

Quantitative Bias Analysis for a Misclassified Confounder

A Comparison Between Marginal Structural Models and Conditional Models for Point Treatments

Linda Nab,^a Rolf H. H. Groenwold,^{a,b} Maarten van Smeden,^a and Ruth H. Keogh^c

Abstract: Observational data are increasingly used with the aim of estimating causal effects of treatments, through careful control for confounding. Marginal structural models estimated using inverse probability weighting (MSMs-IPW), like other methods to control for confounding, assume that confounding variables are measured without error. The average treatment effect in an MSM-IPW may however be biased when a confounding variable is error prone. Using the potential outcome framework, we derive expressions for the bias due to confounder misclassification in analyses that aim to estimate the average treatment effect using a marginal structural model estimated using inverse probability weighting (MSM-IPW). We compare this bias with the bias due to confounder misclassification in analyses based on a conditional regression model. Focus is on a point-treatment study with a continuous outcome. Compared with bias in the average treatment effect in a conditional model, the bias in an MSM-IPW can be different in magnitude but is equal in sign. Also, we use a simulation study to investigate the finite sample performance of MSM-IPW and conditional models when a confounding variable is misclassified. Simulation results indicate that confidence intervals of

the treatment effect obtained from MSM-IPW are generally wider, and coverage of the true treatment effect is higher compared with a conditional model, ranging from overcoverage if there is no confounder misclassification to undercoverage when there is confounder misclassification. Further, we illustrate in a study of blood pressure-lowering therapy, how the bias expressions can be used to inform a quantitative bias analysis to study the impact of confounder misclassification, supported by an online tool.

Keywords: Inverse probability weighting; Marginal structural models; Misclassification; Point-treatment study; Quantitative bias analysis

(*Epidemiology* 2020;31: 796–805)

The aim of many observational epidemiologic studies is to estimate a causal relation between an exposure and an outcome, through careful control for confounding. In the case of a point-treatment, that is, estimating the effect of a treatment at a single time point on a subsequent outcome, many methods exist that aim to estimate average treatment effects. These include traditional conditional regression analysis and marginal structural models estimated using inverse probability weighting (MSMs-IPW).^{1,2} Unlike conditional regression, MSMs extend to estimation of joint treatment effects over multiple time points in longitudinal settings with time-dependent confounding.^{1,3}

To obtain valid inference, MSMs-IPW, like other methods to control for confounding, assume that confounding variables are measured without error, an assumption hardly ever warranted in observational epidemiologic research.^{4–7} A type of measurement error is classification error, which occurs when categorical variables are misclassified. For instance, smoking status (smoker versus nonsmoker) is prone to classification error but has been used as a confounding variable in studies investigating dialysis on mortality⁸ and iron supplement use during pregnancy on anemia at delivery.⁹ Another example of the use of a potentially misclassified confounding variable is alcohol use during pregnancy (yes versus no) in studies investigating associations between exposure to triptans during fetal life and risk of externalizing and internalizing behaviors in children.¹⁰ In all aforementioned examples, MSMs were used to estimate the exposure–outcome relation, but the assumption of error-free confounding variables is

Submitted December 11, 2019; accepted July 17, 2020.

From the ^aDepartment of Clinical Epidemiology, Leiden University Medical Center, Leiden, The Netherlands; ^bDepartment of Biomedical Data Sciences, Leiden University Medical Center, Leiden, The Netherlands; and ^cDepartment of Medical Statistics, London School of Hygiene and Tropical Medicine, London, United Kingdom.

L.N., R.H.H.G., and M.v.S. were supported by grants from the Netherlands Organization for Scientific Research (ZonMW-Vidi project 917.16.430) and Leiden University Medical Center. R.H.K. was supported by a Medical Research Council Methodology Fellowship (MR/M014827/1) and a UK Research and Innovation Future Leaders Fellowship (MR/S017968/1). This work received support from Stichting Jo Kolk Studiefonds and Leids Universiteits Fonds in the form of a travel grant

The authors report no conflicts of interest.

SDC Supplemental digital content is available through direct URL citations in the HTML and PDF versions of this article (www.epidem.com).

Data and code availability: The data and code used for the simulation study and quantitative bias analysis have been made publicly and can be accessed via www.github.com/LindaNab/memsm.

Correspondence: Linda Nab, Department of Clinical Epidemiology, Leiden University Medical Center, Postzone C7-P, P.O. Box 9600, 2300 RC Leiden, The Netherlands. E-mail: lnab@lumc.nl.

Copyright © 2020 The Author(s). Published by Wolters Kluwer Health, Inc. This is an open access article distributed under the Creative Commons Attribution License 4.0 (CCBY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ISSN: 1044-3983/20/3106-0796

DOI: 10.1097/EDE.0000000000001239

possibly violated and may lead to bias in the treatment effect estimator.

There is a substantial literature on bias due to measurement error in confounding variables in conditional regression analyses,^{11–15} but the impact of measurement error in confounding variables in causal inference methods, such as MSMs-IPW, has not received much attention. One exception is a study by Regier et al¹⁶ that showed by means of a simulation study that measurement error in continuous confounding variables can introduce bias in the average treatment effect estimator in a point-treatment study. McCaffrey et al¹⁷ proposed a weighting method to restore the treatment effect estimator when covariates are measured with error.

We provide a discussion of measurement error in a confounding variable. In addition, we derive expressions that quantify the bias in the average treatment effect estimator if a dichotomous confounding variable is misclassified, focusing on a point-treatment study with a continuous outcome. These expressions allow us (1) to quantify the bias due to classification error in a confounding variable in MSMs-IPW and to compare this with the bias from a conditional regression analysis and (2) to inform quantitative bias analyses.^{18–20} We use simulation results to study the finite sample performance of an marginal structural model estimated using inverse probability weighting (MSM-IPW) compared with that of conditional regression models if classification error in a confounding variable is present. We illustrate our quantitative bias analysis in a study of the effect of blood pressure-lowering drugs on blood pressure.

SETTINGS AND IMPACT OF MEASUREMENT ERROR, NOTATION, AND ASSUMPTIONS

Let A denote the treatment indicator and Y the outcome. Let there be a variable L that confounds the association between treatment and outcome and suppose that, instead of confounding variable L , the error-prone confounding variable L^* is observed. We consider two settings in which measurement error in confounding variables may occur and discuss the impact of measurement error in both settings.

Settings and Impact of Measurement Error

The directed acyclic graph (DAG) in Figure 1 illustrates setting 1. In this setting, treatment initiation is based on the error-prone confounding variable. Consider, for example, a study investigating the relation between the use of

antidepressant drugs (A) and the risk of a hip fracture (Y).²¹ Benzodiazepine use may be a confounding variable but is prone to classification error because only prescription information may be available and over-the-counter use is often unknown. The clinician initiating the antidepressant drugs might not know their patient’s over-the-counter use and initiates treatment based on the observed error-prone benzodiazepine use (L^*) instead of actual use (L), as depicted in Figure 1A. Here, conditioning on the error-prone L^* will block the backdoor path from treatment A to outcome Y . Thus, it is sufficient to control for the error-prone confounding variable to estimate the causal effect of treatment on outcome. This means that measurement error in a confounding variable will not always lead to bias.

