Behavior Genetics, Vol. 35, No. 4, July 2005 (© 2005)
DOI: 10.1007/s10519-004-1484-9

Environmental Mediation and The Twin Design

Shaun Purcell[1] and Karestan C. Koenen[2]

Received 7 July 2004—Final 23 Nov. 2004

Behavior genetic twin designs are increasingly used to study the effects of a measured environment whilst controlling for genetic variation. In this research note, we show that, in the context of the classical twin design, (1) when the environmental variable is necessarily shared between twins, the notion of controlling for genetic influence is logically flawed and (2) when the environmental variable varies between twins in the same family, partial control for genetic influence is possible, but only if appropriate analytic models are used, which is commonly not the case. Based on a simple simulation study, recommendations are given as to which methods should be applied and which should be avoided.

KEY WORDS: Environmental mediation; maximum likelihood; regression; simulation; twins; variance components.

ENVIRONMENTAL MEDIATION AND THE TWIN DESIGN

A centrally important insight for behavior genetics was that the environment can show genetic influence: the so-called *nature of nurture* (Plomin and Bergeman, 1991). As a consequence, behavior genetic designs have been promoted as being uniquely positioned to address questions of *environmental mediation* as well as genetic influence: that is, to investigate the effect of an environment on a trait whilst controlling for common genetic influence. This is important because an association between an environment and an outcome may arise due to a third variable, namely common genetic liability.

This genetic third variable confound was described in detail by DiLalla and Gottesman (1991), in responding to the current wisdom regarding the role of child abuse in the intergenerational transmission of violence (Widom, 1989). In particular, they noted that children who are abused may also have increased genetic risk for antisocial behavior, if antisocial parents are more likely to abuse their children and antisocial behavior is heritable. As such, childhood abuse does not cause antisocial behavior in any sense: the environment does not mediate the association.

Whether or not an association is environmentally mediated might be important for informing intervention. For example, if the association between child abuse and antisocial behavior is environmentally mediated, then interventions aimed at preventing child abuse may also prevent antisocial behavior. If the association is due to a genetic third variable, then preventing abuse, although clearly important for its own sake, may well have no impact on risk for antisocial behavior.

Twin designs

The basic logic of using the twin design to test environmental mediation is as follows: if the test for association between an environment and an outcome can be placed within a genetically-informative context in which genetic variance is explicitly modelled, then a significant result should reflect true environmental mediation. That is, an association between environment and outcome can be estimated that is "uncontaminated" by genetic influence. However, as we shall show, twin data alone cannot offer as much of a control for common genetic influence as is

[1] Psychiatry and Neurodevelopmental Genetics Unit, Center for Human Genetics Research, Massachusetts General Hospital, Boston, MA, USA.
[2] Department of Society, Human Development and Health, Harvard School for Public Health, Boston, MA, USA.

sometimes claimed; furthermore, even in situations where they potentially could, inappropriate analytic methods are often used.

Environments may be either *obligatory-shared* between twins (e.g., neighborhood or parental socio-economic status, SES) or *individual-specific* (e.g., diet or suffering some trauma). Researchers have tested environmental mediation models within behavior genetic designs both when the environment is obligatory-shared (Miller *et al*., 2001) and when it is not (Jaffee *et al*., 2004). Depending on the type of environment, we argue that environmental mediation models are either logically flawed (in the case of obligatory-shared environments) or often performed in a biased manner (in the case of individual-specific environments).

In this article we present a number of analytic strategies that have been used with the aim of testing for environmental mediation unconfounded by genetic influence; we review some of the details of methods and apply simple simulation study approaches to evaluate them. Throughout this article, the term common will be used only in the bivariate sense (i.e., a genetic factor that is common to both X and Y) and not in the family-wide sense (i.e., the term *shared environment* will always be used with respect to the latent variable C) or the frequency sense (i.e., a rare or common genetic variant).

OBLIGATORY-SHARED ENVIRONMENTS

To illustrate these methods most easily, we assume that outcome trait Y and environment X are quantitative, normally-distributed measures. The most simple analytic baseline is the regression of trait on environment, in a sample of unrelated individuals

$$Y = b_0 + b_1 X + \varepsilon. \quad (1)$$

The test for an association between trait and environment is performed with the 1 df test. $H_0 : b_1 = 0$. A significant b_1 is consistent with environmental mediation; it is also consistent with common genetic factors. This phenotypic regression method is labelled *PH* in the simulations below.

