Causal mediation analysis in the context of clinical research

Zhongheng Zhang\textsuperscript{1,2}, Cheng Zheng\textsuperscript{3}, Chanmin Kim\textsuperscript{4}, Sven Van Poucke\textsuperscript{5}, Su Lin\textsuperscript{6}, Peng Lan\textsuperscript{7}

\textsuperscript{1}Department of Emergency Medicine, Sir Run-Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou 310016, China; \textsuperscript{2}Jinhua Hospital of Zhejiang University, Jinhua 321000, China; \textsuperscript{3}Joseph. J. Zilber School of Public Health, University of Wisconsin-Milwaukee, Milwaukee, Wisconsin, USA; \textsuperscript{4}Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA 02115, USA; \textsuperscript{5}Department of Anesthesiology, Emergency Medicine, Critical Care Medicine and Pain Therapy. Ziekenhuis Oost-Limburg, Genk, Belgium; \textsuperscript{6}Liver Research Center of the First Affiliated Hospital of Fujian Medical University, Fuzhou, China; \textsuperscript{7}Department of Critical Care Medicine, Sir Run-Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou 310016, China

Correspondence to: Zhongheng Zhang, MMed. No 3, East Qingchun Road, Hangzhou 310016, China. Email: zh_zhang1984@hotmail.com.

Abstract: Clinical researches usually collected numerous intermediate variables besides treatment and outcome. These variables are often incorrectly treated as confounding factors and are thus controlled using a variety of multivariable regression models depending on the types of outcome variable. However, these methods fail to disentangle underlying mediating processes. Causal mediation analysis (CMA) is a method to dissect total effect of a treatment into direct and indirect effect. The indirect effect is transmitted via mediator to the outcome. The mediation package is designed to perform CMA under the assumption of sequential ignorability. It reports average causal mediation effect (ACME), average direct effect (ADE) and total effect. Also, the package provides visualization tool for these estimated effects. Sensitivity analysis is designed to examine whether the results are robust to the violation of the sequential ignorability assumption since the assumption has been criticized to be too strong to be satisfied in research practice.

Keywords: Causal mediation analysis (CMA); mediator; clinical research; sensitivity analysis

Submitted Aug 02, 2016. Accepted for publication Sep 25, 2016.
doi: 10.21037/atm.2016.11.11
View this article at: http://dx.doi.org/10.21037/atm.2016.11.11

Introduction

In clinical research, a large number of variables may be collected and the relationship between each variable can be complex. The traditional method to explore the relationships between explanatory variables and outcomes is by using multivariable regression model. However, this method fails to disentangle mechanisms underlying the association between explanatory variable and outcome, while investigators are sometimes interested in such underlying mediating pathphysiological processes (1,2). In these situations, the investigators may have prior knowledge that an explanatory variable exerts its effect on outcome via direct and indirect pathways. In the indirect pathway, there is a mediator that transmits the causal effect. For example, suppose that corticosteroids are effective for reducing mortality rate of patients with acute respiratory distress syndrome (ARDS), and this effect is partly mediated by suppressing inflammatory response (3). In this scenario, corticosteroid treatment is an exposure, mortality is a binary outcome and inflammatory response is a mediator. It is interesting to explore how much of the total effect is transmitted via inflammatory response. The example motivates the causal mediation analysis (CMA). In this article, we will introduce a commonly used method for CMA proposed by Imai et al. (4). Its implementation with R package will be described in a step-by-step approach.

Understanding CMA

Suppose we have variables X and Y indicating the explanatory variable (or treatment variable) and outcome variable, respectively (Figure 1). Mediation in its simplest form is to add a mediator M between X and Y (5). Model-
based mediation analysis is implemented in the mediation package (6). The sequential ignorability assumption should be satisfied for validity of the method used for this package. The assumption states that “the treatment (explanatory variable X) is first assumed to be ignorable given the pre-treatment covariates, and then the mediator variable (M) is assumed to be ignorable given the observed value of the treatment as well as the pre-treatment covariates.” (7). The first part is often satisfied by randomization while the second part implies that there are no unmeasured confounding between mediator and outcome, which is often questionable even in randomized trial. Moreover, this assumption is not verifiable in practice and sensitivity analysis should be performed to quantify the degree to which violation of the assumption would change the results. Sensitivity analysis will be described in detail in following sections.

