

Pervasive Downward Bias in Estimates of Liability Scale Heritability in GWAS Meta-Analysis: A Simple Solution

Andrew D. Grotzinger^{1,2*}, Javier de la Fuente^{3,4}, Michel G. Nivard,⁵ Elliot M. Tucker-Drob^{3,4}

¹ Institute for Behavioral Genetics, University of Colorado Boulder, Boulder, CO USA

10 ²Department of Psychology and Neuroscience, University of Colorado Boulder, Boulder, CO
11 USA

¹² ³ Department of Psychology, University of Texas at Austin, Austin, TX USA

13 ⁴Population Research Center, University of Texas at Austin, Austin, TX USA

14 ⁵Department of Biological Psychology, VU University Amsterdam, Amsterda

14. Department of Biological Phylogeny, VU University Amsterdam, Amsterdam, the Netherlands

15 Correspondence to Andrew D. Grotzinger (Andrew.Grotzinger@colorado.edu)

20

22

23

24

25

26

28

Abstract

29 SNP heritability is a fundamental quantity in the genetic analysis of complex traits. For binary
30 phenotypes, in which the continuous distribution of risk in the population is unobserved, observed-scale
31 heritabilities must be transformed to the more interpretable liability-scale. We demonstrate here that the
32 field standard approach for performing the liability conversion can downwardly bias estimates by as much
33 as ~20% in simulation and ~30% in real data. These attenuated estimates stem from the standard approach
34 failing to appropriately account for varying levels of ascertainment across the cohorts comprising the
35 meta-analysis. We formally derive a simple procedure for incorporating cohort-specific ascertainment
36 based on the summation of effective sample sizes across the contributing cohorts, and confirm via
37 simulation that it produces unbiased estimates of liability-scale heritability.

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54 *SNP heritability (h_{SNP}^2)* quantifies the proportion of total variance in a phenotype within a
55 population that is attributable to the additive effect of tagged genetic variants. For continuously measured
56 quantitative traits, in which phenotypic variation is directly observed, h_{SNP}^2 estimates produced from
57 standard methods such as LD score regression (LDSC)¹ are directly interpretable. However, when the
58 measured phenotypes are binary (e.g., for case-control disease traits) conventional estimates of h_{SNP}^2 are
59 not easily interpreted for two reasons. The first is due to the binarized scale of the data, where h_{SNP}^2 is
60 most interpretable when taking into account the continuous distribution of risk in the population. The
61 second relates to the fact that GWAS of disease traits are often performed on ascertained samples, in
62 which affected individuals are overrepresented so as to increase statistical power for rare disorders. The
63 standard transformation for binary traits then uses a liability threshold model to convert observed-scale
64 SNP heritability (h_o^2) to liability-scale SNP heritability (h_l^2) in order to produce an estimate that both
65 accounts for the continuous distribution of risk in the population and is not biased by ascertainment. In
66 practice, h_l^2 is commonly estimated with summary-based methods such as LDSC using results from
67 GWAS meta-analysis across many different samples, varying in their levels of ascertainment.

68 Here we highlight a critical error in the standard practice for calculating h_l^2 from GWAS meta-
69 analysis that can cause substantial downward bias due to variation in cohort-specific ascertainment, and
70 we formally derive a simple procedure for obtaining unbiased h_l^2 estimates. We report results from
71 simulations that illustrate the extent of the downward bias across a variety of conditions, and showcase
72 the unbiased nature of the proposed procedure within these same conditions. We go on to quantify the
73 extent of this bias in 12 recent GWAS meta-analyses of case/control psychiatric and neurological traits. It
74 appears that the biased approach has been used for h_l^2 estimates for nearly all meta-analytic GWAS of
75 binary traits, with the exception of Mullins et al. (2021)² who implemented an unbiased estimator of h_l^2
76 akin to the procedure developed here. We report the corrected estimates of h_l^2 under the same population
77 prevalences assumed by the original GWAS reports and we observe that the field standard approach for
78 computing h_l^2 attenuates estimates by as much as 31%.

79 Observed scale heritability is estimated within univariate LDSC as:¹

$$80 \quad \mathbb{E}[Z_j^2] = \mathbb{E}[\chi_j^2] = N \frac{h_o^2}{M} \ell(j) + a + 1 \quad (1)$$

81 where N is the sample size, h_o^2 is the observed scale heritability, $\ell(j)$ is the total LD-score for SNP j , M is
82 the total number of SNPs used to calculate the LD-scores, and a is a term representing unmeasured
83 sources of confounding such as population stratification.

