Published in final edited form as:

Hum Genet. 2013 December ; 132(12): . doi:10.1007/s00439-013-1334-z.

How meaningful are heritability estimates of liability?

Penny H. Benchek and Nathan J. Morris

Department of Epidemiology and Biostatistics, School of Medicine, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, OH 44106, USA

Penny H. Benchek: pennybenchek@gmail.com; Nathan J. Morris: njm18@case.edu

Abstract

It is commonly acknowledged that estimates of heritability from classical twin studies have many potential shortcomings. Despite this, in the post-GWAS era, these heritability estimates have come to be a continual source of interest and controversy. While the heritability estimates of a quantitative trait are subject to a number of biases, in this article we will argue that the standard statistical approach to estimating the heritability of a binary trait relies on some additional untestable assumptions which, if violated, can lead to badly biased estimates. The ACE liability threshold model assumes at its heart that each individual has an underlying liability or propensity to acquire the binary trait (e.g., disease), and that this unobservable liability is multivariate normally distributed. We investigated a number of different scenarios violating this assumption such as the existence of a single causal diallelic gene and the existence of a dichotomous exposure. For each scenario, we found that substantial asymptotic biases can occur, which no increase in sample size can remove. Asymptotic biases as much as four times larger than the true value were observed, and numerous cases also showed large negative biases. Additionally, regions of low bias occurred for specific parameter combinations. Using simulations, we also investigated the situation where all of the assumptions of the ACE liability model are met. We found that commonly used sample sizes can lead to biased heritability estimates. Thus, even if we are willing to accept the meaningfulness of the liability construct, heritability estimates under the ACE liability threshold model may not accurately reflect the heritability of this construct. The points made in this paper should be kept in mind when considering the meaningfulness of a reported heritability estimate for any specific disease.

Introduction

Heritability may be defined as the proportion of the total variability in an observed trait attributable to genetic effects in a given population at a particular point in time (Lewontin 1974; Visscher et al. 2008). It serves as a measure of the extent to which a disease or trait is genetically determined. Heritability studies have heavily influenced our understanding of disease etiology by showing us evidence that numerous diseases and traits are largely genetic in nature. However, heritability studies have also been at the center of numerous historical controversies about nature vs. nurture (Plomin and Spinath 2004; Galton 1869; Kendler et al. 1993; Pam et al. 1996; Joseph 2000). Most recently there has been much dismay and some controversy regarding the amount of heritability explained by genome wide association studies (GWAS) (Goldstein 2009; Eichler et al. 2010; Maher 2008; Manolio et al. 2009; Zuk et al. 2012). Still, before we bemoan or applaud our ability to explain heritability, we should be confident about the heritability estimates we are adopting. We may be trying to explain heritability that does not exist or we may believe that we have

adequately explained heritability when indeed there is much more. While there are numerous designs for studying heritability, here we focus on the classic twin study design. However, our work is broadly applicable to other study designs that utilize a liability model.

Studying how a trait aggregates in families offers insight regarding the genetic basis of a trait. However, simply observing an increased occurrence of a trait in families does not tell us whether that increase is due to genes or due to some aspect of the family's common environment. Classic twin studies (Boomsma et al. 2002) involving monozygotic (MZ) and dizygotic (DZ) twins are one common approach to disentangling the effects due to genes from those due to the common environment. There are numerous well-known limitations to twin studies. The main limitations stem from the following assumptions: MZ and DZ twin pairs share their environments to the same extent; gene-environment correlations and interactions are minimal for the trait; random mating occurs in the population; and twins are no different from the general population in terms of the trait (Rijsdijk and Sham 2002). Newer research has also raised the possibility that MZ twins are not even truly identical (Baranzini et al. 2010; Bruder et al. 2008). Many researchers are aware of these limitations but still believe that twin studies provide a valuable source of information about heritability. Here, we will assume for the sake of argument that the standard twin study assumptions are reasonable. Unfortunately, when dealing with binary traits (e.g., affected vs. unaffected), there is an additional set of assumptions, which are commonly used. In this article, we show that even moderate deviations from the most fundamental of these additional assumptions may lead to wildly incorrect results.

