Inference for additive interaction under exposure misclassification

BY TYLER J. VANDERWEELE

*Departments of Biostatistics and Epidemiology, Harvard School of Public Health, Boston,
Massachusetts 02115, U.S.A.
tvanderw@hsph.harvard.edu*

SUMMARY

Results are given concerning inferences that can be drawn about interaction when binary exposures are subject to certain forms of independent nondifferential misclassification. Tests for interaction, using the misclassified exposures, are valid provided the probability of misclassification satisfies certain bounds. Results are given for additive statistical interactions, for causal interactions corresponding to synergism in the sufficient cause framework and for so-called compositional epistasis. 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. For additive statistical interaction, a method to correct estimates and confidence intervals for misclassification is described. The consequences for power of interaction tests under exposure misclassification are explored through simulations.

Some key words: Causal inference; Epistasis; Interaction; Misclassification; Sufficient cause; Synergism.

1. INTRODUCTION

Misclassification and measurement error are ubiquitous in studies of effects of exposures and are often 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 outcome or, by symmetry, a binary exposure, was subject to nondifferential misclassification, then tests of association between the outcome and the misclassified exposure are valid, in that they 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. [Garcia-Closas \(1998\)](#), [Zhang et al. \(2008\)](#), [Cheng & Lin \(2009\)](#) and [Lindström et al. \(2009\)](#) consider only multiplicative interaction, whereas additive interaction is generally considered more important for public health purposes and for assessing biologic synergism ([Greenland et al., 2008](#)).

Here we provide results on the implications of nondifferential misclassification for standard interaction tests and estimates on the linear additive scale, generalizing Bross's result to interaction parameters. We extend these results further to causal interaction corresponding to synergism in Rothman's sufficient cause framework ([Rothman, 1976](#); [VanderWeele & Robins, 2008](#)) or compositional epistasis ([Cordell, 2009](#); [VanderWeele, 2010a, b](#)) in genetics.

2. REVIEW OF ADDITIVE AND CAUSAL INTERACTIONS

Let X_1 and X_2 denote two binary exposures and let D denote a binary outcome. Let $p_{x_1x_2} = \text{pr}(D = 1 \mid x_1, x_2)$. The standard condition for positive interaction on the linear 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, that is, $(p_{11} - p_{00}) > (p_{10} - p_{00}) + (p_{01} - p_{00})$.

Let $D_{x_1x_2}(\omega)$ denote the counterfactual outcome (Rubin, 1974) for D for individual ω if X_1 and X_2 had been set to x_1 and x_2 , with $D_{x_1x_2}$ treated as a random variable over the population. VanderWeele & Robins (2008) said a sufficient cause interaction between X_1 and X_2 was present if there was some ω such that $D_{11}(\omega) = 1$ but $D_{10}(\omega) = D_{01}(\omega) = 0$ and showed that if there were such an individual, then synergism between X_1 and X_2 within Rothman's (1976) sufficient cause framework must be present. They 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. This is a stronger notion of interaction insofar as (2) is more stringent than (1).

We say that the effects of X_1 and X_2 are positive monotonic if $D_{x_1x_2}(\omega)$ is nondecreasing in x_1 and x_2 for all ω . Negative monotonic effects are defined analogously. VanderWeele & Robins (2008) noted that if the effects of X_1 and X_2 were positive monotonic, then (1) would suffice to draw the conclusion of a sufficient cause interaction (cf. Greenland et al., 2008); testing (2) would be necessary without such monotonicity (VanderWeele & Robins, 2007, 2008). Tests for synergism for the absence of the outcome or for sufficient causes involving the absence of exposures can be obtained by recoding (Greenland & Poole, 1988; VanderWeele & Knol, 2011); in such cases negative monotonic effects allow for less stringent tests. We will focus on positive monotonic effects, without recoding exposure or outcome. We will use the terms positive monotonic and monotonic interchangeably.

VanderWeele (2010a, b) discussed empirical tests for an even stronger notion of interaction. In the genetics literature, when gene-gene interactions are considered, compositional epistasis (Cordell, 2009) is said to be present if for some ω , $D_{11}(\omega) = 1$ and $D_{10}(\omega) = D_{01}(\omega) = D_{00}(\omega) = 0$, so that the effect of one genetic factor is masked unless the other is present. VanderWeele (2010a, b) 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, that is, the presence of compositional epistasis. In contrast with (1), we subtract rather than add p_{00} in (3). If $D_{x_1x_2}(\omega)$ is nondecreasing in one of x_1 or x_2 for all ω , then (2) suffices for compositional epistasis; if $D_{x_1x_2}(\omega)$ is nondecreasing in both x_1 and x_2 for all ω , then (1) implies compositional epistasis.

