

Interaction tests under exposure misclassification

BY TYLER J. VANDERWEELE

Departments of Epidemiology and Biostatistics, Harvard School of Public Health

677 Huntington Avenue, Boston, Massachusetts, 02115, U.S.A.

tvanderw@hsph.harvard.edu

SUMMARY.

Results are given concerning inferences that can be drawn about interaction when the exposures of interest are misclassified with independent non-differential misclassification. It is shown that interaction tests for interactions using the misclassified exposures give valid tests for interaction between the true exposures provided the probability of misclassification satisfies certain bounds. Results are given for standard statistical interactions, for causal interactions corresponding to synergism in the sufficient cause framework and for what is sometimes referred to as "compositional epistasis" in the genetics literature. Both two-way and three-way interactions are considered. The results require only that the probability of misclassification be no larger than $1/2$ or $1/4$ depending on the test i.e. the predictive value positive and negative need only be above $1/2$ or $3/4$. The consequences for power of interaction tests under exposure misclassification are explored through simulations. The implications of the results for epistasis and linkage disequilibrium between genetic factors are discussed and the results are illustrated by an example from the genetics literature.

Key words: Causal inference; Epistasis; Gene-gene; Interaction; Misclassification; Sufficient cause; Synergism.

1. INTRODUCTION

Misclassification and measurement error is a ubiquitous problem in studies of environmental exposures and one that is often simply ignored. Part of the justification for setting aside issues of measurement error arguably arises from a classic result of Bross (1954) who showed that if a binary exposure had non-differential misclassification then tests of association between the outcome and the misclassified exposure constitute valid tests (i.e. provide conservative type I error rates) of the association between the outcome and the true exposure. In more complex settings, such as when the exposure is not binary, analogous results do not hold (Dosemeci et al., 1990). Relatively little is known about misclassification in the context of interaction studies. The small literature on the topic that does exist (Garcia-Closas et al., 1998; Cheng & Lin, 2008; Zhang et al., 2008; Lindström et al., 2009) considers only multiplicative interaction, whereas interaction on the additive scale is generally considered more important for public health purposes and for assessing biologic synergism (Rothman & Greenland, 2008).

We will first provide a result on the implications of non-differential misclassification for standard interaction tests on the linear additive. We provide a generalization of Bross' classic result (Bross, 1954) and extend its conclusions concerning main effects to interaction parameters. We moreover will also provide analogous results on the implications of misclassification for tests for causal interactions corresponding to synergism in Rothman's sufficient cause framework (Rothman, 1976; VanderWeele & Robins, 2008) or "compositional epistasis" (Phillips, 2008; Cordell, 2009; VanderWeele, 2010ab) in genetics. In each case we will give bounds on the misclassification probabilities (positive predictive value and negative predictive value of the misclassified exposures) such that the type I error of tests are preserved (i.e. are conservative).

The contributions of this paper are effectively two-fold. First we provide new insights into the relatively undeveloped area of misclassification for interaction analy-

ses. The results are of intrinsic theoretical interest and, as is discussed in the application section, are also important for reasoning about the consequences of linkage disequilibrium in the context of the study of gene-gene and gene-environment interactions. Second, only very recently has there been any work on causal inference in the context of measurement error (Goetghebeur & Vansteelandt, 2005; Lewbel, 2007; Vansteelandt et al., 2009; Babanezhad et al., 2010). The results here concerning synergism within the sufficient cause framework and concerning "compositional epistasis" will relate misclassification to causal concepts and will thus contribute to this relatively new and emerging subfield regarding how measurement error affects our inferences within the counterfactual framework.

2. REVIEW OF TESTS FOR INTERACTIONS

Let X_1 and X_2 denote two binary exposures of interest and let D denote a binary outcome. Let $p_{x_1x_2} = E(D|X_1 = x_1, X_2 = x_2)$. The standard test for an interaction on the additive scale is

$$p_{11} - p_{10} - p_{01} + p_{00} > 0. \quad (1)$$

If (1) is satisfied then the effects of X_1 and X_2 together exceed the sum of the effects of X_1 and X_2 considered individually i.e. $(p_{11} - p_{00}) > (p_{10} - p_{00}) + (p_{01} - p_{00})$. In the following section we will consider the implications of exposure misclassification for tests of additive statistical interaction as given in (1).

In the following section we will also consider tests for more causal notions of interaction. Let Ω denote the sample space of individuals and let $D_{x_1x_2}(\omega)$ denote the counterfactual outcome (or potential outcome, Neyman, 1923; Rubin, 1974) for D for individual ω if, possibly contrary to fact, X_1 had been set to x_1 and X_2 had been set to x_2 . VanderWeele and Robins (2008) said that a sufficient cause interaction was present between X_1 and X_2 if there was some individual ω such that $D_{11}(\omega) = 1$ but $D_{10}(\omega) = D_{01}(\omega) = 0$ and they showed that if there were such an individual then synergism within Rothman's sufficient cause framework (Rothman,

1976) must be present. VanderWeele and Robins (2008) furthermore showed that if the effects of X_1 and X_2 on D were unconfounded (in the sense that $D_{x_1x_2}$ is independent of (X_1, X_2)) then

$$p_{11} - p_{10} - p_{01} > 0 \quad (2)$$

would imply the presence of a sufficient cause interaction. Note that this is in some sense a stronger notion of interaction insofar as (2) is a more stringent condition than (1); in (2) we are no longer adding p_{00} in the probability contrast. VanderWeele and Robins (2008) also noted that if the effects of X_1 and X_2 were monotonic in the sense that $D_{x_1x_2}(\omega)$ is non-decreasing in x_1 and x_2 for all ω then (1) would in fact suffice to draw the conclusion of the presence of a sufficient cause interaction (VanderWeele, 2008; cf. Rothman & Greenland, 1998); testing (2) would be necessary without such monotonicity (VanderWeele & Robins, 2007, 2008).

VanderWeele (2010ab) discussed empirical tests for an even stronger notion of interaction. In the genetics literature, when gene-gene interactions are in view, "compositional epistasis" (Phillips, 2008; Cordell, 2009) is said to be present if there is some individual ω such that $D_{11}(\omega) = 1$ and $D_{10}(\omega) = D_{01}(\omega) = D_{00}(\omega) = 0$. In such cases the effect of the first genetic factor X_1 is masked unless X_2 is present (and likewise the effect of the X_2 is masked unless X_1 is present). VanderWeele (2010ab) noted that if the effects of X_1 and X_2 on D were unconfounded then

$$p_{11} - p_{10} - p_{01} - p_{00} > 0 \quad (3)$$

would imply an "epistatic interaction" i.e. the presence of compositional epistasis. In contrast with (1) we subtract rather than add p_{00} . If $D_{x_1x_2}(\omega)$ is non-decreasing in one of x_1 or x_2 for all ω then (2) suffices for compositional epistasis; if $D_{x_1x_2}(\omega)$ is non-decreasing in both x_1 and x_2 for all ω then (1) implies compositional epistasis. In the following section we will give results concerning exposure misclassification for

each of the contrasts in (1), (2) and (3).