Using twin data, several researchers have adapted the DeFries-Fulker (DF) regression model (DeFries and Fulker, 1985, 1988) to attempt to model the association whilst controlling for genetic factors:

$$Y_1 = b_0 + b_1 Y_2 + b_2 R + b_3 R Y_2 + b_4 X_1 + \varepsilon, \quad (2)$$

where the subscripts on X and Y distinguish between twins in a pair and R is the coefficient of genetic relatedness (1 for MZ pairs; 0.5 for DZ pairs). In this case, we assume the environment is obligatory-shared so $X_1 = X_2$; this model is also applicable for individual-specific environments, as discussed in the next section. The proposed test of environmental mediation is the 1 df test $H_0 : b_4 = 0$. This method is labelled DF_1 in the simulations below. Despite claims that this formulation controls for genetic influence, this is not the case, as will be shown below.

Alternatively, variance components approaches have been used to address the question of environmental mediation. If the environment is obligatory-shared, a standard univariate **ACE** model for Y can be constructed, with an additional single-headed arrow path from observed obligatory-shared environment X to a single latent **C** variable (Caspi *et al*., 2000). The coefficient for this extra path is labelled m and the model is the **ACEm** model. The test involves fixing m to zero: this procedure is equivalent to testing the effect of an obligatory-shared environment entered as a predictor variable in the means model (Koenen *et al*., 2003). Whereas the standard **ACEm** model is applicable only to obligatory-shared environments, the means model formulation is also applicable to individual-specific environments. However, both approaches are in fact estimating latent genetic factors whilst controlling for the measured environment, rather than estimating the effect of the measured environment whilst controlling for latent genetic factors. That is, considering for example the means model approach, although both the phenotypic mean and variance are modelled, this is of the form $Y \sim N(\mu + \beta X, \sigma^2)$ where σ^2 represents the residual variation around $E(Y \mid X)$, where X is the measured environment. Although it is possible to further partition σ^2 (e.g., into additive genetic and environmental components) the variation estimated by σ^2 is only variation around the expected value of the phenotype conditional on X.

No matter which analytic method or model is applied, it is logically impossible to test for an environmental effect of an obligatory-shared variable whilst controlling for genetic effects using only twin data. If the environment is obligatory-shared, then within the context of the twin study, the genetic and nonshared environmental correlations r_G and r_E between environment and outcome are not defined or identified. The question of controlling for genetic influence is therefore moot. Although it is perfectly possible for an obligatory-shared environment to show genetic influence, this genetic influence would never be detectable using a twin design. For example,

it is quite possible that parents' genes influence their SES. One could, in theory, perform a molecular genetic study to detect a correlation between measured parental genotype and SES (whether this is a function of paternal and/or maternal genotype). Furthermore, one could also perform a study to find a (smaller, due to segregation variance) correlation between measured offspring genotype and measured parental SES. However, a twin study of the variable 'parental SES' would clearly yield a heritability estimate of zero, as $r_{MZ} = r_{DZ} = 1.0$ (and so all variance would be attributed to the shared environment). That is, despite being a genetically-influenced trait, SES is a property of the parents, not of the twins. Even if SES were completely determined by genes, parental SES would still contribute to **C** and not **A** in a twin model, as there is no difference between the extent to which MZ and DZ twins share parental genotype.

We performed a simple proof-of-principle simulation study, using the model outlined in Appendix I. The simulations approximately gave the following scenarios: an MZ correlation of 0.50; a DZ correlation of 0.25; a parent-offspring correlation 0.25; an offspring trait and family-wide environment correlation of 0.15. Two sets of simulations were performed: in the first instance, the family-wide environment was a function of *parental phenotypes*; in the second case, the family-wide environment was a function of *parental genotypes*. This second scenario represents the perfectly plausible case in which the genes that influence, say, parental antisocial behavior also influence parental SES (i.e., mediated via personality for instance), while the environmental influences on parental antisocial behavior are different from the environmental influences on parental SES. In both cases, there is no environmental mediation.

Applying the variance components model described above to test for environmental mediation, a significant reduction in fit at the 5% significance level was observed in 96% of the replicates in the first scenario (parental phenotypes influence the environment) and in 97% of the replicates in the second scenario (parental genotypes influence the environment). In other words, rather than the expected 5% error rates, this test suggests that environmental mediation is present when in fact the association between environment and outcome is due to common genetic factors.