In the case of three exposures of interest, let $D_{x_1x_2x_3}(\omega)$ denote the counterfactual outcome D for individual ω if X_1 , X_2 and X_3 had been set to x_1 , x_2 and x_3 , respectively. We say that there is a three-way sufficient cause interaction between X_1 , X_2 and X_3 if for some ω , $D_{111}(\omega) = 1$ and $D_{110}(\omega) = D_{101}(\omega) = D_{011}(\omega) = 0$, which implies, in the sufficient cause framework, synergism between X_1 , X_2 and X_3 (VanderWeele & Robins, 2008). If we let $p_{x_1x_2x_3} = \text{pr}(D = 1 \mid x_1, x_2, x_3)$, then if the effects of X_1 , X_2 and X_3 on D are unconfounded, then

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

implies a three-way sufficient cause interaction (VanderWeele & Robins, 2008). If $D_{x_1x_2x_3}$ is nondecreasing in x_1 , x_2 and x_3 , then any of the following three conditions (VanderWeele & Robins, 2008) implies a three-way sufficient cause interaction:

$$\begin{aligned} p_{111} - p_{110} - p_{101} - p_{011} + p_{100} + p_{010} &> 0, \\ p_{111} - p_{110} - p_{101} - p_{011} + p_{100} + p_{001} &> 0, \\ p_{111} - p_{110} - p_{101} - p_{011} + p_{010} + p_{001} &> 0. \end{aligned} \quad (5)$$

For all of the above inequalities the contrasts on the left-hand side also constitute a lower bound for the prevalence of individuals exhibiting the corresponding causal interaction. 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 be made conditional on C . The issue of modelling and robustness to control for some set of confounding variables C is considered elsewhere (Vansteelandt et al., 2008, 2012; VanderWeele, 2009). 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. ADDITIVE AND CAUSAL INTERACTIONS UNDER EXPOSURE MISCLASSIFICATION

Let X_i denote the true exposure and let X_i^* denote the observed exposure. Let $d_i = \text{pr}(X_i = 0 \mid X_i^* = 1)$ and let $u_i = \text{pr}(X_i = 1 \mid X_i^* = 0)$. These misclassification probabilities are 1 minus the positive and negative predictive values, respectively, of the measured exposures for the true exposures. We say that the misclassification is nondifferential if (i) $\text{pr}(D = 1 \mid x_1, x_2, x_3, x_1^*, x_2^*, x_3^*) = \text{pr}(D = 1 \mid x_1, x_2, x_3)$. Nondifferential 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 predicted probabilities are independent if (ii) $\text{pr}(x_1, x_2, x_3 \mid x_1^*, x_2^*, x_3^*) = \text{pr}(x_1 \mid x_1^*)\text{pr}(x_2 \mid x_2^*)\text{pr}(x_3 \mid x_3^*)$; a sufficient condition for this is that $\text{pr}(x_1^*, x_2^*, x_3^* \mid x_1, x_2, x_3) = \text{pr}(x_1^* \mid x_1)\text{pr}(x_2^* \mid x_2)\text{pr}(x_3^* \mid x_3)$ and X_1, X_2, X_3 are mutually independent. All results will make assumptions (i) and (ii). Let $p_{x_1 x_2}^* = \text{pr}(D = 1 \mid X_1^* = x_1, X_2^* = x_2)$ and $p_{x_1 x_2 x_3}^* = \text{pr}(D = 1 \mid X_1^* = x_1, X_2^* = x_2, X_3^* = x_3)$. We then have the following results. The proofs are given in the Supplementary Material.