84 When summary data are derived from a single case-control GWAS (either of a single sample or
85 of raw data that have been combined across multiple samples prior to GWAS), the observed scale
86 heritability (h_o^2) can be converted to the liability scale (h_l^2) as follows:^{3–5}

87
$$h_l^2 = h_o^2 \frac{P^2(1-P)^2}{\phi^2 v(1-v)} \quad (2)$$

88 where v is the sample prevalence, P is the population prevalence, and ϕ is the height of the standard
89 normal probability density function at the threshold corresponding to P . Combining *Eqs. 1* and *2* produces
90 the reduced form LDSC equation for binary traits:

91
$$\mathbb{E}[\chi_j^2] = \frac{\phi^2}{P^2(1-P)^2} v(1-v) N \frac{h_l^2}{M} \ell(j) + a + 1 \quad (3)$$

92 In the **Supplementary Note**, we show that when GWAS summary data are derived from meta-
93 analysis of summary statistics from multiple, individual case-control GWAS, the appropriate reduced
94 form equation for estimating h_l^2 is:

95
$$\mathbb{E}[\chi_j^2] = \frac{\phi^2}{P^2(1-P)^2} (\sum v_k(1-v_k)n_k) \frac{h_l^2}{M} \ell(j) + a + 1 \quad (4)$$

96 which resembles *Eq. 3* for summary statistics derived from a single GWAS, with the key difference being
97 that $v(1-v)N$ is replaced by $(\sum v_k(1-v_k)n_k)$.

98 Importantly, currently available software does not allow for direct entry of $(\sum v_k(1-v_k)n_k)$,
99 and standard practice in LDSC analysis of meta-analytic GWAS summary data has been to compute a
100 single meta-analytic v as the total sample prevalence (i.e. aggregate number of cases across all samples
101 divided by the aggregate sample size, and enter this quantity into *Eq. 3*). When samples are differentially
102 ascertained, as is nearly always the case in empirical settings, such an approach is not equivalent to the
103 correct approach given by *Eq. 4*, i.e.:

104
$$\sum v_k(1-v_k)n_k \neq \frac{(\sum v_k n_k)(\sum(1-v_k)n_k)}{\sum n_k} (\sum n_k) = \left(\frac{N_{Cases\ Total}}{N_{Total}} \right) \left(\frac{N_{Controls\ Total}}{N_{Total}} \right) N_{Total} \quad (5)$$

105 We refer to the quantity on the left of the inequality as the summation of cohort-specific ascertainment,
106 and the quantities on the right of the inequality as the total sample ascertainment. In the **Supplemental**
107 **Note** we describe a simple procedure in which the correct estimate of h_l^2 (i.e. *Eq. 3*) can be obtained for
108 meta-analytic summary statistics using standard software (e.g., LDSC,¹ Genomic SEM,⁶ MTAG⁷
109 LDAK⁸). First, the effective sample size, $EffN$ is computed for each study, k , as:

110
$$EffN_k = 4v_k(1-v_k)n_k \quad (6)$$

111 where $EffN$ represents the sample size for an equivalently powered GWAS within a balanced sample
112 (i.e. 50% cases, 50% controls). Because the $EffN$ values are directly comparable across GWAS samples
113 they can be summed. The sum of $EffN$ across all contributing GWAS ($\sum EffN$) is then entered for N
114 along with the $v=.5$, so as to represent the balanced nature of the design. The population prevalence from
115 collateral epidemiological data is entered for P as usual.

116 In addition to being pragmatically appealing given currently available software, this approach has
117 two additional advantages. First, a number of statistical pipelines adopted by major, genomic research
118 consortia (e.g., the RICOPILI pipeline implemented by the Psychiatric Genomics Consortium [PGC]¹)
119 default to outputting $\sum EffN$ in the GWAS summary statistics output. Second, this allows the user to
120 account for both cohort-specific and SNP-specific information. That is, when participant sample size
121 varies by SNP, as is often the case given different genotyping platforms utilized by cohorts, it is
122 preferable to use this SNP-specific information.

123 We used simulation to empirically evaluate the extent of potential bias associated with the field
124 standard, total ascertainment approach, and to confirm the unbiased nature of the cohort-specific
125 ascertainment approach. For all simulations, the liability scale heritability (h_l^2) was 15% in the generating
126 population, the population prevalence (P) was 1%, and the meta-analytic GWAS consisted of ten cohorts,
127 each with a sample size of 5,000 (i.e., total meta-analytic $N = 50,000$). All cohorts were specified to
128 genetically correlate at 1 to reflect meta-analysis of the same phenotype. Each cohort was also specified
129 to consist of either 50% cases and 50% controls (representing samples with high ascertainment) or 10%
130 cases and 90% controls (representing samples with relatively low ascertainment). These parameters were
131 chosen to reflect a combination of realistic scenarios generally observed for univariate GWAS of binary
132 traits.

133 We conducted simulations for 11 conditions that iteratively changed with respect to the number
134 of contributing cohorts with 50% cases and controls. More specifically, Condition 1 consisted of 10
135 cohorts with 10% cases and 90% controls, Condition 2 consisted of 9 cohorts with 10% cases and 90%
136 controls and 1 cohort with 50% cases and controls, and so on up through Condition 11 consisting of 10
137 cohorts with 50% cases and controls. For each of the 11 conditions, we ran 100 simulations, each
138 comprised of 10 cohort-level sets of summary statistics (i.e., 11,000 simulated cohorts). Summary
139 statistics were generated directly from the LD-score regression equation, subsequently combined across
140 the 10 cohorts using inverse variance weighted (IVW) meta-analysis, and finally analyzed in LDSC in
141 one of two ways (see **Method** for details). The first was using the standard procedure of inputting the
142 total sample prevalence (v_{Total}) and total sample size (N_{Total}). The second was using $\sum EffN$ across the
143 cohorts and a v of .5, as we propose here.