One of the most well-known approaches to analyzing binary trait twin data is to use a liability threshold model (Falconer 1965). Curnow (1972) has pointed out that the threshold model is mathematically equivalent to a probit-like risk model. Liability threshold models assume that each individual has a hypothetical continuous liability composed of latent genetic and environmental factors. In a liability model, the observed binary trait (Y) is related to the continuous liability (Z) by way of a threshold (T). That is,

$$Y = \begin{cases} 1 & \text{if } Z > T \\ 0 & \text{otherwise} \end{cases}.$$

Thus, the probability that an individual is affected is the probability that their liability exceeded some threshold. Put more mathematically: $P(Y = 1) = P(Z > T)$, where Z follows some specified liability distribution. Because the liabilities for two twins are dependent, the joint distribution of each twin pair type (i.e., MZ or DZ) must be specified. Even if we are willing to accept the meaningfulness of the liability construct, the validity of heritability estimates from the liability model is dependent on correctly specifying the underlying distribution of the liability. For example, Kidd and Cavalli-Sforza (1973) noted that very different heritability estimates of schizophrenia could be obtained by assuming a single underlying gene or by assuming a polygenic model (Elston 1977).

The ACE liability model (Neale and Cardon 1992) is the most frequently used liability model. The ACE model holds that the liability is made up of three components: additive genetic factors 'A', common environmental factors 'C', and random environmental factors 'E'. It is assumed that the interactions between the components, as well as factors within the components, are minimal. Each of the factors within each component is believed to have a small and additive effect on the liability; thus, each component is assumed to follow a normal distribution. Because the total liability is equal to the summation of the three normally distributed components, it follows that a single individual's total liability under the

ACE model is also normally distributed; that is, $Z \sim N(\mu, \sigma_A^2 + \sigma_C^2 + \sigma_E^2)$, where the total variability in liability is partitioned according to the variability explained by each of the components. Additionally, the joint liability distribution for twin pairs under the ACE model is assumed to be multivariate normal (MVN). The joint twin pair distribution under the ACE model is

$$\begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} = \text{MVN} \left(\begin{bmatrix} \mu \\ \mu \end{bmatrix}, \begin{bmatrix} 1 & \phi\sigma_A^2 + \sigma_C^2 \\ \phi\sigma_A^2 + \sigma_C^2 & 1 \end{bmatrix} \right),$$

where the expected proportion of shared genes (ϕ) is set to 1 for MZ twins and 1/2 for DZ twins, the mean liability (μ) is set to 0, and total variability in liability (σ_T^2) is set to 1 (i.e., $\sigma_T^2 = \sigma_A^2 + \sigma_C^2 + \sigma_E^2 = 1$). Assigning values to the mean (μ) and total variance (σ_T^2) is required to fix the scale for the liability, avoiding the inherent problem of identifiability (Huang 2005). It does not represent a problem because we are only interested in the proportion of the total variability that each component explains. Curnow (1972) points out that there always exists a transformation which would make the underlying liability model marginally normal. Unfortunately, marginal normality does not imply multivariate normality.

We have no assurance that the MVN liability assumption matches reality. We cannot test the assumption of a MVN liability because the liability and its contributing factors are unobservable. Therefore, if we are to believe the heritability estimates derived from the ACE liability model, we should know just how robust the model is to violations of the MVN assumption.

Methods

Heritability estimation assuming the ACE model

Maximum likelihood is one relatively straightforward approach to estimating heritability using MZ and DZ twin data under the ACE liability model (Rijsdijk and Sham 2002). In this

approach, an estimate ($\hat{\theta}_{\text{ML}}$) is achieved by finding the parameters that maximize the log-likelihood function based on the multinomial distribution,

$$\log L(\theta | \hat{\mathbf{p}}, \mathbf{n}) = \sum_j n_j \sum_i \hat{p}_{ij} \ln(p_{ij}(\theta)), \quad (1)$$

where $\theta = (\sigma_A^2, \sigma_C^2, \tau_{\text{MZ}}, \tau_{\text{DZ}})$ are the parameters to be estimated; $j \in \{\text{MZ}, \text{DZ}\}$ defines twin zygosity; $i \in \{0, 1, 2\}$ indicates the number of affected individuals in the twin pair; n_j is the number of MZ or DZ twin pairs; $\hat{p}_{ij}$ is the proportion of twin pairs with i affected individuals among the twins with zygosity j ; and $p_{ij}(\theta)$ is the probability that a pair of twins will have i affected individuals given that they have zygosity j . The $p_{ij}(\theta)$ in Eq. (1) is calculated by numerically integrating over the joint twin pair distribution under the ACE model. The threshold values (τ_{MZ} and τ_{DZ}) are used as upper or lower limits of integration, depending on which $p_{ij}(\theta)$ is being calculated. These thresholds are determined by the prevalence of the trait under the standard normal distribution. Here, we set the prevalence to be the same for both MZ and DZ twins.

