The Sensitivity of Variance Component Estimates to Underreporting: Method and Application to Substance Abuse Data

Paul W. Andrews, Kenneth S. Kendler, Nathan Gillespie, and Michael C. Neale

Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University, Richmond, Virginia, United States of America

Many studies of human behavior and psychological constructs rely on subjects' willingness to disclose information about themselves. This is problematic for phenotypes that require the disclosure of sensitive information, such as sexual behavior or illicit drug use, which are likely to be underreported. We describe a method for evaluating how sensitive variance component estimates are to underreporting. The method involves estimating, by maximum likelihood, the original population proportions of the response classes, and adjusting them for a set of hypothesized underreporting parameters. If the true values of the underreporting parameters were known, the researcher could estimate the variance components based on these values. Usually, underreporting levels are not known with certainty. However, it is possible to assume a specific value for the underreporting rate, obtain response pattern proportions adjusted for this rate, and then to conduct the analyses on these revised estimates. By repeating the procedure across the range of plausible underreporting values, the researcher can assess how sensitive the variance component estimates are to variation in underreporting. We apply this method to a sample of male-male twin pairs who reported on themselves and their co-twins for illicit drug abuse and dependence (DAD). We show how underreporting influences estimates of additive genetic, common environment, and specific environment variance components (A, C, and E) obtained for DAD in a classical twin design.

Researchers are often interested in the degree to which population-level variation in a particular phenotype may be attributed to differences in genes or environmental experiences. The accurate partitioning of phenotypic variation into genetic and environmental components depends, of course, on measuring the phenotype accurately. Some phenotypes can only be assessed by asking people to report about themselves, and such information may be subject to distortion, bias, or incorrect recall. Such self-report measures are likely to be particularly problematic for phenotypes that involve sensitive information, such as sexual or criminal behavior (e.g., the use of illicit drugs). In such situations, a common concern for researchers is underreporting, in which a proportion of subjects who are positive for the phenotype fail to disclose that they are positive.

It is not immediately clear how underreporting influences heritability estimates. In the typical twin study, heritability is estimated from the difference between the phenotypic correlation in monozygotic (MZ) and dizygotic (DZ) twins. In principle, heritability estimates may accrue positive or negative bias if underreporting is present. Underreporting is a form of measurement bias. If it is uncorrelated between twins, it may cause specific environmental variance to be overestimated, and additive genetic and shared environment variance to be underestimated. If, however, underreporting is a heritable trait, then it may result in the overestimation of additive genetic and shared environment variance.

One way of attempting to deal with underreporting is to use multiple sources. For instance, informant paradigms use information from personal interviews and from interviews of family members or other knowledgeable people (Heath et al., 1992; Kendler et al., 2002; Neale & Stevenson, 1989). In prior work, we have assumed that subjects who report being negative for a sensitive phenotype are more likely to be underreporting if their co-twin reports that the subject is positive for the phenotype (Kendler et al., 2002). Thus, we used the co-twin's report as another piece of information about the subject's true phenotype. A limitation of that method is that such instances may not always be underreporting by the subject — they could be mistakes by the co-twin.

In this article, we describe a different method for dealing with underreporting, which explicitly models its effects. We are thus able to investigate how sensitive estimates of variance components and other model

Received 22 June, 2007; accepted 31 July, 2007.

Address for correspondence: Paul W. Andrews, Ph.D., Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University, P.O. Box 980126, Richmond, VA 23298-0126. E-mail: pandrews@vcu.edu

parameters are to underreporting. While the main focus of the analyses we report is how underreporting influences heritability estimates, it has broader application, including the estimation of other variance components, such as the effects of quantitative trait loci or environmental factors or their interaction.

The information used to assess variance components of binary traits, such as presence or absence of substance use, is contained in the contingency tables of responses. For example, in the classical twin study, the relative proportion of twin pairs who are concordant positive (++), concordant negative, or discordant for substance use, may be used to estimate a tetrachoric correlation for twin pairs. We refer to the possible outcomes (– –, –+, +– or ++) as response patterns. If underreporting occurs, it will increase the proportion of negative outcomes relative to the positive ones. In principle, if the rates of underreporting were known, it would be possible to reverse the effects of underreporting by incrementing the proportions of positive outcomes. The researcher could then conduct the analyses on these adjusted proportions. In practice, it is rarely possible for the researcher to obtain accurate estimates of the underreporting rates for sensitive behavior. However, it is possible to assume a specific value for the underreporting rate, obtain response pattern proportions adjusted for this rate, and then to conduct the analyses on these revised estimates. By repeating the procedure across a range of plausible underreporting values, it is possible to assess how sensitive variance component estimates are to hypothesized levels of underreporting.

