*Biometrika* (2013), **100**, 1, *pp.* 241–248
doi: 10.1093/biomet/ass054

Advance Access publication 6 December 2012
*Printed in Great Britain*

# Bias attenuation results for nondifferentially mismeasured ordinal and coarsened confounders

BY ELIZABETH L. OGBURN AND TYLER J. VANDERWEELE

*Departments of Biostatistics and Epidemiology, Harvard School of Public Health, Boston, Massachusetts 02115, U.S.A.*

ogburn@post.harvard.edu    tvanderw@hsph.harvard.edu

SUMMARY

Suppose we are interested in the effect of a binary treatment on an outcome where that relationship is confounded by an ordinal confounder. We assume that the true confounder is not observed but, rather, we observe a nondifferentially mismeasured version of it. We show that, under certain monotonicity assumptions about its effect on the treatment and on the outcome, an effect measure controlling for the mismeasured confounder will fall between the corresponding crude and true effect measures. We also present results for coarsened and, under further assumptions, multiple misclassified confounders.

*Some key words*: Bias; Confounder; Measurement error; Misclassification.

## 1. INTRODUCTION

Accurately measuring and adjusting for confounders is a crucial step in every observational study. Confounder measurement error is a pervasive problem that plagues such studies and has been investigated extensively (Ahlbom & Steineck, 1992; Carroll et al., 2006; Greenland, 1996, 2004, 2005; Gustafson, 2004; Kelsey, 1996; Lash & Fink, 2003; Rothman et al., 2008; Stefanski & Carroll, 1985; Willett, 1989). In this paper we focus on the nondifferential misclassification of polytomous confounders in studies of the effect of a binary treatment. Greenland (1980) argued that adjustment by a binary nondifferentially misclassified confounder gives a partially adjusted measure of effect, reducing but not eliminating the bias due to the confounder and therefore producing a partially adjusted effect measure that falls between the crude, unadjusted measure and the true measure, which is adjusted by the true confounder. We hereafter refer to this as the partial control result. Ogburn & VanderWeele (2012) gave a counterexample to the partial control result for binary confounders and proved that the result holds under certain conditions. Brenner (1993) proved by counterexample that adjustment by a nondifferentially misclassified polytomous confounder may induce greater bias than failing to adjust for the confounder at all, or it may change the direction of the bias.

To our knowledge, no analytic results exist in the literature that characterize the bias due to adjustment by a polytomous confounder subject to nondifferential misclassification. We present assumptions under which the partial control result holds for polytomous confounders for both the average treatment effect and the treatment effect among the treated. Our remarks and results apply equally to effects on the mean difference, mean ratio and odds ratio scales.

## 2. BACKGROUND AND NOTATION

Let $A$ be an indicator of treatment or exposure and $Y$ an outcome, and suppose that the effect of $A$ on $Y$ is confounded by an ordinal confounder $C$ with levels $1, \ldots, K$. We assume that $C$ is unobserved and that $C'$, an imperfect measure of $C$, is observed instead, where $C'$ is determined by the misclassification probabilities $p_{ij} = \mathrm{pr}(C' = i \mid C = j)$, $i, j \in \{1, \ldots, K\}$. Then $\mathrm{pr}(C' = i) = \sum_{j=1}^{K} p_{ij}\mathrm{pr}(C = j)$.

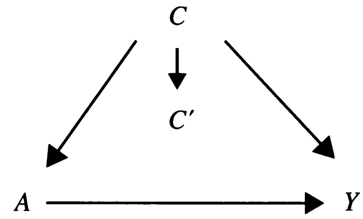


Fig. 1. Relationships between treatment $A$, outcome $Y$, confounder $C$, and an imperfect measure $C'$ of $C$.

We assume throughout that the misclassification of $C$ is nondifferential with respect to $A$ and $Y$; that is, $\mathrm{pr}(C' = c' \mid C = c, Y = y, A = a) = \mathrm{pr}(C' = c' \mid C = c, Y = y', A = a')$ for all $y, y'$ in the support of $Y$ and all $a, a'$ in the support of $A$. If $p_{ij} \leqslant p_{ik}$ and $p_{ji} \leqslant p_{ki}$ for $j < k < i$, and $p_{il} \geqslant p_{im}$ and $p_{li} \geqslant p_{mi}$ for $i < l < m$, then we say that the misclassification probabilities are tapered. Roughly, this means that the probability of correct classification is at least as great as the probability of misclassification into any one level and that, for a fixed level $i$ of either $C$ or $C'$, the misclassification probabilities are nonincreasing in each direction away from $i$. Requiring tapered misclassification probabilities is a weaker condition than requiring $p_{ij}$ to be decreasing with the distance between $i$ and $j$. We begin by assuming that $C$ is the only confounder of the effect of $A$ on $Y$, so that adjustment by $C$ allows for identification of the true effect. Figure 1 depicts the relationships between the variables. Generalizations that allow for multiple confounders and coarsened confounders will be discussed further in § 3·2.

