Handling missing data, selection bias, and measurement error in observational studies

Choi, J.

Chapter 4

A comparison of different methods for handling measurements affected by medication use

Ready to be submitted

Jungyeon Choi, Olaf M. Dekkers, Saskia le Cessie

Abstract

In epidemiological research, it is common to encounter measurements affected by medication use, such as blood pressure lowered by antihypertensive drugs. When one is interested in the relation between the variables not affected by medication, ignoring medication use can cause bias. Several methods have been proposed, but the problem is often ignored or handled with generic methods, such as excluding individuals on medication or adjusting for medication use in the analysis.

This study aimed to investigate methods for handling measurements affected by medication use when one is interested in the relation between the unaffected variables and to provide guidance for how to handle the problem optimally. We focused on linear regression and distinguished between the situation where the affected measurement is an exposure, confounder, or outcome. In the Netherlands Epidemiology of Obesity study and several simulated settings, we compared generic and more advanced methods; such as substituting or adding a fixed value to the treated values, regression calibration, censored normal regression, Heckman's treatment model, and multiple imputation methods.

For an exposure affected by medication, restricting the analysis to untreated individuals could yield unbiased estimates. Regression calibration is an alternative, but the mean and standard deviation of the medication effect should be known. For an outcome affected by medication, adding the mean medication effect, censored normal regression, and imputation using censored regression worked well. For a confounder affected, selecting untreated individuals worked well, as well as adjusting for medication use, adding mean medication effect, and censored normal regression imputation. In conclusion, methods for handling medication effects should be carefully chosen based on which variable is affected by medication and available information of the clinical setting.

1. Introduction

Measurements affected by medication use are commonly encountered in epidemiological research. Examples are glucose levels lowered by glucose-lowering medications or blood pressure relieved by antihypertensive drugs. Depending on the research questions, these measurements can be an outcome of interest or covariates.

Although researchers often are interested in the effect of certain drugs, the relation between the values not affected by medication can also be the primary scientific interest. However, the value of a variable had an individual not been treated is often not available. Using the values affected by medication instead may lead to biased results. In clinical research, however, medication use is often ignored or handled with naïve methods such as excluding medication users or adjusting for medication use. For outcomes affected by medication use, these naïve methods may introduce bias (1-4).

Several methods have been proposed to handle measurements affected by medication use. Relatively simple methods are adding an expected medication effect to treated values or substituting the treated values for other values (1, 4, 5). More sophisticated methods include censored normal regression, Heckman's treatment model, quantile regression, measurement error methods, or advanced imputation techniques (1, 2, 6, 7). However, these methods are seldom used in applied research. Additionally, many of the suggested methods are limited to outcomes affected by medication, and little has been known about how to handle exposures or confounders affected by medication.

This study aims to investigate methods for handling measurements affected by medication use when the unaffected values are of interest. We focused on etiological studies where effects are estimated by linear regression. We discuss different methods and compare these methods in a large cross-sectional study of the Netherlands Epidemiology of Obesity (NEO) study and several simulation scenarios generated based on the NEO data. The scenarios vary on whether the exposure, confounder, or outcome is affected by medication use. Based on the results of the simulation study, we provide guidance on how to handle the medication effect optimally.

2. Methods to handle measurements affected by medication use

We will consider the situation where for some individuals a variable is affected by medication use (e.g., blood pressure affected by antihypertensive drugs), while the relation between variables when no one is affected by medication is of interest. For convenience, we assume that medication is taken when values are high, aiming to lower

the values. Depending on the research question, the variable affected by medication use can be the exposure, a confounder, or the outcome in an analysis.

The problem of measurements affected by medication can be viewed from different perspectives; it can be viewed as a missing data problem, because for people on medication, their untreated values are unobserved. It may be viewed as a measurement error problem, as the observed values differ systematically from the values had the treated individuals not been treated. It could also be viewed as a censoring problem if we assume that the unobserved untreated values are at least as high as the observed values under treatment. Depending on how one approaches the problem, methods for missing data, measurement error, or censored observations can be used.

Table 1 summarizes methods for handling measurements affected by medication use. The methods can be categorized as generic methods [M1-M5], a method for the exposure affected by medication [M6], methods for the outcome affected by medication [M7-M10], and multiple imputation approaches [M11-M13]. Detailed descriptions of each method and underlying assumptions are available in Appendix 1. All methods are applied to empirical and simulated data in the following sections.

Table 1. Overview of methods for Handling Measurements Affected by Medication use

	Methods	**Description**
Generic methods	**[M1]** Ignoring medication use	Medication use is ignored.
	[M2] Restricting to untreated individuals	The analysis is performed in the subgroup of individuals who are not receiving medication.
	[M3] Binary adjustment for medication use	An indicator for medication use is added as a covariate in the regression model.
	[M4] Substituting measurement of treated individuals with a fixed value	Measurements affected by medication are substituted with a prespecified value.
	[M5] Adding a constant value to observations of treated individuals	A prespecified treatment effect is added to the observed measurements of treated individuals.
For exposures affected by medication	**[M6]** Regression calibration	Measurement error methods are used. Based on the expected mean treatment effect and its standard deviation, the observed measurements affected by medication are corrected.

Table 1. Overview of methods for Handling Measurements Affected by Medication use *(continued)*

	Methods	**Description**
For outcomes affected by medication	**[M7]** Inverse probability weighting	Treated individuals are removed from the analysis, and a reweighted analysis is performed where more weight is given to individuals who are untreated but have a similar profile as treated individuals
	[M8] Quantile regression	The method assumes that the untreated values of individuals on medication would have been above the median, conditional on covariates. The median outcome is modelled as a function of covariates.
	[M9] Censored normal regression	Measurements of treated individuals are considered to be censored observations, where the untreated values are assumed to be at least as high as the observed values affected by treatment, or in more complex censoring mechanisms, at least as high as the observed values and a clinical guideline at which treatment is prescribed.
	[M10] Heckman's treatment model	Treatment assignment is assumed to be dependent on the untreated values, and the treatment results in a "structural shift" of the mean outcome.
Multiple imputation approaches	**[M11]** Predictive mean matching	A default multiple imputation option in commonly used statistical software. It assumes that the observations of treated individuals are missing at random.
	[M12] Censored normal imputation	Censored normal regression is used in the imputation algorithm to predict the untreated values of those on treatment.
	[M13] Heckman's model imputation	Heckman's model is used in the imputation algorithm to predict the untreated values of those on treatment

3. Example: the Netherlands Epidemiology of Obesity Study

The Netherlands Epidemiology of Obesity (NEO) study is a population-based study designed to investigate pathways that lead to obesity-related diseases. From 2008 to 2012, 6,671 individuals aged 45–65 years were included in the study. Participants brought all medication they were using to the NEO study site, which was coded using the Anatomical Therapeutic Chemical Classification (8). Details can be found elsewhere (9). The NEO study data includes several measurements affected by medication; for example, 31% of the participants used antihypertensive medication, and 15% used lipid-lowering medication.