We will moreover also give results concerning exposure misclassification for three-way sufficient cause interactions. If we let $D_{x_1x_2x_3}$ denote the counterfactual outcome D for individual ω if X_1 , X_2 , and X_3 had been set, possibly contrary to fact, to x_1 , x_2 and x_3 respectively, we say that there is a 3-way sufficient cause interaction between X_1 , X_2 , and X_3 if there is an individual ω such that $D_{111}(\omega) = 1$ and $D_{110}(\omega) = D_{101}(\omega) = D_{011}(\omega) = 0$; a 3-way sufficient cause interaction implies in Rothman's sufficient cause framework the presence of a sufficient cause which contains all of X_1 , X_2 , and X_3 (VanderWeele & Robins, 2008). VanderWeele and Robins (2008) noted that if we let $p_{x_1x_2x_3} = E(D|X_1 = x_1, X_2 = x_2, X_3 = x_3)$ then if the effects of X_1 , X_2 and X_3 on D are unconfounded and if

$$p_{111} - p_{110} - p_{101} - p_{011} > 0 \quad (4)$$

then a sufficient cause interaction must be present between X_1 , X_2 and X_3 . Finally if $D_{x_1x_2x_3}$ is non-decreasing in x_1 , x_2 and x_3 then any of the following three conditions (VanderWeele & Robins, 2008) imply that a sufficient cause interaction is present between X_1 , X_2 and X_3 :

$$p_{111} - p_{110} - p_{101} - p_{011} + p_{100} + p_{010} > 0 \quad (5)$$

$$p_{111} - p_{110} - p_{101} - p_{011} + p_{100} + p_{001} > 0$$

$$p_{111} - p_{110} - p_{101} - p_{011} + p_{010} + p_{001} > 0.$$

If the effects of $\{X_1, X_2\}$ or $\{X_1, X_2, X_3\}$ on D are unconfounded conditional on some set of covariates C , then conditions (1)-(5) can also be made conditional on C . The issue of modeling and robustness to control for some set of confounding variables C is considered elsewhere (Vansteelandt et al., 2008; VanderWeele, 2009; VanderWeele et al., 2010). In the next section we show that conditions (1)-(5) as applied to the potentially misclassified exposures imply that the analogous conditions

also hold for the true exposures provided that the probabilities of misclassification satisfy certain bounds.

3. EMPIRICAL TESTS FOR INTERACTIONS UNDER EXPOSURE MISCLASSIFICATION

We derive bounds on the probabilities that observed exposures are misclassified such that conditions (1)-(5) applied to the observed exposures, rather than the true exposures, still imply that the interaction contrasts (1)-(5) are satisfied for the true exposures. Let X_i denote the true exposure and let X_i^* denote the observed exposure. Let $d_i = P(X_i = 0|X_i^* = 1)$ and let $u_i = P(X_i = 1|X_i^* = 0)$. These misclassification probabilities are essentially 1 minus the predictive value positive and the predictive value negative of the measured exposures for the true exposures. Following terminology in the literature, we say that the misclassification is non-differential if $P(D|X_1 = i, X_2 = j, X_3 = k, X_1^* = l, X_2^* = m, X_3^* = n) = P(D|X_1 = i, X_2 = j, X_3 = k)$. Non-differential misclassification implies that the measured exposures gives no information about the outcome beyond the information it gives about the true exposures. We say that the misclassification is independent if the events of X_1 , X_2 and X_3 being misclassified are independent i.e. that $P(X_1 = i, X_2 = j, X_3 = k|X_1^* = l, X_2^* = m, X_3^* = n) = P(X_1 = i|X_1^* = l)P(X_2 = j|X_2^* = m)P(X_3 = k|X_3^* = n)$. Let $p_{x_1x_2}^* = E(D|X_1^* = x_1, X_2^* = x_2)$ and $p_{x_1x_2x_3}^* = E(D|X_1^* = x_1, X_2^* = x_2, X_3^* = x_3)$.

We then have the following results. The proofs of Theorems 1-5 are given in the Appendix.

THEOREM 1. If misclassification of X_1 and X_2 is non-differential and independent and if $d_i < 1/2$ and $u_i < 1/2$ then

$$p_{11}^* - p_{10}^* - p_{01}^* + p_{00}^* > 0 \tag{6}$$

implies $p_{11} - p_{10} - p_{01} + p_{00} > 0$.

If condition (6) is satisfied for the measured exposures then the true exposure also exhibit statistical interaction. If in addition the effects of the true exposures on the outcome are unconfounded and monotonic then the true exposure exhibit a sufficient cause interaction and an epistatic interaction (VanderWeele & Robins, 2008; VanderWeele 2010ab). It can be verified that Theorem 1 holds if for each i only one of the two inequalities $d_i \leq 1/2$ and $u_i \leq 1/2$ holds strictly. Theorem 1 is a generalization for interaction of the classic result of Bross (1954) for main effects.

Theorem 2 gives an analogous result for the empirical test for a 2-way sufficient cause interactions when at least one of the exposures does not have a monontonic effect on the outcome.

THEOREM 2. If misclassification of X_1 and X_2 is non-differential and independent and if $d_i < 1/2$ and $u_i < 1/4$ then

$$p_{11}^* - p_{10}^* - p_{01}^* > 0 \tag{7}$$

implies $p_{11} - p_{10} - p_{01} > 0$.

If condition (7) is satisfied for the measured exposures and the effects of the true exposures on the outcome are unconfounded then the true exposures exhibit a sufficient cause interaction; if, in addition, at least one of the exposures has a monontonic effect on the outcome then the true exposures also exhibit an epistatic interaction (VanderWeele, 2010ab). It can be verified from the proof of Theorem 2 that the theorem would still hold if only one of the four inequalities $d_1 \leq 1/2$, $d_2 \leq 1/2$, $u_1 \leq 1/4$, $u_2 \leq 1/4$ holds strictly.

Theorem 3 below gives an analogous result for the empirical test for a 2-way epistatic interactions when no assumptions are made about monotonicity.

THEOREM 3. If misclassification of X_1 and X_2 is non-differential and indepen-

dent and if $d_i \leq 1/3$ and $u_i \leq 1/4$ then

$$p_{11}^* - p_{10}^* - p_{01}^* - p_{00}^* > 0 \quad (8)$$

implies $p_{11} - p_{10} - p_{01} - p_{00} > 0$.

If condition (8) is satisfied for the measured exposures and the effects of the true exposures on the outcome are unconfounded then the true exposures exhibit an epistatic interaction, without any assumptions about monotonicity. For a three-way sufficient cause interaction, without the monotonicity assumption, more stringent inequalities on the misclassification probabilities are needed; all the misclassification probabilities d_i and u_i are required to be than $1/4$ as stated in Theorem 4.

THEOREM 4. If misclassification of X_1 , X_2 , and X_3 is non-differential and independent and if $d_i < 1/4$ and $u_i < 1/4$ then

$$p_{111}^* - p_{110}^* - p_{101}^* - p_{011}^* > 0 \quad (9)$$

implies $p_{111} - p_{110} - p_{101} - p_{011}$.

If condition (9) is satisfied for the measured exposures and the effects of the true exposures on the outcome are unconfounded then the true exposures exhibit a 3-way sufficient cause interaction, without any assumptions about monotonicity. Theorem 4 also holds if just one of the six inequalities on the misclassification probabilities holds strictly. If the effects of X_1 , X_2 and X_3 on D are monotonic then the bounds on the misclassification probabilities which are required in order to be able to conclude the presence of a 3-way sufficient cause interaction are no more stringent than the bounds in the case of a 2-way interaction, namely all the misclassification probabilities must be less than $1/2$.