This section focuses on the model based mediation analysis with continuous outcome, which is based on three linear equations:

\[ Y = i_1 + cX + e_1 \]  
\[ Y = i_2 + c'X + bM + e_2 \]  
\[ M = i_3 + aX + e_3 \]

where \( i_1, i_2 \) and \( i_3 \) denote intercepts, Y is the outcome variable, X is the explanatory variable, M is the mediator, \( c \) is coefficient linking X and Y (total causal effect), \( c' \) is the coefficient for the effect of X on Y adjusting for M (direct effect), \( b \) is the effect of M on Y adjusting for explanatory variable, \( a \) is the coefficient relating to the effect of X on M. \( e_1, e_2 \) and \( e_3 \) are residuals (7) that are uncorrelated with the variables in the right hand side of the equation and are independent to each other. Under this specific model, the causal mediation effect (CME) is represented by the product coefficient of \( ab \). Of note, Eq. [3] can be substituted into Eq. [2] to eliminate the term M:

\[ Y = i_2 + bi_3 + (c' + ab)X + e_2 + be_3 \]  

It appears that the parameters related to direct (\( c' \)) and indirect effect (\( ab \)) of X on Y is different from that of its total effect. That is, to testing the null hypothesis \( c=0 \) is unnecessary since CME can be nonzero even when the total causal effect is zero (i.e., direct and indirect effects can be opposite) (4,7), which reflects the effect cancellation from different pathways.

However, the above-mentioned equations within linear structural equation modeling (LSEM) framework are not directly applicable to binary outcome. LSEM equation needs extension to accommodate non-linear settings (8,9). In this article we will focus on Imai’s method that is applicable to non-linear settings (4).

**Worked example**

Next, a dataset is simulated using the motivating example described in the introduction section. Suppose investigators want to explore the effect of corticosteroids on mortality outcome, and it is valuable to notice if part of the total effect is mediated via reducing inflammatory response. C-reactive protein (CRP) is a biomarker of inflammatory response and there is evidence that higher CRP is associated with adverse clinical outcomes (10). Note that the example is created to demonstrate the use of CMA into the clinical context, and there is no clinical relevance of the results.

```r
> set.seed(888)
> treat_flg<-rbinom(1000,1,0.3) #binary treatment variable with binomial distribution
> crp<-round(abs(40*treat_flg+rnorm(1000,100,30)),1)
#regress crp on treatment
> lp<-10*treat_flg+0.02*crp-5 # linear prediction
> link_lp <- exp(lp)/(1 + exp(lp)) #link function
> mort<-(runif(1000) < link_lp)
```

Figure 1 Causal mediation analysis in its simplest form. X, Y and M are treatment, outcome and mediator variables. \( c \) is coefficient linking X and Y (total causal effect), \( c' \) is the coefficient for the effect of X on Y adjusting for M (direct effect), \( b \) is the effect of M on Y adjusting for explanatory variable, \( a \) is the coefficient relating to the effect of X on M. \( e_1, e_2 \) and \( e_3 \) are residuals.
CMA

This article implements CMA using the R mediation package (6). This package also contains functions for CMA with certain designs or conditions such as multiple outcome/treatment/mediator combinations, multiple causal mechanisms, treatment non-compliance, and dataset with missing values. In addition, this package allows researchers to implement sensitivity analysis for certain parametric models. In the following example, the authors only applied the simplest form of CMA. Let’s first install the package and then perform CMA with a few lines of codes.

```R
> install.packages("mediation")
> library(mediation)
> model.m <- lm(crp~treat_flg,data=df[1:100,])
> model.y <- glm(mort~treat_flg+crp,family = binomial,data=df[1:100,])
> med.out <- mediate(model.m, model.y, treat = "treat_flg", mediator = "crp",robustSE = TRUE, sims = 100)
```

The mediate() function requires two fitted model objects: one is for the mediator and the other is for the outcome. In the example, the mediator model is a linear model that regresses `crp` on `treat_flg`. The outcome model is a logistic regression model that regresses `mort` on `treat_flg` and `crp`. Note that these two regression models allow other confounding factors to be added if they are observed. In the example, only the first 100 observations are used for analysis. The `treat` argument takes a character string indicating the name of the treatment variable (`treat_flg`) used in the models. Similarly, the `mediator` argument takes a character string indicating the name of the mediator variable (`crp`) used in the model.