Let $Y_1$ be the counterfactual outcome under treatment $A = 1$, that is, the outcome we would have observed if, possibly contrary to fact, a subject had received treatment $A = 1$. Let $Y_0$ be the counterfactual outcome under treatment $A = 0$. We make the consistency assumption that $Y_a = Y$ whenever $A = a$. The true effect measures are formulated in terms of the distributions of these counterfactuals; for example, the true risk difference for binary $Y$ is $\mathrm{RD}_{\mathrm{true}} \equiv E(Y_1) - E(Y_0)$. For no subject do we observe both $Y_1$ and $Y_0$; we observe only the counterfactual corresponding to the subject's actual treatment. However, within levels of $C$, subjects are effectively randomized with respect to treatment by the assumption that $C$ is the sole confounder, so $E(Y_1 \mid A = 0, C = c) = E(Y_1 \mid A = 1, C = c) = E(Y \mid A = 1, C = c)$. Therefore, both $E(Y_1)$ and $E(Y_0)$ can be calculated by standardizing by $C$:

$$E(Y_a) = \sum_{k=1}^{K} E(Y \mid A = a, C = k)\, \mathrm{pr}(C = k). \tag{1}$$

The corresponding expression standardized by $C'$ instead of $C$ is

$$E_{C'}(Y \mid A = a) \equiv \sum_{k=1}^{K} E(Y \mid A = a, C' = k)\, \mathrm{pr}(C' = k). \tag{2}$$

We will call effect measures that adjust for $C'$ observed adjusted measures (Brenner, 1993). Effect measures that adjust for $C$ are said to be true, and those that collapse over $C$ are said to be crude. The observed adjusted measures are formulated in terms of the means standardized by $C'$; for example, the observed adjusted risk difference is $\mathrm{RD}_{\mathrm{obs}} \equiv E_{C'}(Y \mid A = 1) - E_{C'}(Y \mid A = 0)$. The crude measures are formulated in terms of the unstandardized conditional means; for example, $\mathrm{RD}_{\mathrm{crude}} \equiv E(Y \mid A = 1) - E(Y \mid A = 0)$.

We present analytic results comparing observed adjusted effect measures to the corresponding true and crude measures. Our results hold for any effect measure that can be written as $h[g\{E(Y_1)\} - g\{E(Y_0)\}]$ where $g$ and $h$ are monotonic functions. For example, for the effect on the mean difference scale, $g$ and $h$ are both the identity function; for the effect on the mean ratio scale, $g$ is the logistic function and $h$ the exponential function; and for the odds ratio, $g$ is the logit function and $h$ the exponential function.

Our results hold also for measures of the treatment effect among the treated and, by symmetry, among the untreated, which are defined analogously to the general treatment effect measures above but restricted to the subpopulation of subjects with $A = 1$. We will denote such measures by a superscript T. For example, the true treatment effect among the treated on the risk difference scale for binary $Y$ is $\text{RD}^{\text{T}}_{\text{true}} \equiv E(Y_1 \mid A = 1) - E(Y_0 \mid A = 1)$, where, if $C$ is the only confounder, $E(Y_a \mid A = 1) = \sum_c E(Y \mid A = a, C = c)\text{pr}(C = c \mid A = 1)$ is the mean of $Y$ given $A = a$ standardized by the distribution of $C$ among the treated. We define the analogous observed adjusted expressions as $E_{C'|A=1}(Y \mid A = a) \equiv \sum_c E(Y \mid A = a, C' = c)\text{pr}(C' = c \mid A = 1)$, which is the mean of $Y$ given $A = a$ standardized by the distribution of $C'$ among the treated. Then the observed adjusted treatment effect among the treated on the mean difference scale is $\text{RD}^{\text{T}}_{\text{obs}} \equiv E_{C'|A=1}(Y \mid A = 1) - E_{C'|A=1}(Y \mid A = 0)$. The crude correlates of the mean outcomes among the treated are simply $E(Y \mid A = 1)$ and $E(Y \mid A = 0)$, because the premise of the crude calculations is that the effect of $A$ on $Y$ is unconfounded, and therefore the mean counterfactual outcome for the treated had they not received treatment simply equals the observed mean outcome among the untreated.