The DAG in Figure 1B illustrates setting 2, in which treatment initiation is based on L , but only a proxy of L is observed (L^*). An example here might be a study investigating the effect of influenza vaccination (A) on mortality (Y) in the elderly population.²² Frailty (L) possibly confounds the association between influenza vaccination and mortality. Frailty is observed by a clinician, but only a proxy of frailty (L^*) may be available in electronic health records, as depicted in Figure 1B. Here, conditioning on L^* will not fully adjust for confounding by L , because conditioning on L^* does not block the backdoor path from A to Y via L .

Notation and Assumptions

We will now continue investigating the impact of classification error in setting 2, by focusing on the setting where L is a dichotomous confounding variable and Y a continuous outcome. We use the potential outcomes framework.^{23,24} Let $Y^{a=0}$ denote the outcome that an individual would have had if treatment A was set to $a=0$, and let $Y^{a=1}$ denote the outcome if treatment A was set to $a=1$. We assume that L^* is nondifferentially misclassified with respect to the outcome ($L^* \perp\!\!\!\perp Y|L$) and to the treatment ($L^* \perp\!\!\!\perp A|L$). Let p_1 denote the sensitivity of L^* and $1-p_0$, the specificity of L^* (i.e., $P(L^*=1|L=1) = p_1$). We also denote the probability of treatment given the level of L by $P(A=1|L=l) = \pi_l$ and the prevalence of L by $P(L=1) = \lambda$. Here, we assume that $0 < \lambda < 1$ because we are not interested in populations where L is present or absent in everyone. Finally, we assume no measurement error in exposure and outcome.

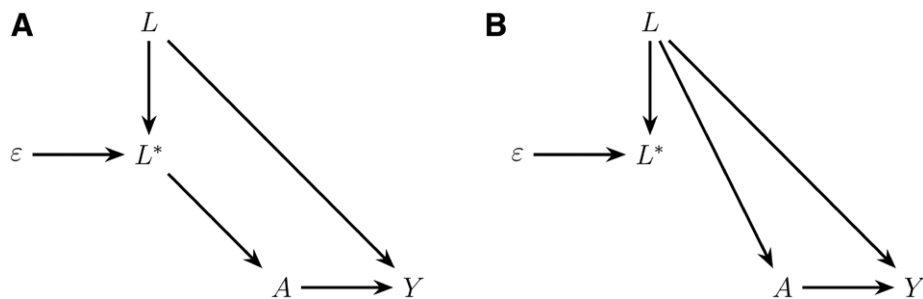


FIGURE 1. Measurement error ϵ in variable L that confounds the association between treatment A and outcome Y in two settings illustrated in directed acyclic graphs. A, Setting 1: Treatment A is initiated based on the error-prone confounding variable L^* . B, Setting 2: Treatment A is initiated based on confounding variable L .

We also assume that the following causal assumptions are satisfied to recover the causal effect of treatment on the outcome. Under the consistency assumption, we require that we observe $Y = Y^{a=0}$ if the individual is not exposed or $Y = Y^{a=1}$ if the individual is exposed.²⁵ Further, we assume that the potential outcome Y^a for an individual does not depend on treatments received by other individuals and that there are not multiple versions of treatment, also referred to as Stable Unit Treatment Value Assumption.²⁶ Additionally, we assume conditional exchangeability, i.e., given any level of L , if the untreated group had in fact received treatment, then their expected outcome would have been the same as that in the treated, and vice versa.²⁵ In notation, $A \perp\!\!\!\perp Y^a | L$, for $a=0, 1$. Finally, we assume $\pi_L > 0$ for $L=0, 1$ (positivity).²⁷

For causal contrasts, we compare expected potential outcomes (i.e., counterfactual outcomes) under the two different treatments. The average causal effect of the treatment on the outcome is $\beta = \mathbb{E}[Y^{a=1}] - \mathbb{E}[Y^{a=0}]$. Under the above defined assumptions, the conditional effect of treatment A on outcome Y can be defined through the following linear model:

$$\mathbb{E}[Y^a | L] = \mathbb{E}[Y | A = a, L] = \alpha + \beta a + \gamma L. \tag{1}$$

Estimates for β in the above model can be obtained by fitting a conditional regression model. Alternatively, the effect of treatment A on outcome Y may be estimated by fitting an MSM:

$$\mathbb{E}[Y^a] = \alpha_{\text{msm}} + \beta a, \text{ where } \alpha_{\text{msm}} = \alpha + \gamma \mathbb{E}[L]. \tag{2}$$

Estimates for β in the above model can be obtained by IPW estimation: by fitting a linear regression model for Y on A where the contribution of each individual is weighted by 1 over the probability of that individual's observed treatment given L , estimating the marginal treatment effect.² Because our focus is on linear models and we make the simplifying assumption that the effect of A on Y does not vary between strata of L , the conditional and marginal treatment effects, denoted by β in model equations 1 and 2, respectively, are equal. This is not generally true for nonlinear models due to noncollapsibility.² We assume that the effect of A on Y does not vary between strata of L , to derive bias expressions that are easier to use in practice and require fewer parameters.²⁸

QUANTIFICATION OF BIAS DUE TO CLASSIFICATION ERROR IN A CONFOUNDING VARIABLE

Our aim is to study the effect of using the misclassified confounding variable L^* in place of the confounding variable L in the conditional regression model or in the model for the weights used to fit the MSM on the average treatment effect estimator in the setting where L , not L^* , influences treatment initiation (setting 2 above).

Conditional Model

By the law of total expectation, the expected value of the outcome Y given treatment A and L^* is (see eAppendix 1; <http://links.lww.com/EDE/B698> section 1 for further detail),

$$\begin{aligned} \mathbb{E}[Y | A = a, L^*] &= \mathbb{E}_{L|A=a,L^*}[\mathbb{E}[Y | A = a, L^*, L]] = \{\alpha + \gamma\phi_{00} + \delta u_0\} \\ &+ \{\beta + \gamma(\phi_{10} - \phi_{00}) + \delta u_A\} a \\ &+ \{\gamma(\phi_{01} - \phi_{00}) + \delta u_{L^*}\} L^* \end{aligned}$$

where $\phi_{al^*} = P(L = \mathbb{1} | A = a, L^* = l^*)$, $\delta = \mathbb{E}[Y | A = 1, L^* = 1] = \gamma \times (\phi_{11} - \phi_{10} - \phi_{01} + \phi_{00})$ and u_0, u_A, u_{L^*} represent the coefficients of the linear model $\mathbb{E}[AL^* | A, L^*] = u_0 + u_A A + u_{L^*} L^*$, modeling the mean of A times L^* (i.e., AL^*) given A and L^* (see next paragraph for an explanation of why these appear). The coefficient for treatment A in the above model is $\beta + \gamma(\phi_{10} - \phi_{00}) + \delta u_A$, and is therefore biased for the parameter of interest (i.e., β). By rewriting u_A in terms of λ , π_0 , π_1 , P_0 and P_1 (see eAppendix 1 section 1), we find that the bias due to classification error in L^* in the average treatment effect in a conditional regression model is as follows:

$$\begin{aligned} \text{Bias}_{\text{cm}}(\beta) &= \gamma(\phi_{10} - \phi_{00}) \left(1 - \ell \times \left\{ \frac{\pi_1^*(1 - \pi_1^*)}{\pi_1^*(1 - \pi_1^*)\ell + \pi_0^*(1 - \pi_0^*)(1 - \ell)} \right\} \right) \\ &+ \gamma(\phi_{11} - \phi_{01}) \left(\ell \times \left\{ \frac{\pi_1^*(1 - \pi_1^*)}{\pi_1^*(1 - \pi_1^*)\ell + \pi_0^*(1 - \pi_0^*)(1 - \ell)} \right\} \right), \tag{3} \end{aligned}$$

where $\pi_{l^*}^* = P(A = \mathbb{1} | L^* = l^*)$, $\ell = P(L^* = 1)$ (see eAppendix 1 section 1 for a derivation).