THEOREM 5. If $p_{x_1 x_2 x_3}$ is non-decreasing in x_1 , x_2 and x_3 and $d_i < 1/2$ and

$u_i < 1/2$ then any of the following three conditions

$$\begin{aligned}
p_{111}^* - p_{110}^* - p_{101}^* - p_{011}^* + p_{100}^* + p_{010}^* &> 0 \\
p_{111}^* - p_{110}^* - p_{101}^* - p_{011}^* + p_{100}^* + p_{001}^* &> 0 \\
p_{111}^* - p_{110}^* - p_{101}^* - p_{011}^* + p_{010}^* + p_{001}^* &> 0.
\end{aligned} \tag{10}$$

implies that the analogous inequality for the true exposures in (5) is satisfied.

If one of the conditions in (10) is satisfied for the measured exposures and the effects of the true exposures on the outcome are unconfounded then the true exposures exhibit a 3-way sufficient cause interaction. Theorem 5 also holds if for each i only one of $d_i \leq 1/2$ and $u_i \leq 1/2$ holds strictly. Theorems 1-5 apply also if the effects of X_1 and X_2 or X_1 , X_2 and X_3 on D are unconfounded conditional on some set of covariates C . Conditions (6)-(10) can then be made conditional on $C = c$ and we would define $u_{ic} = P(X_i = 1|X_i^* = 0, C = c)$ and $d_{ic} = P(X_i = 0|X_i^* = 1, C = c)$ and require that $P(X_1 = i, X_2 = j, X_3 = k|X_1^* = l, X_2^* = m, X_3^* = n, C = c) = P(X_1 = i|X_1^* = l, C = c)P(X_2 = j|X_2^* = m, C = c)P(X_3 = k|X_3^* = n, C = c)$.

Often, out of convenience, logistic or log-linear models are fit to data with a dichotomous outcome. The results above have implications for such models. It can be shown that for two dichotomous exposures, if a log-linear model is fit to the observed data, $\log(p_{x_1x_2}^*) = \log\{P(D = 1|X_1^* = x_1, X_2^* = x_2, C = c)\} = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_1x_2 + \beta_4'c$ then provided β_1 and β_2 are non-negative, it follows that $\beta_3 > 0$ implies (6), $\beta_3 > \log(2)$ implies (7), and $\beta_3 > \log(3)$ implies (8) (VanderWeele, 2010ab). From Theorems 1-3, it follows that if we fit the log-linear model to the observed, potentially misclassified, data then provided β_1 and β_2 are non-negative, $\beta_3 > 0$ implies (1), $\beta_3 > \log(2)$ implies (2), and $\beta_3 > \log(3)$ implies (3). We can draw conclusions about various forms of statistical and causal interaction from the misclassified data and log-linear model; the same results hold approximately if a logistic model is used and the outcome is rare so that $\log(p_{x_1x_2}^*) \approx \text{logit}(p_{x_1x_2}^*)$.

4. SIMULATIONS

We use simulations to examine the effects of misclassification of the exposure variables on power for tests to detect interaction. For each scenario we generated 5,000 simulated data sets with two binary exposures X_1 and X_2 and binary outcome D . Scenarios vary according to sample size ($n = 250, 500, 1000, 2500, 5000$) and the extent of misclassification. Specifically we consider six misclassification scenarios: (i) $d_1 = u_1 = 0, d_2 = u_2 = 0$ (no misclassification); (ii) $d_1 = u_1 = 0.1, d_2 = u_2 = 0$ (moderate misclassification of X_1); (iii) $d_1 = u_1 = 0.25, d_2 = u_2 = 0$ (substantial misclassification of X_1); (iv) $d_1 = u_1 = 0.1, d_2 = u_2 = 0.1$ (moderate misclassification of X_1 and X_2); (v) $d_1 = u_1 = 0.25, d_2 = u_2 = 0.1$ (substantial misclassification of X_1 , moderate misclassification of X_2); (vi) $d_1 = u_1 = 0.25, d_2 = u_2 = 0.25$ (moderate misclassification of X_1 and X_2). We assume $P(X_1 = i, X_2 = j) = 0.25$ for all i, j for the true exposures, X_1 and X_2 . In generating the outcome D , we consider two different scenarios for the magnitude of the standard interaction contrast, $p_{11} - p_{10} - p_{01} + p_{00}$: either moderate interaction ($p_{11} - p_{10} - p_{01} + p_{00} = 0.1$ with $p_{00} = 0.1, p_{10} = 0.3, p_{01} = 0.2, p_{11} = 0.5$) or substantial interaction ($p_{11} - p_{10} - p_{01} + p_{00} = 0.2$ with $p_{00} = 0.1, p_{10} = 0.3, p_{01} = 0.2, p_{11} = 0.6$). Note that for the underlying distribution of the outcome under the true exposures we have that $p_{11} - p_{10} - p_{01} + p_{00} > 0$. We assume data is only available on the misclassified exposures, X_1^* and X_2^* . For each simulated data, we use the misclassified exposures to test $p_{11}^* - p_{10}^* - p_{01}^* + p_{00}^* > 0$ using the test statistic:

$$Z = \frac{\hat{p}_{11}^* - \hat{p}_{10}^* - \hat{p}_{01}^* + \hat{p}_{00}^*}{\sqrt{\frac{\hat{p}_{11}^*(1-\hat{p}_{11}^*)}{n_{11}} + \frac{\hat{p}_{10}^*(1-\hat{p}_{10}^*)}{n_{10}} + \frac{\hat{p}_{01}^*(1-\hat{p}_{01}^*)}{n_{01}} + \frac{\hat{p}_{00}^*(1-\hat{p}_{00}^*)}{n_{00}}}} \sim N(0, 1)$$

where $\hat{p}_{ij}^*$ is the sample estimate of $P(D = 1|X_1^* = i, X_2^* = j)$ and n_{ij} is the number in the sample with $X_1^* = i, X_2^* = j$. The tables below report the proportion of the 5,000 simulated datasets for which $Z > 1.96$. Note that by Theorem 1, if $p_{11}^* - p_{10}^* - p_{01}^* + p_{00}^* > 0$ then $p_{11} - p_{10} - p_{01} + p_{00} > 0$. Table 1 reports the

power estimates for the moderate interaction scenario ($p_{11} - p_{10} - p_{01} + p_{00} = 0.1$). The first column presents power estimates without misclassification ($d_1 = u_1 = 0$, $d_2 = u_2 = 0$). The remaining five columns present power estimates under differing misclassification scenarios. With a moderate interaction, power decreases noticeably even when one of the exposures is subject to moderate misclassification. When both exposures are subject to substantial misclassification, power is poor even at a sample size of $n = 5000$.

Table 2 reports the power estimates for the substantial interaction scenario ($p_{11} - p_{10} - p_{01} + p_{00} = 0.2$). With a substantial interaction, power is fairly good at sample size of $n = 1000$ and above in the scenarios in which one or both exposures are subject to moderate misclassification ($d_1, u_1, d_2, u_2 \leq 0.1$). If one of the exposures is subject to substantial misclassification ($d_1 = u_1 = 0.25$), power remains relatively high so long as the sample size is $n = 2500$ and above. Note, however, that when both exposures are subject to substantial misclassification ($d_1 = u_1 = d_2 = u_2 = 0.25$) power is below 50% even with a sample size of $n = 5000$.