Although it is beyond the scope of the classical twin design and this research note, it is of interest to note that some authors have suggested incorporating parental phenotypes into the models that attempt to test for environmental mediation with obligatory-shared environments (Kendler *et al.,* 1996). The basic model used by Kendler *et al.,* (in which parental phenotypes are modelled as extra dependent variables) should approximately correspond to entering both parental phenotypes as covariates in the means model. Re-running the simulations in this way, then only 5% of the replicates (instead of 96%) are significant when testing for environmental mediation in the first scenario (parental phenotypes influence the environment). That is, the test for environmental mediation is no longer massively liberal as correct type I error rates are obtained when the null hypothesis is true. However, in the second scenario, where parental genotypes influence the environment, entering parental phenotypes as covariates does not help: 60% of the replicates are significant. Kendler *et al*., acknowledge the distinction between these two scenarios, in that they explicitly present an alternative model that assumes that parental genotypes (not phenotypes) directly influence the environment. Such a model would presumably give correct type I error rates under the second scenario, but not necessarily the first. In any case, it would be of interest to investigate the power of these models and the ability to distinguish between them. If the power to distinguish between models is low, then, in practice, it seems that models including parental phenotypes will end up having to make strong and unwarranted assumptions about the nature of the bivariate relationship between parental trait and environment.

Finally, it is worth noting that if a similar model is applied to singleton (i.e., non-twin) data with parental phenotypes entered as covariates, then similar error rates are obtained in the two scenarios (5% and 40%). That is, at least for the first parental phenotype model described above, using a twin design does not offer any additional information, with respect to testing for environmental mediation, over a simpler non-twin design. However, it will be worth investigating other designs, such as the children-of-twins design, which may prove useful to study environmental mediation with obligatory-shared variables (D'Onofrio *et al*., 2003).

INDIVIDUAL-SPECIFIC ENVIRONMENTS

The rest of this article concentrates on individual-specific environments. Some of the methods previously discussed in the context of obligatory-shared environments can also be applied to individual-specific environments: for example, the basic phenotypic

regression model (*PH*) and DF model shown in Eq. 2 (DF_1). A number of other regression-based methods exist to assess the effects of individual-specific environments. Rodgers *et al*., (1994) adapted the DF model specifically for 'nonshared' environmental variables. These authors proposed the model

$$Y_1 = b_0 + b_1 Y_2 + b_2 R + b_3 R Y_2 + b_4(X_1 - X_2) + \varepsilon. \quad (3)$$

The test of environmental effects is the 1 df test $H_0 : b_4 = 0$. This method is labelled NS_1 in the simulations below. This model attempts to specifically focus on the nonshared component of the individual-specific environment, by using difference scores to predict each individual's phenotype. This type of between-family/within-family partitioning is also sometimes used in molecular genetic association studies to control for population stratification (i.e. as confounding due to stratification will only influence the between-family component). However, in the present scenario, it should be noted that additive genetic factors in DZ twins will also influence the within–family component, as these influences do not work uniformly at the family-wide level, and therefore the confounding effect of genes cannot be controlled for by using a difference measure in this way. As a consequence, this approach does not provide a valid test of environmental mediation, as Appendix III illustrates further.

Rodgers et al (1994) also extended their basic model to allow for the case where the ". . . difference score interacts with the genetic coefficient reflecting the level of relatedness of the pairs, . . . [which would] suggest that the type of nonshared influence has a genetic component" as follows:

$$Y_1 = b_0 + b_1 R + b_2 Y_2 + b_3 R Y_2 + b_4(X_1 - X_2) + b_5 R(X_1 - X_2), \quad (4)$$

where the test for environmental mediation is a above, $H_0 : b_4 = 0$. The tests of b_4 and b_5 are labeled NS_2 and NS_3 in the simulations below, which illustrate that this reformulation does not provide valid tests of environmental mediation either.