144 Simulation results are presented in Figure 1 and Table 1. These results reveal three primary
145 findings. First, the field standard approach of using v_{Total} can produce substantial, downward bias in
146 liability-scale heritability estimates, with bias increasing as a function of the degree of variability in
147 ascertainment across contributing cohorts (Figure 1A). Thus, bias was greatest for those conditions where
148 there was an even mixture of 50%/50% and 10%/90%, case/control cohorts; in these instances, the
149 downward bias was ~20% (Table 1). Second, both approaches produce the same estimates when
150 ascertainment is equivalent across all cohorts (Figure 1B and 1L). Importantly, the standard procedure of
151 using total sample prevalence (v_{Total}) is not biased as a function of the overall degree of ascertainment.
152 Rather, the bias is attributable to the level of ascertainment differences across cohorts. Third, our
153 proposed procedure of using $\sum EffN$ removes this bias, producing accurate estimates of the population-
154 level, liability scale heritability. Having established that $\sum EffN$ produces an accurate estimate of h_l^2 , we
155 went on to examine the difference across using v_{Total} and $\sum EffN$ in real data.

156 We compared the field standard procedure of using v_{Total} versus our proposed approach of using
157 $\sum EffN$ for 12 major, binary traits for which sufficient cohort-level information was available. We used
158 the same population prevalences from the original GWAS publications from which the summary statistics
159 were derived. We quantify bias here as the proportional difference across
160 $\sum EffN$ and v_{Total} (i. e., $\frac{h_l^2 \text{ Estimate for } v_{Total}}{h_l^2 \text{ estimate for } \sum EffN} - 1$). Consistent with simulation findings, real data
161 results revealed that in all cases using v_{Total} produced a deflated estimate of liability scale heritability
162 relative to $\sum EffN$. This bias ranged from as little as 1.3% for autism spectrum disorder⁹ to as much as
163 28.1% for alcohol use disorder¹⁰ and 31.8% for bipolar disorder² (Table 2). In all but one instance, the
164 heritability estimates reported in the corresponding manuscripts most closely matched those produced
165 from using v_{Total} (Supplementary Table 1). The exception was the most recent release (Freeze 3) of the
166 PGC bipolar summary statistics² which report a liability scale heritability consistent with using $\sum EffN$.

167 SNP heritability is a fundamental quantity in complex trait genetics. As such, SNP heritability
168 estimates from GWAS summary statistics are standard results to report in any major GWAS meta-
169 analysis effort. For binary traits, such as case-control disease traits, SNP heritability estimates must be
170 converted to the liability scale in order to be meaningfully interpreted. We demonstrate here that the de
171 facto field standard approach for estimating liability scale heritability from meta-analytic GWAS
172 summary data can downwardly bias these estimates by as much as ~20% in simulation and ~30% in real
173 data. We have introduced a simple procedure for obtained unbiased estimates of liability scale SNP
174 heritability in these contexts.

175 Downwardly biased estimates of SNP heritability will propagate to produce downwardly biased
176 estimates of genetic covariance, which may in turn bias methods that rely on these estimates (e.g.,
177 MTAG,¹¹ Genomic SEM⁶). Importantly, genetic correlations are expected to be unaffected by this bias as
178 they standardize genetic covariance estimates relative to heritability estimates, thereby cancelling out the
179 bias. Another issue which merits further investigation is the presence of ascertainment differences that
180 stratify by meaningful covariates across cohorts. For example, it is currently unknown how estimates may
181 be biased when ascertainment varies across GWAS cohorts more for one sex than the other.

182 Genomic-relatedness matrix restricted maximum likelihood (GREML)^{3,12} is a major alternative to
183 LDSC that estimates heritability using raw genotypes among unrelated individuals. While LDSC has the
184 advantage of requiring only summary-level data, and is thus especially applicable to GWAS meta-
185 analysis, GREML is often considered preferable when raw data are available,^{13,14} as it is often found to
186 produce larger SNP heritability estimates than those obtained from LDSC.¹⁵ One potential explanation for
187 this discrepancy includes the possibility that, because LDSC is typically applied to meta-analytic GWAS
188 data, it will only detect the portion of heritable signal that is consistent across contributing GWAS
189 datasets. A second potential explanation for this discrepancy is that LDSC may produce attenuated
190 heritability estimates due to discrepancies between LD structure in the reference data used to construct the
191 LD scores and the samples from which the GWAS estimates were derived. While these issues are still
192 likely to be at play, the present findings highlight another source of discrepancy across LDSC and
193 GREML for binary traits that is easily corrected.