Understanding summary output with binary outcome variable

The results of CMA can be accessed with summary() and plot() function.

```R
> summary(med.out)
```

Causal Mediation Analysis
Quasi-Bayesian Confidence Intervals

The results show the ACME and average direct effect (ADE) for the treated and control groups. For a single index, researchers can use the average ACME and ADE to determine the average mediation strength. Note that the ACMEs are different for the treated and control groups. This can be understood in the counterfactual framework. Let’s first define ACME with mathematical equation according to Imai and colleagues (7):

\[
\delta(t) = Y_i[t, M_i(1)] - Y_i[t, M_i(0)]
\]

\(\delta(t)\) is the CME under treatment \(t\), which represents the indirect effect of treatment on outcome \(Y(t)\) through mediator. The subscript \(i\) indicates the individual patient. The counterfactual framework can be understood with the following question: what changes would occur to the outcome if the mediator is changed from the value under treatment \(t=1\) \([M_i(1)]\) to the value that would be observed under treatment \(t=0\) \([M_i(0)]\), while holding treatment at \(t\). While the outcome of the form \(Y_i[t, M(t)]\) is observable, \(Y_i[t, M(t-1)]\) is unobservable for a given patient. In the example, CME (treated, \(\delta(1)\)) is the difference between two potential mortality outcomes for patient \(i\) who received corticosteroids. For the patient, \(Y_i[1, M_i(1)]\) is observed outcome if he or she receives corticosteroids, whereas \(Y_i[1, M_i(0)]\) represents the mortality outcome under the condition that the patient still receives corticosteroids but change the mediator value that would result without treatment. This definition with
counterfactual framework reflects the intuitive notion of mediation that the treatment indirectly influences outcome via mediator. Similarly, direct effect can be defined as:

$$\zeta_i(t) \equiv Y_i[1, Mi(t)] - Y_i[0, Mi(t)]$$  \[6\]

In the example, $\zeta_i(1)$ represents the direct effect of corticosteroids on patient $i$’s mortality outcome while holding the $crp$ at the level that would be observed under treatment. The causal effect of the treatment on the outcome for unit $i$ now can be defined as the following relationship

$$Total = \delta_i(t) - \zeta_i(1-t)$$  \[7\]

While CME is at the individual level, the ACME is the expected value for the whole study population:

$$\bar{\zeta}(t) = IE \{Y_i[t, Mi(t)] - Y_i[t, Mi(0)]\}$$  \[8\]

It appears that ACMEs of both the treated and the control are statistically non-significant. This may be due to the limited sample size. Recall that we only used the first 100 patients for analysis. The results can be visualized with plot() function.

> plot(med.out)

The simple function automatically produces a plot of mediation analysis. Estimates for both treated and control group were depicted because they are different in the example. The dashed line represents the control and solid line represents the treated. It appears that most of the total effect is explained by the direct effect (Figure 2).

The next question that often confuses people is the interpretation of coefficients. In normal linear model, all variables are linked in the same scale and the interpretation is straightforward. However, in CMA with binary outcome variables, the estimated coefficients from above links to the probability scale, rather than that obtained from logistic regression model that its exponentiation gives the odds ratio (OR). Next, let's take a look at how estimation of mediation effects can be different using different methods: the standard LSEM approach proposed by Baron (11) and the method by Imai.

Imai’s method consists of two steps which are based on Eqs. [5] and [6]. The first step is to fit regression models for mediator and outcome, which has been done in the above R codes. The second is to simulate the unobserved potential mediator and then compute the average potential outcome from the fitted model.