## 3. Results

### 3·1. *Main result*

If $E(Y \mid A = a, C = i) \leqslant E(Y \mid A = a, C = j)$ for all $i < j$ and for both $a = 0$ and $a = 1$, then we say that $E(Y \mid A, C)$ is nondecreasing in $C$. If $E(Y \mid A = a, C = i) \geqslant E(Y \mid A = a, C = j)$ for all $i < j$ and for both $a = 0$ and $a = 1$, then we say that $E(Y \mid A, C)$ is nonincreasing in $C$. If $E(Y \mid A, C)$ is either nonincreasing or nondecreasing in $C$, then it is monotonic in $C$. Similarly, if $E(A \mid C)$ is either nonincreasing or nondecreasing in $C$, then it is monotonic in $C$. Monotonicity of $E(Y \mid A, C)$ in $C$ requires the outcome to exhibit a dose response to the confounder or, in the degenerate case of no confounding by $C$, to be unaffected by it. It also requires the confounder to affect the outcome in the same direction among the treated and untreated. If $C$ has a protective effect among one treatment group and a harmful effect among the other, then monotonicity of $E(Y \mid A, C)$ will be violated.

Our first result says that under monotonicity, the partial control result holds for an ordinal confounder of the effect of $A$ on $Y$. Proofs of all the theorems are given in the Appendix.

Theorem 1. *Let A be binary, and let C be ordinal and nondifferentially misclassified with respect to A and Y, with tapered misclassification probabilities. Suppose that $E(Y \mid A, C)$ and $E(A \mid C)$ are monotonic in C. Then the observed adjusted average treatment effect and the observed adjusted effect of treatment on the treated will lie between the corresponding true and crude effects for any effect measure that can be written as $h[g\{E(Y_1)\} - g\{E(Y_0)\}]$ where g and h are monotonic functions.*

If $C$ is binary, then monotonicity of $E(A \mid C)$ holds by default and need not be assumed; furthermore, the assumption of tapered misclassification probabilities is not required for the partial control result to hold. When $C$ is binary, the partial control result for the average treatment effect requires the assumption that $E(Y \mid A, C)$ is monotonic in $C$, but the result for the effect of treatment on the treated requires no monotonicity assumptions (Ogburn & VanderWeele, 2012).

For $C$ with at least three levels, counterexamples are easily constructed to show that for both the average treatment effect and the treatment effect on the treated, if either $E(Y \mid A, C)$ or $E(A \mid C)$ is not monotonic in $C$, then the observed adjusted effect need not lie between the crude and the true effects. Table 1 gives a counterexample for the average treatment effect and the treatment effect on the treated when $E(Y \mid A, C)$ is monotonic in $C$ but $E(A \mid C)$ is not. The full data are represented by the true $2 \times 2 \times 3$ table in the middle, the crude data are collapsed over $C$, and the observed $2 \times 2 \times 3$ table was generated from the true one by the tapered misclassification probabilities $\text{pr}(C' = 1 \mid C = 1) = 1{\cdot}0$, $\text{pr}(C' = 2 \mid C = 1) = \text{pr}(C' = 3 \mid C = 1) = 0$, $\text{pr}(C' = 1 \mid C = 2) = 0$, $\text{pr}(C' = 2 \mid C = 2) = 0{\cdot}6$, $\text{pr}(C' = 3 \mid C = 2) = 0{\cdot}4$, $\text{pr}(C' = 1 \mid C = 3) = 0$, $\text{pr}(C' = 2 \mid C = 3) = 0{\cdot}4$ and $\text{pr}(C' = 3 \mid C = 3) = 0{\cdot}6$. Monotonicity holds for $E(Y \mid A, C)$ because $E(Y \mid A = 1, C = c)$ and $E(Y \mid A = 0, C = c)$ are both nondecreasing in $c$: $E(Y \mid A = 1, C = 1) = 0{\cdot}83$ and $E(Y \mid A = 1, C = 2) = E(Y \mid A = 1, C = 3) = 0{\cdot}85$, while $E(Y \mid A = 0, C = 1) = E(Y \mid A = 0, C = 2) = 0{\cdot}33$ and $E(Y \mid A = 0, C = 3) = 0{\cdot}83$. On the other hand, $E(A \mid C)$ is

Table 1. *Counterexample showing that the partial control result may be violated when $E(A \mid C)$ is not monotonic in $C$*

Crude

| | $A=0$ | $A=1$ | | | | |
|---|---|---|---|---|---|---|
| $Y=0$ | 221 | 126 | $\text{RD}_{\text{crude}}$ | 0·492 | $\text{RD}^{\text{T}}_{\text{crude}}$ | 0·492 |
| $Y=1$ | 115 | 633 | $\text{RR}_{\text{crude}}$ | 2·437 | $\text{RR}^{\text{T}}_{\text{crude}}$ | 2·437 |
| | | | $\text{OR}_{\text{crude}}$ | 9·654 | $\text{OR}^{\text{T}}_{\text{crude}}$ | 9·654 |