THEOREM 1. *Under assumptions (i) and (ii), if $d_i + u_i < 1$, then*

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

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

If condition (6) is satisfied for the measured exposures, then the true exposures also exhibit additive statistical interaction. If the effects of the true exposures on the outcome are also unconfounded and monotonic, then the true exposures exhibit a sufficient cause interaction and an epistatic interaction (VanderWeele & Robins, 2008; VanderWeele, 2010a, b). The condition $d_i + u_i < 1$ will hold if for each i the misclassification probabilities are such that $d_i < 1/2$ and $u_i < 1/2$. Theorem 1 is a generalization for interaction of the classic result of Bross (1954) for main effects. In fact, as shown in the Supplementary Material,

$$(p_{11} - p_{10} - p_{01} + p_{00}) = (p_{11}^* - p_{10}^* - p_{01}^* + p_{00}^*) / \{(1 - d_1 - u_1)(1 - d_2 - u_2)\}.$$

For known misclassification probabilities (d_1, u_1, d_2, u_2), one can thus obtain estimates for additive interaction, corrected for exposure misclassification, by dividing additive interaction estimates using the observed data, $(p_{11}^* - p_{10}^* - p_{01}^* + p_{00}^*)$, by the factor $(1 - d_1 - u_1)(1 - d_2 - u_2)$. Because of the deterministic relationship, confidence intervals for the true interaction could also be obtained by dividing both limits of the confidence interval for $(p_{11}^* - p_{10}^* - p_{01}^* + p_{00}^*)$ by the factor $(1 - d_1 - u_1)(1 - d_2 - u_2)$. For unknown misclassification probabilities, the parameters (d_1, u_1, d_2, u_2) could be varied in a sensitivity analysis.

Results analogous to Theorem 1 hold for sufficient cause and epistatic interactions and for three-way sufficient cause interactions with and without monotonicity.

THEOREM 2. *Under assumptions (i) and (ii), if $d_i < 1/2$ and $u_i < 1/4$, then*

$$p_{11}^* - p_{10}^* - p_{01}^* > 0 \quad (7)$$

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

THEOREM 3. Under assumptions (i) and (ii), 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 that $p_{11} - p_{10} - p_{01} - p_{00} > 0$.

THEOREM 4. Under assumptions (i) and (ii), if $d_i < 1/4$ and $u_i < 1/4$, then

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

implies that $p_{111} - p_{110} - p_{101} - p_{011} > 0$.

THEOREM 5. Under assumptions (i) and (ii), If $p_{x_1x_2x_3}$ is nondecreasing in x_1 , x_2 and x_3 and $d_i < 1/2$ and $u_i < 1/2$, then any of

$$\begin{aligned} p_{111}^* - p_{110}^* - p_{101}^* - p_{011}^* + p_{100}^* + p_{010}^* &> 0, \\ p_{111}^* - p_{110}^* - p_{101}^* - p_{011}^* + p_{100}^* + p_{001}^* &> 0, \\ p_{111}^* - p_{110}^* - p_{101}^* - p_{011}^* + p_{010}^* + p_{001}^* &> 0 \end{aligned} \quad (10)$$

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

Theorems 1–5 apply also if the effects of 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 $u_{ic} = \text{pr}(X_i = 1 \mid X_i^* = 0, c)$ and $d_{ic} = \text{pr}(X_i = 0 \mid X_i^* = 1, c)$ and require that $\text{pr}(x_1, x_2, x_3 \mid x_1^*, x_2^*, x_3^*, c) = \text{pr}(x_1 \mid x_1^*, c)\text{pr}(x_2 \mid x_2^*, c)\text{pr}(x_3 \mid x_3^*, c)$.

The misclassification results above require an assumption of independence. This might be plausible for a genetic and an environmental factor or for two genetic factors on different chromosomes. Genetic exposures generally have a low probability of misclassification; but a genetic marker may serve as a proxy for a true causal variant and could be conceived of as a misclassified version of the causal variant with which it is associated due to linkage disequilibrium. The independence assumption may be less plausible in settings with two environmental or behavioural exposures, especially if assessed by retrospective self-report. Whether the assumption is reasonable cannot be known definitively unless data are available on a gold standard measurement for the exposures.

4. CASE-CONTROL DATA AND LOGISTIC REGRESSION

For data with a dichotomous outcome, logistic or log-linear models are often fitted; logistic models are also used to accommodate a case-control design. The results above have implications for such models. For two dichotomous exposures, we have (VanderWeele, 2010a, b) that under a log-linear model for the observed data, $\log(p_{x_1x_2}^*) = \log\{\text{pr}(D = 1 \mid X_1^* = x_1, X_2^* = x_2, C = c)\} = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_1x_2 + \beta_4c$, provided β_1 and β_2 are nonnegative, $\beta_3 > 0$ implies (6), $\beta_3 > \log(2)$ implies (7) and $\beta_3 > \log(3)$ implies (8). 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 nonnegative, $\beta_3 > 0$ 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 $\log(p_{x_1x_2}^*) \approx \text{logit}(p_{x_1x_2}^*)$.