Asymptotic Bias in the ACE model estimator when the assumptions are not met

If the ACE liability model is “true,” then, as the sample size becomes large, the maximum likelihood estimate should become very close to the “true” value. Put more mathematically, under standard conditions, the estimate converges in probability to the truth: $\hat{\theta}_{ML} \xrightarrow{P} \theta_{true}$. Of more interest is the question: what happens for large samples when the ACE model is “false”? This question can easily be answered using the tools of theoretical statistics (White 1982) as discussed in the “Appendix”. Using those results, we were able to calculate the large sample bias in the heritability estimates.

Under the ACE model, each of the three components contributing to the liability (i.e., additive genetic, common environmental, and random environmental) is assumed to follow a normal distribution. Thus, if any of the components deviate from the assumed normal distribution, then the MVN assumption is likely to be violated. We investigated the biases of the ACE liability model maximum likelihood estimator under the following hypothetical scenarios:

1. The true liability consists of common environmental and random effects that are normally distributed; however, the additive genetic effect is due to a single diallelic gene with three distinct genotypes.
2. The true liability consists of additive genetic and random effects that are normally distributed; however, the common environmental effect follows a Bernoulli distribution due to a single dichotomous common environmental exposure (e.g., secondhand smoke), where both twins are exposed or neither twin is exposed.
3. The true liability consists of additive genetic and random effects that are normally distributed; however, the common environmental effects follow a t distribution.
4. The true liability consists of additive genetic and common environmental effects that are normally distributed; however, the random effects follow a mixture of normal distributions, such that the marginal probability distribution of each twin is similar to the distribution of systolic blood pressure (SBP) found in NHANES data (Centers for Disease Control and Prevention (CDC) 2009).

When we refer to the “true” liability above, we are referring to the data generative model as opposed to the data analysis model. The first three scenarios were chosen to represent a wide spectrum of deviations from the component normality assumption. The first two are relatively strong deviations from normality while the third scenario is relatively subtle. The fourth scenario was chosen as an example of a real trait that does not follow a normal distribution. The bias in percentage heritability was calculated as

$$\frac{\sigma_A^2(\hat{\theta}_{ML}) - \sigma_A^2(\theta_{true})}{\sigma_A^2(\theta_{true})} \times 100, \quad (2)$$

where $\sigma_A^2(\theta_{true})$ represents the true additive genetic variance of the liability and $\sigma_A^2(\hat{\theta}_{ML})$ represents the value to which the maximum likelihood estimate will converge. We calculated the bias in heritability according to Eq. (2) with fixed parameter values σ_A^2 and σ_C^2 across varying scenario specific parameter values (e.g., prevalence, gene frequency, exposure probability and degrees of freedom). Threshold values for trait liability were estimated assuming known population prevalence of the trait.

Non-asymptotic bias in the ACE model estimator when the assumptions are met

We further explored bias in the ACE model estimator through simulations derived under the ACE model. Each sample consisted of an equal mix of MZ and DZ twin pairs. For example, a sample size of 100 consists of 50 MZ twin pairs and 50 DZ twin pairs. Here we fixed sample size at 100, 500, 1,000, and 10,000. For each sample size, we fixed trait prevalence at 0.10, 0.20 and 0.30. For each of the combinations of sample size and prevalence, we fixed parameter values σ_A^2 and σ_C^2 to reflect high heritability/low common environment and vice versa (i.e., $\sigma_A^2=0.60$, $\sigma_C^2=0.15$ and $\sigma_A^2=0.15$, $\sigma_C^2=0.60$) We ran 5,000 simulations based on each of these configurations and calculated bias as in Eq. (2). Here it is not assumed that the population prevalence is known, thus threshold values are determined by the sample prevalence under each simulated data set.

Results

Asymptotic bias in the ACE model estimator when a single causal diallelic gene is present

For the scenario where the additive genetic effects do not follow a normal distribution due to the presence of a single causal diallelic gene (Fig. 1), we mostly see that heritability is underestimated. However, under the combination of low gene frequency and low prevalence (Fig. 1a), we see an overestimation of heritability (e.g., for gene frequency and prevalence both at 0.025, the bias is 84 %). We also see that at low gene frequencies, the magnitude of bias is generally the greatest. This is most pronounced at the highest prevalence levels, for example (Fig. 1d), for gene frequency = 0.025 and prevalence = 0.50, the bias is -89 %. The same is true, although to a lesser extent, for low prevalence. The effect of low prevalence is most pronounced at the highest and lowest gene frequencies. Bias is generally absent for the combination of higher prevalence and higher gene frequencies (Fig. 1a-c). However, this unbiased region decreases as the true heritability increases (Fig. 1d).