5. APPLICATION TO EPISTASIS AND LINKAGE DISEQUILIBRIUM

The results above on misclassification of exposures in inference for interaction have important implications in the interpretation of interaction studies in genetic epidemiology. Environmental exposures are often misclassified; furthermore, correctly measured continuous exposures with a threshold effect are often dichotomized and are not always dichotomized at the biologically relevant threshold, also giving rise to potential misclassification. Although genetic information is now available with relatively high accuracy, misclassification of genetic factors still occurs. The results above imply that provided the misclassification of factors in interaction analyses is independent and non-differential, tests for interaction using data with mismeasured exposures will still constitute valid tests of interaction for the true exposures.

The implications of results above are of perhaps even greater importance in addressing issues of linkage disequilibrium. Linkage disequilibrium is the phenomenon

whereby genetic variants at loci on the same chromosome will often be statistically dependent, because loci close together will generally not be separated during the crossover of recombination events. As a result of linkage disequilibrium, genetic variants at loci on the same chromosome close to one another are generally correlated. In part, this aids research in genetics because markers at loci near the true causal variant will be correlated with the true causal variant and it will often be possible to detect a signal even if there is no marker at the locus of the causal variant. The phenomenon of linkage disequilibrium does, however, make it more difficult to assess whether a particular genetic variant that is correlated with some outcome is in fact causal since non-causal variants, associated with the causal variant, will also be correlated with the outcome. Gene-gene interactions (also sometimes referred to as "epistasis" or "statistical epistasis") are thought by some to explain a substantial proportion of the heritability that our current genetic risk prediction measures miss (Maher, 2008; Moore & Williams, 2009). However, when tests of gene-gene interaction are of interest, ensuring that both genetic factors represent the true causal variants is challenging and generally lies beyond our current state of knowledge.

Instead, when a particular genetic risk factor is in linkage disequilibrium with the true causal genetic factor, we might conceptualize the genetic factor as a misclassified version of the true causal factor. In cases in which two genetic risk factors under study are on separate chromosomes (serving as proxies for the true causal variants on each chromosome), this misclassification by linkage disequilibrium will likely be independent. The results in this paper then imply that a test for interaction between genetic factors for which data is available will constitute a valid test for gene-gene interaction between the true causal variants provided the predictive value positive and negative are e.g. at least 0.75.

We illustrate these remarks and the results in this paper by considering an example drawn from the existing literature on gene-gene interactions. The analysis here is given for illustrative purposes only; a more definitive analysis would require

careful control for confounding by population stratification (Hoggart et al., 2003). We consider data presented by Yang et al. (2005) who assess possible interaction between ALDH2 and ADH2 polymorphisms on esophageal cancer. Specifically, let X_1 denote the presence of *Arg* variants for the AHD2 genotype and let X_2 denote the *Glu/Glu* versus *Glu/Lys* ALDH2 genotype. As AHD2 is located on chromosome 4 and ALDH2 is located on chromosome 12, the genetic factors will likely be independent of one another and if the factors are considered as potentially misclassified versions of the true causal variants, this misclassification will likely be independent. Odds ratios obtained from the number of cases and controls in Yang et al. (2005) are summarized in Table 3.

Because data from Yang et al. (2005) were obtained from a case-control study, we do not have data on risks and cannot test the conditions in sections 2 and 3 directly. However, because esophageal cancer is rare, odds ratios approximate risk ratios and dividing condition (6) by p_{00}^* we obtain

$$RR_{11}^* - RR_{10}^* - RR_{01}^* + 1 > 0$$

where $RR_{x_1x_2}^* = p_{x_1x_2}^*/p_{00}^*$. This quantity is sometimes referred to as *REI* or the relative excess risk due to interaction and thus condition (6) could be expressed as $REI > 0$. It is relatively straightforward to obtain confidence intervals for the relative excess risk due to interaction (Hosmer & Lemeshow, 1992; Richardson & Kaufman, 2009). Likewise conditions (7) and (8) can be expressed respectively as $REI > 1$ and $REI > 2$. Using the data from Yang et al. (2005) we obtain $RR_{11}^* - RR_{10}^* - RR_{01}^* + 1 = 7.20 - 1.40 - 3.52 + 1 = 3.28$ (95% CI: 0.40, 6.16). Provided the predictive value positive and negative for the genetic factors are at least 0.5, by Theorem 1 we have evidence for a statistical interaction between the effects on esophageal cancer of the true causal variants on chromosome 4 and 12 associated with AHD2 and ALDH2 respectively since both the point estimate and

the confidence interval would imply that $p_{11}^* - p_{10}^* - p_{01}^* + p_{00}^* > 0$. If monotonicity assumptions are thought reasonable for the true genetic factors then this would also give evidence for a sufficient cause interaction and for compositional epistasis as discussed in sections 2 and 3. Even without the monotonicity assumptions, the point estimate for $RR_{11}^* - RR_{10}^* - RR_{01}^* + 1$ of 3.28 suggests evidence (by Theorems 2 and 3) of a sufficient cause interaction and compositional epistasis since the estimate satisfies $REI > 1$ and $REI > 2$ but note that the confidence interval for $RR_{11}^* - RR_{10}^* - RR_{01}^* + 1$ does include the values 1 and 2 and we could thus not reject the null that conditions (7) and (8) are not satisfied. The analyses here are from a single study and would require replication and control for population stratification to clearly establish the biologic conclusions.

6. CONCLUDING REMARKS

The results in this paper demonstrate that provided the probabilities of misclassification lie within very reasonable bounds, less than 1/4 or 1/2 depending on the context, conclusions from interaction tests including additive statistical interaction, synergism within the sufficient cause framework (VanderWeele and Robins, 2008) and compositional epistasis in genetics (Phillips, 2008; Cordell, 2009; VanderWeele, 2010ab), still hold even if the exposures are misclassified. The results here could be extended in a number of directions. First, although we have not assumed that the exposures are themselves independent, we have assumed that the misclassification of the two exposures occurs independently. Work could be done examining what, if anything, could be concluded in settings in which the misclassification of one exposure was associated with the misclassification of another exposure. Second, the results given above concern the tests for 2-way and 3-way interactions. In other work, we have given empirical conditions for n -way sufficient cause interactions. Future work could be done investigating the necessary bounds on misclassification probabilities for these more general empirical conditions. In particular, it would be of interest to know if the bounds that are required without the monotonicity

assumption continue to get smaller as n increases; it would also be of interest to know if, when the monotonicity assumption holds, the bounds on misclassification probabilities stay fixed at $1/2$ as n increases. A final extension to the present work that may be of interest would be accommodating possible misclassification of the outcome. In most epidemiologic studies the quality of the outcome data is generally much better than the quality of the data for the exposures. Nevertheless, it may be of interest to examine to the extent to which the empirical conditions for sufficient cause interactions apply also to settings in which the outcome data is also misclassified.

ACKNOWLEDGEMENTS. This research was funded by National Institutes of Health, U.S.A.