We focused on a model for Y conditional on A and L^* which includes only main effects of A and L^* , as this is typically done in practice when replacing L with L^* . In fact, it can be shown that when the model for Y given A and L includes only main effects of A and L , the implied correctly specified model for Y given A and L^* also includes an interaction between A and L^* , explaining the appearance of u_0, u_A , and u_{L^*} in the above because the interaction is not modeled. See eAppendix 1 section 1 for the bias in case an interaction is modeled

Marginal Structural Model Estimated Using Inverse Probability Weighting

An MSM-IPW proceeds by fitting a linear regression for outcome Y on treatment A , where the contribution of each individual is weighted by 1 over the probability of that individual's observed treatment given misclassified L^* .² An estimator for the average treatment effect β is as follows:

$$\hat{\beta} = \frac{\sum_{i=1}^n \frac{1}{P(A_i | L_i^*)} (Y_i - \bar{Y}_w)(A_i - \bar{A}_w)}{\sum_{i=1}^n \frac{1}{P(A_i | L_i^*)} (A_i - \bar{A}_w)^2} \text{ where } \bar{Y}_w = \frac{\sum_{i=1}^n Y_i / P(A_i | L_i^*)}{\sum_{i=1}^n 1 / P(A_i | L_i^*)}$$

$$\text{and } \bar{A}_w = \frac{\sum_{i=1}^n A_i / P(A_i | L_i^*)}{\sum_{i=1}^n 1 / P(A_i | L_i^*)}$$

It can be shown that $E[\hat{\beta}] = \beta + \gamma(\phi_{10} - \phi_{00})(1 - \ell) + \gamma(\phi_{11} - \phi_{01})\ell$. Consequently, the bias in the average treatment effect in an MSM-IPW is as follows:

$$\text{Bias}_{\text{msm}}(\beta) = \gamma(\phi_{10} - \phi_{00})(1 - \ell) + \gamma(\phi_{11} - \phi_{01})\ell. \quad (4)$$

We refer to eAppendix 1 section 2 for a derivation of the above formula.

EXPLORATION OF BIAS

To study the bias due to misclassification from the conditional model and MSM-IPW, we explore bias expressions (equations 3 and 4).

Null Bias

To confirm the derived bias expressions, we consider three trivial conditions where bias in the average treatment effect estimator is expected to be null, in line with general understanding of causal inference.²⁹ (1) If there is no classification error in L^* , i.e., specificity is 1 ($p_0 = 0$) and sensitivity is 1 ($p_1 = 1$), it follows that L corresponds to L^* , irrespective of treatment level (i.e., $\phi_{10} = 0$, $\phi_{00} = 0$, $\phi_{11} = 1$, and $\phi_{01} = 1$). (2) If the true relation between L and Y is null (i.e., γ is zero, thus there is no arrow from L to Y in Figure 1B). (3) If L does not affect

the probability of receiving treatment (i.e., $\pi_0 = \pi_1$, thus there is no arrow from L to A in Figure 1B), the probability that L is 1 depends on the value of L^* but no longer on A (i.e., $\phi_{00} = \phi_{10}$ and $\phi_{01} = \phi_{11}$). Bias is null under these conditions for both models (MSM-IPW and conditional model). Because the bias expressions are strictly monotonic, the bias in an MSM-IPW cannot be negative if the bias in the conditional model is positive and vice versa (i.e., the bias will be in the same direction for both models).

Equal Biases

The bias in the average treatment effect in a conditional regression analysis is equal to that in an MSM-IPW if bias in both models is null (see above). We also see that bias expressions (equations 3 and 4) show that bias for the two methods is equal if the term between curly brackets in equation 3 is equal to 1, which is the case if (1) $\ell = 1$; (2) $\pi_0^* = \pi_1^*$; and (3) $\pi_0^* = 1 - \pi_1^*$. If conditions (1) and/or (2) are met, there is no bias in an MSM-IPW nor in a conditional model. Under condition (3), bias is generally non-null (except if, for example, $\gamma = 0$, see Null Bias).

Sign and Magnitude of Bias

Figures 2–4 illustrate the contributions to bias in the average treatment effect due to misclassification components (sensitivity and specificity) and due to confounding components (prevalence of confounding variable, strength of association between confounding variable and treatment and outcome) in a conditional model and an MSM-IPW, obtained by using the bias expressions.

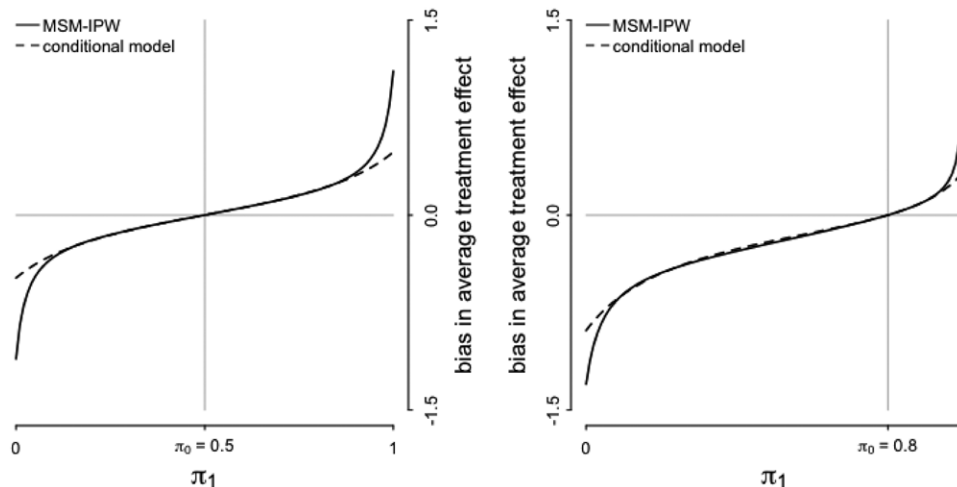


FIGURE 2. Visualization of the direction and magnitude of the bias in the average treatment effect estimator in relation to the prevalence of treatment among individuals with the confounding variable present. In this visualization, the confounding variable L is misclassified with a sensitivity of 0.9 and specificity of 0.95. Consequently, the average treatment effect estimated in an MSM-IPW or conditional regression model is biased, independent of true average treatment effect. The prevalence of L is 50% (i.e., $P(L=1) = 0.5$). The direction and magnitude of the bias depend on (1) the strength and direction of the association between L and treatment (denoted by $\pi_1 = P(\text{treatment}=1 | L=1)$ and $\pi_0 = P(\text{treatment}=1 | L=0)$, here set at $\pi_0 = 0.5$ in the left-hand-side plot and $\pi_0 = 0.8$ in the right-hand-side plot); and (2) the strength and direction of the association between L and the outcome (denoted by γ in the text and here set at $\gamma = 2$). Larger values of γ will result in steeper curves; $\gamma = -2$ will mirror the graph in $\gamma = 0$.