We illustrate the method with twin data on lifetime diagnoses of DSM-IV drug abuse or drug dependence (DAD; APA, 1994). There is substantial unreliability in the measurement of psychiatric disorders that are assessed by interview (Aneshensel et al., 1987; Bromet et al., 1986; Fendrich et al., 1990; Kendler et al., 1999; Kendler et al., 1993; Kendler et al., 2002; Prusoff et al., 1988). We chose this phenotype because underreporting is likely to be an important cause of measurement uncertainty in DAD (Chen et al., 2006; Colon et al., 2001; Johnson & Fendrich, 2005). Moreover, it seems unlikely that people would overreport lifetime DAD. The current social and legal climate probably makes people feel that admission of drug use is potentially risky, even if they are otherwise told that the information is only being collected for research purposes and will be kept confidential. And while memory problems may sometimes lead a few people to overreport recent drug use (Johnson & Fendrich, 2005), the problems that are associated with DAD are likely to be much more salient in memory. By explicitly modeling underreporting, we can assess the degree to which published estimates of heritability of DAD may have been biased by such processes. For instance, in prior work, we estimated that the heritability of lifetime DAD was 74%, based on the self-reports of male–male twin pairs (Kendler et al., 2002).

In principle, the method we describe could be applied to other phenotypes where underreporting may be a problem, such as sexual behavior or having been the victim of child sexual abuse. It could also be applied to studies of children's behavior where there are multiple raters (e.g., children and parents) who are often discordant with each other.

Methods

Sample

The twin data used in this analysis came from the second wave of interviews of male–male and male–female twin pairs ascertained from the Virginia Twin Registry, which is now part of the Mid-Atlantic Twin Registry. The details of these interviews are discussed elsewhere (Kendler & Prescott, 2006).

The eligible sample consisted of twins who were matched to state records, members of a multiple birth in which one was a male, Caucasian, and born between 1940 and 1974. Of the total meeting the eligibility requirements, 72.4% were successfully interviewed in the first wave, and 82.6% of those participating in the first wave were interviewed in the second wave. The second wave interviews were conducted between 1994 and 1998, and at this time subjects were between 20 and 58 years old (mean = 36.8, *SD* = 9.1). Each member of a twin pair was interviewed by a different interviewer, who was blind to the clinical information about the co-twin. The zygosity of each twin pair was based on a discriminant function analysis of the answers to six standard zygosity questions. Development of the algorithm was conducted on 227 twin pairs genotyped with at least eight highly polymorphic DNA markers (Kendler & Prescott, 2006).

For each twin pair, we labeled one of them twin 1 and the other twin 2. Each provided self-report information obtained at personal interview about DAD. In addition, each subject provided information about DAD in his co-twin. Thus, for every twin-pair, there were four pieces of information: (1) twin 1's self-report; (2) twin 1's report on twin 2; (3) twin 2's self-report; and (4) twin 2's report on twin 1. To simplify our analyses, we restricted our sample to male–male twin pairs for which we had all four reports. This resulted in 707 MZ male–male twin pairs and 488 DZ male–male twin pairs.

Diagnostic Methods

Each twin in a family was interviewed by a different interviewer, who had a master's degree in a mental health related field, or a bachelor's degree with two years of clinical experience. The interviewer of each twin was blind to the clinical information about the co-twin.

The Subject's Self-Report

For simplicity, we refer to the DAD information provided by each twin on himself as the self-report.

https://doi.org/10.1375/twin.10.5.721 Published online by Cambridge University Press

Self-reported lifetime DAD was assessed in the wave 2 personal interview by items adapted from the Structured Clinical Interview for DSM-III-R (Spitzer & Williams, 1985). DSM-IV criteria for DAD at any time over the lifetime was assessed separately for cannabis, sedatives, stimulants, cocaine, opiates, hallucinogens, inhalants, and 'over the counter' medications. If the self-report was positive for all the DSM-IV criteria required for a diagnosis of abuse/dependency for one or more of these substances, it was scored as 1, and we refer to it as a positive self-report. Otherwise, it was scored as 0, and we refer to it as a negative self-report.

The Subject's Report on the Co-Twin

In wave 2, we also asked each subject to report DAD information on their co-twin, which we refer to as the co-twin report. To assess DAD in the co-twin, we used the relevant portion of the Family-History Research Diagnostic Criteria (Endicott et al., 1978). If the co-twin report was positive for abuse or dependence using the Family-History Research Diagnostic Criteria, it was coded as 1, and we refer to it as a positive co-twin report. Otherwise, it was coded as 0, and we refer to it as a negative co-twin report.