> ###average over different simulation runs to obtain mediation effects, direct effects and total effect.
> ###M is simulation times, this is sims=100 in Imai’s method, when M is larger, the result is less sensitive to the seed use.
> M<-100
> acme0<-acme1<-rep(NA,M)
> total<-rep(NA,M)
> direct1<-direct0<-rep(NA,M)
> for (i in 1:M){
> ###impute potential mediators
> m0<-predict(model.m,data.frame(treat_flg=0))+summary(model.m)$sigma*rnorm(100,0,1)
> m1<-predict(model.m,data.frame(treat_flg=1))+summary(model.m)$sigma*rnorm(100,0,1)
> ###impute potential outcomes
> y0_m0<-predict(model.y.data.frame(treat_flg=0,crp=m0),type="response")
> y0_m1<-predict(model.y.data.frame(treat_flg=0,crp=m1),type="response")
> y1_m0<-predict(model.y.data.frame(treat_flg=1,crp=m0),type="response")
> y1_m1<-predict(model.y.data.frame(treat_flg=1,crp=m1),type="response")
>

![Figure 2](mediate.jpg)

**Figure 2** Visualization of results from mediate() function. Estimates for both treated and control group were depicted because they are different in the example. The dashed line represents the control and solid line represents the treated. It appears that most of the total effect is explained by the direct effect.
> y1_m1<-predict(model.y,data.frame(treat_flg=1,crp=m1),type="response")
cacme0[i]<-mean(y0_m1-y0_m0)
cacme1[i]<-mean(y1_m1-y1_m0)
total[i]<-mean(y1_m1-y0_m0)
direct1[i]<-mean(y1_m1-y0_m1)
direct0[i]<-mean(y1_m0-y0_m0)
}

> ###point estimate for ACME0, ACME1 compare to
mediation package output
> c(mean(acme0),med.out$d0)
[1] 0.1929078 0.1842675
> c(mean(acme1),med.out$d1)
[1] 7.032967e-09 1.131077e-08

> ###point estimate for total effect, compare to
mediation package output
> c(mean(total),med.out$tau.coef)
[1] 0.9447887 0.9396576
> ###point estimate for direct effects, compare to
mediation package output
> c(mean(direct0),med.out$z0)
[1] 0.9447887 0.9396576
> c(mean(direct1),med.out$z1)
[1] 0.7518808 0.7553901

The value 0.193 approximate the ACME(0) reported in tabular output by summary() function. A small difference is that we perform limited number of simulations. If we compute acme0 with our code 1,000 times and take the average, the value would be more approximate. The coefficients estimated with mediate() function can be better understood with above explicit computation. Because the type argument takes a string variable “response” within the predict() function, y0_m0 is the probability of death for a given patient in the control group and a crp level being observed. y0_m1 is the probability of death for a given patient in the control group and a crp level that would be observed supposing the patient had received treatment. Thus, y0_m1 is unobservable in reality but only in counterfactual framework. The difference between y0_m0 and y0_m1 gives the ACME in the control group that can be interpreted as the probability of death being changed by varying crp values, while holding the treatment status constant. The total effect of 0.94 can be interpreted as the increased probability of death comparing treatment versus control group. The model coefficients estimated from Imai’s nonparametric method have a direct link to the probability scale and thus is more clinically relevant.

Now we will show that if we ignore the fact that some of the equations are fit with nonlinear link and still use the formula derived from linear setting, i.e., the total effect \( c = c' + a*b \), we will not be able to yield correct result. Eqs. [2] with logit link and [3] have been fitted and stored in objects model.y and model.m, respectively. We only have to fit the Eq. [1] with logit link and store it in an object called model.t.

> model.t<-glm(mort~treat_flg,family = binomial,data=df[1:100,])
> c<-model.t$coef[[2]]
> c.prime<-model.y$coef[[2]]
> b<-model.y$coef[[3]]
> a<-model.m$coef[[2]]
> c(c,c.prime+a*b)
[1] 23.38447 24.45443

The results show that the total effect estimated from regression model.t approximate that estimated from regression models.y and model.m. For nonlinear case, the two quantities might approximately the same under some special cases, for example, weak indirect effect (a*b), log link function or rare outcome. Our example is the one with weak indirect effect. However, the values of these estimated coefficients are not equal to that obtained via mediate() function, because the later coefficients are identified with methods as described above (12). The coefficients estimated from logistic regression model are in the linear prediction function scale that its exponentiation gives the odd ratio (OR).

> exp(c(a*b,c.prime,c))

[1] 7.082856e+00 5.891355e+09 1.431347e+10

The total OR is the sum of direct effect OR and indirect effect OR. The results show that there is negligible indirect effect, and the direct effect takes nearly all of the total effect.