True

| $C=1$ | $A=0$ | $A=1$ | $C=2$ | $A=0$ | $A=1$ | $C=3$ | $A=0$ | $A=1$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | $\text{RD}_{\text{true}}$ | 0·486 | $\text{RD}^{\text{T}}_{\text{true}}$ | 0·484 |
| $Y=0$ | 200 | 120 | $Y=0$ | 20 | 2 | $Y=0$ | 1 | 4 | $\text{RR}_{\text{true}}$ | 2·398 | $\text{RR}^{\text{T}}_{\text{true}}$ | 2·380 |
| $Y=1$ | 100 | 600 | $Y=1$ | 10 | 11 | $Y=1$ | 5 | 22 | $\text{OR}_{\text{true}}$ | 9·430 | $\text{OR}^{\text{T}}_{\text{true}}$ | 9·311 |

Observed adjusted

| $C'=1$ | $A=0$ | $A=1$ | $C'=2$ | $A=0$ | $A=1$ | $C'=3$ | $A=0$ | $A=1$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | $\text{RD}_{\text{obs}}$ | 0·495 | $\text{RD}^{\text{T}}_{\text{obs}}$ | 0·496 |
| $Y=0$ | 200 | 120 | $Y=0$ | 12·4 | 2·8 | $Y=0$ | 8·6 | 3·2 | $\text{RR}_{\text{obs}}$ | 2·459 | $\text{RR}^{\text{T}}_{\text{obs}}$ | 2·468 |
| $Y=1$ | 100 | 600 | $Y=1$ | 8 | 15·4 | $Y=1$ | 7 | 17·6 | $\text{OR}_{\text{obs}}$ | 9·801 | $\text{OR}^{\text{T}}_{\text{obs}}$ | 9·844 |

Table 2. *Counterexample showing that the partial control result may be violated for dichotomized $C$ when $E(A \mid C)$ is not monotonic in $C$*

Observed adjusted

| $C'=1$ | $A=0$ | $A=1$ | $C'=2$ | $A=0$ | $A=1$ | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | $\text{RD}_{\text{obs}}$ | 0·495 | $\text{ATT}^{\text{RD}}_{\text{obs}}$ | 0·496 |
| $Y=0$ | 200 | 120 | $Y=0$ | 21 | 6 | $\text{RR}_{\text{obs}}$ | 2·461 | $\text{ATT}^{\text{RR}}_{\text{obs}}$ | 2·470 |
| $Y=1$ | 100 | 600 | $Y=1$ | 15 | 33 | $\text{OR}_{\text{obs}}$ | 9·809 | $\text{ATT}^{\text{OR}}_{\text{obs}}$ | 9·856 |

not monotonic in $C$, as $E(A \mid C=1) = 0{\cdot}71$, $E(A \mid C=2) = 0{\cdot}30$ and $E(A \mid C=3) = 0{\cdot}81$. Rather than following the partial control ordering, the observed adjusted measure is greater than the crude measure, which is greater than the true measure on all three scales, for both the average treatment effect and the effect of treatment on the treated. The observed adjusted effect measures exhibit more bias than do the crude effect measures.

In the Supplementary Material we provide two additional counterexamples: one in which the misclassification probabilities are tapered and $E(A \mid C)$ is monotonic but $E(Y \mid A, C)$ is not, and another in which $E(A \mid C)$ and $E(Y \mid A, C)$ are both monotonic but the misclassification probabilities are not tapered.

Suppose that instead of observing $C$ with nondifferential error we observe a coarsening of $C$. Then we have the following analogous result:

THEOREM 2. *Let $A$ be binary, and let $C$ be ordinal and coarsened such that $C' = j$ if $k_j \leqslant C \leqslant l_j$ for $k_j, l_j \in \{1, \dots, K\}$ and $j \in \{1, \dots, L\}$ for some $L < K$. If $E(Y \mid A, C)$ and $E(A \mid C)$ are monotonic in $C$, then the observed adjusted average treatment effect and the observed adjusted effect of treatment on the treated will lie between the corresponding true and crude effects for any effect measure that can be written as $h[g\{E(Y_1)\} - g\{E(Y_0)\}]$ where $g$ and $h$ are monotonic functions.*

The full data given in Table 1 serve as a counterexample to Theorem 2 when the monotonicity assumption is violated, if instead of mismeasuring the true $C$ we dichotomize it by letting $C' = 1$ if $C = 1$ and $C' = 2$ if $C > 1$. Then the true and crude data and the effect measures will remain the same as in Table 1; the new observed data and effect measures are given in Table 2. Again, the observed adjusted measure is greater than the crude effect measure, which in turn is greater than the true measure on all three scales, for both the average treatment effect and the effect of treatment on the treated.