The Underreporting Parameters

We can think of no reason why the self-reports of MZ and DZ twins about their own DAD would differ in the mean level of underreporting. We therefore assume that underreporting in self-report does not vary as a function of zygosity.

However, we think it plausible that the chance a subject is underreporting may depend on what his co-twin says about him. The rationale for using informant paradigms is that informants often have substantial information about each other, and twins will often have access to highly privileged information about each other. For this reason, if a twin says that he has not had any DAD, the chance that he is underreporting may be greater if his co-twin reports that the twin is positive for DAD.

Although in principle it is possible to allow for differences between twin 1 and twin 2 for underreporting rates, which may be appropriate if they are of opposite sex or otherwise ordered, in these analyses we constrain the rates to be equal across twins. We therefore allow two underreporting parameters: (1) the rate of underreporting by a twin when the co-twin reports that the twin is negative (p); and (2) the rate of underreporting by a twin when the co-twin reports that the twin is positive (q).

Formatting the Data

The information about each twin pair can be defined by each twin's self-report and each twin's report of their co-twin. This leads to 16 possible twin-pair categories. The pattern variable in Table 1 lists these categories. Consider, for instance, twin-pairs that fall into the category where the pattern variable = 0110. The first column describes twin 1's self-report; column 2 is twin 2's report on twin 1; column 3 is twin 2's self-report; and column 4 is twin 1's report on twin 2. Thus, the value corresponding to pattern = 0110 refers to twin pairs in which twin 1 self-reports no DAD (first column = 0), but twin 2 reports that twin 1 is positive for DAD (second column = 1), and twin 2 self-reports DAD (third column = 1), but twin 1 reports that twin 2 is negative for DAD (fourth column = 0). This format allows us to simplify the data while retaining information about nonindependence. The observed frequencies and proportions of the MZ and the DZ twin pairs falling into each of these categories are also provided in Table 1.

Adjustment of Predicted Cell Frequencies for Underreporting

Table 1 shows the observed cell frequencies — it reflects what subjects reported about their own DAD state and what they reported about the DAD state of their co-twins. But, due to underreporting, some subjects who had DAD probably failed to disclose information that would have led us to give them a diagnosis of DAD. Ideally, we would conduct variance component analyses on the true cell frequencies, i.e., subjects' actual DAD states, which could be recovered if we knew the underreporting rates. Table 2 is the transition matrix modeling how the estimate of the true twin-pair frequencies in each of the categories (on the rows) are reapportioned, via the underreporting parameters, among the categories (on the columns). Mathematically, this is simply achieved by application of the matrix formula: $E = DY$, where D is a 16 by 1 vector containing the estimates of the true population cell frequencies, Y is the 16 by 16 transition matrix (given in Table 2), and E is the (16 by 1) vector of the estimated cell frequencies following the effects of underreporting.

The logic for deriving the transition matrix in Table 2 is simple. For example, consider twin pairs whose true state falls in the pattern category of 1011. This category corresponds to twin pairs in which twin 1 has DAD, but twin 2 reports that twin 1 is negative for DAD, and twin 2 has DAD, and twin 1 also reports that twin 2 is positive for DAD. Since we assumed that underreporting in subjects' self-reports may depend on what their twins say about them, twin 1 underreports at rate p (because twin 2's report on twin 1 is negative) and twin 2 underreports at rate q (because twin 1's report on twin 2 is positive), Thus, underreporting makes proportion $p \times q$ of the true number of twin pairs in category 1011 go into observed category 0001, proportion $p \times (1\text{-}q)$ go into observed category 0011, proportion $(1\text{-}p) \times q$ go into 1001, and proportion $(1\text{-}p) \times (1\text{-}q)$ remain in 1011. In Table 2, this partitioning can be seen by noting how twin pairs in the row corresponding to true category 1011 are proportioned into the pertinent observed category columns. The rest of the matrix can be filled out

https://doi.org/10.1375/twin.10.5.721 Published online by Cambridge University Press

Table 1

The Observed Frequencies and Proportions in Each of the 16 Possible Twin Pair Categories, Broken Down by Zygosity