Alternatively, an environmental difference score can be used to predict a phenotypic difference score. In particular, difference score approaches have often been used in studies looking at differences between MZ twins only:

$$Y_1 - Y_2 = b_0 + b_1(X_1 - X_2) + \varepsilon. \quad (5)$$

This is a natural extension of the powerful discordant MZ pair design. This method is labelled Δ_1 in the simulations below. To include DZ twins in the same design, one might use the model (see Appendix II)

$$Y_1 - Y_2 = b_0 + b_1(X_1 - X_2) + b_2(1 - R)(X_1 - X_2) + \varepsilon. \quad (6)$$

The tests of the coefficients b_1 and b_2 are labelled Δ_2 and Δ_3 in the simulations below.

Other approaches have focused on using residuals instead of differences (see Turkheimer and Waldron, (2000) for a review). For example, regressing the trait of twin 2 on the trait of twin 1 and taking the residuals forms the dependent variable. The predictor variable is formed by regressing the environmental measure for twin 2 on the environmental measure for twin 1 and taking the residuals. The test for environmental mediation regresses the residual trait values on the residual environmental values, say in MZ twin pairs only. Although this approach is taken to be similar to the methods using differences, this is not the case: residuals have different statistical properties than differences. Whereas the difference between two bivariate normal measures will be uncorrelated with the sum, the residuals generated after regressing one twin's score on the other will still be correlated with sum of the twins' scores (to an extent the depends on the twin correlation). As such, correlating residual measures will not only index factors that are nonshared between twins. Indeed, the simulations below, in which this method is labelled *RES*, indicate that using residuals does not lead to valid tests of environmental mediation. The test, for MZ pairs only, is based on

$$Y_1 - E(Y_1|Y_2) = b_0 + b_1(X_1 - E(X_1|X_2)) + \varepsilon. \quad (7)$$

An individual-specific environment can be incorporated into a bivariate variance components model as the second dependent variable. In the standard bivariate Cholesky **ACE** twin model, the path from latent variable V_i to observed variable j is denoted v_{ij} (where X and Y are traits 1 and 2, respectively). In this context, controlling for common genetic effects between X and Y involves estimating the a_{12} parameter (i.e., so that the genetic correlation between traits is allowed to vary) under both alternate and null models. Under the null, one could constrain either $H_0 : e_{12} = 0$ to provide a 1 df test of common nonshared environmental influences, or $H_0 : c_{12} = e_{12} = 0$ to provide a 2 df test of common shared and nonsh-

ared environmental influences. These two tests are labelled ML_1 and ML_2, respectively. Note that testing $e_{12}=0$ also provides a test for measurement error that is correlated between X and Y.

When the environmental variable is individual - specific, the covariation due to common genetic influence can be partly estimated, in the same manner as the standard bivariate twin model. It is worth noting that any C component of an individual - specific measured environment could still reflect 'parental genetic influence' in the same way as for obligatory-shared environments, and this component of covariation would not be controlled for by any of the designs described above. For the variance components models, therefore, the more restrictive ML_1 test is guaranteed to be free of common genetic influence in a way that ML_2 (which also tests for common shared environmental influence) is not. (This issue is not addressed in the simulations below.)

Finally, as shown in the Appendix and subsequent simulations, the standard DF model which incorporates a measured environmental term (Eq. 2) does not in fact control for mediating genetic influence. One possible revised model which correctly controls for genetic influence is

$$Y_1 = b_1 + b_2R + b_3Y_2 + b_4RY_2 + b_5X_1 + b_6X_2 + b_7RX_2, \quad (8)$$

where a 1 df test for environmental effects controlling for genetic effects is $H_0 : b_5 = b_7$ (see Appendix III). This model is labelled DF_2 in the simulations below.

We conducted a simulation study to illustrate the different properties of the above methods. Further detail regarding the simulation procedure is given in Appendix IV. There were 4 conditions; 500 replicate samples were generated for each condition and analyzed using all the methods outlined above. The results we present below are all for single-entry datasets; as expected, the same pattern of results was obtained re-running the regression-based methods using double-entry and correcting the standard errors of the estimates for the non-independence of the observations by use of the Huber-White method for correlated responses. Table 1 shows the different methods as rows, the different simulation scenarios as columns labelled A, B, C and D. The table entries represent the proportion of replicates significant at the $p=0.05$ level. Scenario A represents no environmental mediation and no common genetic effects; scenario B represents no environmental mediation but common genetic effects; scenario C represents environmental mediation but no common genetic effects; scenario D represents both environmental mediation and common genetic effects. All tests of environmental mediation should therefore show error rates not significantly different from the nominal 5% error rate in scenarios A and B; for scenarios C and D the equivalent figures represent the power to detect environmental mediation. In particular, we would expect that tests of environmental mediation that do not in fact control for common genetic influence will show values significantly greater than 5% for scenario B.