Alternatively, to assess additive interaction with case-control data, the relative excess risk due to interaction is often used. Define $\text{RERI}^* = \text{RR}_{11}^* - \text{RR}_{10}^* - \text{RR}_{01}^* + 1$ where $\text{RR}_{x_1x_2}^* = p_{x_1x_2}^*/p_{00}^*$. Provided the outcome is rare, we can approximate these risk ratios, $\text{RR}_{x_1x_2}^*$, by the odds ratios, $\text{OR}_{x_1x_2}^* = \{p_{x_1x_2}^*/(1 - p_{x_1x_2}^*)\}/\{p_{00}^*/(1 - p_{00}^*)\}$, estimated from the observed case-control data. It is straightforward to obtain confidence intervals for the relative excess risk due to interaction (Hosmer & Lemeshow, 1992; Richardson & Kaufman, 2009). Dividing (6)–(8) by p_{00}^* , we obtain $\text{RERI}^* > 0$, $\text{RERI}^* > 1$ and $\text{RERI}^* > 2$, respectively. If one of these is satisfied, then the corresponding condition (6), (7) or (8) will hold. From this it would follow under the assumptions of Theorems 1–3, that inequalities (1), (2) or (3), respectively,

would hold. We can draw conclusions about various forms of statistical and causal interaction from the misclassified data using RERI*. Similar remarks hold for risk ratios conditional on covariates.

5. SIMULATIONS

We use simulation to examine the effects of misclassification of the exposure variables on power for tests to detect interaction and on coverage probabilities for uncorrected and corrected confidence intervals for additive interaction. For each scenario we generated 5000 datasets with 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. We consider six 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 $\text{pr}(X_1 = i, X_2 = j) = 0.25$ 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 $p_{00} = 0.1, p_{10} = 0.3, p_{01} = 0.2, p_{11} = 0.6$. For the underlying distribution of the outcome under the true exposures, we have $p_{11} - p_{10} - p_{01} + p_{00} > 0$. We assume that data are only available on the misclassified exposures, X_1^* and X_2^* . For each simulated data, we use the misclassified exposures to estimate $p_{11}^* - p_{10}^* - p_{01}^* + p_{00}^*$ and test whether this is positive using the test statistic

$$Z = \frac{\hat{p}_{11}^* - \hat{p}_{10}^* - \hat{p}_{01}^* + \hat{p}_{00}^*}{\{\hat{p}_{11}^*(1 - \hat{p}_{11}^*)/n_{11} + \hat{p}_{10}^*(1 - \hat{p}_{10}^*)/n_{10} + \hat{p}_{01}^*(1 - \hat{p}_{01}^*)/n_{01} + \hat{p}_{00}^*(1 - \hat{p}_{00}^*)/n_{00}\}^{1/2}} \sim N(0, 1),$$

where $\hat{p}_{ij}^*$ is the sample estimate of $\text{pr}(D = 1 | X_1^* = i, X_2^* = j)$ and n_{ij} is the number in the sample with $X_1^* = i, X_2^* = j$. We report coverage probabilities for the confidence intervals for additive interaction estimates using the observed data and also for the corrected confidence interval using the method described above, assuming the misclassification probabilities are known. The first part of Table 1 reports estimates for the moderate interaction scenario, $p_{11} - p_{10} - p_{01} + p_{00} = 0.1$. The first column presents estimates without misclassification, $d_1 = u_1 = 0, d_2 = u_2 = 0$. The remaining five columns present power estimates and coverage probabilities for uncorrected and corrected confidence intervals 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 $n = 5000$. Coverage probabilities for corrected estimates are in all cases close to 95%.