194 The discrepancy between a heritability estimate from GWAS data and those reported from
195 family-based (e.g., twin) studies is referred to as missing heritability. Missing heritability is in and of
196 itself often used as a meaningful metric to the extent that it provides a measure of the genetic signal left to
197 uncover, for example, in the rare variant end of the allelic spectrum.¹⁶ Liability scale heritability estimates
198 that are downwardly biased to the degree observed here will likely shift the level of missing heritability
199 for binary outcomes. Indeed, current reports of SNP-based heritability recapturing only 30%-50% of the
200 heritability estimates from twin studies of psychiatric disorders (i.e., 70%-50% missing heritability)¹⁷ may
201 need to be revisited for binary traits from which heritability estimates are derived from GWAS meta-
202 analysis. To this end, we strongly recommend using $\sum EffN$ for liability correction, as we have shown
203 this to produce an unbiased estimate of SNP-based heritability in the population.

204

205

206

207

Method

208 **Simulation of SNP Heritability**

209 *Simulation of Summary Statistics*

210 Each simulation began by generating genome-wide summary statistics for binary traits for 10 individual
211 cohorts. All simulations specified a population prevalence of 1%, a liability scale heritability of 15% in
212 the population, cross-trait intercepts of 0 to reflect no sample overlap across the 10 cohorts, and a
213 univariate intercept of 1.04 to reflect minimal, uncontrolled for population stratification. Each cohort was
214 specified to have a sample prevalence of either 10% (low ascertainment) or 50% (high ascertainment),
215 with the balance of cohorts with low and high ascertainment varying across 11 simulation conditions (see
216 Table 1 for details on each condition). For each condition, 100 sets of 10 cohorts were simulated (i.e.,
217 1,000 cohort level summary statistics per condition for a total of 11,000 simulated cohorts across the 11
218 conditions). Note that when liability scale heritability is equal, but sample prevalence differs across
219 cohorts, observed scale heritability will differ across cohorts. Data were simulated using European
220 population LD scores provided by the original LDSC developers¹⁸ for 1,184,461 HapMap3 SNPs,
221 excluding the MHC region, according to simulation procedures first described in de la Fuente et al.
222 (2021).¹⁹ More specifically, summary statistics were simulated following the multivariate LDSC equation:

$$223 [Z_{1j}, Z_{2j}, \dots, Z_{10j}] \sim N([0, 0, \dots, 0], cov(Z_{1j}, Z_{2j}, \dots, Z_{10j})) \quad (7)$$

224 where

$$225 cov(Z_{1j}, Z_{2j}, \dots, Z_{10j}) =$$

$$226 \begin{bmatrix} N_1 \frac{h_1^2}{M} \ell(j) + 1 + a_1 & & & & \\ \sqrt{N_1 N_2} \frac{\sigma_{g1,2}}{M} \ell(j) + \frac{\rho_{1,2} N_{s1,2}}{\sqrt{N_1 N_2}} & N_2 \frac{h_2^2}{M} \ell(j) + 1 + a_2 & & & \\ \vdots & \vdots & \ddots & & \\ \sqrt{N_1 N_{10}} \frac{\sigma_{g1,10}}{M} \ell(j) + \frac{\rho_{1,10} N_{s1,10}}{\sqrt{N_1 N_{10}}} & \sqrt{N_2 N_{10}} \frac{\sigma_{g2,10}}{M} \ell(j) + \frac{\rho_{2,10} N_{s2,10}}{\sqrt{N_2 N_{10}}} & \dots & N_{10} \frac{h_{10}^2}{M} \ell(j) + 1 + a_{10} \end{bmatrix} \quad (8)$$

227 and $[Z_{1j}, Z_{2j}, \dots, Z_{10j}]$ reflects the Z statistics for the 10 GWAS cohorts (expressed in condensed form, not
228 depicting cohorts 3-9 from the current simulations for display reasons), M is the number of SNPs from the
229 LD file (1,184,461), N_s is the number of overlapping individuals (0), N is the sample size of the individual
230 GWAS (5,000), ρ is the phenotypic correlation (0.5), $\ell(j)$ is the LD score of SNP j , and $a+1$ reflects the
231 univariate LDSC intercept that picks up on unmeasured confounds, such as population stratification
232 (1.04). The bivariate LDSC intercept, expressed as $\frac{\rho_{1,2} N_{s1,2}}{\sqrt{N_1 N_2}}$ for cohorts 1 and 2, was 0 owing to setting

233 the sample overlap (N_s) to 0 for all simulations. GWAS Z-statistics were simulated following the equation
234 above and utilizing the *mvrnorm* R function from the *MASS* package for each SNP.

235 From the simulated cohort-level GWAS Z-statistics, we computed logistic betas as follows:

$$236 \quad b\text{logit}_{k,j} = \frac{Z_{k,j}}{\sqrt{n_k v_k (1-v_k) \sigma_{SNP,j}^2}} \quad (9)$$

237 where v_k and n_k reflects the cohort-specific sample prevalence and sample size, respectively, and $\sigma_{SNP,j}^2$
238 reflects the variance of a given SNP j calculated as $2 * MAF * (1 - MAF)$, where *MAF* is the minor allele
239 frequency. The logistic standard errors for a given SNP j and cohort k were then calculated as:

$$240 \quad SE_{b\text{logit}_{k,j}} = \frac{b\text{logit}_{k,j}}{Z_{k,j}} \quad (10)$$

241 These logistic betas and standard errors were used to calculate the inverse-variance weighted (IVW)
242 meta-analytic beta across the 10 contributing cohorts as described in the **Supplementary Note**. This
243 procedure then produced a single summary statistics file reflecting the meta-analyzed output across the 10
244 simulated cohorts. This summary statistics file was finally analyzed in LDSC in one of two ways, as
245 described in the section below.

