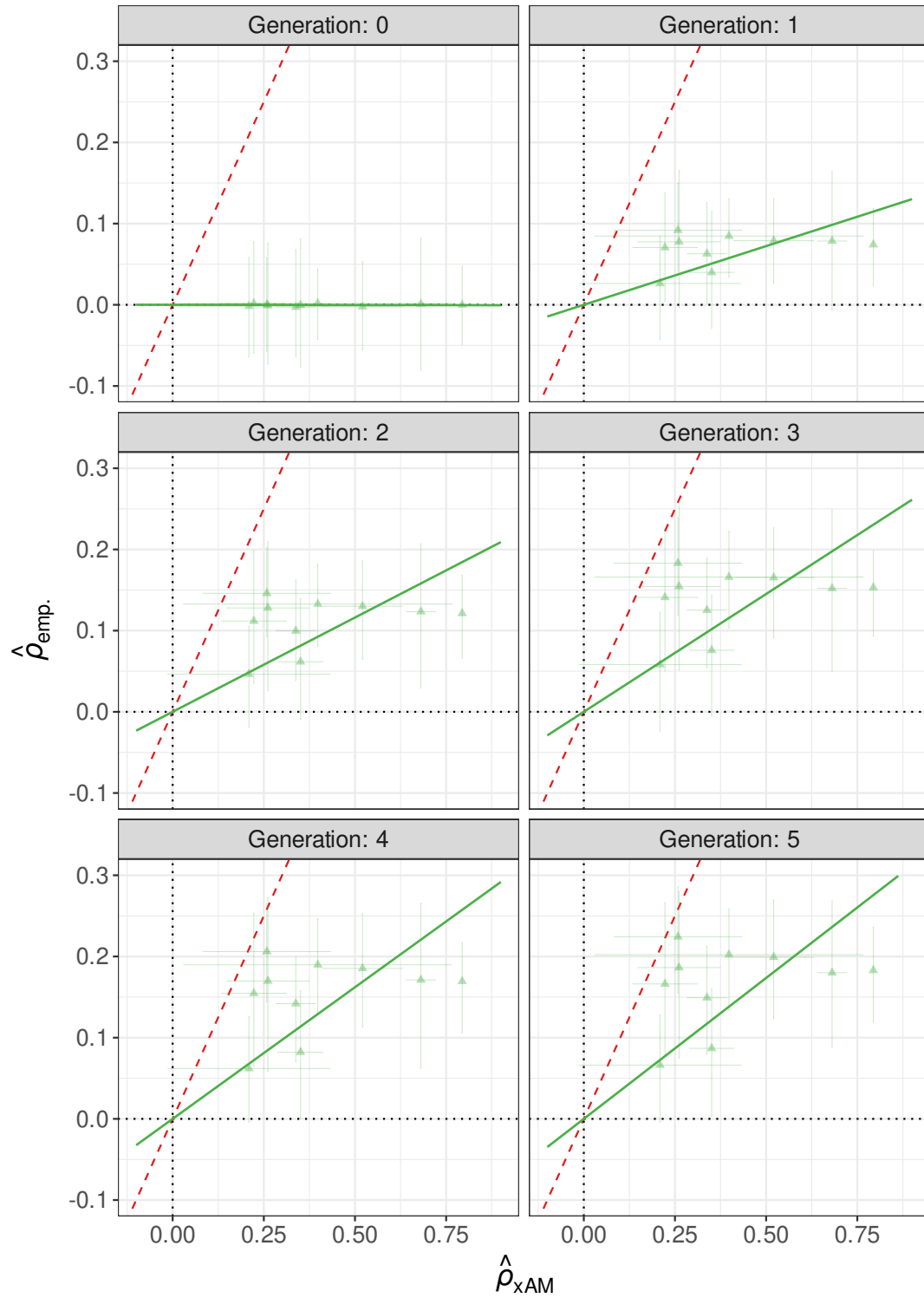


**Figure S9:** Projected  $\hat{\rho}_\beta$  estimates relative to empirical estimates for UK Biobank trait pairs (2 of 2).