The second part of Table 1 reports the power estimates and coverage probabilities for uncorrected and corrected confidence intervals for the substantial interaction scenario, $p_{11} - p_{10} - p_{01} + p_{00} = 0.2$. With a substantial interaction, power is fairly good when $n = 1000$ 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 when $n = 2500$ and above. When both exposures are subject to substantial misclassification, $d_1 = u_1 = d_2 = u_2 = 0.25$, power is only about 50% even when $n = 5000$. Additional simulations suggest that power declines for rarer exposures; case-control study designs generally result in greater power.

6. ILLUSTRATIONS

We illustrate the results with two examples. [Canonicao et al. \(2008\)](#) tested for additive interaction between presence of a non-O blood type and the use of oral oestrogen on the risk of venous thromboembolism among postmenopausal women. They used data from a case-control study with 271 cases and 610 controls, adjusting for age, centre and admission date, and obtained a measure of $\text{RERI} = 5.4, p < 0.05$.

Table 1. Power estimates for interaction tests and coverage probabilities for uncorrected and corrected 95% confidence intervals for different misclassification probabilities and sample sizes

	$d_1 = u_1 = 0$ $d_2 = u_2 = 0$	$d_1 = u_1 = 0.1$ $d_2 = u_2 = 0$	$d_1 = u_1 = 0.25$ $d_2 = u_2 = 0$	$d_1 = u_1 = 0.1$ $d_2 = u_2 = 0.1$	$d_1 = u_1 = 0.25$ $d_2 = u_2 = 0.1$	$d_1 = u_1 = 0.25$ $d_2 = u_2 = 0.25$
Moderate interaction ($p_{11} - p_{10} - p_{01} + p_{00} = 0.1$)						
$n = 250$	16, 95, 95	12, 94, 95	7, 93, 95	9, 94, 95	6, 91, 95	4, 90, 95
$n = 500$	28, 95, 95	18, 95, 95	13, 92, 94	13, 92, 94	8, 88, 95	5, 84, 95
$n = 1000$	47, 95, 95	33, 94, 95	22, 89, 94	22, 89, 94	11, 82, 96	7, 74, 95
$n = 2500$	85, 95, 95	64, 90, 95	46, 70, 96	46, 82, 95	21, 57, 96	11, 43, 95
$n = 5000$	99, 96, 96	91, 87, 95	52, 46, 95	74, 68, 95	37, 33, 95	17, 15, 95
Substantial interaction ($p_{11} - p_{10} - p_{01} + p_{00} = 0.2$)						
$n = 250$	48, 94, 94	33, 93, 94	15, 85, 95	22, 90, 95	11, 82, 95	6, 73, 95
$n = 500$	76, 95, 95	54, 92, 95	25, 76, 95	39, 86, 95	17, 68, 95	8, 55, 95
$n = 1000$	97, 95, 95	84, 89, 95	43, 57, 95	65, 75, 95	30, 43, 95	14, 24, 95
$n = 2500$	100, 95, 95	100, 79, 95	80, 18, 95	96, 46, 95	61, 7, 95	29, 2, 94
$n = 5000$	100, 95, 95	100, 63, 95	98, 2, 95	100, 16, 95	89, 0, 95	52, 0, 95

Numbers in each cell are the power estimate, coverage probability for uncorrected 95% confidence interval and coverage probability for corrected 95% confidence interval, all as per cent.

Data on blood type was self-reported and probably subject to misclassification; oestrogen use is less likely to be misclassified. From the results above, if the exposures and misclassification are independent, and provided that the positive and negative predicted values for non-O blood group are at least 0.5 we could conclude, even from the analysis subject to misclassification, that there is true additive interaction.

VanderWeele et al. (2011) examined additive interaction between smoking and high well-water-arsenic exposure in producing pre-malignant skin lesions using a cohort of 11 746 persons in Bangladesh. They dichotomized arsenic at 100 $\mu\text{g/l}$ and considered conclusions that could be drawn about the underlying continuous measure when dichotomizing an exposure. A measure of additive interaction for risk, using dichotomized arsenic and adjusting for and marginalizing over covariates, was estimated to be 0.035 with 95% confidence interval (0.0003, 0.070). Well water arsenic is subject to measurement error, so the dichotomized mismeasured value may constitute a misclassified version of the true dichotomized value. Self-reported smoking is also potentially subject to misclassification. Measured exposures have correlation close to 0 and misclassification of the two exposures is likely independent. Using the correction method above, if the positive and negative predictive values for smoking and dichotomized arsenic, conditional on covariates, were all 0.95, then the corrected additive interaction estimate would be 0.043 with confidence interval (0.0004, 0.086). If the positive predictive values for smoking and dichotomized arsenic were 0.85 and the negative predictive values were 0.93, the corrected estimate would be 0.058 with confidence interval (0.0005, 0.115). Evidence for additive interaction remains, with somewhat higher estimates after adjusting for misclassification. See VanderWeele et al. (2011) for the implications of such additive interaction estimates concerning dichotomized true exposures for conclusions about underlying continuous measures.

