Published in final edited form as:
Biometrika. 2013 ; 100(1): 241–248. doi:10.1093/biomet/ass054.

Bias attenuation results for nondifferentially mismeasured ordinal and coarsened confounders

ELIZABETH L. OGBURN and **TYLER J. VANDERWEELE**
Departments of Biostatistics and Epidemiology, Harvard School of Public Health, Boston, Massachusetts 02115

ELIZABETH L. OGBURN: ogburn@post.harvard.edu; TYLER J. VANDERWEELE: 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, 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 its corresponding crude and the true effect measures. We present results for coarsened, and, under further assumptions, for multiple misclassified confounders.

Keywords

Bias; Confounding; 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 studied 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 characterizing 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, …, 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 $p_{ij} = \text{pr}(C' = i \mid C = j)$, $i, j \in \{1, \ldots, K\}$. Then $\text{pr}(C'=i)=\sum_{j=1}^{K} p_{ij}\text{pr}(C=j)$. We assume throughout that the misclassification of C is nondifferential with respect to A and Y. That is, we assume that $\text{pr}(C' = c' \mid C = c, Y = y, A = a) = \text{pr}(C' = c' \mid C = c, Y = y', A = a')$ for all y, y' in the support of Y and all a, a' in the support of A. If $p_{ij} \le p_{ik}$ and $p_{ji} \le p_{ki}$ for $j < k < i$, and if $p_{il} \ge p_{im}$ and $p_{li} \ge p_{mi}$ for $i < l < m$, then we say that that the misclassification probabilities are tapered. Roughly, this means that the probability of correct classification is at least as big 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 in 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 the identification of the true effect. Figure 1 depicts the relationships among the variables. We discuss generalizations allowing for multiple confounders and for coarsened confounders further below.

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' 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 $\text{RD}_{\text{true}} \equiv E(Y_1) - E(Y_0)$. For no subject do we observe both Y_1 and Y_0; we observe only the counterfactual corresponding to the subject's actual treatment. But 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:

$$E(Y_a)=\sum_{k=1}^{K} E(Y|A=a, C=k)\,\text{pr}(C=k). \quad (1)$$

The corresponding expression standardized by C' instead of C is

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

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

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 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 with a superscript T. For example, the true treatment effect among the treated on the risk difference scale for binary Y is $\mathrm{RD}^T_{\text{true}} = E(Y_1|A=1) - E(Y_0|A=1)$, where, if C is the only confounder, $E(Y_a \mid A = 1) = \Sigma_c E(Y \mid A = a, C = c)\ \mathrm{pr}(C = c \mid A = 1)$ 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 $E_{C'|A=1}(Y \mid A = a) \equiv \Sigma_c E(Y \mid A = a, C' = c)\ \mathrm{pr}(C' = c \mid A = 1)$, 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 $\mathrm{RD}^T_{\text{obs}} \equiv E_{C'|A=1}(Y|A=1) - E_{C'|A=1}(Y|A=0)$. 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 $E(Y \mid A = a, C = i) \le E(Y \mid A = a, C = j)$ 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 $E(Y \mid A = a, C = i) \ge E(Y \mid A = a, C = j)$ 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 results 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 and 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 be 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, and 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 is 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 pr $(C' = 1 \mid C = 1) = 1{\cdot}0$, pr $(C' = 2 \mid C = 1) =$ pr $(C' = 3 \mid C = 1) = 0$, pr $(C' = 1 \mid C = 2) = 0$, pr $(C' = 2 \mid C = 2) = 0{\cdot}6$, pr $(C' = 3 \mid C = 2) = 0{\cdot}4$, pr $(C' = 1 \mid C = 3) = 0$, pr $(C' = 2 \mid C = 3) = 0{\cdot}4$, and 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 not monotonic in C: $E(A \mid C = 1) = 0{\cdot}71$, $E(A \mid C = 2) = 0{\cdot}30$, and $E(A \mid C = 3) = 0{\cdot}81$. Instead of following the partial control ordering, the observed adjusted measure is greater than the crude, 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 online supplement 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. We have the following analogous result:

Theorem 2—Let A be binary and C ordinal and coarsened such that $C' = j$ if $k_j \le C \le l_j$ for $k_j, l_j \in (1, \ldots, K)$ and $j \in (1, \ldots, L)$ 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 be 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 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 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.