						MZ Twin Pairs		DZ Twin Pairs	
Category	T1 on T1	T2 on T1	T2 on T2	T1 on T2	Pattern	Observed frequency	Observed proportion	Observed frequency	Observed proportion
1	0	0	0	0	0000	506	.7157	292	.5984
2	0	0	0	1	0001	2	.0028	6	.0123
3	0	0	1	0	0010	51	.0721	56	.1148
4	0	0	1	1	0011	6	.0085	9	.0184
5	0	1	0	0	0100	3	.0042	3	.0061
6	0	1	0	1	0101	0	.0	0	.0
7	0	1	1	0	0110	2	.0028	1	.0020
8	0	1	1	1	0111	0	.0	1	.0020
9	1	0	0	0	1000	42	.0594	56	.1148
10	1	0	0	1	1001	2	.0028	4	.0082
11	1	0	1	0	1010	51	.0721	26	.0533
12	1	0	1	1	1011	18	.0255	10	.0205
13	1	1	0	0	1100	5	.0071	13	.0266
14	1	1	0	1	1101	0	.0	0	.0
15	1	1	1	0	1110	15	.0212	7	.0143
16	1	1	1	1	1111	4	.0057	4	.0082
						$N = 707$	$\Sigma = 1$	$N = 488$	$\Sigma = 1$

Note: T1 on T1 = Twin 1's DAD self-report; T1 on T2 = Twin 1's report on the DAD status of Twin 2; T2 on T2 = Twin 2's DAD self-report; T2 on T1 = Twin 2's report on the DAD status of Twin 1.
Each twin-pair category is described by a 4-column pattern variable, and the value of each column corresponds to the values for T1 on T1, T1 on T2, T2 on T2, and T2 on T1, respectively.

using similar logic, which we leave to the reader. Thus, the rows in Table 2 describe how twin pairs in a specific true category are proportioned into separate observed categories, where the proportions in each row sum to 1. The columns describe the contributions of different true categories to a specific observed category.

Model Fitting

The structural equation model that we fit to the data is provided in Figure 1.[1] We allow the latent DAD phenotype for each twin to influence what their co-twins say about them (the ctr paths) as well as their self-reports (the sr paths). Each report has residual error, and we allow for the possibility that the residual errors in the twin's self-report and his report on his co-twin are correlated by virtue of the fact that they come from the same informant.

Underreporting is incorporated into the model in the following way. The estimates of the free parameters (the A, C, and E components of variance of the latent true phenotype, the loadings on the latent phenotypes, and the residual error terms, along with the population thresholds) generate an estimate of the true cell frequencies, which may be described as a row vector *D*. Given an estimate of the true frequencies and the values of the underreporting parameters that we have set, we can estimate, via the transition matrix Y shown in Table 2, the overall frequencies of DAD states that the twins are expected to report, which we call the expected reported frequencies. In other words, an estimate of the free parameters generates an estimate of the true cell frequencies, which, in turn, generates an estimate of the expected reported frequencies for given values of the underreporting variables. Thus, if it is given a set of values for the underreporting variables, Mx can search the free parameter space to find, through maximum likelihood, the values for which the expected reported frequencies come closest to matching the actual observed frequencies. We applied a range of hypothesized values of the underreporting parameters p and q to examine their effects on the free parameters of the model, especially the A, C, and E variance components.

Results

Descriptive Statistics

Table 3 shows the cross-tabulation of the twins' DAD self-reports by what their co-twins said about them. Among MZ pairs, twins were more likely to self-report DAD (20.1% or 284/1414) than they were to report it in their co-twin (4.3% or 61/1414). DZ twins were also more likely to self-report DAD (24.0% or 234/976) than they were to report it in their co-twin (6.5% or 63/976). This suggests that both MZ and DZ twins may have had more knowledge about themselves than their co-twins. MZ twins appear to be about as likely to self-report DAD (20.1%) as DZ twins (24.0%), which suggests that zygosity may not

https://doi.org/10.1375/twin.10.5.721 Published online by Cambridge University Press

Table 2

The Transition Matrix Describing How Twin Pairs in Each True Category (Given in Rows) are Proportioned, via Underreporting, to the Observed Categories (Given in Columns)

True categories	Observed categories															
	0000	0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011	1100	1101	1110	1111
0000	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
0001	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—
0010	p	—	1-p	—	—	—	—	—	—	—	—	—	—	—	—	—
0011	—	q	—	1-q	—	—	—	—	—	—	—	—	—	—	—	—
0100	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—
0101	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—
0110	—	—	—	—	p	—	1-p	—	—	—	—	—	—	—	—	—
0111	—	—	—	—	—	q	—	1-q	—	—	—	—	—	—	—	—
1000	p	—	—	—	—	—	—	—	1-p	—	—	—	—	—	—	—
1001	—	p	—	—	—	—	—	—	—	1—p	—	—	—	—	—	—
1010	p^2	—	p(1-p)	—	—	—	—	—	p(1-p)	—	$(1-p)^2$	—	—	—	—	—
1011	—	pq	—	p(1-q)	—	—	—	—	—	q(1-p)	—	(1-p)(1-q)	—	—	—	—
1100	—	—	—	—	q	—	—	—	—	—	—	—	1-q	—	—	—
1101	—	—	—	—	—	q	—	—	—	—	—	—	—	1-q	—	—
1110	—	—	—	—	pq	—	q(1-p)	—	—	—	—	—	p(1-q)	—	(1-p)(1-q)	—
1111	—	—	—	—	—	q^2	—	q(1-q)	—	—	—	—	—	q(1-q)	—	$(1-q)^2$