REFERENCES

- BABANEZHAD, M., VANSTEEELANDT, S. & GOETGHEBEUR, E. (2010). Comparison of causal effect estimators under exposure misclassification. *Journal of Statistical Planning and Inference*, 140:1306-1319.
- BROSS, I. (1954). Misclassification in 2x2 tables. *Biometrics*, **10**, 478-486.
- CHENG, K.F. & LIN, W.J. (2009). The effects of misclassification in studies of gene-environment interactions. *Human Heredity*, **67**, 77-87.
- CORDELL, H.J. (2009). Detecting gene-gene interaction that underlie human diseases. *Nat Rev Genet*, **10**, 392-404.
- DOSEMECI, M., WACHOLDER, S. & LUBIN, J.H. (1990). Does nondifferential misclassification of exposure always bias a true effect towards the null value? *Am. J. Epidemiol.*, **132**, 746-748.

- GARCIA-CLOSAS, M., THOMPSON, W.D. & ROBINS, J.M. (1998). Differential misclassifications and the assessment of gene-environment interactions. *Am. J. Epidemiol.*, **147**, 426-433.
- GOETGHEBEUR, E. & VANSTEELANDT, S. (2005). Structural mean models for compliance analysis in randomized clinical trials and the impact of errors on measures of exposure. *Statist. Meth. Med. Res.*, **14**, 397-415.
- HOGGART, C.J., PARRA, E.J., SHRIVER, M.D., BONILLA, C., KITTLES, R.A., CLAYTON, D.G., & MCKEIGUE, P.M. (2003). Control of confounding of genetic associations in stratified populations. *Am. J. Hum. Genet.*, **72**, 1492-1504.
- HOSMER, D.W. & LEMESHOW, S. (1992). Confidence interval estimation of interaction. *Epidemiology*; **3**, 452-56.
- LEWBEL, A. (2007). Estimation of Average Treatment Effects with Misclassification. *Econometrica*, **75**, 537-551.
- LINDSTRÖM, S., YEN, Y.-C., SPIEGELMAN, D. & KRAFT, P. (2009). The impact of gene-environment dependence and misclassification in genetic association studies incorporating gene-environment interactions. *Human Heredity*, **68**, 171-181.
- MAHER, B. (2008). The case of the missing heritability. *Nature*, **456**(7218): 18-23.
- MOORE, J.H. & WILLIAMS S.M. (2009) Epistasis and its implications for personal genetics. *American Journal of Human Genetics*, **85**, 309-320.
- NEYMAN, J. (1923). Sur les applications de la th  orie des probabilit  s aux experiences Agaricales: Essay des principe. Excerpts reprinted (1990) in English (D. Dabrowska and T. Speed, Trans.) in *Statistical Science*, **5**, 463-472.
- PHILLIPS, P.C. (2008). Epistasis – the essential role of gene interactions in the structure and evolution of genetic systems. *Nat. Rev. Genet.*, **9**, 855-867.

- RICHARDSON, D.B. AND KAUFMAN, J.S. (2009). Estimation of the relative excess risk due to interaction and associated confidence bounds. *Am. J. Epidemiol.*, **169**, 756-60.
- ROTHMAN, K.J. (1976). Causes. *Am. J. Epidemiol.*, **104**, 587-92.
- ROTHMAN, K.J. & GREENLAND, S. (1998). *Modern Epidemiology*. Philadelphia: Lippincott-Raven.
- RUBIN, D.B. (1974). Estimating causal effects of treatments in randomized and non-randomized studies. *Journal of Educational Psychology*, **66**, 688-701.
- VANDERWEELE, T.J. (2009). Sufficient cause interactions and statistical interactions. *Epidemiology*, **20**, 6-13.
- VANDERWEELE, T.J. (2010). Empirical tests for compositional epistasis. *Nature Reviews Genetics*, **11**, 166.
- VANDERWEELE, T.J. (2010). Epistatic interactions. *Statistical Applications in Genetics and Molecular Biology*, **9**, Article 1, 1-22.
- VANDERWEELE, T.J. AND ROBINS J.M. (2007). The identification of synergism in the sufficient-component cause framework. *Epidemiol.*, **18**, 329-339.
- VANDERWEELE, T.J. & ROBINS, J.M. (2008). Empirical and counterfactual conditions for sufficient cause interactions. *Biometrika*, **95**, 49-61.
- VANDERWEELE, T.J., VANSTEELANDT, S. & ROBINS, J.M. (2010). Marginal structural models for sufficient cause interactions. *American Journal of Epidemiology*, **171**, 506-514.
- VANSTEELANDT, S., BABANEZHAD, M. AND GOETGHEBEUR, E. (2008). Correcting instrumental variables estimators for systematic measurement error. *Statistica Sinica*, **19**, 1223-1246.

VANSTEELANDT, S., VANDERWEELE, T.J., TCHETGEN, E.J., ROBINS, J.M. (2008). Multiply robust inference for statistical interactions. *Journal of the American Statistical Association*, **103**, 1693–1704.

YANG, C.-X., MATSUO, K., ITO, H., KIROSE, K., WAKAI, K., SAITO, T., SHINODA, M., HATOOKA, S., MIZUTANI, K. & TAJIMA, K. (2005). Esophageal cancer risk by ALDH2 and ADH2 polymorphisms and alcohol consumption: exploration of gene-environment and gene-gene interactions. *Asian Pacific Journal of Cancer Prevention*, **6**, 256-262.

ZHANG, L., MUKHERJEE, B., GHOSH, M., GRUBER, S., & MORENO, V. (2008). Accounting for error due to misclassification of exposures in case-control studies of gene-environment interaction. *Statistics in Medicine*, **27**, 2756–2783.

APPENDIX.

Proof of Theorem 1. We can re-write $p_{11}^* - p_{10}^* - p_{01}^* + p_{00}^*$ as follows

$$\begin{aligned}
& +(1-d_1)(1-d_2)p_{11} - (1-d_1)u_2p_{11} - u_1(1-d_2)p_{11} + u_1u_2p_{11} \\
& +(1-d_1)d_2p_{10} - (1-d_1)(1-u_2)p_{10} - u_1d_2p_{10} + u_1(1-u_2)p_{10} \\
& +d_1(1-d_2)p_{01} - d_1u_2p_{01} - (1-u_1)(1-d_2)p_{01} + (1-u_1)u_2p_{01} \\
& +d_1d_2p_{00} - d_1(1-u_2)p_{00} - (1-u_1)d_2p_{00} + (1-u_1)(1-u_2)p_{00} \\
& = (1-d_1-u_1)(1-d_2-u_2)(p_{11} - p_{10} - p_{01} + p_{00})
\end{aligned}$$

Thus if $d_i < 1/2$ and $u_i < 1/2$ and $p_{11}^* - p_{10}^* - p_{01}^* + p_{00}^* > 0$ then $p_{11} - p_{10} - p_{01} + p_{00} > 0$.