246 *Analysis of Simulated Summary Statistics*

247 We compared the ability to recover the population liability scale heritability (h_l^2) for two
248 approaches: 1) the standard procedure of inputting the total sample prevalence (v_{Total}) and the total sample
249 size (N_{Total}) versus, 2) our proposed approach of inputting the sum of effective sample sizes ($\sum EffN$) and
250 a population prevalence (P) of .5 to reflect the fact that the effective sample size equation already
251 accounts for cohort-specific sample ascertainment. For each simulation condition and liability correction
252 approach we report the mean liability scaled heritability estimate, standard deviation across the 100
253 simulations, the range of parameter estimates, and the mean proportional bias relative to the population
254 generating parameter, calculated as:

$$255 \quad Mean \% Bias = \frac{1}{100} \sum_{r=1}^{100} \left(\frac{\widehat{h}_{l,r}^2}{h_{l,True}^2} - 1 \right) \quad (11)$$

256 where $\widehat{h}_{l,r}^2$ is the parameter estimate for a given run, r , and $h_{l,True}^2$ was the population generating value of
257 15%.

258 **Analysis of Real Data**

259 We examined liability scale heritability estimates for publicly available, European only summary
260 statistics for 12 major disorders: attention-deficit hyperactivity disorder (ADHD),²⁰ alcohol dependence
261 (ALCH)¹⁰ Alzheimer's disease (ALZ),²¹ anorexia nervosa (AN),²² autism-spectrum disorder (ASD),⁹
262 bipolar disorder (BIP),² cannabis use disorder (CUD),²³ major depressive disorder (MDD),²⁴ obsessive-
263 compulsive disorder (OCD),²⁵ post-traumatic stress disorder (PTSD),²⁶ schizophrenia (SCZ),²⁷ and
264 Tourette's syndrome (TS).²⁸ For each set of summary statistics, we followed the standard quality control
265 procedure of filtering out SNPs with an imputation quality (INFO) score < .9 and minor allele frequency
266 (MAF) < 1% and filtering to SNPs present in the HapMap3 file excluding the MHC region. In line with
267 prior work for ALZ, we also removed the APOE region prior to calculating heritability. In addition, for
268 ALZ we confirmed that the GERAD consortium was analyzed as a single cohort while the remaining
269 contributing consortiums (ADGC, CHARGE, and EADI) reflected meta-analyzed summary statistics
270 obtained from individual cohorts. Thus, a single $\sum EffN$ was calculated for GERAD while $\sum EffN$ was
271 calculated for each of the contributing cohorts for the other consortiums prior to summing them all
272 together to produce a single $\sum EffN$ for ALZ. For all traits, the liability scale heritability was then
273 calculated using either our proposed approach of inputting $\sum EffN$ or the field standard approach of
274 using v_{Total} . For $\sum EffN$, the SNP-specific sum of effective sample sizes was used when available.
275 Similarly, when using v_{Total} the SNP-specific total sample size was utilized when this information was
276 available. Bias was calculated here as the proportion of the $\sum EffN$ estimate captured by v_{Total} :

$$277 \frac{h_l^2 \text{ Estimate for } v_{Total}}{h_l^2 \text{ estimate for } \sum EffN} - 1 \quad (12)$$