All our results hold for $Y$ binary, ordinal or continuous, as none of the proofs depend on the distribution of $Y$. Because these results are entirely nonparametric, they apply to any estimators with probability limits given by the analytic expressions in (1) and (2), for example to regression, propensity score weighting methods, and inverse probability weighting methods. The results also hold for the ordering of contrasts of $E(Y \mid A = a)$, contrasts of the analytic expression given in (1), and contrasts of the analytic expression given in (2), even when $\sum_{k=1}^{K} E(Y \mid A=a, C=k)\mathrm{pr}(C=k)$ is not equal to $E(Y_a)$, for instance if there are unmeasured confounders of the effect of $A$ on $Y$ in addition to $C$.

### 3·2. *Extensions*

If $C$ is categorical rather than ordinal, then Theorem 1 will hold if any ordering of $C$ can be found such that $E(Y \mid A, C)$ and $E(A \mid C)$ are both monotonic in $C$ and the misclassification probabilities are tapered with respect to that ordering. Theorem 2 will hold if $C$ is coarsened with respect to the monotonic ordering.

All the results may be extended to effects conditional on additional covariates $X$, provided that the misclassification of $C$ is tapered and nondifferential with respect to $A$ and $Y$ conditional on $X$ and that $E(Y \mid A, C, X=x)$ and $E(A \mid C, X=x)$ are monotonic in $C$ for each value $x$ in the support of $X$. This extension is immediate because under these provisions our proofs hold conditional on $X$. Furthermore, if $E(Y \mid A, C, X)$ and $E(A \mid C, X)$ are monotonic in $C$, that is, if $E(Y \mid A=a, C=c, X=x)$ is either nonincreasing or nondecreasing in $c$ for all $a$ and all $x$, and $E(A \mid C=c, X=x)$ is either nonincreasing or nondecreasing in $c$ for all $x$, then the ordering of the crude, observed adjusted, and true effects will be the same in each level of $X$ and so the results will also hold standardized over $X$.

In some settings, we can also extend our results to scenarios with multiple misclassified confounders.

THEOREM 3. *Let $C$ be a vector of ordinal confounders, $C=(C_1,\ldots,C_m)$. Suppose that:* (i) *the components of $C$ are mutually independent;* (ii) *for each $i \in \{1,\ldots,m\}$, either $C_i$ is nondifferentially misclassified with respect to $A$, $Y$ and $C_j$ for $j \neq i$ with tapered misclassification probabilities, or $C_i$ is coarsened; and* (iii) *$E(A \mid C)$ and $E(Y \mid A, C)$ are monotonic in each component of $C$. Then the observed adjusted average treatment effect and the observed adjusted effect of treatment on the treated will lie between the corresponding crude and true effects for any effect measure that can be written as $h[g\{E(Y_1)\} - g\{E(Y_0)\}]$ where $g$ and $h$ are monotonic functions.*

## 4. DISCUSSION

The results in this paper imply that when the effects of an ordinal confounder on the treatment and on the outcome are both monotonic, bias will be reduced by adjusting for a coarsened or nondifferentially misclassified confounder with tapered misclassification probabilities. In the absence of monotonicity, adjusting for a misclassified or coarsened confounder may increase bias. Future work is needed to determine how often the assumptions of monotonicity are likely to be violated in practice, and how often such violations will lead to violations of the partial control result. These results are important because the use of proxy or misclassified confounders is common in practice, and because misclassification of confounders is nearly impossible to rule out.

## ACKNOWLEDGEMENT

Both authors were supported by the National Institutes of Health, U.S.A. The authors are grateful to the referees and the editor for extremely helpful comments which have resulted in substantial improvements to this work.

## SUPPLEMENTARY MATERIAL

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

## APPENDIX

In order to prove each theorem, we must show that the conditions of Lemma A1 hold; Lemmas A2–A5 will help to establish this in the various cases.