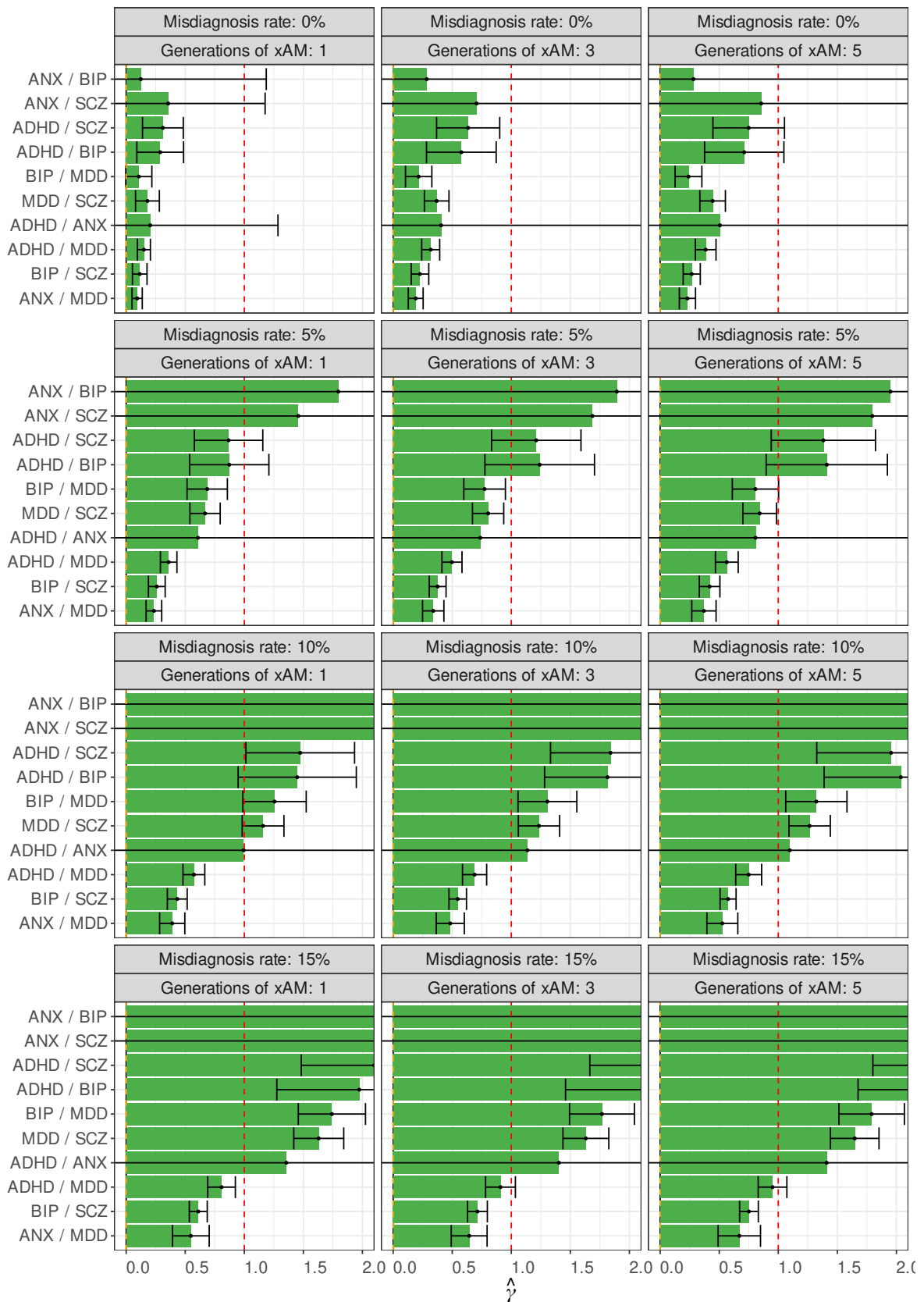




**Figure S10:** Projected versus empirical  $\hat{\rho}_\beta$  for psychiatric trait pairs



**Figure S11:** Projected  $\hat{\rho}_\beta$  estimates relative to empirical estimates for psychiatric trait pairs under xAM and misdiagnosis rates



**Figure S12:** Cross-chromosome polygenic score correlations and cross-mate phenotypic correlations for varying  $p$ -value thresholds