Note: See the text for details.

affect the mean level of DAD. Finally, the rate of discrepancies between the twins' self-reports and their reports on their co-twins was about twice as high for DZs (20.5% or 201/976) as for MZs (10.0% or 141/1414). This suggests that subjects from MZ twin pairs had more accurate knowledge about their co-twins than those from DZ twin pairs.

ACE Estimates

We estimated the A, C, and E components of the variance in DAD for various levels of underreporting in the self-report (Table 4). For all the underreporting models we ran, we also calculated confidence intervals for the estimates of A, C, and E. We first estimated A, C, and E under the assumption of no underreporting. The estimates of A = 77%, C = 0%, and E = 23% are very similar to the estimates in our prior work (Kendler et al., 2002).[2]

If the level of underreporting in the self-report is not conditional on the co-twin report, then p and q are equal. To assess how constant underreporting affected the ACE estimates, we set p and q equal to each other, and incrementally increased their values from zero in steps of 0.1. This analysis indicated that increasing p and q had complex effects on ACE estimates (see Table 4). In particular, E tended to decrease as p and q increased, C tended to increase, and A had a quadratic effect (peaking at intermediate values of p and q).

We then assessed how p and q separately influenced the ACE estimates (Table 4). First, we fixed q to zero, and incrementally increased p from zero in steps of 0.1. This analysis indicated that E tended to decrease as p increased, C increased at high underreporting levels, and A had a quadratic effect (peaking at an intermediate value of p). Next, we fixed p to zero, and incrementally increased q from zero in steps of 0.1. This analysis indicated that A, C, and E were relatively unaffected by variation in q.

Discussion

In this article, we have described a transition method for determining how sensitive heritability estimates are to underreporting. However, the method is potentially useful for any dataset where subjects may have been reluctant to disclose sensitive information.

The influence of underreporting on the estimate of heritability in DAD depends on its type. If underreporting is modeled as a single parameter, in which the rate is assumed to be constant throughout the population, heritability shows a quadratic effect. The heritability estimate varies from .73 to .91 over the range in underreporting that we have reported, with the peak occurring approximately where the underreporting rate is 0.3.

We also explored the independent effects of the rate of underreporting by subjects with negative co-twin reports (p) and those with positive co-twin reports (q). Underreporting by subjects with negative co-twin reports (p) influenced the heritability of DAD in a quadratic fashion, peaking at intermediate underreporting levels. However, overall family resemblance increased monotonically, because the estimate of C increased with underreporting by subjects with negative co-twin reports. On the other hand, the estimates of A, C, and E were largely unaffected by underreporting by subjects

https://doi.org/10.1375/twin.10.5.721 Published online by Cambridge University Press