To illustrate the effect of different methods for handling medication use, we use data collected at baseline and consider three research questions:

i) The effect of systolic blood pressure (SBP) on the intima-media thickness (IMT), where the exposure is affected by medication.
ii) The effect of BMI on SBP, where the outcome is affected by medication.
iii) The effect of BMI on IMT, adjusted for SBP, where the confounder is affected by medication.

All methods described in Table 1 were applied to estimate the regression models corresponding to the three research questions stated above. The analyses were adjusted for potential confounders: BMI, sex, age, education level, and smoking status.

In the Netherlands, physicians prescribe blood pressure medication generally aiming at values below 140 mmHg (10). Therefore, we replaced treated SBP values with 150 mmHg in the substitution method [M4] and repeated it using 170 mmHg. For adding medication effect [M5], we followed previous literature using the values 10 mmHg and 15 mmHg (4, 11). For regression calibration [M6], the assumed mean treatment effect was 15 mmHg; SD=10 mmHg.

For inverse probability weighting [M7], logistic regression was used to estimate the probability of medication use based on 21 covariates (see Appendix 2 for details). The same covariates were used in the probit part of Heckman's treatment model [M10] and in the multiple imputation approaches [M11-M13]. For quantile regression [M8], the values of treated individuals were replaced by 150, 170, and 190 mmHg. For censored regression [M9] and imputation [M12], we used 140 mmHg and 160 mmHg as a clinical threshold for treatment prescription. For research questions i) and iii) the outcome variable IMT was added to the imputation models (12). Ten imputed datasets were created in each imputation.

All analyses were performed using R version 3.6.1, with packages Survival v3.1-8 for (13) [M9], SampleSelection v1.2-6 (function treatreg) (14) for [M10], Quantreg v5.54 (15) for [M11], MICE v3.7.0 (16) with default options for [M11] and miceMNAR (17) for [M13]. R code for [M12] is provided in Appendix 2.

Figure 1 presents effect estimates from the different methods for the three research questions. The results show that different methods can lead to quite different effect estimates in all three considered situations. This signals that choosing an appropriate method for handling measurements affected by medication use is essential for the validity of study results.

4. Simulation studies

To understand the results of the NEO study and provide recommendations, we performed several so-called real-life simulation studies. To mimic the NEO study as closely as possible, we used the baseline variables of the NEO participants (BMI, sex, age education, and LDL cholesterol). We simulated SBP, antihypertensive medication prescription, and IMT values based on the other baseline variables directly from the NEO study. We generated different scenarios where blood pressure could be the exposure (scenario 1), the outcome (scenario 2), or the confounder (scenario 3). In each scenario, we considered the research questions i), ii), and iii) of Section 3, respectively.

4.1 Simulation setting 1: Medication effect on the exposure

In this simulation setting, we are interested in the effect of SBP on IMT, with SBP affected by antihypertensive drugs in some individuals. The *untreated SBP* depended linearly on *BMI, sex* (*man*=0, *women*=1), *age,* and *education* (*low*=0, *high*=1), with parameter values closely corresponding to observed values in the NEO study:

$$Untreated\ SBP\ = 90 + 0.8\,BMI - 8.0\,Sex + 0.6\,Age$$

with the residual error ε_{SBP} normally distributed with mean 0 and SD 15.9 mmHg. The probability of receiving medication depended on *BMI*, *sex*, *education,* and the *untreated SBP* values:

$$logit\big(pr(Medication = 1)\big) =$$
$$-16 + 0.01\,BMI - 0.5\,Sex - 0.3\,Education + 0.1\,Untreated\ SBP$$

In this way, approximately 28% of the participants were treated for high SBP. For a SBP of 150 mmHg, the probability of receiving medication was approximately 11%, while for 180 mmHg, the probability was 88%. The *Observed SBP* was lowered when medication was used:

$$Observed\ SBP = Untreated\ SBP - medication\ effect, \quad if\ Medication = 1$$
$$Observed\ SBP = Untreated\ SBP, \quad if\ Medication = 0,$$

where the *medication effect* was generated from a normal distribution (30 mmHg, SD=10 mmHg). The outcome, IMT was generated as:

$$IMT = 31 + 0.2\ Untreated\ SBP + 0.3\ BMI + 2.8\ Sex + 0.4\ Age + 0.8\ Fasting$$

with ε_{IMT} following a normal distribution (0, SD= 9.2 mm). The relation between medication use and IMT is confounded by sex and BMI.

4.2 Simulation setting 2: Medication effect on the outcome

In Simulation setting 2, we consider the effect of BMI on untreated SBP. BMI was taken directly from the NEO data. Untreated SBP, medication prescription, and the observed SBP were generated in the same way as in Simulation setting 1.

4.3 Simulation setting 3: Medication effect on a confounder

Here, we consider the effect of BMI on IMT measurement when adjusted for SBP. Untreated and observed SBP, medication prescription, and IMT were generated as in Simulation setting 1.

4.4 Alternative simulation scenarios

Simulation setting 1, 2, and 3 were repeated while changing three parameters: i) The size of the mean treatment effect decreased from 30 mmHg to 10 mmHg. In this simulation, 16% of the treated individuals' SBP increased after medication. ii) The standard deviation of the treatment effect changed from 10 mmHg to 1 mmHg. iii) The percentage of individuals on medication increased from approximately 28% to 50% by changing the intercept of the logistic model for medication use.

4.5 Analysis

All methods [M1-M13] were applied to the simulated data sets in the same way as described in Section 3, except we used 20 mmHg and 30 mmHg to add to the treated SBP in [M5]. Analyses were adjusted for BMI, sex, age, education level, and smoking status. Each simulation was repeated 1000 times. The estimates obtained from using untreated SBP values were considered a reference. Mean bias and mean squared error were calculated as an overall measure of performance.

ε_{IMT},

5. Results

5.1 Simulation setting 1: Medication effect on the exposure

Figure 2 (left) and Table 2 display the results of simulation setting 1. The results show that medication use cannot be ignored [M1]. Restricting the analysis to untreated individuals [M2] yielded estimates very close to the true values. In this setting, medication use was affected by the exposure and several covariates, in which case one should adjust for all variables both affecting medication use and the outcome to prevent selection bias (18). Furthermore, there was no effect modification, meaning that the effect of SBP on the outcome in the subgroup of untreated individuals is the same as in the total population.

Binary adjustment for medication use [M3] did not work well. In our simulation, the medication effect was generated with large variability. This random variability in medication effect attenuated the association between SBP and IMT in the *treated* individuals and led to a bias toward the null in the overall effect. The method worked better when the variance of the medication effect was smaller (Appendix 3). Substituting treated values [M4] did not perform well in any scenarios. The method cannot reconstruct the original distribution of the exposure and, therefore in general, will yield biased results.

Adding 30 mmHg [M5], which was the true mean medication effect in our simulations, did not perform well either. The reason is that the medication effect was generated with SD=10 mmHg. Therefore, by adding 30 mmHg to all treated SBP values, we reconstruct untreated SBP with random measurement error. Random measurement error in exposures will bias the estimates in a regression model (14). The method performed better when the random variation of the medication effect was smaller (Appendix 3). Regression calibration [M6] yielded unbiased results in all our simulations scenarios, assuming that true medication effect and standard deviation are known.