The results in Table I illustrate that many of the methods that claim to control for common genetic influence in fact do not. That is, for scenario B, we see that DF_1, NS_1 and NS_2, and RES all detect 'environmental mediation' when it is not

Table I. Results of a simulation study comparing methods for detecting environmental mediation with individual-specific environments

Test	A	B	C	D
PH	0.064	0.996	0.802	1.000
DF_1	0.056	0.914	0.626	1.000
NS_1	0.060	0.686	0.692	0.998
NS_2	0.060	0.768	0.120	0.912
NS_3	0.052	0.532	0.062	0.480
RES	0.048	0.644	0.687	0.988
ML_1	0.050	0.042	0.520	0.668
ML_2	0.062	0.054	0.410	0.574
DF_2	0.042	0.052	0.150	0.186
Δ_1	0.046	0.050	0.480	0.502
Δ_2	0.016	0.054	0.416	0.402
Δ_3	0.036	0.646	0.040	0.654

Proportion of replicates significant at the $p=0.05$ level. The columns represent alternative scenarios: if E represents the true presence of environmental mediation and G represents the true presence of common genetic effects between trait and environment, then the four scenarios A, B, C and D correspond to $\{E-, G-\}$, $\{E-, G+\}$, $\{E+, G-\}$ and $\{E+, G+\}$ respectively. PH is a simple phenotypic regression (Eq. 1, $b_1=0$); DF_1 is the DF model incorporating a measured environmental term (Eq. 2, $b_4=0$); NS_1 is the Rodgers *et al.* basic nonshared test (Eq. 3, $b_4=0$); NS_2 is the first term in the Rodgers *et al.* extended nonshared test (Eq. 4, $b_4=0$); NS_3 is the second term in the Rodgers *et al.* extended nonshared test (Eq. 4, $b_5=0$); RES is a residuals-based regression (Eq. 7, $b_1=0$); ML_1 is the test of common nonshared environmental variance in the bivariate Cholesky variance components model ($e_{12}=0$); ML_2 is the test of common shared and nonshared environmental variance in the bivariate Cholesky variance components model ($c_{12}=e_{12}=0$); DF_2 is a revised DF model incorporating a measured environmental term (Eq. 8, $b_5=b_7$); Δ_1 is the basic MZ differences regression (Eq. 5, $b_1=0$); Δ_2 is the first term of the MZ/DZ differences regression (Eq. 6, $b_1=0$); Δ_3 (a test of common genetic influence rather than environmental mediation) is the second term of the MZ/DZ differences regression (Eq. 6, $b_2=0$).

present, at very high rates. In contrast, the revised DF_2 model, the difference regression models and the variance components models correctly control for common genetic effects when testing for environmental mediation.

The variance components models and the difference regression tests of environmental mediation show roughly equivalent power: these values will differ depending on the genetic and environmental architecture of the outcome and environment. The revised DF_2 model shows consistently low power however. In addition, the revised model no longer has the analytic transparency that perhaps attracts many researchers to these models over variance components.

SUMMARY

A number of analytic strategies, including ones previously used by the current authors, have claimed in one way or another to control for genetic influence and allow tests of environmental mediation within the context of the classical twin design. We have demonstrated that a number of these approaches are flawed. Assuming the classical twin design, we would first advise that efforts to determine environmental mediation should only be made for individual-specific environmental variables. The results here suggest that for obligatory-shared environments, designs other than the classical twin study need to be considered. Similar conclusions regarding the application of the basic twin design to obligatory-shared environments have also been reached by other reseachers (Turkheimer et al., under review).

Secondly, in the case of individual-specific environments, bivariate variance components or regression analysis using between- twin differences should be applied. Variance components approaches have the advantage of being easily extended to include other designs and models, whereas difference regression methods have the advantage of simplicity. Preliminary work suggests that the simple difference regression model can actually be more powerful than the variance components in a number of situations – in particular, the MZ only difference method is a very efficient design for this specific test (results from simulations, not shown). However, a full investigation of the properties of these methods in terms of statistical power under the alternate hypothesis is beyond the scope of this research note.