278

279

280

281

282

283

284

285

286

287

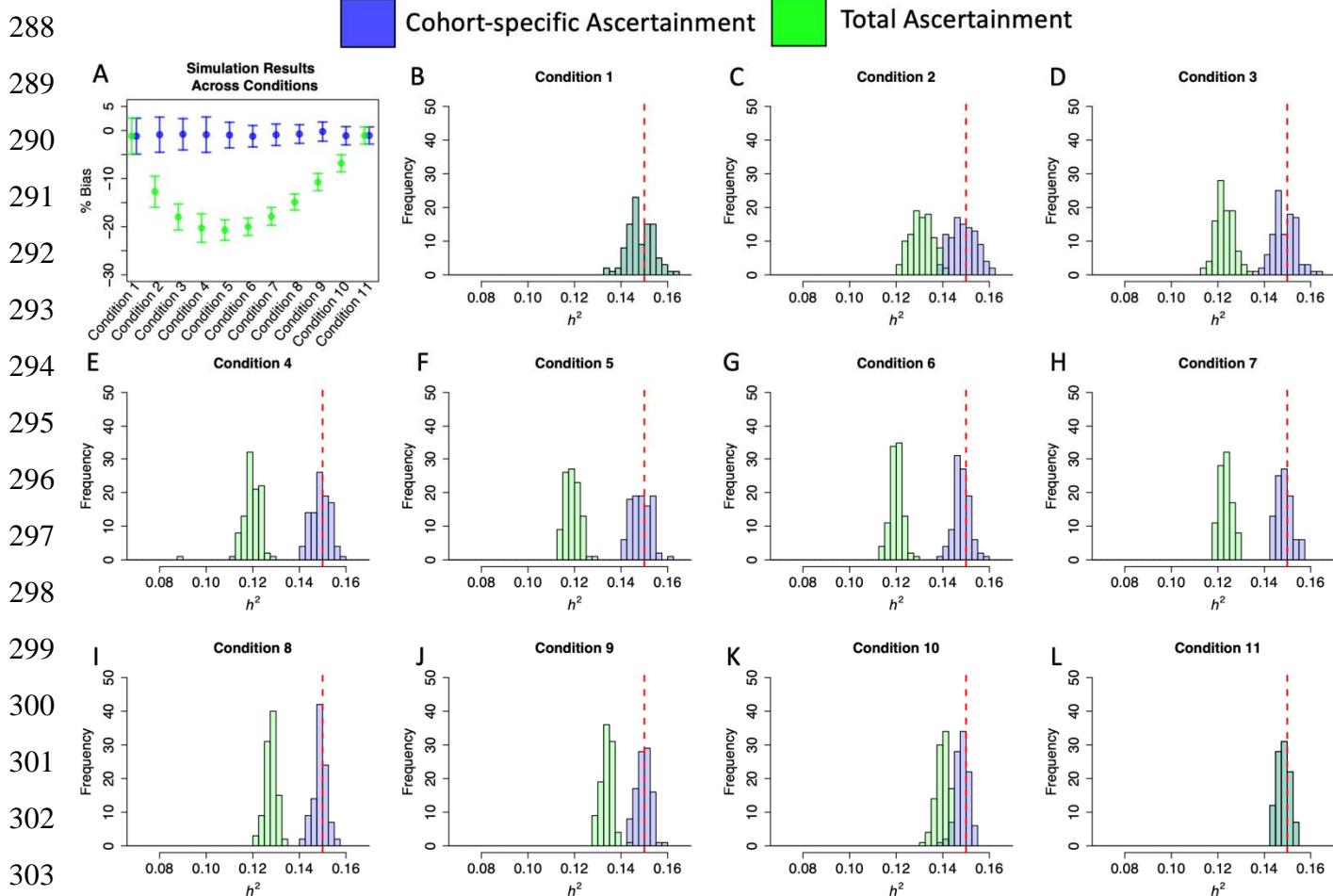


Figure 1. **Simulation Results across Conditions.** Panel A depicts the mean percentage bias on the y-axis across the 11 simulation conditions on the x-axis. Error bars depict +/- 1 SD. Panels B-L depict the individual point estimates from the 100 simulations per condition across the 11 conditions. The red dashed line indicates the liability scale h^2 in the population of 15%. All panels depict in green the results from using v_{Cohort} to account for cohort-specific ascertainment and in blue the results from using v_{Total} .

Table 1. Simulation Results across Conditions

Condition	50% Cases /	10% Cases /	Proposed Approach: $\sum EffN$			Field Standard Approach: v_{Total}		
	50% Controls	90% Controls	<i>Mean h^2 (SD)</i>	<i>h^2 Range</i>	<i>Mean % Bias (SD)</i>	<i>Mean h^2 (SD)</i>	<i>h^2 Range</i>	<i>Mean % Bias (SD)</i>
<i>Condition 1</i>	0	10	14.83 (0.56)	13.31 : 16.33	-1.18 (3.74)	14.83 (0.56)	13.31 : 16.33	-1.18 (3.74)
<i>Condition 2</i>	1	9	14.87 (0.55)	13.78 : 16.09	-0.87 (3.65)	13.09 (0.48)	12.13 : 14.17	-12.72 (3.22)
<i>Condition 3</i>	2	8	14.88 (0.49)	13.76 : 16.37	-0.78 (3.25)	12.30 (0.40)	11.37 : 13.53	-17.99 (2.69)
<i>Condition 4</i>	3	7	14.87 (0.55)	11.02 : 15.95	-0.87 (3.68)	11.96 (0.44)	8.87 : 12.83	-20.27 (2.90)
<i>Condition 5</i>	4	6	14.86 (0.40)	14.09 : 16.02	-0.95 (2.67)	11.89 (0.32)	11.27 : 12.83	-20.71 (2.13)
<i>Condition 6</i>	5	5	14.82 (0.33)	13.91 : 15.77	-1.19 (2.22)	12.00 (0.27)	11.26 : 12.76	-20.01 (1.88)
<i>Condition 7</i>	6	4	14.87 (0.33)	14.27 : 15.62	-0.89 (2.22)	12.32 (0.28)	11.83 : 12.95	-17.85 (1.88)
<i>Condition 8</i>	7	3	14.89 (0.29)	14.03 : 15.68	-0.74 (1.93)	12.77 (0.25)	12.03 : 13.44	-14.90 (1.66)
<i>Condition 9</i>	8	2	14.97 (0.29)	14.26 : 15.93	-0.23 (1.99)	13.39 (0.27)	12.76 : 14.26	-10.71 (1.78)
<i>Condition 10</i>	9	1	14.86 (0.28)	13.99 : 15.38	-1.09 (1.89)	13.97 (0.27)	13.18 : 14.49	-6.82 (1.78)
<i>Condition 11</i>	10	0	14.84 (0.27)	14.25 : 15.41	-1.05 (1.78)	14.84 (0.27)	14.25 : 15.41	-1.05 (1.78)

Note. The **50% cases / 50% controls** and **10% cases / 90% controls** columns denote the total number of cohorts with this case/control split for each condition. The **Proposed Approach** columns denote the simulation results when using $\sum EffN$ for the liability correction. The **Field Standard Approach** columns report simulation results when using v_{Total} .