**Sensitivity analysis**

CMA described above is performed under sequential
ignorability assumption. However, this assumption is untestable in practice. One approach is to perform sensitivity analysis to examine whether the results are robust to the violation of the assumption. Violation to assumption indicates that there exists an unmeasured confounder that is related to both the mediator and the outcome. Note that this unmeasured confounder is not affected by the treatment. Sensitivity analysis uses certain statistics to quantify how strong the confounder would have to be to change the conclusion being drawn about the direct and indirect effect (13). Imai and colleagues proposed a correlation parameter ($\rho$) reflecting the existence of omitted variables that were related to the mediator and outcome even after conditioning on treatment, and the parameter was added to the calculations of ACME. Sensitivity analysis is to vary $\rho$ values and compute corresponding ACMEs. There are two methods to calculating $\rho$. One is to compute $\rho$ based on residual correlation:

$$\rho \equiv \text{cor}(e_2, e_1)$$

The other way is to calculate correlation parameter $\rho$ based on coefficients of determination (i.e., a number that indicates the proportion of the variance in the dependent variable that is predictable from the independent variable):

$$\rho = \frac{\text{sgn}(\lambda_2 \lambda_3) R^2_M R^2_Y}{\sqrt{(1 - R^2_M)(1 - R^2_Y)}}$$

$$e_i = \lambda_i U + e'_i$$

$$e_i = \lambda_i U + e'_i$$

where $R^2_M$ and $R^2_Y$ represent the proportion of previously unexplained variance (mediator and outcome) that is explained by the unobserved confounders. $\overline{R}^2_M$ and $\overline{R}^2_Y$ are the proportion of variance that is explained by the unobserved confounders. U is unobserved confounding variable. $R^2_M$ and $R^2_Y$ are the coefficients of determination for the mediation and outcome regression models. $\lambda_2$ and $\lambda_3$ are coefficients for unobserved variable in the equation regressing residual $e_i$ or $e_1$ on the unobserved variable. We only care about the direction (sign) of the effect. In this framework, investigators can specify values of $(R^2_M, R^2_Y)$ and $(\overline{R}^2_M, \overline{R}^2_Y)$, and the sign of $(\lambda_2, \lambda_3)$ to estimate ACME (4,14).

Next I proceed to perform sensitivity analysis with medsens() function. Because the function in its current version only support probit link, the mediation model needs to be updated.
$\rho$ values for the treatment and control groups. Two tables are similar in interpretation, so we only take a look at the first table (control group). The first column is the $\rho$ values with an increment of 0.1. The second column shows the corresponding ACMEs, and the third and forth columns are confidence intervals. The last two columns are quantities based on coefficients of determination. Because the $\rho$ value at which ACME becomes zero is of interest, it is displayed following the table. A graphical display of the results can be more intuitive and helpful in sensitivity analysis.

```r
> par(mfrow=c(2,1))
> plot(sens.out, sens.par = "rho", main = "Mortality")
```

The `sens.par` argument specifies which sensitivity parameter, residual correlation or coefficients of determination, to be displayed. The dashed line represents the ACME value without correlation ($\rho=0$), which is computed under the assumption of sequential ignorability. We need to set $\rho \geq 0.3$ for the ACME to become negative (Figure 3). A large critical $\rho$ value is needed to reverse the sign of ACME indicates the robustness of the result to the violation of ignorability assumption. It requires subject knowledge to judge whether the value of 0.3 is large or not. In practice, observed confounding variables are often used to determine the upper bound of sensitivity parameters.

Sensitivity analysis based on $R^2$ statistic is performed by assign “R2” to `sens.par`. The type of $R^2$ to be used is specified in `r.type` argument. A string variable “residual” indicates that the effects are plotted against the proportions of the residual variances that are explained by the unobserved confounder. On the other hand, a value of “total” indicates that the proportions of the total variances are used as sensitivity parameters. When sensitivity analysis is performed with $R^2$ statistic, the user must specify whether the unobserved confounder affects the mediator ($\lambda_3$) and outcome ($\lambda_2$) in the same direction (sgn[\lambda_2\lambda_3]=1) or opposite direction (sgn[\lambda_2\lambda_3]=-1). If the effects are in the same direction, users assign `sign.prod = “positive”`. Otherwise, the `sign.prod` argument takes the string variable “negative”.

```r
> par(mfrow=c(2,2))
> plot(sens.out, sens.par = "rho", main = "Mortality")
```