LEMMA A1. *If $E(Y \mid A=1) \geqslant E_{C'}(Y \mid A=1) \geqslant E(Y_1)$ and $E(Y \mid A=0) \leqslant E_{C'}(Y \mid A=0) \leqslant E(Y_0)$, or if $E(Y \mid A=1) \leqslant E_{C'}(Y \mid A=1) \leqslant E(Y_1)$ and $E(Y \mid A=0) \geqslant E_{C'}(Y \mid A=0) \geqslant E(Y_0)$, then the observed adjusted effect falls between the crude and true effects for any effect measure that can be written as $h[g\{E(Y_1)\} - g\{E(Y_0)\}]$ where $g$ and $h$ are monotonic functions.*

LEMMA A2. *If $E(Y \mid A, C)$ and $E(A \mid C)$ are either both nonincreasing or both nondecreasing in $C$, then $E(Y \mid A=1) \geqslant E(Y_1)$ and $E(Y \mid A=0) \leqslant E(Y_0)$. If one of $E(Y \mid A, C)$ and $E(A \mid C)$ is nonincreasing and the other nondecreasing in $C$, then $E(Y \mid A=1) \leqslant E(Y_1)$ and $E(Y \mid A=0) \geqslant E(Y_0)$.*

LEMMA A3. *Suppose that $C$ is nondifferentially misclassified with respect to $A$ and $Y$. If $E(Y \mid A, C)$ and $E(A \mid C)$ are both nondecreasing or both nonincreasing in $C$, then $E_{C'}(Y \mid A=1) \geqslant E(Y_1)$ and $E_{C'}(Y \mid A=0) \leqslant E(Y_0)$. If one of $E(Y \mid A, C)$ and $E(A \mid C)$ is nondecreasing and the other nonincreasing in $C$, then $E_{C'}(Y \mid A=1) \leqslant E(Y_1)$ and $E_{C'}(Y \mid A=0) \geqslant E(Y_0)$.*

LEMMA A4. *Suppose that $C$ is nondifferentially misclassified with respect to $A$ and $Y$ with tapered misclassification probabilities. If $E(Y \mid A, C)$ and $E(A \mid C)$ are both nondecreasing or both nonincreasing in $C$, then $E_{C'}(Y \mid A=1) \leqslant E(Y \mid A=1)$ and $E_{C'}(Y \mid A=0) \geqslant E(Y \mid A=0)$. If one of $E(Y \mid A, C)$ and $E(A \mid C)$ is nondecreasing and the other nonincreasing in $C$, then $E_{C'}(Y \mid A=1) \geqslant E(Y \mid A=1)$ and $E_{C'}(Y \mid A=0) \leqslant E(Y \mid A=0)$.*

LEMMA A5. *Suppose that $A$ is binary and $C$ is ordinal and coarsened. If $E(Y \mid A, C)$ and $E(A \mid C)$ are both nondecreasing or both nonincreasing in $C$, then $E(Y \mid A=1) \geqslant E_{C'}(Y \mid A=1) \geqslant E(Y_1)$ and $E(Y \mid A=0) \leqslant E_{C'}(Y \mid A=0) \leqslant E(Y_0)$. If one of $E(Y \mid A, C)$ and $E(A \mid C)$ is nonincreasing and the other nondecreasing, then $E(Y \mid A=1) \leqslant E_{C'}(Y \mid A=1) \leqslant E(Y_1)$ and $E(Y \mid A=0) \geqslant E_{C'}(Y \mid A=0) \geqslant E(Y_0)$.*

*Proof of Theorem* 1. For the average treatment effect, Lemmas A2–A4 establish the conditions of Lemma A1, and Theorem 1 then follows. The condition of tapered misclassification probabilities is required for Lemma A4, which describes the relation between the observed adjusted and crude expectations, but not for Lemma A3, which describes the relation between the observed adjusted and true expectations. For the effect of treatment on the treated, we will argue that under monotonicity, $E_{C'|A=1}(Y \mid A=0)$ lies between $E(Y_0 \mid A=1)$ and $E(Y \mid A=0)$. This, together with the fact that $E(Y_1 \mid A=1) = E_{C'|A=1}(Y \mid A=a) = E(Y \mid A=1)$ when $C$ is the only confounder of the relationship between $A$ and $Y$, proves the result.

Because $E(Y \mid A=0)$ does not depend on the distribution of $C$, and because $\text{pr}(C'=i \mid C=j, A=1) = p_{ij}$ by the assumption of nondifferential misclassification, we can replace the marginal distribution of $C$ with the conditional distribution of $C$ given $A=1$ in the proof of Theorem 1 to show that the same relationships must hold among $E_{C'|A=1}(Y \mid A=0)$, $E(Y_0 \mid A=1)$ and $E(Y \mid A=0)$ as among $E_{C'}(Y \mid A=0)$, $E(Y_0)$ and $E(Y \mid A=0)$. Therefore $E_{C'|A=1}(Y \mid A=0)$ lies between $E(Y_0 \mid A=1)$ and $E(Y \mid A=0)$. □

*Proof of Theorem* 2. Lemma A5 establishes the conditions of Lemma A1, and Theorem 2 for the average treatment effect follows from that lemma. The proof for the effect of treatment on the treated is analogous to the proof of Theorem 1 applied to the effect of treatment on the treated. □