APPENDICES

I. Simulation of twin data with an obligatory-shared environment

The basic framework for simulation is as follows: lowercase letters represent coefficients and uppercase letters represent simulated values. For each family, we first simulate 11 random standard normal deviates: paternal and maternal additive genetic components A_P and A_M; the family-wide shared environmental component C, paternal and maternal nonshared environmental components E_P and E_M; the component R_s which influences S; the additive genetic components unique to the two twins, R_1 and R_2, which ensure the additive genetic components are correlated 0.5 between first-degree relatives; an environmental component shared only between the two twins T; finally, two nonshared environmental components for the two twins, E_1 and E_2. From these, the following quantities are calculated:

$$
\begin{aligned}
P &= aA_P + cC + eE_p \\
M &= aA_M + cC + eE_M \\
S &= f(P+M)/2 + R_S \\
A_1 &= (A_p + A_M)/\sqrt{4} + R_1/\sqrt{2} \\
A_2 &= A_1 && \text{for MZ twins} \\
A_2 &= (A_p + A_M)/\sqrt{4} + R_2/\sqrt{2} && \text{for DZ twins} \\
Y_1 &= aA_1 + cC + tT + eE_1 + sS \\
Y_2 &= aA_2 + cC + tT + eE_2 + sS,
\end{aligned}
$$

where a is the additive genetic path; c is a family-wide shared environmental path; t is a twin-specific shared environmental path; e is a nonshared environmental path; f the loading of average parental phenotype on S; s is the loading of S on twin phenotypes. The (unobserved) paternal and maternal phenotypes are P and M, respectively. The observed variables are trait values for twins 1 and 2 (Y_1 and Y_2) and the measured obligatory-shared environment S.

It is an easy matter to construct different scenarios where, for example, S depends not on the average parental phenotype, but on the average parental genotype $S = f(A_p + A_M)/2 + R_s$ or maternal genotype $S = fA_M + R_s$; or the shared environment $fC + R_s$. So long as S depends to at least some extent on A_P and/or A_M, then confounding due to common genetic variation can occur.

For the simulations reported in the main text, the parameters were set at $a = 1$, $t = c = 0$, $e = 1$, $f = 0.5$ and $s = 0$ (i.e., there is no environmental mediation). In the

first scenario $S = f(P + M)/2 + R_s$; in the second scenario the definition of S is changed to $f(A_P + A_M)/2 + R_S$. For each of 500 replicates, 200 MZ and 200 DZ twin pairs were generated. The regression-based analyses were performed using the freely-available statistical package R; the variance components models were fitted using Mx (Neale, 1997).

II. Difference regression

From the Cholesky model, using path tracing rules, the variable Y_1 can be expressed as

$$\begin{aligned} Y_1 = {} & (a_{11}a_{12} + c_{11}c_{12} + e_{11}e_{12})X_1 \\ & + (Ra_{11}a_{12} + c_{11}c_{12})X_2 \\ & + (R(a_{12}a_{12} + a_{22}a_{22}) + c_{12}c_{12} + c_{22}c_{22})Y_2 \end{aligned}$$

and therefore

$$\begin{aligned} Y_2 = {} & (a_{11}a_{12} + c_{11}c_{12} + e_{11}e_{12})X_2 \\ & + (Ra_{11}a_{12} + c_{11}c_{12})X_1 \\ & + (R(a_{12}a_{12} + a_{22}a_{22}) + c_{12}c_{12} + c_{22}c_{22})Y_1 \end{aligned}$$

then taking the difference between the two regression equations gives the difference model equation

$$\begin{aligned} Y_1 - Y_2 = {} & (a_{11}a_{12} + c_{11}c_{12} + e_{11}e_{12})X_1 \\ & + (Ra_{11}a_{12} + c_{11}c_{12})X_2 \\ & + (R(a_{12}a_{12} + a_{22}a_{22}) + c_{12}c_{12} + c_{22}c_{22})Y_2 \\ & - (a_{11}a_{12} + c_{11}c_{12} + e_{11}e_{12})X_2 \\ & - (Ra_{11}a_{12} + c_{11}c_{12})X_1 \\ & - (R(a_{12}a_{12} + a_{22}a_{22}) + c_{12}c_{12} + c_{22}c_{22})Y_1 \\ = {} & ((1 - R)a_{11}a_{12} + e_{11}e_{12})X_1 \\ & - ((1 - R)a_{11}a_{12} + e_{11}e_{12})X_2 \\ & - (R(a_{12}a_{12} + a_{22}a_{22}) + c_{12}c_{12} + c_{22}c_{22})Y_1 \\ & + (R(a_{12}a_{12} + a_{22}a_{22}) + c_{12}c_{12} + c_{22}c_{22})Y_2 \\ = {} & ((1 - R)a_{11}a_{12} + e_{11}e_{12})(X_1 - X_2) \\ & + 0 \times R(Y_1 + Y_2) \\ = {} & a_{11}a_{12}(1 - R)(X_1 - X_2) + e_{11}e_{12}(X_1 - X_2) \end{aligned}$$