Asymptotic bias in the ACE model estimator when a single dichotomous common environmental exposure is present

Consider the scenario where the common environmental effects follow a Bernoulli distribution (Fig. 2). Under the combination of low exposure probability and low prevalence, we see an overestimation of heritability (Fig. 2a, b). For example, for probability and prevalence both at 0.025, the bias is 231 % (Fig. 2a). We also see that for either low probability or low prevalence, the magnitude of bias is greatest. The effect of low prevalence is most pronounced at the highest probabilities and the effect of low probability is most pronounced at the highest prevalence values. For example (Fig. 2a), for probability = 0.025 and prevalence = 0.50, the bias is 381 %. Bias is generally absent for the combination of higher prevalence and higher exposure probabilities (Fig. 2d). However, this unbiased region decreases as the true variability due to common environmental effects increases (Fig. 2a-c).

Asymptotic bias in the ACE model estimator when the common environmental effects follow a t distribution

In the scenario where the common environmental effects follow a t distribution (Fig. 3), the magnitude of bias is greatest for the lowest degrees of freedom. This is especially true under the combination of low degrees of freedom and low or high prevalence. As an example of the combination of low degrees of freedom and low prevalence (Fig. 3a), we see that with $df = 3$ and prevalence = 0.025, the bias is -60 %. As an example of the combination of low degrees of freedom and high prevalence (Fig. 3a), we see that with $df = 3$ and prevalence = 0.50, the bias is 76 %. Bias is generally absent for the combination of higher prevalence and higher degrees of freedom.

Asymptotic bias in the ACE model estimator when the random environmental effects follow a mixture of normal distributions

We looked at systolic blood pressure (SBP) for males 55+, obtained from the NHANES (2009–2010) datasets (Centers for Disease Control and Prevention (CDC) 2009). Adjustments for treated blood pressure were made by adding 10 mmHg to the corresponding SBP values (Cui et al. 2003). The resulting kernel density plot (Fig. 4a) and QQ plot (Fig. 4b) suggest that the data deviate from normality, particularly in the right tail of the distribution. Therefore, we derived a mixture model of four normal distributions using the R package ‘mixtools’ (Benaglia et al. 2009). Comparing the empirical CDF to the CDF of the mixture model, we see that the mixture model fits the data well (Fig. 4c). Although we used the mixture model, numerous other models would fit the data equally well. We simply wanted to select a distribution that mimics as closely as possible the true distribution of SBP for males 55+. We investigated the situation where the continuous SBP variable was converted to a binary hypertension trait using a cutoff value. If the genetic and common environmental components are normally distributed while the random component involves a mixture of normal distributions, then it is straightforward to calculate the asymptotic bias.

For this scenario, where the random environmental effects follow a mixture model distribution, we looked at two fixed combinations for σ_A^2 and σ_C^2 . The first was with $\sigma_A^2=0.4$ and $\sigma_C^2=0.1$ the second was with $\sigma_A^2=0.3$ and $\sigma_C^2=0.2$. These choices of heritability reflect current estimates of the heritability of blood pressure (Cowley et al. 2012). Both fixed parameter combinations showed similar bias (Fig. 4). The magnitude of bias is greatest for the lowest and highest cutoff values ($\approx 60\%$). Under high cutoff values, the ACE model estimator tends to underestimate heritability and under low cutoff values heritability tends to be overestimated. For the first combination, the bias estimates for commonly used cutoff values (120, 140, and 160) were 48, 29, and -14% , respectively. For the second combination, the corresponding bias estimates were 50, 32, and -11% . These cutoff values represent the boundaries between prehypertension, stage 1 hypertension, and stage 2 hypertension, respectively, as defined by the American Heart Association (Kurtz et al. 2005).

Non-asymptotic bias in the ACE model estimator when the assumptions are met

For the ACE model simulated data, heritability tends to be underestimated for $\sigma_A^2=0.60$ and $\sigma_C^2=0.15$ (Fig. 5a) and tends to be overestimated for $\sigma_A^2=0.15$ and $\sigma_C^2=0.60$ (Fig. 5b). Bias is greatest for the lowest sample size ($n=100$) and the lowest prevalence (prevalence = 0.10) (Fig. 5). For $\sigma_A^2=0.60$ and $\sigma_C^2=0.15$ and sample size 100, the magnitude of bias decreases as trait prevalence increases, however, not substantially (i.e., prevalence = 0.10, 0.20, and 0.30 gives bias = $-24, -19, -16\%$, respectively). For $\sigma_A^2=0.15$ and $\sigma_C^2=0.60$, the magnitude of bias is even higher for sample size of 100 (i.e., 81, 67, and 61 %, respectively). In general, bias decreases as either sample size or prevalence increases. Prevalence at 0.10 shows the highest bias for sample sizes 100, 500, and 1,000. Surprisingly, bias is still at 16 % for $\sigma_A^2=0.15$ and $\sigma_C^2=0.60$ and sample size 1,000.