ACKNOWLEDGEMENT

The author thanks the editor, associate editor and three referees for helpful comments. This research was funded by National Institutes of Health, U.S.A.

SUPPLEMENTARY MATERIAL

Supplementary material available at *Biometrika* online includes proofs of all theorems.

REFERENCES

- BROSS, I. (1954). Misclassification in 2×2 tables. *Biometrics* **10**, 478–86.
- CANONICAO, M., OLIÉ, V., CARCAILLON, L., TUBERT-BITTER, P. & SCARABIN, P.-Y. (2008). Synergism between non-O blood group and oral estrogen in the risk of venous thromboembolism among postmenopausal women. The ESTHER Study. *Thromb. Haemost.* **99**, 246–8.
- CHENG, K. F. & LIN, W. J. (2009). The effects of misclassification in studies of gene-environment interactions. *Hum. Hered.* **67**, 77–87.
- CORDELL, H. J. (2009). Detecting gene-gene interactions that underlie human diseases. *Nature 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**, 746–8.
- GARCIA-CLOSAS, M., THOMPSON, W. D. & ROBINS, J. M. (1998). Differential misclassifications and the assessment of gene-environment interactions. *Am. J. Epidemiol.* **147**, 426–33.
- GREENLAND, S. & POOLE, C. (1988). Invariants and noninvariants in the concept of interdependent effects. *Scand. J. Work Envir. Health* **14**, 125–9.
- GREENLAND, S., LASH, T. L. & ROTHMAN, K. J. (2008). Concepts of interaction. In *Modern Epidemiology*, Ed. K. J. Rothman, S. Greenland and T. L. Lash, 3rd ed., pp. 71–83. Philadelphia: Lippincott Williams & Wilkins.
- HOSMER, D. W. & LEMESHOW, S. (1992). Confidence interval estimation of interaction. *Epidemiology* **3**, 452–6.
- 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. *Hum. Hered.* **68**, 171–81.
- RICHARDSON, D. B. & 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.
- RUBIN, D. B. (1974). Estimating causal effects of treatments in randomized and non-randomized studies. *J. Educ. Psychol.* **66**, 688–701.
- VANDERWEELE, T. J. (2009). Sufficient cause interactions and statistical interactions. *Epidemiology* **20**, 6–13.
- VANDERWEELE, T. J. (2010a). Empirical tests for compositional epistasis. *Nature Rev. Genet.* **11**, 166.
- VANDERWEELE, T. J. (2010b). Epistatic interactions. *Statist. Appl. Genet. Molec. Biol.* **9**, 1–22.
- VANDERWEELE, T. J. & KNOL, M. J. (2011). Remarks on antagonism. *Am. J. Epidemiol.* **173**, 1140–7.
- VANDERWEELE, T. J. & ROBINS, J. M. (2007). The identification of synergism in the sufficient-component cause framework. *Epidemiology* **18**, 329–39.
- VANDERWEELE, T. J. & ROBINS, J. M. (2008). Empirical and counterfactual conditions for sufficient cause interactions. *Biometrika* **95**, 49–61.
- VANDERWEELE, T. J., CHEN, Y. & AHSAN, H. (2011). Inference for causal interactions for continuous exposures under dichotomization. *Biometrics* **67**, 1414–21.
- VANSTEELANDT, S., VANDERWEELE, T. J., TCHETGEN, E. J. & ROBINS, J. M. (2008). Multiply robust inference for statistical interactions. *J. Am. Statist. Assoc.* **103**, 1693–704.
- VANSTEELANDT, S., VANDERWEELE, T. J. & ROBINS, J. M. (2012). Semiparametric tests for sufficient cause interactions. *J. R. Statist. Soc. B* **74**, 223–44.
- 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. *Statist. Med.* **27**, 2756–83.

[Received November 2010. Revised January 2012]