None of the multiple imputation methods [M11-M13] yielded valid results. A possible explanation is that the imputation models included the outcome, which does not correspond to how medication use was generated in our simulations.

5.2 Simulation setting 2: Medication effect on the outcome

Figure 2 (middle) and Table 2 show the results of Simulation setting 2. Ignoring medication use [M1], restricting to untreated subgroup [M2], and binary adjustment for medication use [M3] yielded biased results. As the outcome determines medication use directly, adjusting or selecting based on medication use [M2 & M3] implies selection based on outcome values, which will generally lead to selection bias (19, 20).

Substituting method [M4] using 150 mmHg led to a large underestimation. It performed better when 170 mmHg was used, which was slightly higher than the mean untreated SBP in the treated individuals (164 mmHg). Regardless of the substituting values, however, the method cannot reconstruct the original distribution of the outcome.

Adding 30 mmHg [M5] yielded unbiased results in all simulation settings (Appendix 4). Unlike in Simulation setting 1, adding the true mean medication effect yields valid results irrespective of the amount of variance in the medication effect.

Inverse probability weighting [M7] resulted in a large bias. Quantile regression [M8] performed poorly for all replacement values. In our simulation setting, more than 50% were using antihypertensive drugs among individuals with very high BMI. Therefore, the median SBP conditional on high BMI was affected by the substituting values.

Censored normal regression [M9] performed reasonably well when the simple censoring method was used or when clinical guideline set to 140 mmHg was applied. However, in alternative scenarios with a smaller medication effect, the results were off (Appendix 4). One reason is that the treated SBP was sometimes higher in these scenarios than the untreated SBP. This violates the assumption that untreated values are at least as high as untreated values (1). Heckman's treatment model [M10] performed less well in our main scenario, which contrasts with the results reported by Spieker et al. (2, 7). Heckman's treatment model assumes that the residual variances of two linear regression models, one for untreated individuals and the other for treated individuals, are equal. This assumption was violated in our main simulation scenario, as we simulated a medication effect with large random variability. This reflects the reported instability of Heckman's treatment (21, 22). In the scenarios with a smaller variance in the medication effect, Heckman's treatment model outperformed the censored regression (Appendix 4).

Multiple imputation with predictive mean matching [M11] resulted in bias. Results of multiple imputation with censored regression [M12] were only slightly biased, but for smaller medication effects, the method performed less well. Multiple imputation with Heckman's model [M13] sometimes yielded a large underestimation of the effect.

5.3 Simulation setting 3: Medication effect in a confounder

Figure 2 (right) and Table 2 show the results of Simulation setting 3. Ignoring the medication effect [M1] resulted in bias. Restricting to untreated individuals [M2] performed well, which is the same as adjusting for confounding by restriction. The method will yield valid estimation under the conditions as in Simulation setting 1, that is, with proper adjustment for variables affecting both medication use and the outcome. Binary adjustment for medication use [M3] yielded results close to the truth. Substitution methods [M4] were biased, because the distribution of untreated SBP could not be correctly reconstructed.

Adding 30 mmHg [M5] yielded a very small upward bias. This is due to the random measurement error introduced by the method. It has been known random measurement error in exposures attenuates the effect, while random measurement in confounders can lead to overestimation (23, 24).

Multiple imputation with censored regression [M12] yielded results close to the truth, especially when clinical guideline information was incorporated, and performed better than multiple imputation with Heckman's model [M13]. All results were consistent in the alternative simulation scenarios (Appendix 5).

Table 2. Mean Coefficient, Standard Deviation, Bias, and Mean Squared Error (MSE) for Three Main Simulation Settings

	Variable affected by medication use:											
	Exposure affected[1]				Outcome affected[2]				Confounder affected[3]			
Methods	**Mean**	**SD**	**Bias**	**MSEx1000**	**Mean**	**SD**	**Bias**	**MSEx10**	**Mean**	**SD**	**Bias**	**MSEx100**
All untreated values known (true coefficient)	0.200	0.006	0.000	0.004	0.800	0.053	0.000	0.028	0.292	0.025	0.000	0.063
Generic methods												
[M1] Ignoring medication use	0.165	0.007	-0.035	0.127	0.483	0.048	-0.317	1.028	0.372	0.025	0.080	0.703
[M2] Restricting to untreated individuals	0.200	0.008	0.000	0.006	0.564	0.054	-0.236	0.586	0.295	0.030	0.003	0.091
[M3] Binary adjustment for medication use	0.182	0.007	-0.018	0.037	0.533	0.048	-0.267	0.736	0.302	0.025	0.010	0.073
[M4] Substituting treated values												
to 150 mmHg	0.223	0.007	0.023	0.058	0.532	0.040	-0.268	0.734	0.333	0.026	0.041	0.236
to 170 mmHg	0.180	0.006	-0.020	0.044	0.743	0.052	-0.057	0.060	0.318	0.025	0.026	0.130
[M5] Adding a constant value												
20 mmHg	0.200	0.006	0.000	0.004	0.694	0.051	-0.106	0.138	0.313	0.025	0.021	0.107
30 mmHg (true value)	0.187	0.006	-0.013	0.021	0.800	0.056	0.000	0.031	0.302	0.025	0.010	0.073
Methods for exposure affected												
[M6] Regression calibration	0.199	0.006	-0.001	0.004	-	-	-	-	-	-	-	-
Methods for outcome affected												
[M7] Inverse probability weighting	-	-	-	-	0.521	0.056	-0.279	0.810	-	-	-	-

Table 2. Mean Coefficient, Standard Deviation, Bias, and Mean Squared Error (MSE) for Three Main Simulation Settings *(continued)*

	Variable affected by medication use:											
	Exposure affected[1]				**Outcome affected[2]**				**Confounder affected[3]**			
[M8] Quantile regression												
k = 150 mmHg	-	-	-	-	0.429	0.037	-0.371	1.390	-	-	-	-
k = 170 mmHg	-	-	-	-	0.948	0.070	0.148	0.268	-	-	-	-
k = 190 mmHg	-	-	-	-	1.052	0.100	0.252	0.735	-	-	-	-
[M9] Censored normal regression												
standard censoring	-	-	-	-	0.749	0.054	-0.051	0.055	-	-	-	-
with guideline at 140 mmHg	-	-	-	-	0.756	0.054	-0.044	0.049	-	-	-	-
with guideline at 160 mmHg	-	-	-	-	0.879	0.063	0.079	0.102	-	-	-	-
[M10] Heckman's treatment model	-	-	-	-	0.660	0.083	-0.140	0.265	-	-	-	-
Multiple imputation methods												
[M11] Predictive mean matching	0.208	0.008	0.008	0.013	0.555	0.056	-0.245	0.632	0.332	0.026	0.040	0.228
[M12] Censored normal regression												
standard censoring	0.221	0.007	0.021	0.049	0.750	0.055	-0.050	0.055	0.286	0.025	-0.006	0.066
with guideline at 140 mmHg	0.212	0.007	0.012	0.019	0.756	0.055	-0.044	0.050	0.291	0.025	-0.001	0.063
[M13] Heckman's model	0.182	0.008	-0.018	0.039	0.737	0.127	-0.063	0.201	0.311	0.030	0.019	0.126

[1]Scenario 1: Effect of systolic blood pressure on IMT measurement. [2] Scenario 2: Effect of BMI on systolic blood pressure. [3] Scenario 3: Effect of BMI on IMT measurement, where systolic blood pressure is one of the confounders. In all scenarios, systolic blood pressure was the variable affected by medication.