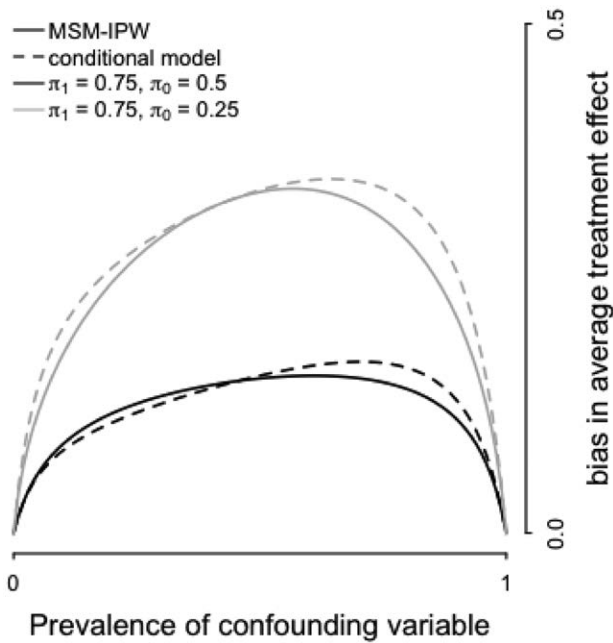


FIGURE 3. Visualization of the magnitude of the bias in the average treatment effect estimator in relation to the prevalence of a confounding variable. In this visualization, the confounding variable L is misclassified with a sensitivity of 0.9 and specificity of 0.95. Consequently, the average treatment effect estimated in an MSM-IPW or conditional regression model is biased, independent of true average treatment effect. The confounding variable is positively associated with treatment (i.e., here $\pi_1 > \pi_0$, where $\pi_1 = P(\text{treatment}=1|L=1)$ and $\pi_0 = P(\text{treatment}=1|L=0)$), and outcome (denoted by γ in the text and here set at $\gamma=2$). The magnitude of the bias depends on the prevalence of the confounding variable (i.e., $P(L=1)$). Larger values of γ will result in steeper curves.

Figure 2 shows that (1) the bias is positive if both the association between L and treatment and L and outcome are positive (i.e., $\pi_1 > \pi_0$ and $\gamma = 2$, respectively) and (2) the bias is greater if the difference between π_1 and π_0 is greater (i.e., if the strength of the association between L and treatment is greater). In contrast, the bias is negative if $\pi_1 < \pi_0$, whereas γ is positive. In case $\gamma = -2$, Figure 2 is mirrored in $y = 0$, and consequently, bias is negative if $\pi_1 > \pi_0$ and positive if $\pi_1 < \pi_0$. An increment in γ will result in greater bias and steeper curves in Figure 2. Figure 3 shows that the magnitude of the bias depends on the prevalence of L . Further, it shows that bias is greater if the strength of association between L and treatment is greater. Figure 4 shows that, generally, the bias is greater if L^* has lower specificity and sensitivity. Moreover, for a fixed sensitivity, bias is minimal if specificity equals 1 and is maximal if $1 - \text{specificity}$ equals sensitivity; by fixing specificity, bias is minimal if sensitivity equals 1 and is maximal if sensitivity equals $1 - \text{specificity}$. Figure 4 shows that the bias is greater if the strength of the association between L and treatment is greater. An increment in γ will result in

greater bias and steeper curves in Figure 4. An online application can be used to obtain bias plots for other combinations of the parameters available at <https://lindanab.shinyapps.io/SensitivityAnalysis>.

Simulation Study

We conducted a simulation study to study the finite sample properties of MSMs estimated using IPW and conditional models if there is classification error in the confounding variable. Five thousand data sets were generated with sample sizes of 1,000 and 100, using the following data-generating mechanisms:

$$L \sim \text{Bern}(\lambda), A|L \sim \text{Bern}(\pi_0^{(1-L)} \pi_1^L),$$

$$L^*|L \sim \text{Bern}(p_0^{(1-L)} p_1^L) \text{ and } Y|A, L \sim N(1 + \beta A + \gamma L, 1)$$

We studied five different scenarios, of which the parameters values can be found in Table 1. In all scenarios, the average treatment effect β (estimand) is 1 and the association between the confounding variable L and outcome Y is 2 (i.e., $\gamma = 2$). In scenario 0, we assume no classification error. In scenarios 1–4, we assume that L^* has a specificity of 0.95 (i.e., $p_0 = 0.05$) and a sensitivity of 0.90 (i.e., $p_1 = 0.9$). In scenario 1, bias in the average treatment effect estimator is expected to be negative because the probability of receiving treatment given that L is not present is greater than receiving treatment given that L is present, and the association between L and Y is positive (i.e., $\pi_0 > \pi_1$ and $\gamma = 2$). In contrast, in scenarios 2 and 3, bias in the average treatment effect estimator is expected to be positive, because $\pi_0 < \pi_1$ and $\gamma = 2$. Further, after investigation of Figure 3, we expect that bias in the average treatment effect estimated in a conditional model is greater than that in an MSM-IPW in scenarios 2 and 3. Finally, in scenario 4, we expect that bias in the average treatment effect from the conditional model is equal to that in an MSM-IPW.

Model Estimation and Performance Measures

We obtained the average treatment effect β (estimand) by fitting a conditional model using conditional regression and by fitting an MSM-IPW, both using the misclassified L^* instead of L from the data-generating mechanism. For the MSM-IPW analysis, we used the R package ipw.^{30,31} Performance of both models was evaluated in terms of the bias, the mean squared error of the estimated treatment effect (MSE), the percentages of 95% confidence intervals that contain the true value of the estimand (coverage), the empirical standard deviation of the estimated treatment effects (empSE), and mean model-based standard error of the estimated treatment effect. We estimated robust model-based standard errors of the average treatment effect in an MSM-IPW using the R package survey.³² We calculated Monte Carlo standard errors for all performance measures,³³ using the R package rsimsum.³⁴ Additionally, we calculated the theoretical bias of the average