*Proof of Theorem* 3. Let $C' = (C_1', \ldots, C_m')$ be the vector of observed, misclassified variables. Let $E_{C_1',\ldots,C_m'}(Y \mid A)$ be the expected value of $Y$ given $A$ standardized by $C'$, and let $E_{C_1,C_2',\ldots,C_m'}(Y \mid A)$ be the expected value of $Y$ given $A$ standardized by the true $C_1$ and the mismeasured $C_i'$ for $i = 2, \ldots, m$; let $E_{C_2',\ldots,C_m'}(Y \mid A)$ be the expected value of $Y$ given $A$ collapsed over $C_1$ and standardized by $C_i'$ for $i = 2, \ldots, m$; likewise for $E_{C_3',\ldots,C_m'}(Y \mid A)$ etc.

Because the components of $C$ are mutually independent, $E(Y \mid A, C_i)$ is monotonic in $C_i$ for $i = 1, \ldots, m$. Then

$$E_{C_1, C_2', \ldots, C_m'}(Y \mid A=1) \leqslant E_{C_1', \ldots, C_m'}(Y \mid A=1) \leqslant E_{C_2', \ldots, C_m'}(Y \mid A=1),$$
$$E_{C_1, C_2', \ldots, C_m'}(Y \mid A=0) \geqslant E_{C_1', \ldots, C_m'}(Y \mid A=0) \geqslant E_{C_2', \ldots, C_m'}(Y \mid A=0), \tag{A1}$$

by the monotonicity of $E(Y \mid A, C_1)$ in $C_1$ and because our results hold for $C_1$ standardized by the covariates $C_2', \ldots, C_m'$. In order for the standardized result to hold, the misclassification of $C_1$ must be tapered and nondifferential with respect to $A$ and $Y$ conditional on $C_2', \ldots, C_m'$, and $E(Y \mid A, C_1, C_2', \ldots, C_m')$ and $E(A \mid C_1, C_2', \ldots, C_m')$ must both be monotonic in $C_1$. These two conditions hold by the assumptions of mutually independent components of $C$ and mutually independent misclassification mechanisms. Similarly,

$$E_{C_2, C_3', \ldots, C_m'}(Y \mid A=1) \leqslant E_{C_2', \ldots, C_m'}(Y \mid A=1) \leqslant E_{C_3', \ldots, C_m'}(Y \mid A=1),$$
$$E_{C_2, C_3', \ldots, C_m'}(Y \mid A=0) \geqslant E_{C_2', \ldots, C_m'}(Y \mid A=0) \geqslant E_{C_3', \ldots, C_m'}(Y \mid A=0), \tag{A2}$$

by monotonicity and because our results hold standardized by $C_3', \ldots, C_m'$ and marginalized over $C_1$. Again, the standardized and marginalized results hold by our assumptions of mutually independent components and errors for $C$.

By the same argument as above,

$$E_{C_1, C_2, C_3' \ldots, C_m'}(Y \mid A=1) \leqslant E_{C_1, C_2', \ldots, C_m'}(Y \mid A=1) \leqslant E_{C_1, C_3', \ldots, C_m'}(Y \mid A=1),$$
$$E_{C_1, C_2, C_3', \ldots, C_m'}(Y \mid A=0) \geqslant E_{C_1, C_2', \ldots, C_m'}(Y \mid A=0) \geqslant E_{C_1, C_3', \ldots, C_m'}(Y \mid A=0). \tag{A3}$$

The result holds for $C_2$ standardized by $C_1$ and $C_3', \ldots, C_m'$. Combining (A1), (A2) and (A3) gives

$$\begin{aligned} E_{C_1, C_2, C_3' \ldots, C_m'}(Y \mid A=1) &\leqslant E_{C_1', \ldots, C_m'}(Y \mid A=1) \\ &\leqslant E_{C_2', \ldots, C_m'}(Y \mid A=1) \leqslant E_{C_3', \ldots, C_m'}(Y \mid A=1), \\ E_{C_1, C_2, C_3', \ldots, C_m'}(Y \mid A=0) &\geqslant E_{C_1', \ldots, C_m'}(Y \mid A=0) \\ &\geqslant E_{C_2', \ldots, C_m'}(Y \mid A=0) \geqslant E_{C_3', \ldots, C_m'}(Y \mid A=0), \end{aligned} \tag{A4}$$

where in each case the first inequality follows from (A3) and (A1), the second inequality follows from (A1), and the third inequality follows from (A2).