which corresponds to the simple regression given in Eq. 5 for MZ only pairs and Eq. 6 for MZ and DZ pairs combined, where b_1 is $e_{11}e_{12}$ and b_2 is $a_{11}a_{12}$.

III. A revised DeFries-Fulker model incorporating a measured environment

As stated in Appendix II, if

$$\begin{aligned} Y_1 = {} & (a_{11}a_{12} + c_{11}c_{12} + e_{11}e_{12})X_1 \\ & + (Ra_{11}a_{12} + c_{11}c_{12})X_2 \\ & + (R(a_{12}a_{12} + a_{22}a_{22}) + c_{12}c_{12} + c_{22}c_{22})Y_2. \end{aligned}$$

Adding an intercept term and a term for a mean difference between zygosity, we can rewrite the above equation as

$$\begin{aligned} Y_1 = {} & b_1 + b_2R + b_3Y_2 + b_4RY_2 + b_5X_1 + b_6X_2 \\ & + b_7RX_2 \end{aligned}$$

in which case a 1 df test for environmental effects controlling for genetic effects is $H_0 : b_5 = b_7$. This is because the coefficients for each independent variable are then:

$$\begin{aligned} Y_2 &: b_3 = c_{12}c_{12} + c_{22}c_{22} \\ RY_2 &: b_4 = a_{12}a_{12} + a_{22}a_{22} \\ X_1 &: b_5 = a_{11}a_{12} + c_{11}c_{12} + e_{11}e_{12} \\ X_2 &: b_6 = c_{11}c_{12} \\ RX_2 &: b_7 = a_{11}a_{12} \end{aligned}$$

and so testing $b_5 - b_7 = 0$ is implicitly a test for $c_{11}c_{12} + e_{11}e_{12} = 0$. This also makes clear why simply testing $b_5 = 0$ does not provide a test free from genetic influence.

This also illustrates why the Rodgers et al formulation regressing a difference score on the individual is incorrect: considering the coefficients for the following four variables

$$\begin{aligned} X_1 &: a_{11}a_{12} + c_{11}c_{12} + e_{11}e_{12} \\ X_2 &: c_{11}c_{12} \\ RX_2 &: a_{11}a_{12} \\ RX_1 &: 0 \end{aligned}$$

we see that the coefficient for $(X_1 - X_2)$ is $a_{11}a_{12} + e_{11}e_{12}$ and the coefficient for $R(X_1 - X_2)$ is $-a_{11}a_{12}$ In other words, the association is still confounded by common genetic effects.

Regression-based DF models typically use a double-entry procedure in practice, e.g. entering both {*twin* l, *twin* 2} and {*twin* 2, *twin* l}, which, under the assumption of interchangeability between twin 1 and twin 2, is one way to utilize all the data, i.e. not just the regression of twin 1 on twin 2 as above. Although this practice leads to incorrect standard errors for the parameter estimates (albeit ones which can be easily corrected by use of the Huber-White method to correct for correlated responses, for example), it will not alter the expected values of the parameters, under the assumption of interchangeability. That is, if the regression of twin 1 on twin 2 has an expected value of β for the regression coefficient, then, given the interchangeability of twins, it follows that the

regression of twin 2 on twin 1 must also have an expected value of β. As interchangeability implies that twin 1 and twin 2 are expected to have identical distributions, then it also follows that the same regression performed on the double-entered dataset (i.e. simply the concatenation of the two single-entry datasets) would also have an expected value of β. Under the assumption of interchangeability, double-entry affects only the standard errors, and is therefore only relevant when considering issues of statistical significance and power, not when considering parameter expectations.