Discussion

Here we have focused on the ACE liability model to calculate the heritability of binary traits. There have been other recent efforts, such as GCTA (Lee et al. 2011), to calculate heritability of binary traits based on the liability model utilizing GWAS data. However, these efforts have distinctly different goals from those of the ACE liability model. They attempt to estimate the heritability explained by SNPs as opposed to the total heritability.

While we have not investigated such methods in this work, we suspect that the concerns we have raised about the ACE liability model will have some relevance to threshold modeling approaches using large pedigrees as well as approaches such as GCTA.

Our intention in selecting the scenarios describing asymptotic bias was to provide potent examples of the consequences of both “extreme” and “subtle” deviations from the MVN assumption. It is clear that a single gene or single common environmental exposure, where many genes and many common environmental exposures are assumed, is representative of an “extreme” deviation. It is also clear that a t distributed common environment, when normality is assumed, is representative of a “subtle” deviation, as the t distribution and the normal distribution are quite similar. While we refer here to “extreme” scenarios, we do not mean that these scenarios are unrealistic. For example, scenario 1 is not only representative of a large number of single gene Mendelian disorders, but it also represents situations with multiple variants inherited together on a single haplotype. Likewise, diseases that are caused by many rare variants will behave similarly to our diallelic model because it is unlikely that any given family will have more than one disease variant segregating within it.

For the first three scenarios examined, we saw that heritability estimates, under many parameter combinations, were highly biased. In each of these scenarios, we also observed regions where heritability was estimated well. Among the regions where bias was substantial, the direction of bias was not consistent across parameter combinations. Under some parameter combinations, heritability was overestimated and for others it was underestimated. Bias increased as the variance of the non-normally distributed component (σ_A^2 or σ_C^2) increased. This is intuitive as the ACE model assumed that each of the three components is normally distributed. Therefore, the greater the proportion of the total variability the non-normally distributed component explains, the greater the deviation from the ACE model's assumption of a MVN distribution, and subsequently the greater the bias.

Under the two exemplary scenarios of “severe” deviations, we saw that absolute bias was highest for the lowest parameter values. Even in the case where the proportion of the variance attributable to the associated component is low, we saw that the combination of low prevalence and low gene frequency or low probability of exposure created substantial bias.

Of the three scenarios reviewed, the t distributed common environment scenario was the subtlest in its deviation from the MVN assumption. Even with this subtlety, for low degrees of freedom, the bias was substantial, particularly under low trait prevalence. This effect was most pronounced in the case where the proportion of the variability explained by the common environment was high, yet was still significant when the proportion explained was low.

The fourth scenario was selected as a realistic example of an observable trait that does not follow a normal distribution. We designed the scenario to accurately reflect the non-normal distribution of SBP found in NHANES data (Centers for Disease Control and Prevention (CDC) 2009). It is not uncommon to define a binary hypertension trait based on a cut off value. Thus, our example is not entirely contrived. For our chosen scenario, biased estimates occurred for most cutoff values, with the magnitude and sign dependent on the cutoff.

The underlying biases found in the ACE model simulated results showed that sample size and prevalence level are a concern even when ACE model assumptions are met. Although these biases can be controlled for through increased sample size, the increase in sample size necessary to eliminate bias is dependent on trait prevalence. In cases where the trait prevalence is very low, the necessary increase in sample size to avoid bias may not be obtainable.

Consequently, individuals are often ascertained based on their trait status, which creates a whole new set of biases to be dealt with. However, for small sample sizes, we would also point out that the confidence intervals for the heritability estimates from the ACE liability model are often quite large in practice and may dwarf these biases.

It should be noted that some of the concerns about bias in the second scenario might be avoided if the environmental factor is measured and adjusted for. However, there are likely to be many diseases for which there are rare environmental factors with large effect sizes. Furthermore, reality may be even more complex, because many “environmental” exposures such as smoking may actually encode some genetic information (Boardman et al. 2010) which we would not necessarily want to adjust out. The scenarios that we have presented are only representative of the numerous possible complexities that are likely to underlie real diseases.