Proof of Theorem 2. We can re-write $p_{11}^* - p_{10}^* - p_{01}^*$ as follows

$$\begin{aligned}
& +(1-d_1)(1-d_2)p_{11} - (1-d_1)u_2p_{11} - u_1(1-d_2)p_{11} \\
& +(1-d_1)d_2p_{10} - (1-d_1)(1-u_2)p_{10} - u_1d_2p_{10} \\
& +d_1(1-d_2)p_{01} - d_1u_2p_{01} - (1-u_1)(1-d_2)p_{01} \\
& +d_1d_2p_{00} - d_1(1-u_2)p_{00} - (1-u_1)d_2p_{00}
\end{aligned}$$

Let $V = (1-d_1)(1-d_2) - (1-d_1)u_2 - u_1(1-d_2)$. We then re-write $p_{11}^* - p_{10}^* - p_{01}^*$ as

$$Vp_{11} - (V+u_2)p_{10} - (V+u_1)p_{01} + \{d_1d_2 - d_1(1-u_2) - (1-u_1)d_2\}p_{00}.$$

Thus if $p_{11}^* - p_{10}^* - p_{01}^* > 0$ then

$$V(p_{11} - p_{10} - p_{01}) > u_2p_{10} + u_1p_{01} + \{d_1(1-u_2) + (1-u_1)d_2 - d_1d_2\}p_{00}.$$

If $d_i < 1/2$ and $u_i < 1/4$ then

$$\begin{aligned} d_1(1 - u_2) + (1 - u_1)d_2 - d_1d_2 &= d_1(1 - u_2 - \frac{1}{2}d_2) + d_2(1 - u_1 - \frac{1}{2}d_1) \\ &> d_1(1 - \frac{1}{4} - \frac{1}{2}\frac{1}{2}) + d_2(1 - \frac{1}{4} - \frac{1}{2}\frac{1}{2}) > 0. \end{aligned}$$

Thus $u_2p_{10} + u_1p_{01} + \{d_1(1 - u_2) + (1 - u_1)d_2 - d_1d_2\}p_{00} \geq 0$ and so if $p_{11}^* - p_{10}^* - p_{01}^* > 0$ then $V(p_{11} - p_{10} - p_{01}) > 0$. The result follows if $V > 0$. If $d_i < 1/2$ and $u_i < 1/4$ then

$$\begin{aligned} V &= (1 - d_1)(1 - d_2) - (1 - d_1)u_2 - u_1(1 - d_2) \\ &= (1 - d_1)\{\frac{1}{2}(1 - d_2) - u_2\} + (1 - d_2)\{\frac{1}{2}(1 - d_1) - u_1\} \\ &> (1 - d_1)(\frac{1}{2}\frac{1}{2} - \frac{1}{4}) + (1 - d_2)(\frac{1}{2}\frac{1}{2} - \frac{1}{4}) \geq 0. \end{aligned}$$

Proof of Theorem 3. We can re-write $p_{11}^* - p_{10}^* - p_{01}^* - p_{00}^*$ as follows

$$\begin{aligned} &+(1 - d_1)(1 - d_2)p_{11} - (1 - d_1)u_2p_{11} - u_1(1 - d_2)p_{11} - u_1u_2p_{11} \\ &+(1 - d_1)d_2p_{10} - (1 - d_1)(1 - u_2)p_{10} - u_1d_2p_{10} - u_1(1 - u_2)p_{10} \\ &+d_1(1 - d_2)p_{01} - d_1u_2p_{01} - (1 - u_1)(1 - d_2)p_{01} - (1 - u_1)u_2p_{01} \\ &+d_1d_2p_{00} - d_1(1 - u_2)p_{00} - (1 - u_1)d_2p_{00} - (1 - u_1)(1 - u_2)p_{00} \end{aligned}$$

Let $V = (1 - d_1)(1 - d_2) - (1 - d_1)u_2 - u_1(1 - d_2) - u_1u_2$. Re-write $p_{11}^* - p_{10}^* - p_{01}^* - p_{00}^*$ as

$$Vp_{11} - (V + u_1 + u_2)p_{10} - (V + u_1 + u_2)p_{01} - \{V + 2(d_1 + d_2 + u_1u_2 - d_1d_2 - d_1u_2 - u_1d_2)\}p_{00}.$$

Thus if $p_{11}^* - p_{10}^* - p_{01}^* - p_{00}^* > 0$ then

$$V(p_{11} - p_{10} - p_{01} - p_{00}) > (u_1 + u_2)(p_{10} + p_{00}) + 2(d_1 + d_2 + u_1u_2 - d_1d_2 - d_1u_2 - u_1d_2)p_{00}.$$

Denote the right side of the inequality by H . If $d_i \leq 1/3$ and $u_i \leq 1/4$ then

$$\begin{aligned}
& (d_1 + d_2 + u_1 u_2 - d_1 d_2 - d_1 u_2 - u_1 d_2) \\
&= d_1(1 - u_2 - \frac{1}{2}d_2) + d_2(1 - u_1 - \frac{1}{2}d_1) + u_1 u_2 \\
&\geq d_1(1 - \frac{1}{4} - \frac{1}{2}\frac{1}{3}) + d_2(1 - \frac{1}{4} - \frac{1}{2}\frac{1}{3}) + u_1 u_2 \geq 0.
\end{aligned}$$

Thus H is non-negative and consequently if $p_{11}^* - p_{10}^* - p_{01}^* - p_{00}^* > 0$ then $V(p_{11} - p_{10} - p_{01} - p_{00}) > 0$. The result follows provided $V > 0$. If $d_i \leq 1/3$ and $u_i \leq 1/4$ then

$$\begin{aligned}
V &= (1 - d_1)(1 - d_2) - (1 - d_1)u_2 - u_1(1 - d_2) - u_1 u_2 \\
&= (1 - d_1)\{\frac{3}{8}(1 - d_2) - u_2\} + (1 - d_2)\{\frac{3}{8}(1 - d_1) - u_1\} + \{\frac{1}{4}(1 - d_1)(1 - d_2) - u_1 u_2\} \\
&\geq (1 - d_1)(\frac{3}{8}\frac{2}{3} - \frac{1}{4}) + (1 - d_2)(\frac{3}{8}\frac{2}{3} - \frac{1}{4}) + \{\frac{1}{4}\frac{2}{3}\frac{2}{3} - \frac{1}{4}\frac{1}{4}\} > 0.
\end{aligned}$$

Proof of Theorem 4. We can re-write $p_{111}^* - p_{110}^* - p_{101}^* - p_{111}^*$ as follows

$$\begin{aligned}
& +(1 - d_1)(1 - d_2)(1 - d_3)p_{111} - (1 - d_1)(1 - d_2)u_3p_{111} - (1 - d_1)u_2(1 - d_3)p_{111} \\
& \quad - u_1(1 - d_2)(1 - d_3)p_{111} \\
& +(1 - d_1)(1 - d_2)d_3p_{110} - (1 - d_1)(1 - d_2)(1 - u_3)p_{110} - (1 - d_1)u_2d_3p_{110} - u_1(1 - d_2)d_3p_{110} \\
& +(1 - d_1)d_2(1 - d_3)p_{101} - (1 - d_1)d_2u_3p_{101} - (1 - d_1)(1 - u_2)(1 - d_3)p_{101} - u_1d_2(1 - d_3)p_{101} \\
& +d_1(1 - d_2)(1 - d_3)p_{011} - d_1(1 - d_2)u_3p_{011} - d_1u_2(1 - d_3)p_{011} - (1 - u_1)(1 - d_2)(1 - d_3)p_{011} \\
& +(1 - d_1)d_2d_3p_{100} - (1 - d_1)d_2(1 - u_3)p_{100} - (1 - d_1)(1 - u_2)d_3p_{100} - u_1d_2d_3p_{100} \\
& +d_1(1 - d_2)d_3p_{010} - d_1(1 - d_2)(1 - u_3)p_{010} - d_1u_2d_3p_{010} - (1 - u_1)(1 - d_2)d_3p_{010} \\
& +d_1d_2(1 - d_3)p_{001} - d_1d_2u_3p_{001} - d_1(1 - u_2)(1 - d_3)p_{001} - (1 - u_1)d_2(1 - d_3)p_{001} \\
& +d_1d_2d_3p_{000} - d_1d_2(1 - u_3)p_{000} - d_1(1 - u_2)d_3p_{000} - (1 - u_1)d_2d_3p_{000}
\end{aligned}$$