IV. Simulation of twin data with an individual-specific environment

From the bivariate Cholesky formulation, the paths for latent variable V are represented $\begin{pmatrix} v_{XX} & \\ v_{XY} & v_{YY} \end{pmatrix}$ where v_{ij} represents the path from latent variable V_i to observed variable j. When common genetic effect were generated, the bivariate structure $\begin{pmatrix} 1 & \\ 1 & 0 \end{pmatrix}$ was used, implying complete genetic correlation; otherwise $\begin{pmatrix} 1 & \\ 0 & 1 \end{pmatrix}$ was used. No common shared or nonshared environmental effects were generated in the simulations presented here (other than those induced by the causal, environmental mediation path when present). Both single and double entry approaches were used, although the results are only presented for single-entry (double-entry gave a similar pattern of results). In all cases, 100 MZ and 100 DZ pairs were simulated for 500 replicates. The regression-based analyses were performed using the freely-available statistical package R; the variance components models were fitted using Mx (Neale, 1997).

ACKNOWLEDGMENTS

The authors would particularly like to acknowledge Pak Sham and Fruhling Rijsdijk for suggesting the revised DF_2 model; also Avshalom Caspi and Temi Moffitt for helpful discussion. SP is supported by a UK Medical Research Council Research Training Fellowship in Bioinformatics; KCK is supported by NIMH grant 1K08MH070627.

REFERENCES

Caspi, A., Taylor, A., Moffitt, T., and Plomin, R. (2000). Neighborhood deprivation affects children's mental health: environmental risks identified in a genetic design. *Psychol. Sci.* **11**(4), 338–342.

DeFries, J., and Fulker, D. (1985). Multiple regression analysis of twin data. *Behav. Genet.* **15**(5), 467–473.

DeFries, J., and Fulker, D. (1988). Multiple regression analysis of twin data: etiology of deviant scores versus individual differences. *Acta Genet. Med. Gemel.* **37**:205–216.

DiLalla, L., and Gottesman, I. (1991). Biological and genetic contributors to violence–Widom's untold tale. *Psychol. Bull.* **109**:125–129.

D'Onofrio, B., Turkheimer, E., Eaves, L., Corey, L., Berg, K., Solaas, M., and Emery, R. (2003). The role of the children of twins design in elucidating causal relations between parent characteristics and child outcomes. *J. Child Psychol. Psychiatry* **44**(8), 1130–44.

Jaffee, S. R., Caspi, A., Moffitt, T. E., and Taylor, A. (2004). Physical maltreatment victim to antisocial child: Evidence of an environmentally mediated process. *J. Abnorm. Psychol.* **113**(1), 44–55.

Kendler, K., Neale, M., Prescott, C., Kessler, R., Heath, A., Corey, L., and Eaves, L. (1996). Childhood parental loss and alcoholism in women: a causal analysis using a twin-family design. *Psychol. Med.* **26**(1), 79–95.

Koenen, K., Moffitt, T., Caspi, A., Taylor, A., and Purcell, S. (2003). Domestic violence is associated with environmental suppression of iq in young children. *Dev. Psychopathol.* **15**(2), 297–311.

Miller, P., Mulvery, C., and Martin, N. (2001). Genetic and environmental contributions to educational attainment in Australia. *Econ. Educ. Rev.* **20**:211–224.

Neale, M. (1997). Mx: Statistical modeling. Box 980126 VCU, Richmond VA 23298.

Plomin, R., and Bergeman, C. S. (1991). The nature of nurture: genetic influence on environmental measures. *Behav. Brain Sci.* **14**(3), 373–386.

Rodgers, J. L., Rowe, D. C., and Chengchang, L. (1994). Beyond nature versus nurture: DF analysis of nonshared influences on problem behaviors. *Dev. Psychol.* **30**(3), 374–384.

Turkheimer, E., D'Onofrio, B., Maes, H., and Eaves, L. (Manuscript under reivew).

Turkheimer, E., D'Onofrio, B., Maes, H., Eaves, L., and Waldron, M. (2000). Nonshared environment: a theoretical, methodological and quantitative review. *Psychol. Bull.* **126**(1), 78–108.

Widom, C. (1989). Does violence beget violence? A critical examination of the literature. *Psychol. Bull.* **106**(1), 3–28.

Edited by Stacey Cherny