Table 2. LDSC Heritability Estimates Using Total or Cohort-specific Ascertainment

Trait	Reference	Pop. Prev.	Cases	Controls	Field Standard	Proposed	% Bias
					Approach (v_{Total}): h^2 (SE)	Approach ($\sum EffN$): h^2 (SE)	
ADHD	Demontis <i>et al.</i> (2019) ²⁰	5.0%	19,099	34,194	22.1% (1.4)	23.7% (1.6)	-6.7%
ALCH	Walters <i>et al.</i> (2018) ¹⁰	15.9%	10,206	28,480	10.0% (1.8)	13.9% (2.5)	-28.1%
ALZ	Kunkle <i>et al.</i> (2019) ²¹	4.3%	21,982	41,944	5.8% (0.9)	7.0% (1.1)	-20.7%
AN	Watson <i>et al.</i> (2019) ²²	0.9%	16,992	55,525	13.8% (0.9)	15.5% (1.0)	-10.8%
ASD	Grove <i>et al.</i> (2019) ⁹	1.2%	18,381	27,969	11.7% (1.0)	11.9% (1.0)	-1.3%
BIP	Mullins <i>et al.</i> (2021) ²	2.0%	41,917	371,549	12.8% (0.5)	18.7% (0.8)	-31.8%
CUD	Johnson <i>et al.</i> (2020) ²³	1.0%	14,080	343,726	6.7% (0.6)	7.5% (0.7)	-11.5%
MDD	Wray <i>et al.</i> (2018) ²⁴	15.0%	59,851	113,154	10.2% (0.6)	11.5% (0.7)	-11.8%
OCD	Arnold <i>et al.</i> (2018) ²⁵	2.5%	2,688	7,037	28.5% (4.4)	29.9% (4.6)	-4.7%
PTSD	Nievergelt <i>et al.</i> (2019) ²⁶	30.0%	23,212	151,447	5.3% (0.9)	6.1% (1.1)	-12.9%
SCZ	Ripke <i>et al.</i> (2020) ²⁷	1.0%	53,386	77,258	20.7% (0.7)	22.3% (0.8)	-6.9%
TS	Yu <i>et al.</i> (2019) ²⁸	0.8%	4,819	9,488	21.5% (2.5)	22.4% (2.6)	-4.0%

Note. The **Field Standard Approach** column report liability scale heritability results when using v_{Total} for ascertainment correction. The **Proposed Approach** column reports results when using $\sum EffN$ for ascertainment correction. Liability scale heritabilities were calculated using the SNP-specific total sample sizes, or the SNP-specific $\sum EffN$, when possible. Supplementary Table 1 reports the original heritabilities from the corresponding publication along with heritabilities calculated using both the SNP-specific or total sample sizes for reference. Population prevalences were chosen based on the prevalence reported in the original publication for that trait for comparability purposes. % Bias was calculated as the proportional attenuation in heritability estimates when using total over cohort-specific ascertainment to perform the liability correction; this was calculated using the direct output from LDSC and, therefore, will not exactly match the numbers obtained from using the liability h^2 reported in the table due to rounding. Results are reported for GWAS data that are strictly publicly available (e.g., 23andMe data are not included for MDD). h^2 = LDSC liability scale heritability estimate; SE = standard error; ADHD = attention-deficit/hyperactivity disorder; ALCH = alcohol use disorder; ALZ = Alzheimer's disease; AN = anorexia nervosa; ASD = autism spectrum disorder; BIP = bipolar disorder; CUD = cannabis use disorder; MDD = major depressive disorder; OCD = obsessive compulsive disorder; PTSD = post-traumatic stress disorder; SCZ = schizophrenia; TS = Tourette's Syndrome.

Acknowledgements

ADG was supported by NIH Grants R01MH120219 and RF1AG073593. EMTD was supported by NIH grants R01MH120219 and RF1AG073593 and the Jacobs Foundation. EMTD is a faculty associate of the Population Research Center at the University of Texas, which is supported by NIH grant P2CHD042849. MGN is additionally supported by ZonMW grants 849200011 and 531003014 from The Netherlands Organisation for Health Research and Development, a VENI grant awarded by NWO (VI.Veni.191G.030), NIH grant R01MH120219 and is a Jacobs Foundation Fellow. J.F. is member of the Population Research Center (PRC) and Center on Aging and Population Sciences (CAPS) at The University of Texas at Austin, which are supported by National Institutes of Health (NIH) grants P2CHD042849 and P30AG066614, respectively.