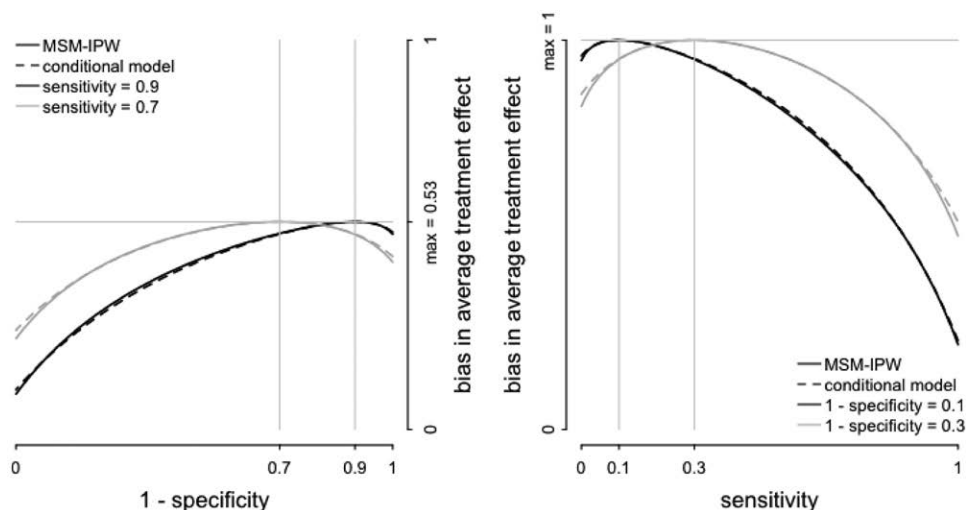


FIGURE 4. Visualization of the magnitude of the bias in the average treatment effect estimator in relation to specificity and sensitivity of a misclassified confounding variable. In this visualization, the prevalence of the confounding variable L is 50% (i.e., $P(L=1)=0.5$), the association between L and treatment (denoted by $\pi_1=P(\text{treatment}=1|L=1)$ and $\pi_0=P(\text{treatment}=1|L=0)$) and outcome is positive (denoted by γ in the text and here set at $\gamma=2$). Given these values, if L is misclassified, the average treatment effect estimated in an MSM-IPW or conditional regression model is biased, independent of true average treatment effect. The magnitude of the bias depends on the specificity and sensitivity of L and is maximal if sensitivity equals $1 - \text{specificity}$. The strength of the association between L and treatment is greater in the right-hand-side plot ($\pi_0=0.25, \pi_1=0.75$) compared with the left-hand-side plot ($\pi_0=0.5, \pi_1=0.75$), and consequently, bias is greater. Larger values of γ will result in steeper curves.

TABLE 1. Values of the Parameters in the Five Different Simulation Scenarios

Scenario	p_0	p_1	λ	π_0	π_1	β	γ
0	0	1	0.50	0.50	0.75	1	2
1	0.05	0.90	0.50	0.90	0.45	1	2
2	0.05	0.90	0.80	0.25	0.75	1	2
3	0.05	0.90	0.80	0.50	0.75	1	2
4	0.05	0.90	0.45	0.50	0.75	1	2

treatment effect in both methods based on the bias expressions (equations 3 and 4).

RESULTS

Table 2 shows the results of the simulation study. Bias found in the simulation study corresponds to the theoretical bias derived from the bias expressions. The empirical standard deviation of the average treatment effect estimates (empSE) from the MSM-IPW is equal to or greater than that from the conditional model. Yet, in the scenarios where bias in the average treatment effect in the MSM-IPW was smaller than bias in the conditional model (scenarios 2 and 3), empSE of both methods was equal, and hence, MSE is smaller for one method if also bias is smaller. Furthermore, the (robust) model-based standard errors of the average treatment effect in an MSM-IPW are conservative and greater than the empirical standard errors, because the uncertainty in estimating the treatment weights is not taken into account. Allowing for the

estimation of the weights will shrink the standard errors.^{2,35} We chose not to use a less conservative standard error estimation for MSM-IPW, such as bootstrapping, because our goal was to frame this simulation as investigating the properties of the commonly used MSM-IPW estimation procedure. Consequently, confidence intervals of the treatment effect obtained in an MSM-IPW are generally wider and coverage of the true treatment effect is higher compared with a conditional model, ranging from overcoverage if there is no classification error to smaller undercoverage when there is classification error.

ILLUSTRATION: QUANTITATIVE BIAS ANALYSIS OF CLASSIFICATION ERROR IN A CONFOUNDING VARIABLE

Quantitative bias analysis provides a tool to incorporate uncertainty in study results due to systematic errors.^{18,20} Using an example study of blood pressure-lowering therapy, we illustrate how the bias expressions (equations 3 and 4) can be used to perform a quantitative bias analysis for misclassification of a confounding variable.

Application

For our illustration we use data of the National Health And Nutritional Examination Survey (NHANES),^{36,37} more details can be found in the Supplement 2; <http://links.lww.com/EDE/B698>. Specifically, we study the effect of diuretic use ($A=1$) in comparison to beta blocker use ($A=0$) on systolic blood pressure (Y) using two approaches: by inverse weighting with the propensity for diuretic or beta blocker use

TABLE 2. Results of Simulation Study Studying the Finite Sample Properties of a marginal structural model estimated using inverse probability weighting (MSM-IPW) and a CM If There Is Classification Error in the Confounding Variable

Method	Sample Size	Scenario ^a	Bias (Formula) ^b	Bias	MSE ^c	Coverage	empSE ^d	modelSE ^e	
MSM-IPW	1,000	0	0.00	0.00 (0.001)	0.00 (0.000)	0.99 (0.001)	0.07 (0.001)	0.10 (0.000)	
		1	-0.42	-0.42 (0.001)	0.18 (0.001)	0.03 (0.002)	0.10 (0.001)	0.11 (0.000)	
		2	0.14	0.14 (0.001)	0.03 (0.000)	0.67 (0.007)	0.08 (0.001)	0.09 (0.000)	
		3	0.29	0.29 (0.001)	0.09 (0.001)	0.08 (0.004)	0.08 (0.001)	0.09 (0.000)	
	100	4	0.15	0.15 (0.001)	0.03 (0.000)	0.68 (0.007)	0.08 (0.001)	0.10 (0.000)	
		0	0.00	0.00 (0.003)	0.05 (0.001)	0.99 (0.001)	0.22 (0.002)	0.31 (0.000)	
		1	-0.42	-0.42 (0.005)	0.29 (0.005)	0.78 (0.006)	0.34 (0.003)	0.37 (0.001)	
		2	0.14	0.14 (0.004)	0.08 (0.002)	0.94 (0.003)	0.25 (0.003)	0.29 (0.000)	
	CM	1,000	3	0.29	0.29 (0.004)	0.15 (0.002)	0.84 (0.005)	0.26 (0.003)	0.28 (0.000)
			4	0.15	0.15 (0.004)	0.08 (0.002)	0.95 (0.003)	0.25 (0.002)	0.31 (0.000)
			0	0.00	0.00 (0.001)	0.00 (0.000)	0.95 (0.003)	0.07 (0.001)	0.07 (0.000)
			1	-0.34	-0.34 (0.001)	0.12 (0.001)	0.02 (0.002)	0.09 (0.001)	0.08 (0.000)
		100	2	0.16	0.16 (0.001)	0.03 (0.000)	0.46 (0.007)	0.08 (0.001)	0.08 (0.000)
			3	0.32	0.32 (0.001)	0.11 (0.001)	0.02 (0.002)	0.08 (0.001)	0.08 (0.000)
			4	0.15	0.15 (0.001)	0.03 (0.000)	0.49 (0.007)	0.08 (0.001)	0.07 (0.000)
			0	0.00	0.00 (0.003)	0.05 (0.001)	0.95 (0.003)	0.22 (0.002)	0.22 (0.000)
	1	-0.34	-0.33 (0.004)	0.19 (0.003)	0.73 (0.006)	0.29 (0.003)	0.27 (0.000)		
	2	0.16	0.16 (0.004)	0.09 (0.002)	0.90 (0.004)	0.25 (0.003)	0.25 (0.000)		
	3	0.32	0.32 (0.004)	0.17 (0.003)	0.74 (0.006)	0.26 (0.003)	0.25 (0.000)		
	4	0.15	0.15 (0.003)	0.08 (0.002)	0.90 (0.004)	0.24 (0.002)	0.24 (0.000)		