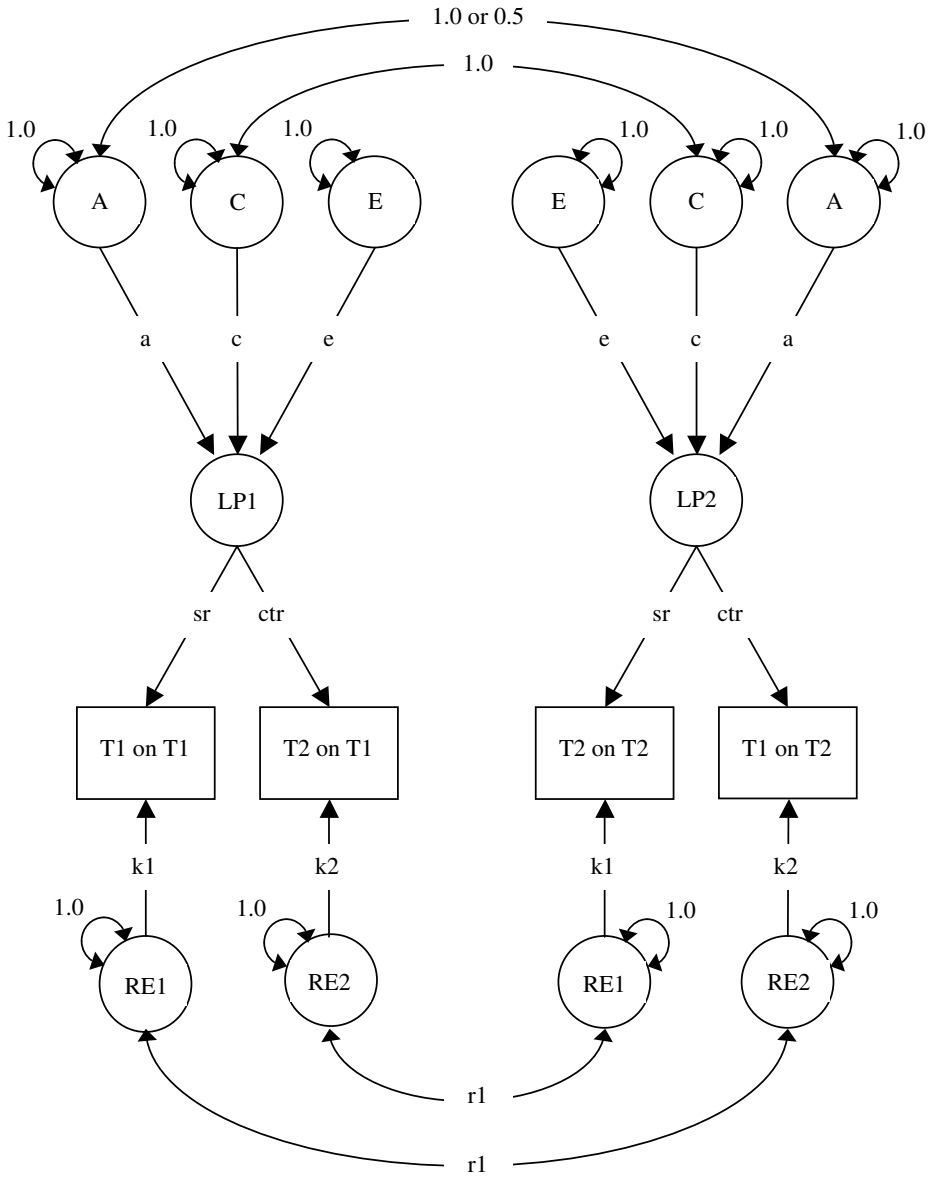

Figure 1

The ACE model for latent phenotype of twin 1 (LP1) and twin 2 (LP2) based on self-reports (T1 on T1, T2 on T2) and co-twin reports (T1 on T2, T2 on T1). In this article the latent phenotype is DAD. The residual errors in the self-reports (RE1) and the co-twin reports (RE2) are allowed to be correlated within Twin 1 (r1) and within Twin 2 (r2). The freely estimated parameters are given in lowercase.

with positive co-twin reports (q), even though the model fits were substantially different. This is because the estimates of the reported cell frequencies varied with q, but they did so in a way that had little impact on the ACE parameters.

We obtained different likelihood estimates for each of the underreporting models that we tried. The most likely model reported in Table 4 is where p = 0.1 and q = 0.0, –2LL = 3046.035. More refined analysis showed that this was indeed very close to the minimum. The most likely values of the underreporting parameters were p = 0.106 and q = 0.000, –2LL = 3046.031. We then calculated the fit of each model as the difference in the likelihood, –2ΔLL, relative to this one (see Table 4). Our results indicate that there may not have been much underreporting about DAD in this population of twins. Most of the underreporting models given in Table 4 that did not have significantly worse fits (i.e., those that were less than 3.84 chi-square units away from the best fitting model) were those with relatively low underreporting values.

The analyses suggest that the underreporting that occurred was probably done mostly by subjects with negative co-twin reports (because the most likely values of p and q were 0.106 and 0.0, respectively). This is perhaps what one might expect, if subjects who are relatively secure about revealing their DAD state to their co-twin are less likely to underreport. From Table 3, there are very few instances in which a subject reported no DAD and his co-twin reported that the

https://doi.org/10.1375/twin.10.5.721 Published online by Cambridge University Press

Table 3

Cross-Tabulation of Twins' Self-Reports and the Co-Twin's Reports on the Twin

		Co-twin's report on twin	
	Twin's self-report	No DAD	DAD
MZ	No DAD	1121	9
	DAD	132	52
DZ	No DAD	727	15
	DAD	186	48

subject did have DAD. The fact that the most likely rate of underreporting by subjects with positive co-twin reports (q) is zero suggests that these instances may have been the result of co-twin error, not underreporting by the subject.

As underreporting rates increase, model fit deteriorates compared to the fit of the most likely model, and this deterioration accelerates. Interestingly, however, the range of the confidence intervals for the estimates of A tend to be fairly stable until underreporting rates become very large (≥ 0.4). This means that the confidence with which A is estimated is relatively unaffected throughout much of the plausible range in underreporting.