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 $\sum_{k=1}^{K} E(Y|A{=}a, C{=}k)\,\mathrm{pr}(C{=}k)$ is not equal to $E(Y_a)$, for example 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 monotonic in C and if 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 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 nonincreasing in c for all a and for all x or is nondecreasing in c for all a and for all x, and $E(A \mid C = c, X = x)$ is either nonincreasing in c for all x or is 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 therefore the results will also hold standardized over X.

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

Theorem 3—Let C be a vector of ordinal confounders, $C = (C_1, \ldots, C_m)$. If the components of C are mutually independent; if for each $i \in \{1, \ldots, m\}$ C_i is nondifferentially misclassified with respect to A, Y, and C_j for $j \neq i$ with tapered misclassification probabilities or if C_i is coarsened; and if $E(A \mid C)$ 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 be 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 will 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.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

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

NIH-PA Author Manuscript

NIH-PA Author Manuscript

References

Ahlbom A, Steineck G. Aspects of misclassification of confounding factors. American Journal of Industrial Medicine. 1992; 21:107–112. [PubMed: 1553979]

Brenner H. Bias due to non-differential misclassification of polytomous confounders. Journal of Clinical Epidemiology. 1993; 46:57–63. [PubMed: 8433115]

Carroll, RJ.; Ruppert, D.; Stefanski, LA.; Crainiceanu, CM. Measurement Error in Nonlinear Models: A Modern Perspective. Boca Raton, Florida: CRC Press; 2006.

Greenland S. The effect of misclassification in the presence of covariates. American Journal of Epidemiology. 1980; 112:564–569. [PubMed: 7424903]

Greenland S. Basic methods for sensitivity analysis of biases. International Journal of Epidemiology. 1996; 25:1107–1116. [PubMed: 9027513]

Greenland S. Interval estimation by simulation as an alternative to and extension of confidence intervals. International Journal of Epidemiology. 2004; 33:1389–1397. [PubMed: 15319402]

Greenland S. Multiple-bias modelling for analysis of observational data. Journal of the Royal Statistical Society: Series A (Statistics in Society). 2005; 168:267–306.

Gustafson, P. Measurement Error and Misclassification in Statistics and Epidemiology: Impacts and Bayesian Adjustments. Boca Raton, Florida: CRC Press; 2004.

Kelsey, JL. Methods in Observational Epidemiology. New York, New York: Oxford University Press; 1996.

Lash TL, Fd AK. Semi-automated sensitivity analysis to assess systematic errors in observational data. Epidemiology. 2003; 14:451–458. [PubMed: 12843771]

Ogburn EL, VanderWeele TJ. On the nondifferential misclassification of a binary confounder. Epidemiology. 2012; 23:433–439. [PubMed: 22450692]

Rothman, KJ.; Greenland, S.; Lash, TL. Modern Epidemiology. Philadelphia, Pennsylvania: Lippincott Williams & Wilkins; 2008.

Stefanski L, Carroll R. Covariate measurement error in logistic regression. The Annals of Statistics. 1985; 13:1335–1351.

Willett W. An overview of issues related to the correction of non-differential exposure measurement error in epidemiologic studies. Statistics in Medicine. 1989; 8:1031–1040. [PubMed: 2678348]

Appendix

In order to prove each result we must show that the conditions of Lemma 1 hold; Lemmas 2 through 5 will help establish this in the various cases. The proofs of all lemmas are given in the online supplement.

Lemma 1

If $E(Y|A=1) \geq E_{C'}(Y|A=1) \geq E(Y_1)$ and $E(Y|A=0) \leq E_{C'}(Y|A=0) \leq E(Y_0)$, or if $E(Y|A=1) \leq E_{C'}(Y|A=1) \leq E(Y_1)$ and $E(Y|A=0) \geq E_{C'}(Y|A=0) \geq E(Y_0)$, 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 2

If $E(Y|A, C)$ and $E(A|C)$ are either both nonincreasing or both nondecreasing in C, then $E(Y|A=1) \geq E(Y_1)$ and $E(Y|A=0) \leq E(Y_0)$. If one of $E(Y|A, C)$ and $E(A|C)$ is nonincreasing and the other nondecreasing in C, then $E(Y|A=1) \leq E(Y_1)$ and $E(Y|A=0) \geq E(Y_0)$.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Lemma 3

Suppose 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 $E_{C'}(Y \mid A=1) \geq E(Y_1)$ and $E_{C'}(Y \mid A=0) \leq E(Y_0)$. If one of $E(Y \mid A, C)$ and $E(A \mid C)$ is nondecreasing and the other nonincreasing in C, then $E_{C'}(Y \mid A=1) \leq E(Y_1)$ and $E_{C'}(Y \mid A=0) \geq E(Y_0)$.