Let Q_i denote the i th line of the above expression (the first two lines being treated as

one). Let

$$V = (1 - d_1)(1 - d_2)(1 - d_3) - (1 - d_1)(1 - d_2)u_3 - (1 - d_1)u_2(1 - d_3) - u_1(1 - d_2)(1 - d_3)$$

so that $Q_1 = Vp_{111}$. Then $Q_2 = (V - u_2d_1 + u_2 - u_1d_2 + u_1)p_{110}$, $Q_3 = (V - u_3d_1 + u_3 - u_1d_3 + u_1)p_{101}$ and $Q_4 = (V - u_2d_3 + u_2 - u_3d_2 + u_3)p_{011}$. So if $p_{111}^* - p_{110}^* - p_{101}^* - p_{011}^* > 0$ then

$$\begin{aligned} V(p_{111} - p_{110} - p_{101} - p_{011}) &> \{u_1(1 - d_3) + u_2(1 - d_1)\}p_{110} + \{u_1(1 - d_3) + u_3(1 - d_1)\}p_{101} \\ &\quad + \{u_2(1 - d_3) + u_3(1 - d_2)\}p_{011} - Q_5 - Q_6 - Q_7 - Q_8 > 0. \end{aligned}$$

We will show that the right hand side of the inequality is non-negative. Clearly each of the terms $\{u_1(1 - d_3) + u_2(1 - d_1)\}p_{110}$, $\{u_1(1 - d_3) + u_3(1 - d_1)\}p_{101}$ and $\{u_2(1 - d_3) + u_3(1 - d_2)\}p_{011}$ are non-negative. If $d_i < 1/4$ and $u_i < 1/4$ then

$$\begin{aligned} -Q_5 &= \{-(1 - d_1)d_2d_3 + (1 - d_1)d_2(1 - u_3) + (1 - d_1)(1 - u_2)d_3 + u_1d_2d_3\}p_{100} \\ &= [(1 - d_1)d_2\{(1 - u_3) - \frac{1}{3}d_3\} + (1 - d_1)d_3\{(1 - u_2) - \frac{1}{3}d_2\} + d_2d_3\{u_1 - \frac{1}{3}(1 - d_1)\}]p_{100} \\ &\geq \{(1 - d_1)d_2(\frac{3}{4} - \frac{1}{3}\frac{1}{4}) + (1 - d_1)d_3(\frac{3}{4} - \frac{1}{3}\frac{1}{4}) + d_2d_3(\frac{1}{4} - \frac{1}{3}\frac{3}{4})\}p_{100} \geq 0. \end{aligned}$$

By symmetry, $-Q_6 \geq 0$ and $-Q_7 \geq 0$. Furthermore,

$$\begin{aligned} -Q_8 &= \{-d_1d_2d_3 + d_1d_2(1 - u_3) + d_1(1 - u_2)d_3 + (1 - u_1)d_2d_3\}p_{000} \\ &= [d_1d_2\{(1 - u_3) - \frac{1}{3}d_3\} + d_1d_3\{(1 - u_2) - \frac{1}{3}d_2\} + d_2d_3\{(1 - u_1) - \frac{1}{3}d_1\}]p_{000} \\ &\geq \{d_1d_2(\frac{3}{4} - \frac{1}{3}\frac{1}{4}) + d_1d_3(\frac{3}{4} - \frac{1}{3}\frac{1}{4}) + d_2d_3(\frac{3}{4} - \frac{1}{3}\frac{1}{4})\}p_{000} \geq 0. \end{aligned}$$

We have thus shown that if $p_{111}^* - p_{110}^* - p_{101}^* - p_{011}^* > 0$ then $V(p_{111} - p_{110} - p_{101} - p_{011}) > 0$.

The result follows provided $V > 0$. If $d_i < 1/4$ and $u_i < 1/4$ then

$$\begin{aligned}
V &= (1-d_1)(1-d_2)(1-d_3) - (1-d_1)(1-d_2)u_3 - (1-d_1)u_2(1-d_3) - u_1(1-d_2)(1-d_3) \\
&= (1-d_1)(1-d_2)\left\{\frac{1}{3}(1-d_3) - u_3\right\} + (1-d_1)(1-d_3)\left\{\frac{1}{3}(1-d_2) - u_2\right\} \\
&\quad + (1-d_2)(1-d_3)\left\{\frac{1}{3}(1-d_1) - u_1\right\} \\
&> (1-d_1)(1-d_2)\left(\frac{1}{3}\frac{3}{4} - \frac{1}{4}\right) + (1-d_1)(1-d_3)\left(\frac{1}{3}\frac{3}{4} - \frac{1}{4}\right) + (1-d_2)(1-d_3)\left(\frac{1}{3}\frac{3}{4} - \frac{1}{4}\right) = 0.
\end{aligned}$$

Proof of Theorem 5.

We give the proof for $p_{111}^* - p_{110}^* - p_{101}^* - p_{011}^* + p_{100}^* + p_{010}^* > 0$; the proof for the other two conditions follows by symmetry. If we have $p_{111}^* - p_{110}^* - p_{101}^* - p_{011}^* + p_{100}^* + p_{010}^* + p_{001}^* - p_{000}^* > 0$ then $p_{111}^* - p_{110}^* - p_{101}^* - p_{011}^* + p_{100}^* + p_{010}^* + p_{001}^* - p_{000}^* > p_{001}^* - p_{000}^*$. It can be verified that

$$p_{111}^* - p_{110}^* - p_{101}^* - p_{011}^* + p_{100}^* + p_{010}^* + p_{001}^* - p_{000}^* = J(p_{111} - p_{110} - p_{101} - p_{011} + p_{100} + p_{010} + p_{001} - p_{000})$$

where

$$J = (1 - d_1 - u_1)(1 - d_2 - u_2)(1 - d_3 - u_3).$$

Thus we have $J(p_{111} - p_{110} - p_{101} - p_{011} + p_{100} + p_{010} + p_{001} - p_{000}) > p_{001}^* - p_{000}^*$ or

$$J(p_{111} - p_{110} - p_{101} - p_{011} + p_{100} + p_{010}) > p_{001}^* - p_{000}^* - J(p_{001} - p_{000}).$$