As we have argued in this report, even if we are willing to accept the meaningfulness of the liability construct, the MVN assumption is a serious weakness of the ACE liability threshold model. Many diseases have known highly penetrant alleles and have very strong environmental risk factors such as smoking. Such diseases will clearly be unlikely to have an underlying liability that adequately satisfies the MVN assumption. By definition, the liability is unobservable and thus the MVN assumption cannot be easily checked. The latency of the components combined with the lack of robustness to the MVN assumption should give pause to consider the meaningfulness of heritability estimates obtained under the ACE liability model for any given disease.

Acknowledgments

We thank Dr. Robert Elston for his helpful comments and suggestions. This work was supported by National Cancer Institute (NCI) Grant R25 CA094186 and by National Heart Lung and Blood Institute (NHLBI) Grant T32 HL007567.

Appendix

Consider the function $f(\hat{\mathbf{p}}) = \arg \max_{\theta} \{\log L(\theta | \hat{\mathbf{p}}, \mathbf{n})\}$, which relates the estimated proportions of twin pairs, with (0,1, or 2) affected in each pair, to the maximum likelihood parameter estimates (i.e., $\hat{\theta}_{ML} = f(\hat{\mathbf{p}})$). As indicated by the notation, $f(\cdot)$ is not a function of $\mathbf{n}$ because the model is saturated. By the law of large numbers, the estimated proportions converge in probability to the true proportions: $\hat{\mathbf{p}} \xrightarrow{P} \mathbf{p}^{true}$. Then, assuming f is continuous at $\mathbf{p}_0$, by Slutsky's Continuity Theorem (Ferguson 1996) we have, $\hat{\theta}_{ML} \xrightarrow{P} f(\mathbf{p}^{true})$. Therefore, even in the case where the model is falsely specified, the maximum likelihood estimate will become very close to a function of the true proportions: $f(\mathbf{p}^{true})$. If the model is correctly specified, then the function of the true proportions will equal the true parameter values: $f(\mathbf{p}^{true}) = \theta_0$, otherwise it may be a biased estimate of the parameter values: $f(\mathbf{p}^{true}) \neq \theta_0$. In all of the scenarios discussed, it is easy to calculate the “true” proportions ($\mathbf{p}^{true}$) using numerical integration. It is also possible to calculate the function $f(\mathbf{p}^{true})$ using numerical optimization techniques combined with numerical integration. Thus, the large sample properties of the estimator under the ACE liability model may be investigated by statistical theory without the use of simulations. We calculated the bivariate normal integral using the method of Genz and Bretz (2002) as implemented in the R package “mvtnorm,” and we performed numerical integration over the t distribution using the adaptive quadrature (Piessens et al. 1983) “integrate” function in R (R Team 2010). We performed numerical optimization using the “optim” function in R (R Team 2010). Also, we calculated the thresholds using the R function “uniroot” (R Team 2010).

As an example of calculating $\mathbf{p}^{\text{true}}$, consider the second scenario with a dichotomous common exposure where the exposure has frequency γ . Let $E \in \{0,1\}$ represent the exposure status of a pair. Consider the probability that neither of the twins is affected:

$$\mathbf{p}_{0j}^{\text{true}} = \gamma P(Z_1 < \tau_j, Z_2 < \tau_j | E=1) + (1 - \gamma) P(Z_1 < \tau_j, Z_2 < \tau_j | E=0). \quad (3)$$

Similar formulas to (3) can be written for $\mathbf{p}_{1j}^{\text{true}}$ and $\mathbf{p}_{2j}^{\text{true}}$, so the probability of 0, 1 and 2 affected in a twin pair can be calculated. If we allow $\beta = \sqrt{\sigma_c^2 / \gamma(1 - \gamma)}$ to represent the effect size for the exposure, then variance of the common environment component is $\text{Var}(\beta E) = \sigma_c^2$. Furthermore, conditional on E we have

$$\begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} = \text{MVN} \left(\begin{bmatrix} \beta E \\ \beta E \end{bmatrix}, \begin{bmatrix} 1 & \phi \sigma_A^2 \\ \phi \sigma_A^2 & 1 \end{bmatrix} \right). \quad (4)$$

Thus, Eq. (4) may be used in conjunction with Eq. (3) to calculate $\mathbf{p}^{\text{true}}$. Very similar equations may be created for the other scenarios discussed in the paper.