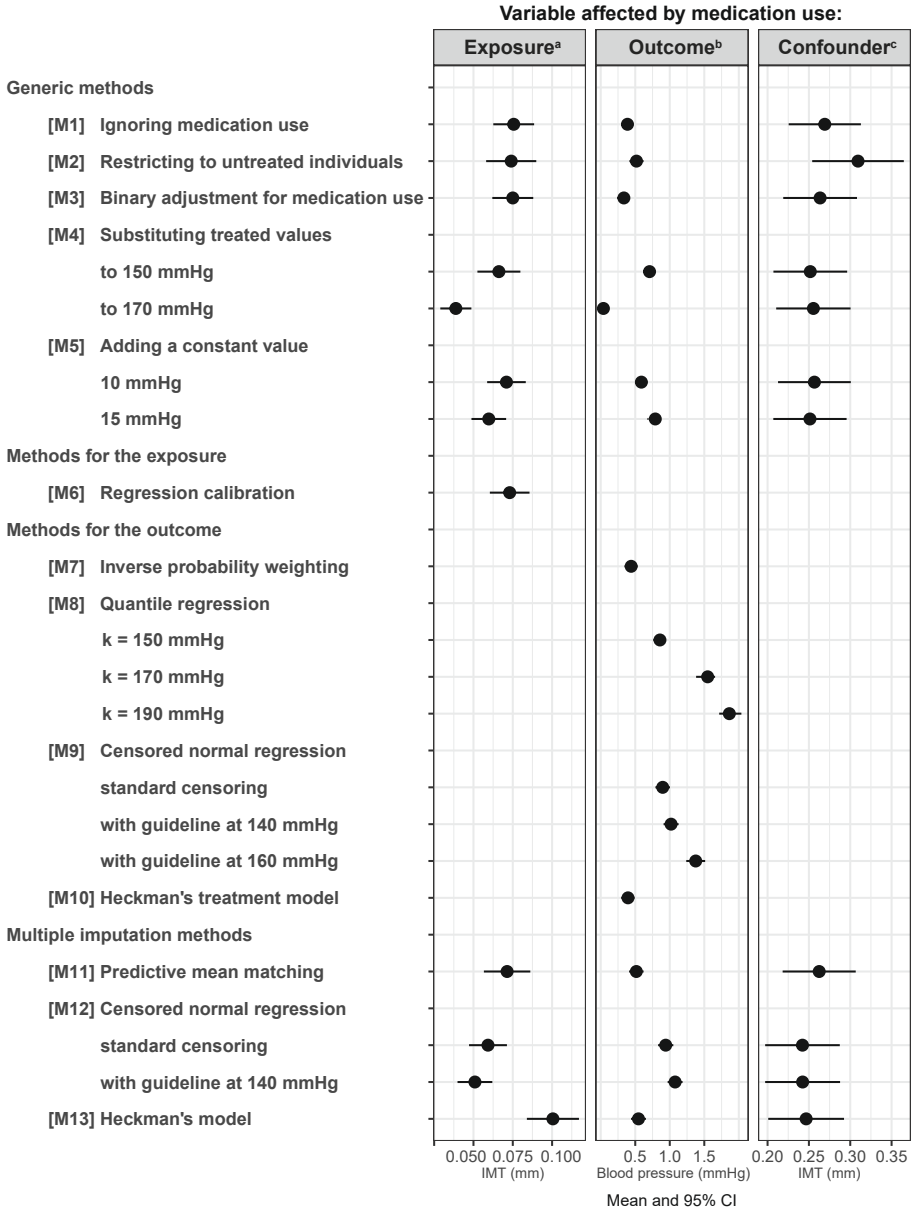


Figure 1. Regression coefficients and their 95% confidence interval estimated from the NEO data using the different methods to handle medication effect. In all analyses, SBP was the variable affected by medication. [a]Question 1: effect of SBP (mmHg) on IMT (mm). [b]Question 2: effect of BMI (kg/m2) on SBP. [C]Question 3: effect of BMI on IMT where SBP is a confounder.