The result follows if $J > 0$ and $p_{001}^* - p_{000}^* - J(p_{001} - p_{000}) \geq 0$. Clearly if $d_i < 1/2$ and $u_i < 1/2$ then $J > 0$. Furthermore, we have that p_{001}^* is equal to:

$$\begin{aligned}
&u_1 u_2 (1 - d_3) p_{111} + u_1 u_2 d_3 p_{110} + u_1 (1 - u_2) (1 - d_3) p_{101} + (1 - u_1) u_2 (1 - d_3) p_{011} \\
&+ u_1 (1 - u_2) d_3 p_{100} + (1 - u_1) u_2 d_3 p_{010} + (1 - u_1) (1 - u_2) (1 - d_3) p_{001} + (1 - u_1) (1 - u_2) d_3 p_{000}
\end{aligned}$$

and p_{000}^* is equal to:

$$\begin{aligned} & u_1 u_2 u_3 p_{111} + u_1 u_2 (1 - u_3) p_{110} + u_1 (1 - u_2) u_3 p_{101} + (1 - u_1) u_2 u_3 p_{011} + u_1 (1 - u_2) (1 - u_3) p_{100} \\ & + (1 - u_1) u_2 (1 - u_3) p_{010} + (1 - u_1) (1 - u_2) u_3 p_{001} + (1 - u_1) (1 - u_2) (1 - u_3) p_{000}. \end{aligned}$$

Also it can be shown that

$$\begin{aligned} J p_{001} &= \{d_1 d_2 (1 - d_3) - d_1 d_2 u_3 - d_1 (1 - u_2) (1 - d_3) - (1 - u_1) d_2 (1 - d_3) \\ &+ d_1 (1 - u_2) u_3 + (1 - u_1) d_2 u_3 + (1 - u_1) (1 - u_2) (1 - d_3) - (1 - u_1) (1 - u_2) u_3\} p_{001} \end{aligned}$$

and

$$\begin{aligned} J p_{000} &= \{d_1 d_2 d_3 - d_1 d_2 (1 - u_3) - d_1 (1 - u_2) d_3 - (1 - u_1) d_2 d_3 + d_1 (1 - u_2) (1 - u_3) \\ &+ (1 - u_1) d_2 (1 - u_3) + (1 - u_1) (1 - u_2) d_3 - (1 - u_1) (1 - u_2) (1 - u_3)\} p_{000}. \end{aligned}$$

Let $H = (1 - u_1) (1 - u_2) (1 - d_3) p_{001} + (1 - u_1) (1 - u_2) d_3 p_{000} - (1 - u_1) (1 - u_2) u_3 p_{001} - (1 - u_1) (1 - u_2) (1 - u_3) p_{000}$. We then have that $p_{001}^* - p_{000}^* - J(p_{001} - p_{000}) = (p_{001}^* - p_{000}^* - H) + \{-J(p_{001} - p_{000}) + H\} = M + Q$ where

$$M = p_{001}^* - p_{000}^* - H$$

and

$$Q = -J(p_{001} - p_{000}) + H.$$

Thus

$$\begin{aligned}
Q &= -Jp_{001} + Jp_{000} + H \\
&= -(p_{001} - p_{000})\{d_1d_2(1 - d_3) - d_1d_2u_3 - d_1(1 - u_2)(1 - d_3) - (1 - u_1)d_2(1 - d_3) \\
&\quad + d_1(1 - u_2)u_3 + (1 - u_1)d_2u_3\} \\
&= -(p_{001} - p_{000})\{d_1d_2(1 - d_3 - u_3) - d_1(1 - u_2)(1 - d_3 - u_3) - (1 - u_1)d_2(1 - d_3 - u_3)\} \\
&= -(p_{001} - p_{000})(1 - d_3 - u_3)\{d_1d_2 - d_1(1 - u_2) - (1 - u_1)d_2\} \\
&= (p_{001} - p_{000})(1 - d_3 - u_3)\{d_1(1 - u_2 - \frac{1}{2}d_2) + d_2(1 - u_1 - \frac{1}{2}d_1)\} \geq 0
\end{aligned}$$

provided $d_i \leq 1/2$ and $u_i \leq 1/2$ since $(p_{001} - p_{000}) \geq 0$ by monotonicity. Moreover,

$$\begin{aligned}
M &= p_{001}^* - p_{000}^* - H \\
&= u_1u_2(1 - d_3 - u_3)p_{111} - u_1u_2(1 - d_3 - u_3)p_{110} + u_1(1 - u_2)(1 - d_3 - u_3)p_{101} \\
&\quad + (1 - u_1)u_2(1 - d_3 - u_3)p_{011} - u_1(1 - u_2)(1 - d_3 - u_3)p_{100} - (1 - u_1)u_2(1 - d_3 - u_3)p_{010} \\
&= u_1u_2(1 - d_3 - u_3)(p_{111} - p_{110}) + u_1(1 - u_2)(1 - d_3 - u_3)(p_{101} - p_{100}) \\
&\quad + (1 - u_1)u_2(1 - d_3 - u_3)(p_{011} - p_{010}) \\
&\geq 0
\end{aligned}$$

provided $d_i < 1/2$ and $u_i < 1/2$ since $(p_{111} - p_{110}) \geq 0$, $(p_{101} - p_{100}) \geq 0$ and $(p_{011} - p_{010}) \geq 0$ by monotonicity. This completes the proof.

Table 1. Power estimates for interaction tests under different misclassification probabilities and sample sizes for a moderate interaction ($p_{11} - p_{10} - p_{01} + p_{00} = 0.1$).

	$d_1 = u_1 = 0$	$d_1 = u_1 = 0.1$	$d_1 = u_1 = 0.25$	$d_1 = u_1 = 0.1$	$d_1 = u_1 = 0.25$	$d_1 = u_1 = 0.25$
	$d_2 = u_2 = 0$	$d_2 = u_2 = 0$	$d_2 = u_2 = 0$	$d_2 = u_2 = 0.1$	$d_2 = u_2 = 0.1$	$d_2 = u_2 = 0.25$
$n = 250$	15.9	11.8	6.9	9.4	5.9	4.4
$n = 500$	26.3	19.1	10.0	13.5	8.0	5.2
$n = 1000$	46.5	32.3	15.3	22.3	10.5	6.7
$n = 2500$	83.7	64.2	29.5	46.1	20.6	10.8
$n = 5000$	98.6	90.6	53.5	73.3	36.3	16.4

Table 2. Power estimates for interaction tests under different misclassification probabilities and sample sizes for a substantial interaction ($p_{11} - p_{10} - p_{01} + p_{00} = 0.2$).

	$d_1 = u_1 = 0$	$d_1 = u_1 = 0.1$	$d_1 = u_1 = 0.25$	$d_1 = u_1 = 0.1$	$d_1 = u_1 = 0.25$	$d_1 = u_1 = 0.25$
	$d_2 = u_2 = 0$	$d_2 = u_2 = 0$	$d_2 = u_2 = 0$	$d_2 = u_2 = 0.1$	$d_2 = u_2 = 0.1$	$d_2 = u_2 = 0.25$
$n = 250$	46.9	33.7	15.5	22.7	10.7	6.6
$n = 500$	76.5	55.1	25.3	37.3	17.7	9.4
$n = 1000$	96.0	84.3	44.3	64.3	30.8	14.0
$n = 2500$	100	99.7	81.6	95.6	60.8	29.4
$n = 5000$	100	100	98.1	99.9	88.8	49.6

Table 3. Cases and Controls from Yang et al. (2005), with odds ratios

	$X_2 = 0$	$X_2 = 1$
$X_1 = 0$	1.0 (20 cases / 154 cont)	3.52 (54 cases / 118 cont)
$X_1 = 1$	1.40 (18 cases / 99 cont)	7.20 (72 cases / 77 cont)