Finally, in the observed data, there was a higher rate of errors in the co-twin reports of DZ twins than those of MZ twins (20.5% vs. 10.0%). This may indicate that MZ twins have greater knowledge of their co-twins' DAD than have DZ twins.

Limitations

The transition matrix given in Table 2 is appropriate for dyads in which both individuals provide information about themselves and each other on a binary trait. Another important limitation is that we assumed that underreporting was not itself a heritable trait. The method can be extended to that situation by incorporating additional parameters into the MZ and DZ transition matrices that represent the degree of concordance between twin pairs for underreporting. The researcher could then explore how sensitive the estimate of heritability of DAD (or other phenotypes) is to the MZ and DZ concordance parameters.

In this article, we assumed that underreporting in the self-report was of primary interest. It seems possible, however, that MZ and DZ twins might differ in the degree of underreporting about their co-twin's drug state. For instance, because MZ twins tend to be closer, they might be more reluctant to disclose information about their co-twins than would DZ twins. However, at least conceptually, the heritability of DAD depends only on self-reported information if people provide accurate information about themselves. Thus, once underreporting in the self-report is taken into account, there is less of a need to model the reluctance of co-twins to report DAD information about their twins, which would otherwise require an alternative transition matrix. In the present analysis, we allowed for potential differences in underreporting in the self-report as a function of the co-twin's report.

The method can also be simplified to deal with the situation where the researcher only has self-reports from each twin pair. This would involve collapsing the 16 by 16 transition matrix to a 4 by 4 matrix. We stress, however, that underreporting is not likely to be homogeneous within a population. It seems reasonable to assume, for instance, that subjects' rate of underreporting in their self-report covaries with what their partners say about them. If so, then the use of informant reports may give the researcher traction for

Table 4

How ACE Estimates and their Confidence Intervals Change With the Rate of Underreporting by Subjects With Negative Co-Twin Reports (p) and the Rate of Underreporting by Subjects With Positive Co-Twin Reports (q)

p	q	A	C	E	–2LL	–2ΔLL	*p* value
.0	.0	.77 (.45–.92)	.00 (.00–.29)	.23 (.08–.31)	3047.880	0.027	*ns*
.1	.1	.80 (.47–.93)	.02 (.00–.32)	.18 (.07–.27)	3049.822	1.969	*ns*
.2	.2	.84 (.49–.93)	.03 (.00–.34)	.14 (.05–.22)	3056.453	8.600	< .05
.3	.3	.89 (.49–1.0)	.04 (.00–.40)	.08 (.04–.18)	3069.987	22.134	< .05
.4	.4	.91 (.51–1.0)	.08 (.00–.47)	.01 (.00–.03)	3090.095	42.242	< .05
.5	.5	.73 (.28–1.0)	.26 (.00–.72)	.01 (.00–.04)	3127.177	79.324	< .05
.1	.0	.82 (.49–.99)	.00 (.00–.31)	.18 (.01–.26)	3047.854	0.001	*ns*
.2	.0	.88 (.52–1.0)	.00 (.00–.33)	.12 (.00–.20)	3048.024	0.171	*ns*
.3	.0	.95 (.57–1.0)	.01 (.00–.39)	.04 (.02–.13)	3048.833	0.980	*ns*
.4	.0	.89 (.49–1.0)	.11 (.00–.46)	.00 (.00–.03)	3050.378	2.525	*ns*
.5	.0	.69 (.29–1.0)	.31 (.00–.36)	.00 (.00–.03)	3065.255	17.402	< .05
.0	.1	.75 (.44–.88)	.02 (.00–.24)	.23 (.12–.31)	3049.893	2.040	*ns*
.0	.2	.75 (.45–.85)	.02 (.00–.29)	.23 (.15–.32)	3056.023	8.170	< .05
.0	.3	.75 (.45–.83)	.01 (.00–.28)	.24 (.17–.32)	3067.542	19.689	< .05
.0	.4	.74 (.45–.82)	.01 (.00–.28)	.25 (.18–.33)	3085.819	37.966	< .05
.0	.5	.74 (.43–.82)	.01 (.00–.28)	.25 (.18–.33)	3112.049	64.196	< .05

Note: The fit of each model is –2LL. The likelihood of each model is significantly worse (p-value < .05) if the difference in fit, –2ΔLL (which, asymptotically, has a chi-square distribution), is 3.84 chi-square units more than the best-fitting model (described in text).

https://doi.org/10.1375/twin.10.5.721 Published online by Cambridge University Press

creating additional underreporting variables that can be used to explore the effects of some of the heterogeneity in underreporting. In some of our prior work, we have assumed that a negative self-report by a subject is more likely to be an underreport if the co-twin reports that the subject is positive (Kendler et al., 2002). Our results indicate that the reverse was more likely — the rate of underreporting by subjects was estimated to be higher if the co-twin said that the subject was negative for DAD than if the co-twin said that the subject was positive. However, this still illustrates how underreporting is likely to be heterogeneous within populations.