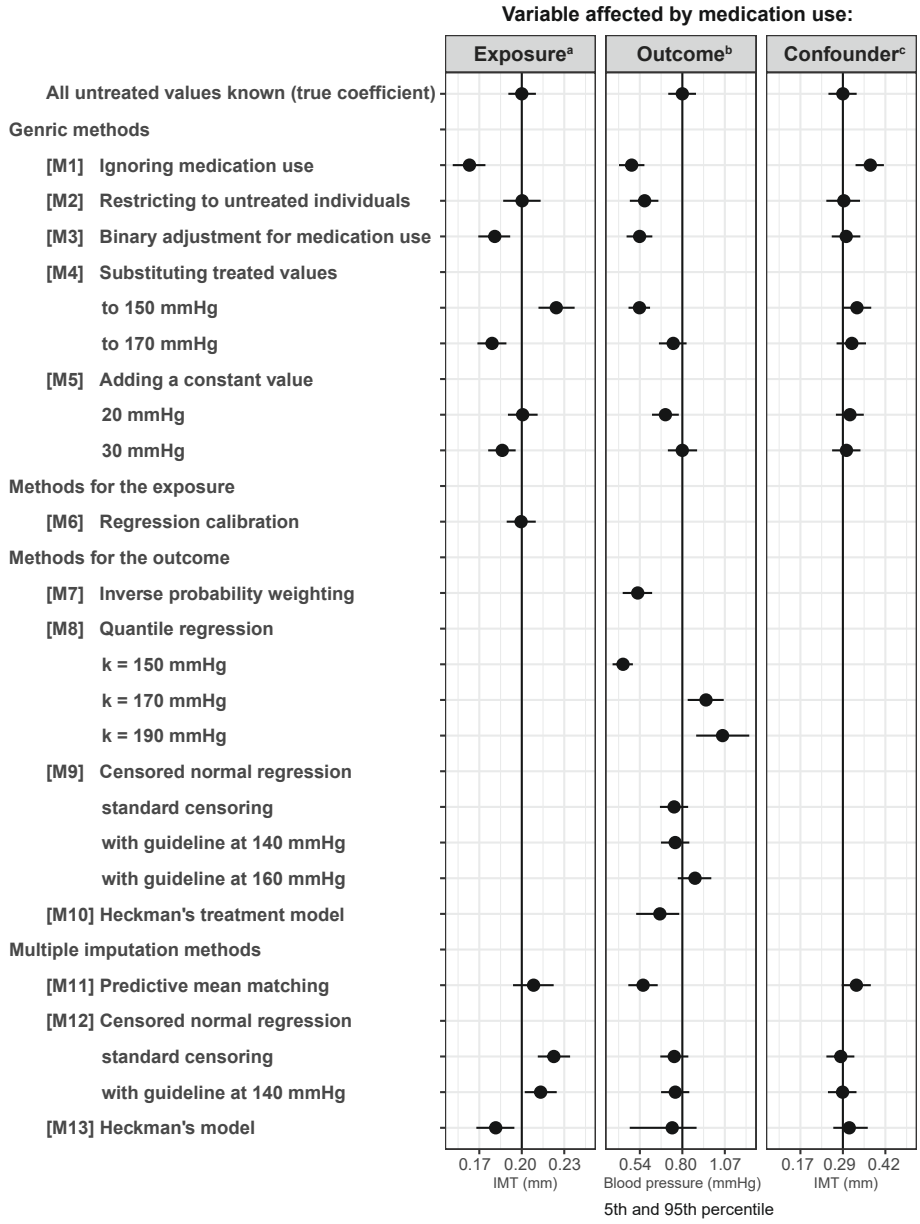


Figure 2. Regression coefficients and their 5th and 95th percentile estimated from simulation setting 1, 2 and 3. Results are standardized based on the mean and standard deviation of the true coefficients in each simulation setting. One grid unit represents 2.5 standard deviation. In all scenarios, SBP was the variable affected by medication use. [a]Question 1: effect of SBP (mmHg) on IMT (mm). [b]Question 2: effect of BMI (kg/m2) on SBP. [c]Question 3: effect of BMI on IMT where SBP is a confounder.

6. Guidance on how to optimally handle measurements affected by medication use

When interest is in the relation between the unaffected variables, ignoring medication use will in general yield biased results regardless of whether the exposure, outcome, or confounder is affected by medication. To obtain valid estimates, adequate methods for handling medication use are needed.

What to do when exposure is affected by medication?

- Performing analysis on the untreated individuals [M2] is a valid approach and will not lead to a large loss in power if the number of treated individuals is relatively low. However, there are two things to consider when applying this method: i) One should adjust for variables that both affect medication use and the outcome. ii) The result cannot be generalized to the total population if the effect of the exposure on the outcome is heterogeneous.
- Regression calibration [M6] may be used but requires an external estimate of the medication effect with its standard deviation.

What to do when the outcome is affected by medication?

- When an estimate of the mean medication effect is available, it could be added to the measurements of treated individuals. This method was also advocated by Tobin et al. (1). Like them, we also highly recommend performing sensitivity analysis with several different values to determine the stability of effect estimates.
- Quantile (median) regression [M8] can be used when less than 50% of the individuals are treated at any value of the exposure. The method does not require knowledge of the medication effect and can yield robust estimates but with lower power (6) than other methods.
- The advantage of censored normal regression [M9] or multiple imputation with censored normal regression [M12] is that no treatment effect needs to be specified. However, the method assumes that the observed values are lower than the untreated values, which could be violated when the treatment is ineffective. Furthermore, the method assumes non-informative censoring, which is likely to be violated in most clinical settings. In our simulation, we relaxed this assumption by incorporating knowledge from a clinical guideline into a censoring mechanism. Both in the study of Tobin et al. and in our main simulation study, the method was rather robust against the violation of the non-informative censoring assumption.
- Heckman's treatment model works well only if the treatment effect has a small variance.

What to do when the confounder is affected by medication?

- Restricting the analysis to untreated individuals [M2] is a valid approach with the same considerations as for the exposure affected by treatment.
- Using a binary indicator [M3] is a reasonable solution.
- Adding the true mean medication effect to the treated individuals [M5] performs relatively well.

7. Discussion

Our simulation study showed that the problem of variables affected by medication use should not be ignored, and proper methods are needed to avoid potential bias. Different methods are needed depending on whether the exposure, the outcome, or a confounder is affected by medication. Additional information, such as medication prescription patterns in clinical settings and the presence of effect heterogeneity, should also be considered carefully. Accordingly, all methods need to be used with caution.

One important consideration is the trade-off between the robustness of a method and the availability of external information. Methods that use external information on the medication effect, such as adding the mean medication effect or regression calibration, performed well when the external information was correct. However, such information is not always available. Other methods, such as censored regression, Heckman's treatment model, or multiple imputation methods, do not require assumptions on the medication effect. However, these methods rely on other assumptions and can perform suboptimally if the assumptions are violated.

We aimed our simulation scenarios to resemble realistic clinical situations instead of creating an ideal scenario for a particular method. Likely, assumptions required for statistical methods will not all be met in clinical data. Therefore, knowing which methods are robust against violation of assumptions is relevant. We encourage researchers to perform real-life simulations more often, as we did when generating simulations based on the NEO study data.

One limitation of our study is that we did not consider situations where more than one variable is affected by medication. Additionally, our study focused on the methods applicable to cross-sectional analyses. Other approaches may be available; for example, when there is an interaction by medication (25), when effect modifiers are associated with medication use (7), in longitudinal settings (26), or in the presence of interaction or mediation by time-varying treatment (27). Furthermore, we focussed on linear regression models, but our recommendations for exposures and confounders will also hold for regression models with a binary or survival outcome.

In summary, the optimal strategy for handling measurements affected by medication depends on whether the medication effect is on the exposure, the outcome, or a confounder. When deciding which strategy to use, we urge researchers to critically consider the processes of medication prescription and what information on medication effects is available.

References

1. Tobin MD, Sheehan NA, Scurrah KJ, et al. Adjusting for treatment effects in studies of quantitative traits: antihypertensive therapy and systolic blood pressure. Statistics in Medicine 2005;24(19):2911-35.
2. Spieker AJ, Delaney JAC, McClelland RL. Evaluating the treatment effects model for estimation of cross-sectional associations between risk factors and cardiovascular biomarkers influenced by medication use. Pharmacoepidemiology and drug safety 2015;24(12):1286-96.
3. Tanamas SK, Hanson RL, Nelson RG, et al. Effect of different methods of accounting for antihypertensive treatment when assessing the relationship between diabetes or obesity and systolic blood pressure. Journal of Diabetes and its Complications 2017;31(4):693-9.
4. Cui JS, Hopper JL, Harrap SBJH. Antihypertensive treatments obscure familial contributions to blood pressure variation. 2003;41(2):207-10.
5. Hunt Steven C, Ellison RC, Atwood Larry D, et al. Genome Scans for Blood Pressure and Hypertension. Hypertension 2002;40(1):1-6.
6. White IR, Koupilova I, Carpenter J. The use of regression models for medians when observed outcomes may be modified by interventions. 2003;22(7):1083-96.
7. Spieker AJ, Delaney JA, McClelland RL. A method to account for covariate-specific treatment effects when estimating biomarker associations in the presence of endogenous medication use. Statistical Methods in Medical Research 2018;27(8):2279-93.
8. NEO codebook baseline measurements. (https://www.lumc.nl/org/neo-studie/researchers/variables/codebook-neo1/medicalhistory/). (Accessed).
9. de Mutsert R, den Heijer M, Rabelink TJ, et al. The Netherlands Epidemiology of Obesity (NEO) study: study design and data collection. European journal of epidemiology 2013;28(6):513-23.
10. NHG-Standaaren. Cardiovasculair risicomanagement (CVRM). 2019. (https://www.nhg.org/standaarden/samenvatting/cardiovasculair-risicomanagement). (Accessed).
11. Neaton JD, Grimm RH, Jr, Prineas RJ, et al. Treatment of Mild Hypertension Study: Final Results. JAMA 1993;270(6):713-24.
12. Moons KGM, Donders RART, Stijnen T, et al. Using the outcome for imputation of missing predictor values was preferred. Journal of Clinical Epidemiology 2006;59(10):1092-101.
13. Therneau TM, Lumley T. Package 'survival'. 2015;128:112.
14. Henningsen A, Toomet O, Henningsen MA, et al. Package 'sampleSelection'. 2019.
15. Koenker R, Portnoy S, Ng PT, et al. Package 'quantreg'. 2019.
16. Buuren Sv, Groothuis-Oudshoorn K. mice: Multivariate imputation by chained equations in R. Journal of statistical software 2010:1-68.
17. Galimard J-E, Chevret S, Protopopescu C, et al. A multiple imputation approach for MNAR mechanisms compatible with Heckman's model. Statistics in Medicine 2016;35(17):2907-20.
18. Hernán M, Robins J. Causal inference: What if. Boca Raton: Chapman & Hall/CRC; 2020.
19. Hernán M, S H-D, Robins J. A Structural Approach to Selection Bias. Epidemiology 2004;15(5):615-25.
20. Cole SR, Platt RW, Schisterman EF, et al. Illustrating bias due to conditioning on a collider. International Journal of Epidemiology 2009;39(2):417-20.