Lemma 4

Suppose 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 $E_{C'}(Y \mid A=1) \leq E(Y \mid A=1)$ and $E_{C'}(Y \mid A=0) \geq E(Y \mid A=0)$. If one of $E(Y \mid A, C)$ and $E(A \mid C)$ is nondecreasing and the other nonincreasing in C, then $E_{C'}(Y \mid A=1) \geq E(Y \mid A=1)$ and $E_{C'}(Y \mid A=0) \leq E(Y \mid A=0)$.

Lemma 5

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

Proof of Theorem 1

For the average treatment effect, Lemmas 2–4 establish the conditions of Lemma 1 and the proof of Theorem 1 follows from that lemma. The condition of tapered misclassification probabilities is required for Lemma 4, describing the relation between observed adjusted and the crude expectations, but not for Lemma 3, describing the relation between the observed adjusted and the true expectations. For the effect of treatment on the treated, we will argue that, under monotonicity, $E_{C' \mid A=1}(Y \mid A=0)$ lies between $E(Y_0 \mid A=1)$ and $E(Y \mid A=0)$. This, together with the fact that $E(Y_1 \mid A=1) = E_{C' \mid A=1}(Y \mid A=a) = E(Y \mid A=1)$ 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 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 prove that the same relationships must hold among $E_{C' \mid A=1}(Y \mid A=0)$, $E(Y_0 \mid A=1)$, and $E(Y \mid A=0)$ as hold among $E_{C'}(Y \mid A=0)$, $E(Y_0)$, and $E(Y \mid A=0)$. Therefore $E_{C' \mid A=1}(Y \mid A=0)$ lies between $E(Y_0 \mid A=1)$ and $E(Y \mid A=0)$.

Proof of Theorem 2

Lemma 5 establishes the conditions of Lemma 1 and the proof of 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 $E_{C_1', \ldots, C_m'}(Y|A)$ be the expected value of Y given A standardized by C', let $E_{C_1, C_2', \ldots, C_m'}(Y|A)$ be the expected value of Y given A standardized by the true C_1 and the mismeasured C_i' for $i = 2, \ldots, m$, and let

$E_{c_2',\ldots,c_m'}(Y|A)$ be the expected value of Y given A collapsed over C_1 and standardized by C_i' for $i = 2, \ldots, m$. Likewise for $E_{c_3',\ldots,c_m'}(Y|A)$, etcetera.

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

$$\begin{aligned} &E_{c_1,c_2',\ldots,c_m'}(Y|A=1) \le E_{c_1',\ldots,c_m'}(Y|A=1) \le E_{c_2',\ldots,c_m'}(Y|A=1), \\ &E_{c_1,c_2',\ldots,c_m'}(Y|A=0) \ge E_{c_1',\ldots,c_m'}(Y|A=0) \ge E_{c_2',\ldots,c_m'}(Y|A=0), \end{aligned} \quad \text{(A1)}$$

by the monotonicity of $E(Y \mid A, C_1)$ in C_1 and because our results hold for C_1 standardized by 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 $E\left(Y|A, C_1, C_2', \ldots, C_m'\right)$ and $E\left(A|C_1, C_2', \ldots, C_m'\right)$ 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,

$$\begin{aligned} &E_{c_2,c_3',\ldots,c_m'}(Y|A=1) \le E_{c_2',\ldots,c_m'}(Y|A=1) \le E_{c_3',\ldots,c_m'}(Y|A=1), \\ &E_{c_2,c_3',\ldots,c_m'}(Y|A=0) \ge E_{c_2',\ldots,c_m'}(Y|A=0) \ge E_{c_3',\ldots,c_m'}(Y|A=0). \end{aligned} \quad \text{(A2)}$$

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,

$$\begin{aligned} &E_{c_1,c_2,c_3',\ldots,c_m'}(Y|A=1) \le E_{c_1,c_2',\ldots,c_m'}(Y|A=1) \le E_{c_1,c_3',\ldots,c_m'}(Y|A=1), \\ &E_{c_1,c_2,c_3',\ldots,c_m'}(Y|A=0) \ge E_{c_1,c_2',\ldots,c_m'}(Y|A=0) \ge E_{c_1,c_3',\ldots,c_m'}(Y|A=0). \end{aligned} \quad \text{(A3)}$$