References

1. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature genetics* 47, 291 (2015).
2. Mullins, N. *et al.* Genome-wide association study of more than 40,000 bipolar disorder cases provides new insights into the underlying biology. *Nat Genet* 53, 817–829 (2021).
3. Lee, S. H., Wray, N. R., Goddard, M. E. & Visscher, P. M. Estimating Missing Heritability for Disease from Genome-wide Association Studies. *Am J Hum Genetics* 88, 294–305 (2011).
4. Dempster, E. R. & Lerner, I. M. Heritability of Threshold Characters. *Genetics* 35, 212–36 (1950).
5. Peyrot, W. J., Boomsma, D. I., Penninx, B. W. J. H. & Wray, N. R. Disease and Polygenic Architecture: Avoid Trio Design and Appropriately Account for Unscreened Control Subjects for Common Disease. *Am J Hum Genetics* 98, 382–391 (2016).
6. Grotzinger, A. D. *et al.* Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. *Nature human behaviour* 3, 513 (2019).
7. Turley, P. *et al.* Multi-trait analysis of genome-wide association summary statistics using MTAG. *Nature genetics* 1 (2018).
8. Speed, D. *et al.* Reevaluation of SNP heritability in complex human traits. *Nat Genet* 49, 986–992 (2017).
9. Grove, J. *et al.* Identification of common genetic risk variants for autism spectrum disorder. *Nature genetics* 51, 431–444 (2019).
10. Walters, R. K. *et al.* Transancestral GWAS of alcohol dependence reveals common genetic underpinnings with psychiatric disorders. *Nature Neuroscience* 21, 1656 (2018).
11. Turley, P. *et al.* Multi-trait analysis of genome-wide association summary statistics using MTAG. *Nature genetics* 50, 229 (2018).
12. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. *The American Journal of Human Genetics* 88, 76–82 (2011).
13. Maier, R. M., Visscher, P. M., Robinson, M. R. & Wray, N. R. Embracing polygenicity: a review of methods and tools for psychiatric genetics research. *Psychol Med* 48, 1055–1067 (2018).
14. Akingbuwa, W. A., Hammerschlag, A. R., Bartels, M. & Middeldorp, C. M. Systematic Review: Molecular Studies of Common Genetic Variation in Child and Adolescent Psychiatric Disorders. *J Am Acad Child Adolesc Psychiatry* (2021) doi:10.1016/j.jaac.2021.03.020.
15. Yang, J. *et al.* Genome-wide genetic homogeneity between sexes and populations for human height and body mass index. *Hum Mol Genet* 24, 7445–7449 (2015).
16. Manolio, T. A. *et al.* Finding the missing heritability of complex diseases. *Nature* 461, 747–753 (2009).
17. Baselmans, B. M., Yengo, L., Rheenen, W. V. & Wray, N. R. Risk in relatives, heritability, SNP-based heritability and genetic correlations in psychiatric disorders: a review. *Biological Psychiatry* (2020).
18. Bulik-Sullivan, B. K. *et al.* An atlas of genetic correlations across human diseases and traits. *Nature genetics* 47, 1236–1241 (2015).
19. de la Fuente, J., Grotzinger, A. D., Marioni, R. E., Nivard, M. G. & Tucker-Drob, E. M. Multivariate Modeling of Direct and Proxy GWAS Indicates Substantial Common Variant Heritability of Alzheimer’s Disease. *Medrxiv* 2021.05.06.21256747 (2021) doi:10.1101/2021.05.06.21256747.

20. Demontis, D. *et al.* Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nature genetics* 51, 63 (2019).
21. Kunkle, B.W. *et al.* Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates A β , tau, immunity and lipid processing. *Nat Genet* 51, 414–430 (2019).
22. Watson, H. J. *et al.* Genome-wide association study identifies eight risk loci and implicates metabo-psychiatric origins for anorexia nervosa. *Nature genetics* 51, 1207–1214 (2019).
23. Johnson, E. C. *et al.* A large-scale genome-wide association study meta-analysis of cannabis use disorder. *Lancet Psychiatry* 7, 1032–1045 (2020).
24. Wray, N. R. *et al.* Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nature genetics* 50, 668 (2018).
25. Arnold, P. D. *et al.* Revealing the complex genetic architecture of obsessive-compulsive disorder using meta-analysis. *Molecular ...* 23, 1181 (2018).
26. Nievergelt, C. M. *et al.* International meta-analysis of PTSD genome-wide association studies identifies sex-and ancestry-specific genetic risk loci. *Nature Communications* 10, 1–16 (2019).
27. Ripke, S., Walters, J. T. & O'Donovan, M. C. Mapping genomic loci prioritises genes and implicates synaptic biology in schizophrenia. *medRxiv* 2020.09.12.20192922 (2020)
doi:10.1101/2020.09.12.20192922.
28. Yu, D. *et al.* Interrogating the genetic determinants of Tourette's syndrome and other tic disorders through genome-wide association studies. *American Journal of Psychiatry* 176, 217–227 (2019).