^aIn all scenarios, the average treatment effect (estimand) is 1 ($\beta=1$) and the effect of the confounding variable on the outcome is 2 ($\gamma=2$). Five thousand data sets were generated. Monte Carlo standard errors are shown between brackets. In scenario 0, there is no classification error (specificity and sensitivity of the misclassified confounding variable are 1, i.e., $p_0=0$ and $p_1=1$). In scenarios 1–4, the specificity of the misclassified confounding variable is 0.95 (i.e., $p_0=0.05$) and the sensitivity is 0.9 (i.e., $p_1=0.9$). The prevalence of the confounding variable (λ) and the probability of receiving treatment if the confounding is not present or present (π_0 and π_1 , respectively) are set as follows in the scenarios: scenario 0: $\lambda=0.5, \pi_0=0.5, \pi_1=0.75$; scenario 1: $\lambda=0.5, \pi_0=0.9, \pi_1=0.45$; scenario 2: $\lambda=0.8, \pi_0=0.25, \pi_1=0.75$; scenario 3: $\lambda=0.8, \pi_0=0.5, \pi_1=0.75$; and scenario 4: $\lambda=0.45, \pi_0=0.5, \pi_1=0.75$.

^bBias based on bias expressions (equations 3 and 4) in the text.

^cMean squared error.

^dEmpirical standard error.

^eModel-based standard error.

given self-reported categorical body mass index (BMI) (L^*) and using a conditional linear regression with adjustment for self-reported categorical BMI. For this illustration, we categorize self-reported BMI into two distinct categories: underweight/normal weight (BMI < 25 ($L^*=0$)) and overweight/obese (BMI ≥ 25 ($L^*=1$)). However, we stress that one should preferably not categorize BMI in most practical applications.³⁸ Moreover, we assume that dichotomizing self-reported BMI does not introduce differential misclassification.⁷

We assume that blood pressure-lowering therapy is initiated based on the true BMI (L) instead of the observed self-reported BMI (setting 2, Figure 1B). Further, we consider BMI the only confounding variable, and treatment and outcome to be measured without error, which is a simplification of reality. Additionally, we assume that the classification error in self-reported BMI category is nondifferential for the subject's treatment or blood pressure (given true BMI category). Expert knowledge is needed to inform this assumption. To quantify how large the bias in the average treatment effect estimator is expected to be due to classification error in self-reported

TABLE 3. Average Treatment Effect of Diuretics Use Compared with Beta Blocker Use on Mean Systolic Blood Pressure in NHANES^{36,37}

Model	Effect Size (CI)
Unadjusted	-4.03 (-6.30, -1.76)
Marginal structural model ^a	-3.52 (-1.21, -5.74)
Conditional model ^b	-3.48 (-1.27, -5.76)

^aEstimated in a marginal structural model, by inverse weighting with the propensity for diuretic or beta blocker use given self-reported categorized body mass index (BMI).

^bEstimated in a conditional regression model with adjustment for self-reported categorical BMI.

BMI category, we perform a quantitative bias analysis using the bias expressions (equations 3 and 4).

Average Treatment Effect

Table 3 shows the average treatment effect of diuretics use in comparison to beta blocker use on mean systolic blood pressure. In an MSM-IPW, we estimated an average treatment effect (95% confidence interval [CI]) of -3.52 (-1.21, -5.74).

In a conditional regression model, we estimated an average treatment effect (95% CI) of -3.48 ($-1.27, -5.76$).

Quantitative Bias Analysis

To inform the quantitative bias analysis, we need to make assumptions on the sensitivity and specificity of the self-reported BMI and that classification errors are nondifferential with respect to blood pressure and treatment. For the purpose of this illustration, we speculate ranges for the sensitivity and specificity of self-reported BMI category of 0.90 to 0.98. In practice, these parameters should be informed by reports in the literature and/or a researcher's expert experience. Researchers may also decide to investigate how extreme the misclassification (measured using sensitivity and specificity) would need to be to change the conclusions of their study. We refer to the Shiny application (introduced in the subsequent section) for other choices for the sensitivity and specificity of self-reported BMI category.

By uniformly sampling from the range of plausible values of p_0 and p_1 and using the bias expressions (equations 3 and 4), a distribution of possible biases is obtained (eAppendix 2; <http://links.lww.com/EDE/B698> for further details). The solid line in Figure 5 shows the distribution of bias in an MSM-IPW. Mean bias is -0.31 , and median bias is -0.30 (interquartile range, -0.40 to -0.20). We also considered sampling p_0 and p_1 from a trapezoidal (with modes at one third and two thirds between the minimum and maximum)

or a symmetrical triangular distribution. Sampling from these distributions results in mean bias approximately equal to when uniform sampling is applied, but with less spread (dashed and dotted line in Figure 5). This result suggests that the results in Table 3 are not affected much by the classification error in self-reported BMI category. In the NHANES, anthropometric measures were also taken by trained technicians. The average treatment effect when BMI measures taken by trained technicians were used instead of self-reported BMI measures is given in eAppendix 2; <http://links.lww.com/EDE/B698>.

SHINY APPLICATION: AN ONLINE TOOL FOR STUDYING THE IMPACT OF A MISCLASSIFIED VARIABLE

We developed an online tool for creating bias plots (Figures 2–4) and performing quantitative bias analyses (illustrated in the previous section), available at <https://lindanab.shinyapps.io/SensitivityAnalysis>. The bias plots can be used to predict the implications of classification error in a confounding variable in specific study settings by varying the strength of association between the confounding variable and treatment and between the confounding variable and outcome; prevalence of the confounding variable; and specificity and sensitivity of the misclassified confounding variable. The quantitative bias analysis can be used for studying the impact of classification error in a confounding variable at the analysis