21. Stolzenberg RM, Relles DA. Theory Testing in a World of Constrained Research Design:The Significance of Heckman's Censored Sampling Bias Correction for Nonexperimental Research. 1990;18(4):395-415.

22. Lillard L, Smith JP, Welch F. What Do We Really Know about Wages? The Importance of Nonreporting and Census Imputation. 1986;94(3, Part 1):489-506.

23. White IR. Commentary: Dealing with measurement error: multiple imputation or regression calibration? International Journal of Epidemiology 2006;35(4):1081-2.

24. Brakenhoff TB, van Smeden M, Visseren FLJ, et al. Random measurement error: Why worry? An example of cardiovascular risk factors. PLOS ONE 2018;13(2):e0192298.

25. Masca N, Sheehan NA, Tobin MD. Pharmacogenetic interactions and their potential effects on genetic analyses of blood pressure. Statistics in Medicine 2011;30(7):769-83.

26. McClelland RL, Kronmal RA, Haessler J, et al. Estimation of risk factor associations when the response is influenced by medication use: An imputation approach. Statistics in Medicine 2008;27(24):5039-53.

27. Schmidt AF, Heerspink HJL, Denig P, et al. When drug treatments bias genetic studies: Mediation and interaction. PLOS ONE 2019;14(8):e0221209.

Appendix 1

Detailed description of methods for handling medication effect

We consider the situation where a linear relationship between variables when no one is affected by medication is of interest and a variable is affected by medication use (e.g., blood pressure affected by antihypertensive drugs) for some individuals. For convenience, we assume people take medication when values are high, aiming to lower the values. Depending on the research question, the variable(s) affected by medication use can be the exposure, a confounder, or the outcome in an analysis. The different methods to handle medication use are :

NAÏVE METHODS

M1. ***Ignoring medication use***
Measurements affected by medication are used in the analysis as they are observed.

M2. ***Selecting untreated individuals***
Only the individuals who are not receiving medication are included in the analysis.

M3. ***Adjusting for medication use by adding a binary indicator variable to the regression model***
An indicator for medication use is added as a covariate in the regression model.

M4. ***Substituting measurements of treated individuals with a fixed value***
As Hunt et al. (1) suggested, measurements affected by medication are substituted with a pre-specified value. For example, when guidelines indicate that antihypertensive durgs should be prescribed for blood pressures over 140 mmHg, a value higher than 140 mmHg can be used as a substitution.

M5. ***Adding a constant value to observations of treated individuals***
When the effect of medication on the variable of interest is approximately known, the mean treatment effect can be added to the observed measurements of treated individuals (2, 3). For blood pressure, for example, some authors added 10 mmHg to the systolic blood pressure and 5 mmHg to the diastolic blood pressure when individuals are using antihypertensive medication (4, 5). These values were based on known average treatment effects from a clinical trial (6). However, this is not a set rule and could be adapted

METHODS FOR A MEDICATION EFFECT ON EXPOSURE

M6. ***Regression calibration***
A vast amount of literature addresses measurement error in the covariates of a regression model (7, 8). A simple method is regression calibration, where the untreated

values of the treated individuals (thus, unobserved) replace the measurements affected by medication. The untreated values are estimated by the observed values and other covariates. The method needs an educated guess of the mean *and* standard deviation of the medication effect. These may be obtained from previous clinical trials or observational studies where the effect of treatment is studied.

For individuals on medication, their observed measurement X is replaced by $\lambda(X-\overline{X})+\overline{X}$ + mean medication effect ; with $\overline{X}$, the mean value of X for those using medication and λ, so-called reliability ratio (9). The reliability ratio is equal to $\lambda = 1 - SD(med)^2/SD(X|Z)^2$; with *SD(med)*, the standard deviation of the medication effect and *SD(X|Z)*, the standard deviation of *X* for the medication users adjusted for *Z*, a set of other covariates in the regression model.

METHODS FOR A MEDICATION EFFECT ON THE OUTCOME

M7. **Inverse probability weighting (Sampling weights)**

In this approach, treated individuals are removed from the analysis, and more weight is given to untreated individuals with a similar profile (10, 11). First, the probability of receiving medication for each individual is estimated by logistic regression. Then the untreated individuals are weighted by 1/(1-probability to receive medication). This creates a pseudo-population with the same characteristics as the original population but where no one is treated.

M8. ***Quantile regression***

White et al. (12) proposed to use quantile regression for outcomes affected by medication use. In this approach, the median outcome is modeled as a function of covariates. The method assumes the untreated values would have been above the median conditional on covariates for individuals on medication. The treated individuals' outcome values are replaced by *k,* that is, any value higher than the conditional median, after which a median regression model can be fitted.

M9. ***Censored normal regression***

An alternative approach is to use methods for censored outcomes (2, 3), such as censored normal regression, which assumes a normal underlying distribution of the untreated outcome. This method is also known as tobit regression. Measurements of treated individuals are considered to be censored observations, where the untreated values are assumed to be at least as high as the observed values affected by treatment. An advantage of this method is that no assumptions on the treatment effect size are needed. However, non-informative censoring is assumed. The non-informative censoring implies that conditional on covariates, the probability of receiving treatment does not depend on the untreated values. This assumption is likely to be invalid, as individuals with higher values are more likely to be treated. Previous simulations showed good

performance in realistic scenarios (2). However, recent literature showed that the method performed poorly under certain scenarios (13).

More complex censoring mechanisms can also be used to resemble realistic clinical settings. For example, when a clinical guideline suggests starting treatment for values above a certain threshold δ this information can be incorporated. In this case, the untreated values are assumed to be higher than the observed measurements *and* higher than the threshold. That is, for the treated observations, we assume that:

$$\begin{cases} Untreated\ value\ \geq \delta, & if\ observed\ value < \delta \\ Untreated\ value\ \geq observed\ value, & if\ observed\ value \geq \delta \end{cases}$$

The threshold value of δ is obtained using external knowledge of the clinical setting.