Finally, it is important to note that we omitted twin pairs in which at least one piece of information was missing. It is not immediately clear how to treat this data. It is possible that missing information is a form of underreporting, where some people who had DAD refused to answer the self-report question. However, it is also possible that some who refused to answer the question did so because they felt that it was none of our business to be asking it. Future work may resolve this issue.

Endnotes

1 The Mx script we used for our analyses can be found at: http://www.vcu.edu/mx/examples.html

2 The results are not identical because we used a more restricted sample. As discussed in the text, we only used twin-pairs for which we had all four pieces of information — 2 self-reports and 2 co-twin reports. In this article, we also used a simpler structural equation model.

Acknowledgments

PWA was supported by a National Research Service Award from the National Institutes of Health, P32 MH-20030 (PI: Michael C. Neale). John Myers provided the data used in this article. We thank two anonymous reviewers for their comments.

References

Aneshensel, C. S., Estrada, A. L., Hansell, M. J., & Clark, V. A. (1987). Social psychological-aspects of reporting behavior: Lifetime depressive episode reports. *Journal of Health and Social Behavior, 28,* 232–246.

American Psychiatric Association (APA). (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Bromet, E. J., Dunn, L. O., Connell, M. M., Dew, M. A., & Schulberg, H. C. (1986). Long-term reliability of diagnosing lifetime major depression in a community sample. *Archives of General Psychiatry, 43*, 435–440.

Chen, W. J., Fang, C. C., Shyu, R. S., & Lin, K. C. (2006). Underreporting of illicit drug use by patients at emergency departments as revealed by two-tiered urinalysis. *Addictive Behaviors, 31*, 2304–2308.

Colon, H. M., Robles, R. R., & Sahai, H. (2001). The validity of drug use responses in a household survey in Puerto Rico: comparison of survey responses of cocaine and heroin use with hair tests. *International Journal of Epidemiology, 30*, 1042–1049.

Endicott, J., Andreasen, N., & Spitzer, R. L. (1978). *Family history research diagnostic criteria*. New York: Biometric Research Department, New York State Psychiatric Institute.

Fendrich, M., Weissman, M. M., Warner, V., & Mufson, L. (1990). 2-year recall of lifetime diagnoses in offspring at high and low-risk for major depression: The stability of offspring reports. *Archives of General Psychiatry, 47*, 1121–1127.

Heath, A. C., Neale, M. C., Kessler, R. C., Eaves, L. J., & Kendler, K. S. (1992). Evidence for genetic influences on personality from self-reports and informant ratings. *Journal of Personality and Social Psychology, 63*, 85–96.

Johnson, T., & Fendrich, M. (2005). Modeling sources of self,report bias in a survey of drug use epidemiology. *Annals of Epidemiology, 15*, 381–389.

Kendler, K. S., Karkowski, L. M., & Prescott, C. A. (1999). Fears and phobias: Reliability and heritability. *Psychological Medicine, 29*, 539–553.

Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C., & Eaves, L. J. (1993). The lifetime history of major depression in women: Reliability of diagnosis and heritability. *Archives of General Psychiatry, 50*, 863–870.

Kendler, K. S., & Prescott, C. A. (2006). *Genes, environment, and psychopathology: Understanding the causes of psychiatric and substance use disorders*. New York: Guilford Press.

Kendler, K. S., Prescott, C. A., Jacobson, K., Myers, J., & Neale, M. C. (2002). The joint analysis of personal interview and family history diagnoses: Evidence for validity of diagnosis and increased heritability estimates. *Psychological Medicine, 32*, 829–842.

Neale, M. C., & Stevenson, J. (1989). Rater bias in the EASI Temperament Scales: A twin study. *Journal of Personality and Social Psychology, 56*, 446–455.

Prusoff, B. A., Merikangas, K. R., & Weissman, M. M. (1988). Lifetime prevalence and age of onset of psychiatric-disorders: Recall 4 years later. *Journal of Psychiatric Research, 22*, 107–117.

Spitzer, R. L., & Williams, J. B. W. (1985). *Structured clinical interview for DSM-III-R (SCID)*. New York: Biometric Research Department, New York State Psychiatric Institute.