References

- Baranzini SE, Mudge J, Van Velkinburgh JC, Khankhanian P, Khrebtukova I, Miller NA, Zhang L, Farmer AD, Bell CJ, Kim RW. Genome, epigenome and RNA sequences of monozygotic twins discordant for multiple sclerosis. *Nature*. 2010; 464:1351–1356. [PubMed: 20428171]
- Benaglia T, Chauveau D, Hunter DR, Young DS. mixtools: An R package for analyzing mixture models. *J Stat Softw*. 2009; 32(6):1–29.
- Boardman JD, Blalock CL, Pampel FC. Trends in the genetic influences on smoking. *J Health Soc Behav*. 2010; 51:108–123. [PubMed: 20420298]
- Boomsma D, Busjahn A, Peltonen L. Classical twin studies and beyond. *Nat Rev Genet*. 2002; 3:872–882. [PubMed: 12415317]
- Bruder CEG, Piotrowski A, Gijsbers AACJ, Andersson R, Erickson S, Diaz de Ståhl T, Menzel U, Sandgren J, von Tell D, Poplawski A. Phenotypically concordant and discordant monozygotic twins display different DNA copy-number-variation profiles. *Am J Human Genet*. 2008; 82:763–771. [PubMed: 18304490]
- National Center for Health Statistics (NCHS). National Health and Nutrition Examination Survey Data. Hyattsville, MD: U.S. Department of Health and Human Services; 2009–2010. Centers for Disease Control and Prevention (CDC).
- Cowley AW Jr, Nadeau JH, Baccarelli A, Berecek K, Fornage M, Gibbons GH, Harrison DG, Liang M, Nathanielsz PW, O'Connor DT. Report of the National Heart, Lung, and Blood Institute Working Group on epigenetics and hypertension. *Hypertension*. 2012; 59:899–905. [PubMed: 22431584]
- Cui JS, Hopper JL, Harrap SB. Antihypertensive treatments obscure familial contributions to blood pressure variation. *Hypertension*. 2003; 41:207–210. [PubMed: 12574083]
- Curnow R. The multifactorial model for the inheritance of liability to disease and its implications for relatives at risk. *Biometrics*. 1972; 28:931–946. [PubMed: 4648797]
- Eichler EE, Flint J, Gibson G, Kong A, Leal SM, Moore JH, Nadeau JH. Missing heritability and strategies for finding the underlying causes of complex disease. *Nat Rev Genet*. 2010; 11:446–450. [PubMed: 20479774]
- Elston RC. Query—estimating heritability of a dichotomous trait. *Biometrics*. 1977; 33:231–233. [PubMed: 843575]
- Falconer DS. Inheritance of liability to certain diseases estimated from incidence among relatives. *Ann Hum Genet*. 1965; 29:51–76.

- Ferguson, TS. A course in large sample theory. Chapman & Hall/CRC; London: 1996.
- Galton, F. Hereditary genius: an inquiry into its laws and consequences. Macmillan and co; London: 1869.
- Genz A, Bretz F. Comparison of methods for the computation of multivariate t probabilities. *J Comput Graph Stat.* 2002; 11:950–971.
- Goldstein DB. Common genetic variation and human traits. *N Engl J Med.* 2009; 360:1696–1698. [PubMed: 19369660]
- Huang GH. Model identifiability. *Encycl Stat Behav Sci.* 2005; 3:1249–1251.
- Joseph J. Not in their genes: a critical view of the genetics of attention-deficit hyperactivity disorder. *Dev Rev.* 2000; 20:539–567.
- Kendler KS, Neale MC, Kessler RC, Heath AC, Eaves LJ. A test of the equal-environment assumption in twin studies of psychiatric illness. *Behav Genet.* 1993; 23:21–27. [PubMed: 8476388]
- Kidd K, Cavalli-Sforza L. An analysis of the genetics of schizophrenia. *Biodemography Soc Biol.* 1973; 20:254–264.
- Kurtz TW, Griffin KA, Bidani AK, Davisson RL, Hall JE. Recommendations for blood pressure measurement in humans and experimental animals Part 2: blood Pressure measurement in experimental animals: a statement for professionals from the Subcommittee of Professional and Public Education of the American Heart Association Council on High Blood Pressure Research. *Hypertension.* 2005; 45:299–310. [PubMed: 15611363]
- Lee SH, Wray NR, Goddard ME, Visscher PM. Estimating missing heritability for disease from genome-wide association studies. *Am J Human Genet.* 2011; 88:294–305. [PubMed: 21376301]
- Lewontin RC. Annotation: the analysis of variance and the analysis of causes. *Am J Hum Genet.* 1974; 26:400. [PubMed: 4827368]
- Maher B. Personal genomes: the case of the missing heritability. *Nature.* 2008; 456:18–21. [PubMed: 18987709]
- Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, et al. Finding the missing heritability of complex diseases. *Nature.* 2009; 461:747–753. [PubMed: 19812666]
- Neale, MC.; Cardon, LR. Methodology for genetic studies of twins and families. Kluwer; Dordrecht: 1992.
- Pam A, Kemker SS, Ross CA, Golden R. The “equal environments assumption” in MZ-DZ twin comparisons: an untenable premise of psychiatric genetics? *Acta geneticae medicae et gemellologiae.* 1996; 45:349–360. [PubMed: 9014000]
- Piessens, R.; Doncker-Kapenga, D.; Uberhuber, C.; Kahaner, D. Springer series in computational mathematics. Vol. 1. Springer-Verlag; Berlin, New York: 1983. Quadpack: a subroutine package for automatic integration.
- Plomin R, Spinath FM. Intelligence: genetics, genes, and genomics. *J Pers Soc Psychol.* 2004; 86:112–129. [PubMed: 14717631]
- Rijsdijk FV, Sham PC. Analytic approaches to twin data using structural equation models. *Brief Bioinforma.* 2002; 3:119–133.
- Team, R. R: a language and environment for statistical computing. R foundation for statistical computing; Vienna, Austria: 2010.
- Visscher PM, Hill WG, Wray NR. Heritability in the genomics era—concepts and misconceptions. *Nat Rev Genet.* 2008; 9:255–266. [PubMed: 18319743]
- White H. Maximum-likelihood estimation of mis-specified models. *Econometrica.* 1982; 50:1–25.
- Zuk O, Hechter E, Sunyaev SR, Lander ES. The mystery of missing heritability: genetic interactions create phantom heritability. *P Natl Acad Sci USA.* 2012; 109:1193–1198.