M10. ***Heckman's treatment effects model***

Heckman's treatment effects model originates from economics and can account for non-random sample selection (13-15). Spieker et al. (13, 15) used this model for handling outcomes affected by medication use. This model assumes that treatment assignment depends on the untreated values where higher values are more likely to be treated and treatment results in a "structural shift" of the mean outcome. In the standard treatment effect model, this treatment effect does not depend on covariates (13), but it is possible to extend this model to incorporate effect modification (15).

Technically, the method assumes that there is an unobserved latent variable that determines treatment. If its value is above 0, treatment is prescribed. The latent variable is correlated with the original untreated values, so people with higher untreated values are more likely to be treated. Parameters are estimated by joint modeling of i) a linear regression model for the effect of exposure on the untreated blood pressure, ii) the same linear regression model for the effect of exposure on the treated blood pressure, with a lower constant term which reflects the effect of treatment and iii) a probit model for the probability of medication prescription (13, 15). Both the linear regression model and the probit model may depend on other covariates.

MULTIPLE IMPUTATION APPROACHES

Untreated values of individuals on treatment can be considered missing, and multiple imputation methods can be used to handle these missing values. The method can be applied in many different ways under different assumptions. We considered three multiple imputation approaches that are based on various assumptions.

M11. ***Multiple imputation with predictive mean matching via a linear regression model***

For a numerical variable with missing values, the default multiple imputation option is chained equations with predictive mean matching via a linear regression model with

the main effects of the covariates. This imputation method is readily available in many standard statistical software packages. Note that the method assumes that the data is missing at random.

M12. ***Multiple imputation with censored normal regression***

Instead of using linear regression as imputation model, censored normal regression may be used to predict missing values (16). This may be done under the different censoring mechanisms we discussed in [M9]. A regular censored normal regression can only be used when medication effect is on the outcome. However, multiple imputation with censored normal regression does not have this restriction.

M13. ***Multiple imputation with Heckman's model***

Galimard et al. developed an imputation approach for missing not at random data using Heckman's model (17). Again, the multiple imputation approach can be used regardless of whether the outcome, exposure, or a confounder is affected.

References

1. Hunt Steven C, Ellison RC, Atwood Larry D, et al. Genome Scans for Blood Pressure and Hypertension. Hypertension 2002;40(1):1-6.
2. Tobin MD, Sheehan NA, Scurrah KJ, et al. Adjusting for treatment effects in studies of quantitative traits: antihypertensive therapy and systolic blood pressure. Statistics in Medicine 2005;24(19):2911-35.
3. Masca N, Sheehan NA, Tobin MD. Pharmacogenetic interactions and their potential effects on genetic analyses of blood pressure. Statistics in Medicine 2011;30(7):769-83.
4. Cui J, Hopper JL, Harrap SBJH. Genes and family environment explain correlations between blood pressure and body mass index. 2002;40(1):7-12.
5. Cui JS, Hopper JL, Harrap SBJH. Antihypertensive treatments obscure familial contributions to blood pressure variation. 2003;41(2):207-10.
6. Neaton JD, Grimm RH, Jr, Prineas RJ, et al. Treatment of Mild Hypertension Study: Final Results. JAMA 1993;270(6):713-24.
7. Carroll RJ. Measurement Error in Epidemiologic Studies. Wiley StatsRef: Statistics Reference Online, 2014.
8. Buonaccorsi JP. Measurement error: models, methods, and applications. Chapman and Hall/CRC; 2010.
9. Carroll RJ, Ruppert D, Stefanski LA, et al. Measurement error in nonlinear models: a modern perspective. Chapman and Hall/CRC; 2006.
10. Seaman S, White I. Inverse Probability Weighting with Missing Predictors of Treatment Assignment or Missingness. Communications in Statistics - Theory and Methods 2014;43(16):3499-515.
11. Hernán M, Robins J. Causal Inference. Harvard TH Chan School of Public Health (July 27, 2019).
12. White IR, Koupilova I, Carpenter J. The use of regression models for medians when observed outcomes may be modified by interventions. 2003;22(7):1083-96.

13. Spieker AJ, Delaney JAC, McClelland RL. Evaluating the treatment effects model for estimation of cross-sectional associations between risk factors and cardiovascular biomarkers influenced by medication use. Pharmacoepidemiology and drug safety 2015;24(12):1286-96.
14. Certo ST, Busenbark JR, Woo H-s, et al. Sample selection bias and Heckman models in strategic management research. Strategic Management Journal 2016;37(13):2639-57.
15. Spieker AJ, Delaney JA, McClelland RL. A method to account for covariate-specific treatment effects when estimating biomarker associations in the presence of endogenous medication use. Statistical Methods in Medical Research 2018;27(8):2279-93.
16. Gartner H, Rässler S. Analyzing the changing gender wage gap based on multiply imputed right censored wages. 2005.
17. Galimard J-E, Chevret S, Protopopescu C, et al. A multiple imputation approach for MNAR mechanisms compatible with Heckman's model. Statistics in Medicine 2016;35(17):2907-20.

Appendix 2

Detailed application of the methods

1) Covariates used for inverse probability weighting [M7] and the imputation methods [M11-M13]

Sex, age, BMI, total body fat, waist circumference, hip circumference, education level, income, smoking status, ethnicity, alcohol intake, the total amount of leisure, glucose, insulin, glycated hemoglobulin A1C, triglycerides, HDL cholesterol, LDL cholesterol and medication use for glucose, lipid, and depression.