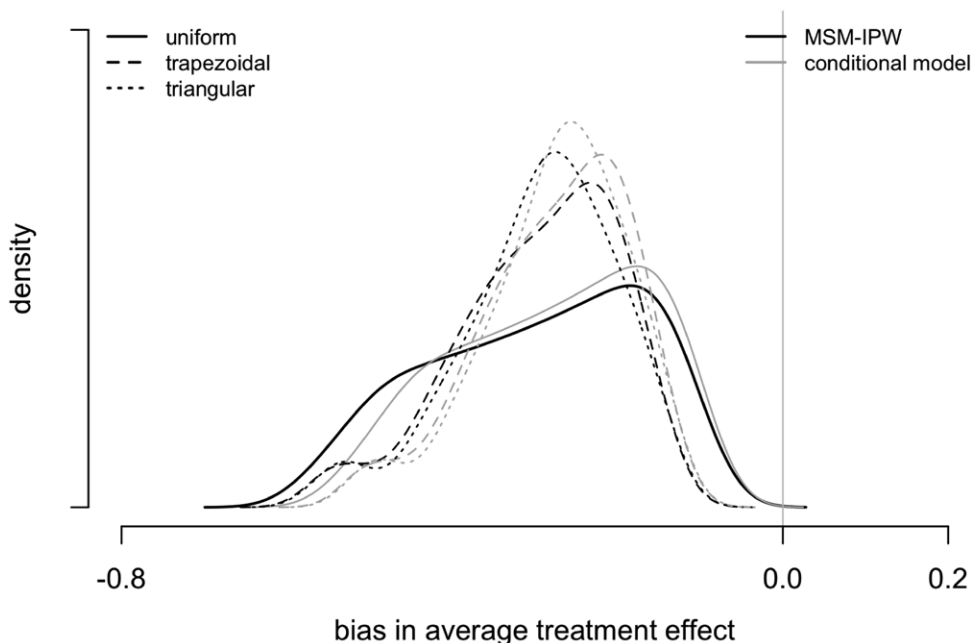


FIGURE 5. Density of predicted bias due to classification error in self-reported BMI category in NHANES.^{36,37} Bias in the average treatment effect of diuretics use compared with beta blocker use on mean systolic blood pressure by inverse weighting with the propensity for diuretic or beta blocker use given self-reported categorical BMI (MSM-IPW), and using a conditional linear regression with adjustment for self-reported categorical BMI. The specificity and sensitivity of self-reported BMI category range from 0.90 to 0.98 and are sampled from a uniform distribution, trapezoidal (with modes on one third and two third), and symmetrical triangular distribution.

stage of a study and to investigate how sensitive conclusions are to the assumption of no classification error. These bias plots can also be used to inform decisions about measurement methods or choice of variables to be extracted in the planning stage of studies.

DISCUSSION

Inverse probability weighting and conditional models are both important and frequently used tools to adjust for confounding variables in observational studies. In this article, we derived expressions for the bias in the average treatment effect in an MSM-IPW and a conditional model. These expressions can inform quantitative bias analyses for bias due to a misclassified confounding variable.

Quantitative bias analysis of misclassified confounding variables is one example of quantitative bias analyses for observational epidemiologic studies. Several approaches exist to assess sensitivity of causal conclusions to unmeasured confounding.^{28,39,40} These aim to quantify the impact of violations of the assumption of no unmeasured confounding, although our approach aims to quantify the impact of violations of the assumption that all confounding variables are measured without error.

Several methods have been proposed to adjust for measurement error in covariates in MSMs-IPW. Pearl⁴¹ developed a general framework for causal inference in the presence of error-prone covariates, which yields weighted estimators in the case of a dichotomous confounding variable measured with error. The framework relies on a joint distribution of the outcome and the confounding variable. Conversely, the weighting method proposed by McCaffrey et al¹⁷ does not require a model for the outcome. Additionally, regression calibration,⁴² simulation-extrapolation,^{43,44} and multiple imputation⁴⁵ have been proposed for correcting for measurement error in covariates of MSMs. These methods assume that the measurement error model is known, which may often be unrealistic. In this context, it is also important to mention previous studies of the impact of measurement error in the exposure or the end point in MSMs, which has been studied by Babanezhad et al⁴⁶ and Shu and Yi,⁴⁷ respectively.

If treatment is allocated based on an error-prone confounding variable, the treatment effect will not be biased (see DAG in Figure 1A). However, investigators should be careful in concluding that covariate measurement error will not affect their analysis. Suppose that there is an unmeasured variable U that acts as a confounding variable between the error-prone covariate L^* and treatment A . Conditioning on L^* will then open a path between A and L via U and thus confound the relation between A and Y .

This article considered classification error in a dichotomous confounding variable in a point-treatment study with a continuous outcome. The same principles apply to measurement error in a categorical or continuous confounding variable or when multiple confounding variables are considered,

although more elaborate assumptions should then be made.⁴⁸ Moreover, we assumed that the relation between exposure and outcome does not vary between strata of the confounding variable, i.e., that there is no treatment effect modification. Future research could extend our bias expressions by relaxing this simplifying assumption, therefore extending our results to more general settings.

MSMs-IPW are increasingly applied to longitudinal data to estimate the joint effects of treatment at multiple time points on a subsequent outcome, including time-dependent outcomes, addressing the problem of time-dependent confounding.^{1,3} There has been little work to understand or correct for the impact of misclassified or mismeasured confounding variables in this more complex setting. Our results extend directly to the time-dependent setting when the aim is to estimate the effect of a current treatment on a time-dependent outcome measured at the next time point.⁴⁹ An area for future work is to extend our results to the setting in which the aim is to estimate the joint effects of treatment at multiple time points and to the time-dependent setting with time-varying treatments and confounding variables. An additional factor to consider in the time-varying setting is the impact of stabilized versus unstabilized weights on the bias if both numerator and denominator of the stabilized weights involve conditioning on an error-prone covariate.

The bias expressions derived in this article can be used to assess bias due to classification error in a dichotomous confounding variable. If classification error in confounding variables is suspected, a quantitative bias analysis provides an opportunity to quantitatively inform readers on the possible impact of such errors on causal conclusions.

REFERENCES

- Hernán MA, Brumback BA, Robins JM. Estimating the causal effect of zidovudine on CD4 count with a marginal structural model for repeated measures. *Stat Med*. 2002;21:1689–1709.
- Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology*. 2000;11:550–560.
- Daniel RM, Cousens SN, De Stavola BL, Kenward MG, Sterne JAC. Methods for dealing with time-dependent confounding. *Stat Med*. 2013;32:1584–1618.
- Rubin DB. For objective causal inference, design trumps analysis. *Ann Appl Stat*. 2008;2:808–840.
- Steiner PM, Cook TD, Shadish WR. On the importance of reliable covariate measurement in selection bias adjustments using propensity scores. *J Educ Behav Stat*. 2011;36:213–236.
- Michels KB. A renaissance for measurement error. *Int J Epidemiol*. 2001;30:421–422.
- van Smeden M, Lash TL, Groenwold RHH. Reflection on modern methods: five myths about measurement error in epidemiological research. *Int J Epidemiol*. 2020;49:338–347.
- Kasza J, Polkinghorne KR, Marshall MR, McDonald SP, Wolfe R. Clustering and residual confounding in the application of marginal structural models: dialysis modality, vascular access, and mortality. *Am J Epidemiol*. 2015;182:535–543.
- Bodnar LM, Davidian M, Siega-Riz AM, Tsiatis AA. Marginal structural models for analyzing causal effects of time-dependent treatments: an application in perinatal epidemiology. *Am J Epidemiol*. 2004;159:926–934.
- Wood ME, Lapane K, Frazier JA, Ystrom E, Mick EO, Nordeng H. Prenatal triptan exposure and internalising and externalising behaviour