Now we need

$$E_{C_1, C_2, C_3, C_4' \ldots, C_m'}(Y \mid A=1) \leqslant E_{C_1', \ldots, C_m'}(Y \mid A=1) \leqslant E_{C_4', \ldots, C_m'}(Y \mid A=1),$$
$$E_{C_1, C_2, C_3, C_4' \ldots, C_m'}(Y \mid A=0) \geqslant E_{C_1', \ldots, C_m'}(Y \mid A=0) \geqslant E_{C_4', \ldots, C_m'}(Y \mid A=0). \tag{A5}$$

Repeating the argument for (A3) gives

$$E_{C_1, C_2, C_3, C_4' \ldots, C_m'}(Y \mid A=1) \leqslant E_{C_1, C_2, C_3', \ldots, C_m'}(Y \mid A=1) \leqslant E_{C_1, C_2, C_4', \ldots, C_m'}(Y \mid A=1),$$
$$E_{C_1, C_2, C_3, C_4' \ldots, C_m'}(Y \mid A=0) \geqslant E_{C_1, C_2, C_3', \ldots, C_m'}(Y \mid A=0) \geqslant E_{C_1, C_2, C_4', \ldots, C_m'}(Y \mid A=0),$$

which, combined with (A4), yields the first inequalities in (A5). Repeating the argument for (A2) gives

$$E_{C_3, C_4', \ldots, C_m'}(Y \mid A=1) \leqslant E_{C_3', \ldots, C_m'}(Y \mid A=1) \leqslant E_{C_4', \ldots, C_m'}(Y \mid A=1),$$
$$E_{C_2, C_4', \ldots, C_m'}(Y \mid A=0) \geqslant E_{C_3', \ldots, C_m'}(Y \mid A=0) \geqslant E_{C_3', \ldots, C_m'}(Y \mid A=0),$$

from which the second inequalities in (A5) follow.

Iterating this argument $K - 3$ more times, we find that

$$E(Y_1) = E_{C_1,\ldots,C_m}(Y \mid A = 1) \leqslant E_{C'_1,\ldots,C'_m}(Y \mid A = 1) \leqslant E(Y \mid A = 1),$$

$$E(Y_0) = E_{C_1,\ldots,C_m}(Y \mid A = 0) \geqslant E_{C'_1,\ldots,C'_m}(Y \mid A = 0) \geqslant (Y \mid A = 0),$$

from which it follows that the observed adjusted effect is between the crude and the true effects.

The proofs for the effect of treatment on the treated and for coarsened rather than misclassified $C'$ are analogous. □

## REFERENCES

AHLBOM, A. & STEINECK, G. (1992). Aspects of misclassification of confounding factors. *Am. J. Indust. Med.* **21**, 107–12.

BRENNER, H. (1993). Bias due to non-differential misclassification of polytomous confounders. *J. Clin. Epidemiol.* **46**, 57–63.

CARROLL, R. J., RUPPERT, D., STEFANSKI, L. A. & CRAINICEANU, C. M. (2006). *Measurement Error in Nonlinear Models: A Modern Perspective*. Boca Raton: CRC Press.

GREENLAND, S. (1980). The effect of misclassification in the presence of covariates. *Am. J. Epidemiol.* **112**, 564–9.

GREENLAND, S. (1996). Basic methods for sensitivity analysis of biases. *Int. J. Epidemiol.* **25**, 1107–16.

GREENLAND, S. (2004). Interval estimation by simulation as an alternative to and extension of confidence intervals. *Int. J. Epidemiol.* **33**, 1389–97.

GREENLAND, S. (2005). Multiple-bias modelling for analysis of observational data. *J. R. Statist. Soc.* A **168**, 267–306.

GUSTAFSON, P. (2004). *Measurement Error and Misclassification in Statistics and Epidemiology: Impacts and Bayesian Adjustments*. Boca Raton: CRC Press.

KELSEY, J. L. (1996). *Methods in Observational Epidemiology*. New York: Oxford University Press.

LASH, T. L. & FINK, A. K. (2003). Semi-automated sensitivity analysis to assess systematic errors in observational data. *Epidemiology* **14**, 451–8.

OGBURN, E. L. & VANDERWEELE, T. J. (2012). On the nondifferential misclassification of a binary confounder. *Epidemiology* **23**, 433–9.

ROTHMAN, K. J., GREENLAND, S. & LASH, T. L. (2008). *Modern Epidemiology*. Philadelphia: Lippincott Williams & Wilkins.

STEFANSKI, L. & CARROLL, R. (1985). Covariate measurement error in logistic regression. *Ann. Statist.* **13**, 1335–51.

WILLETT, W. (1989). An overview of issues related to the correction of non-differential exposure measurement error in epidemiologic studies. *Statist. Med.* **8**, 1031–40.

[*Received September* 2011. *Revised July* 2012]