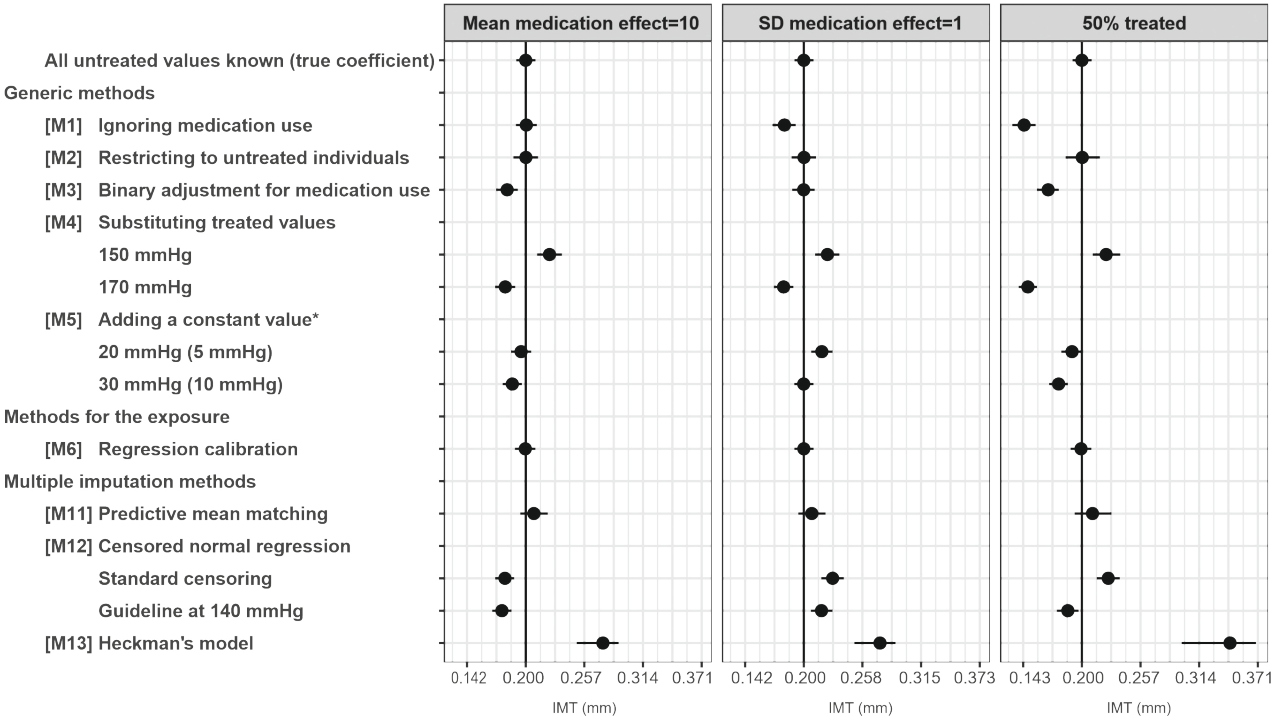
2) R syntax for censored normal imputation [M12]

```
# Function to draw from a truncated normal distribution, range lwb-upb
rnorm.trunc <- function(n,mean,sd, low=-Inf, upp=Inf)
{U <- runif(n,0,1)
qnorm(pnorm(low, mean = mean, sd = sd)+
        (pnorm(upp, mean = mean, sd = sd)-pnorm(low, mean = mean, sd = sd))*U, mean = mean, sd
= sd)
}

# impute censored normal
mice.impute.censnorm <-
  function (y, ry, x, wy = NULL,ycens, ...)
  {
    #1 prepare data
    wy <- !ry # wy= TRUE indicates that value should be imputed
    x <-  as.matrix(x)
    m <- ncol(x)+1

    # 2. estimate coefficients censored model
    fit <- survreg(Surv(ycens, ry) ~ x, dist='gaussian')
    beta <- coefficients(fit)
    sigma <- fit$scale
    #    print(fit)

    #3. generate new beta and sigma for bayesian drawings
    df <- max(length(y[ry]) - ncol(x), 1)
    rv <- t(chol((vcov(fit)[1:m,1:m])))
    beta.star <- beta + rv %*% rnorm(ncol(rv))
    sigma.star <- sqrt(df*sigma^2/rchisq(1, df))

    #4. Draw new observations
    mean.star <- cbind(1,x[wy, , drop = FALSE]) %*% beta.star
    vec<- rnorm.trunc(nrow(mean.star),mean.star,sd=sigma.star, low=ycens[wy])
    return(vec)
  }
```

Appendix 3

Regression coefficients and 5th and 95th percentile estimated from modified scenarios of Simulation setting 1, where we estimated the effect of SBP on IMT measurement. One grid unit represents 2.5 standard deviation of the true coefficient, estimated from 1000 simulation runs. *We added 5 mmHg and 10 mmHg for the scenario where mean medication effect was 10 mmHg,

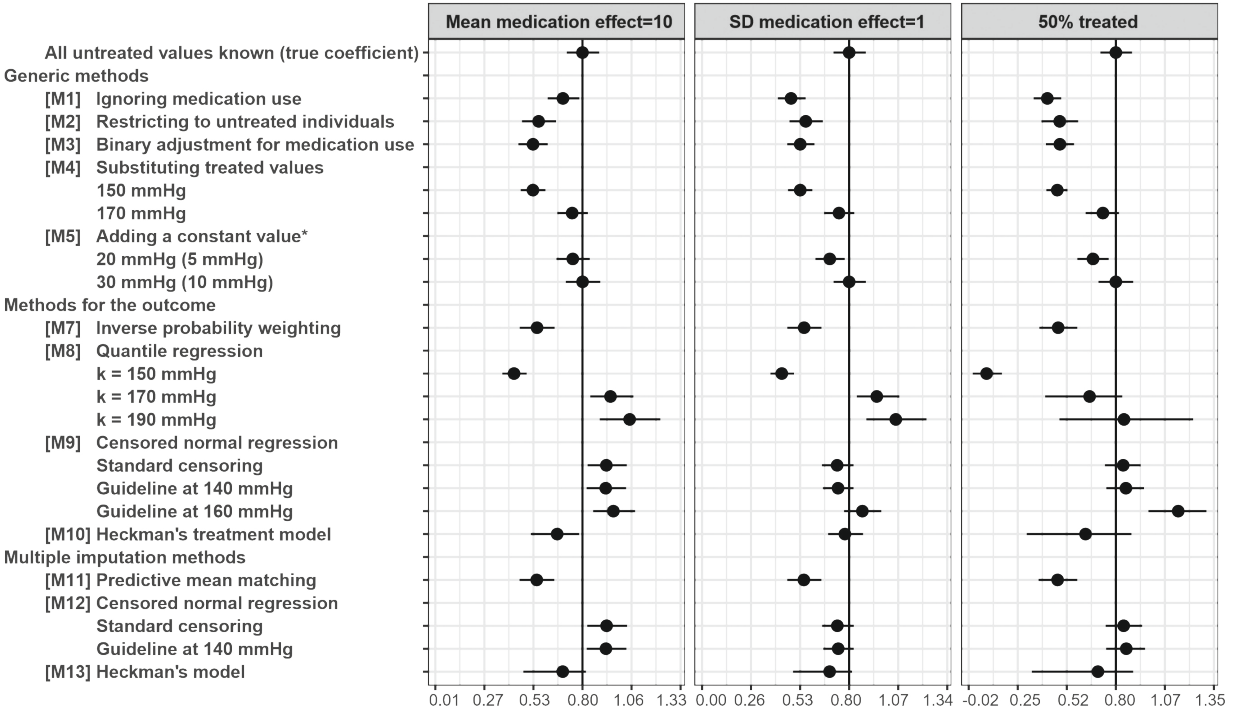
Exposure affected by medication (modified Simulation setting 1)



Appendix 4

Regression coefficients and 5th and 95th percentile estimated from modified scenarios of Simulation setting 2, where we estimated the effect of BMI on SBP. One grid unit represents 2.5 standard deviation of the true coefficient, estimated from 1000 simulation runs. *We added 5 mmHg and 10 mmHg for the scenario where mean medication effect was 10 mmHg,

Outcome affected by medication (modifed Simulation setting 2)



Appendix 5

Regression coefficients and 5th and 95th percentile estimated from modified scenarios of simulation setting 3, where we estimated the effect of BMI on IMT while SBP is one of the confounders. One grid unit represents 2.5 standard deviation of the true coefficient, estimated from 1000 simulation runs. *We added 5 mmHg and 10 mmHg for the scenario where mean medication effect was 10 mmHg,

Confounder affected by medication (modifed Simulation setting 3)