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

$$\begin{aligned} &E_{c_1,c_2,c_3',\ldots,c_m'}(Y|A=1) \le E_{c_1',\ldots,c_m'}(Y|A=1) \le E_{c_2',\ldots,c_m'}(Y|A=1) \le E_{c_3',\ldots,c_m'}(Y|A=1), \\ &E_{c_1,c_2,c_3',\ldots,c_m'}(Y|A=0) \ge E_{c_1',\ldots,c_m'}(Y|A=0) \ge E_{c_2',\ldots,c_m'}(Y|A=0) \ge E_{c_3',\ldots,c_m'}(Y|A=0). \end{aligned} \quad \text{(A4)}$$

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

Now we need

$$\begin{aligned} &E_{c_1,c_2,c_3,c_4',\ldots,c_m'}(Y|A=1) \le E_{c_1',\ldots,c_m'}(Y|A=1) \le E_{c_4',\ldots,c_m'}(Y|A=1), \\ &E_{c_1,c_2,c_3,c_4',\ldots,c_m'}(Y|A=0) \ge E_{c_1',\ldots,c_m'}(Y|A=0) \ge E_{c_4',\ldots,c_m'}(Y|A=0). \end{aligned} \quad \text{(A5)}$$

Repeating the argument for (A3) gives us

$$E_{c_1,c_2,c_3,c'_4,\ldots,c'_m}(Y|A=1) \le E_{c_1,c_2,c'_3,\ldots,c'_m}(Y|A=1) \le E_{c_1,c_2,c'_4,\ldots,c'_m}(Y|A=1),$$
$$E_{c_1,c_2,c_3,c'_4,\ldots,c'_m}(Y|A=0) \ge E_{c_1,c_2,c'_3,\ldots,c'_m}(Y|A=0) \ge E_{c_1,c_2,c'_4,\ldots,c'_m}(Y|A=0), \quad \text{(A6)}$$

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

$$E_{c_3,c'_4,\ldots,c'_m}(Y|A=1) \le E_{c'_3,\ldots,c'_m}(Y|A=1) \le E_{c'_4,\ldots,c'_m}(Y|A=1),$$
$$E_{c_2,c'_4,\ldots,c'_m}(Y|A=0) \ge E_{c'_3,\ldots,c'_m}(Y|A=1) \ge E_{c'_3,\ldots,c'_m}(Y|A=0), \quad \text{(A7)}$$

from which the second inequalities in (A5) follow.

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

$$E(Y_1)=E_{c_1,\ldots,c_m}(Y|A=1) \le E_{c'_1,\ldots,c'_m}(Y|A=1) \le E(Y|A=1),$$
$$E(Y_0)=E_{c_1,\ldots,c_m}(Y|A=0) \ge E_{c'_1,\ldots,c'_m}(Y|A=0) \ge (Y|A=0),$$

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.

NIH-PA Author Manuscript
NIH-PA Author Manuscript

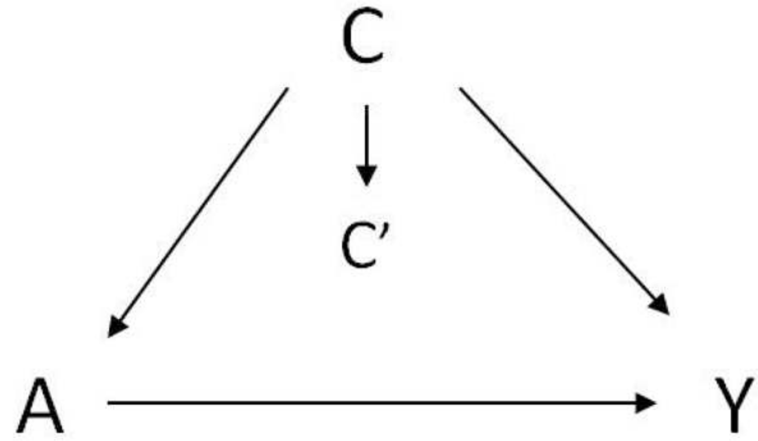

Fig. 1.
Relations among treatment A, outcome Y, confounder C, and an imperfect measure C' of C

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 1

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

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

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

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

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