**Figure 3** Sensitivity analysis based on residual correlation. The dashed line represents the ACME (0.0735) without correlation ($\rho=0$), which is computed under the assumption of sequential ignorability. Note that $\rho \geq 0.2$ is required for the ACME to become negative.
Four plots are produced by varying `r.type` and `sign.prod` arguments within `plot()` function. Note that only the first plot that has displayed a series of ACME values. Each of the other plots has displayed the contour of one ACME value, because the native R contour function automatically adjusts the displaying values.

```r
> plot(sens.out, sens.par = "R2", r.type = "residual", sign.prod = "negative")
> plot(sens.out, sens.par = "R2", r.type = "total", sign.prod = "positive")
```

> `sens.out$d0`

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.913575804</td>
<td>0.661963822</td>
<td>0.453954736</td>
</tr>
<tr>
<td>4</td>
<td>0.314137898</td>
<td>0.220546426</td>
<td>0.156108478</td>
</tr>
<tr>
<td>7</td>
<td>0.110405639</td>
<td>0.077178510</td>
<td>0.052570140</td>
</tr>
<tr>
<td>10</td>
<td>0.034124673</td>
<td>0.020227810</td>
<td>0.009787342</td>
</tr>
<tr>
<td>13</td>
<td>0.002044243</td>
<td>-0.003544247</td>
<td>-0.007382268</td>
</tr>
<tr>
<td>16</td>
<td>-0.009793831</td>
<td>-0.011076604</td>
<td>-0.011559293</td>
</tr>
<tr>
<td>19</td>
<td>-0.011636731</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

> `sens.out$d1`

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.545749e-01</td>
<td>9.531345e-02</td>
<td>5.024507e-02</td>
</tr>
<tr>
<td>4</td>
<td>3.087918e-02</td>
<td>2.067147e-02</td>
<td>1.455696e-02</td>
</tr>
<tr>
<td>7</td>
<td>1.054603e-02</td>
<td>7.727958e-03</td>
<td>5.639027e-03</td>
</tr>
</tbody>
</table>

Figure 4 Sensitivity analysis based on coefficient of determination. Four plots are produced by varying `r.type` and `sign.prod` arguments within `plot()` function. Note that only the first plot that has displayed a series of ACME values. Each of the other plots has displayed the contour of one ACME value, because the native R contour function automatically adjusts the displaying values.
Figure 5 Sensitivity analysis based on coefficient of determination. Series of ACME values are displayed by adjusting levels argument in the plot() function.

\[ \text{ACME}(R_{M}^{2}, R_{Y}^{2}), \text{sgn}(\lambda_2 \lambda_3) = -1 \]

\[ \text{ACME}(R_{M}^{2}, R_{Y}^{2}), \text{sgn}(\lambda_2 \lambda_3) = 1 \]

Note that the ACME values (i.e., the first 9 values of the sens.out$d0 output) for the control group when \( \text{sgn}(\lambda_2 \lambda_3) = -1 \) is increased at a step approximating 0.1 and thus can be displayed without specifying levels argument. However, the ACME values (i.e., the first 9 values of the sens.out$d1 output) for the treated group when \( \text{sgn}(\lambda_2 \lambda_3) = -1 \) is changed at steps significantly different from 0.1, thus we need to specify the levels for displaying more contours.

\[ \text{> par(mfrow=c(2,2))} \]

The bold line represents various combinations of \( R^2 \) statistics where ACME takes the value zero (Figure 5). In the example, the value of ACME under sequential ignorability assumption is not statistically significant, thus the degree of confounding where ACME would be zero is of limited interest. In other situations, where ACME is statistically significant, the critical point of the degree of confounding where ACME takes zero is of great interest because beyond this point the effect size is in opposite direction (i.e., the
sign is changed). However, there is no cutoff value for the statistics of confounding (i.e., ρ and combinations of R² statistics) to judge the robustness of CMA results estimated under ignorability assumption.

Acknowledgements
None.

Footnote
Conflicts of Interest: The authors have no conflicts of interest to declare.

References