Fig. 1.

Percent bias as a function of trait prevalence and gene frequency (gene freq.). Percent bias was calculated according to Eq. (2), where the true distribution of the additive genetic component of the liability is non-normal due to a single causal diallelic gene, and the true parameter values are: **a** $\sigma_A^2 = 0.2, \sigma_C^2 = 0.6$; **b** $\sigma_A^2 = 0.4, \sigma_C^2 = 0.4$; **c** $\sigma_A^2 = 0.6, \sigma_C^2 = 0.2$; **d** $\sigma_A^2 = 0.8, \sigma_C^2 = 0.1$

Fig. 2. Percent bias as a function of trait prevalence and exposure probability. Percent bias was calculated according to Eq. (2), where the true distribution of the common environmental component of the liability is Bernoulli, and the true parameter values are: **a** $\sigma_A^2=0.1$, $\sigma_C^2=0.8$; **b** $\sigma_A^2=0.2$, $\sigma_C^2=0.6$; **c** $\sigma_A^2=0.4$, $\sigma_C^2=0.4$; **d** $\sigma_A^2=0.6$, $\sigma_C^2=0.2$

Fig. 3.

Percent bias as a function of trait prevalence and degrees of freedom (df). Percent bias was calculated according to Eq. (2), where the true liability has a t distributed common environmental component. The true parameter values are: **a** $\sigma_A^2=0.1$, $\sigma_C^2=0.8$; **b** $\sigma_A^2=0.2$, $\sigma_C^2=0.6$; **c** $\sigma_A^2=0.4$, $\sigma_C^2=0.4$; **d** $\sigma_A^2=0.6$, $\sigma_C^2=0.2$

Fig. 4. Systolic blood pressure (SBP) in males 55+. **a** Density plot of SBP; **b** QQ plot of expected quantiles if SBP is normally distributed versus observed quantiles; **c** CDF plot with step function = empirical CDF from SBP data and smooth function = CDF from mixture model; **d** percent bias as a function of cutoff value. Cutoff value determines hypertension (HTN) status. Percent bias was calculated according to Eq. (2), where the true distribution of the random environment is set to be a mixture of normal distributions. The true parameter values are: *continuous lines* $\sigma_A^2=0.4$, $\sigma_C^2=0.1$; *dashed lines* $\sigma_A^2=0.3$, $\sigma_C^2=0.2$

Fig. 5. Percent bias as a function of sample size and prevalence. Percent bias was calculated according to Eq. (2), where the true liability distribution follows the ACE model. The true parameter values are: **a** $\sigma_A^2=0.6, \sigma_C^2=0.15$; **b** $\sigma_A^2=0.15, \sigma_C^2=0.6$