- problems in 3-year-old children: results from the Norwegian Mother and Child Cohort Study. *Paediatr Perinat Epidemiol*. 2016;30:190–200.
11. Armstrong BG. Effect of measurement error on epidemiological studies of environmental and occupational exposures. *Occup Environ Med*. 1998;55:651–656.
 12. Buonaccorsi JP. *Measurement Error: Models, Methods, and Applications*. Boca Raton, FL: Chapman & Hall/CRC; 2010.
 13. Carroll RJ, Ruppert D, Stefanski LA, Crainiceanu CM. *Measurement Error in Nonlinear Models: A Modern Perspective*. 2nd ed. Boca Raton, FL: Chapman & Hall/CRC; 2006.
 14. Gustafson P. *Measurement Error and Misclassification in Statistics and Epidemiology: Impacts and Bayesian Adjustments*. Boca Raton, FL: Chapman & Hall/CRC; 2004.
 15. Fuller WA. *Measurement Error Models*. New York, NY: John Wiley & Sons; 1987.
 16. Regier MD, Moodie EE, Platt RW. The effect of error-in-confounders on the estimation of the causal parameter when using marginal structural models and inverse probability-of-treatment weights: a simulation study. *Int J Biostat*. 2014;10:1–15.
 17. McCaffrey DF, Lockwood JR, Setodji CM. Inverse probability weighting with error-prone covariates. *Biometrika*. 2013;100:671–680.
 18. Greenland S. Basic methods for sensitivity analysis of biases. *Int J Epidemiol*. 1996;25:1107–1116.
 19. Lash TL, Fox MP, Fink AK. *Applying Quantitative Bias Analysis to Epidemiologic Data*. New York, NY: Springer; 2009.
 20. Lash TL, Fox MP, Maclehorse RF, Maldonado G, Mccandless LC, Greenland S. Good practices for quantitative bias analysis. *Int J Epidemiol*. 2014;43:1969–1985.
 21. Ali MS, Groenwold RHH, Belitser SV, et al. Methodological comparison of marginal structural model, time-varying Cox regression, and propensity score methods: the example of antidepressant use and the risk of hip fracture. *Pharmacoepidemiol Drug Saf*. 2016;25(suppl. 1):114–121.
 22. Jackson LA, Jackson ML, Nelson JC, Neuzil KM, Weiss NS. Evidence of bias in estimates of influenza vaccine effectiveness in seniors. *Int J Epidemiol*. 2006;35:337–344.
 23. Rubin DB. Estimating causal effects of treatments in randomized and nonrandomized studies. *J Educ Psychol*. 1974;66:688–701.
 24. Rubin DB. Formal modes of statistical inference for causal effects. *J Stat Plan Inference*. 1990;25:279–292.
 25. Hernán MA, Robins JM. *Causal Inference*. forthcoming ed. Boca Raton, FL: Chapman & Hall/CRC; 2019.
 26. Rubin DB. Randomization analysis of experimental data: the Fisher randomization test comment. *J Am Stat Assoc*. 1980;75:591–593.
 27. Cole SR, Hernán MA. Constructing inverse probability weights for marginal structural models. *Am J Epidemiol*. 2008;168:656–664.
 28. VanderWeele TJ, Arah OA. Bias formulas for sensitivity analysis of unmeasured confounding for general outcomes, treatments, and confounders. *Epidemiology*. 2011;22:42–52.
 29. Steiner PM, Kim Y. The mechanics of omitted variable bias: bias amplification and cancellation of offsetting biases. *J Causal Inference*. 2016;4:20160009.
 30. R Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2018. Available at: <https://www.R-project.org/>.
 31. Van Der Wal WM, Geskus RB. ipw: an R package for inverse probability weighting. *J Stat Softw*. 2011;43:1–23.
 32. Lumley T. Analysis of complex survey samples. *J Stat Softw*. 2004;9:1–19.
 33. Morris TP, White IR, Crowther MJ. Using simulation studies to evaluate statistical methods. *Stat Med*. 2019;38:2074–2102.
 34. Gasparini A. rsumsum: summarise results from Monte Carlo simulation studies. *JOSS*. 2018;3:739.
 35. Robins JM. Marginal structural models versus structural nested models as tools for causal inference. In: Halloran EM, Berry D, eds. *Statistical Models in Epidemiology, the Environment, and Clinical Trials*. 1st ed. New York, NY: Springer-Verlag; 2000:95–133.
 36. Centers for Disease Control and Prevention (CDC). National Center for Health Statistics (NCHS). National Health and Nutrition Examination Survey Data [2011–2012]. Hyattsville, MD: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention. Available at: <https://wwwn.cdc.gov/nchs/nhanes/continuousnhanes/default.aspx?BeginYear=2011>. Accessed 5 December 2019.
 37. Centers for Disease Control and Prevention (CDC). National Center for Health Statistics (NCHS). National Health and Nutrition Examination Survey Data [2013–2014]. Hyattsville, MD: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention. Available at: <https://wwwn.cdc.gov/nchs/nhanes/continuousnhanes/default.aspx?BeginYear=2013>. Accessed 5 December 2019.
 38. Altman DG, Royston P. The cost of dichotomising continuous variables. *BMJ*. 2006;332:1080.
 39. Groenwold RHH, Nelson DB, Nichol KL, Hoes AW, Hak E. Sensitivity analyses to estimate the potential impact of unmeasured confounding in causal research. *Int J Epidemiol*. 2010;39:107–117.
 40. Ding P, VanderWeele TJ. Sensitivity analysis without assumptions. *Epidemiology*. 2016;27:368–377.
 41. Pearl J. On measurement bias in causal inference. Technical report. Corvallis, OR: AUA; 2010.
 42. Cole SR, Jacobson LP, Tien PC, Kingsley L, Chmiel JS, Anastos K. Using marginal structural measurement-error models to estimate the long-term effect of antiretroviral therapy on incident AIDS or death. *Am J Epidemiol*. 2010;171:113–122.
 43. Kyle RP, Moodie EE, Klein MB, Abrahamowicz M. Correcting for measurement error in time-varying covariates in marginal structural models. *Am J Epidemiol*. 2016;184:249–258.
 44. Lockwood JR, McCaffrey DF. Simulation-extrapolation for estimating means and causal effects with mismeasured covariates. *Obs Stud*. 2015;1:241–290.
 45. Webb-Vargas Y, Rudolph KE, Lenis D, Murakami P, Stuart EA. An imputation-based solution to using mismeasured covariates in propensity score analysis. *Stat Methods Med Res*. 2017;26:1824–1837.
 46. Babanezhad M, Vansteelandt S, Goetghebeur E. Comparison of causal effect estimators under exposure misclassification. *J Stat Plan Inference*. 2010;140:1306–1319.
 47. Shu D, Yi GY. Causal inference with measurement error in outcomes: bias analysis and estimation methods. *Stat Methods Med Res*. 2017;28:2049–2068.
 48. Keogh RH, White IR. A toolkit for measurement error correction, with a focus on nutritional epidemiology. *Stat Med*. 2014;33:2137–2155.
 49. Keogh RH, Daniel RM, VanderWeele TJ, Vansteelandt S. Analysis of longitudinal studies with repeated outcome measures: adjusting for time-dependent confounding using conventional methods. *Am J Epidemiol*. 2018;187:1085–1092.