



# HHS Public Access

Author manuscript

*Epidemiology*. Author manuscript; available in PMC 2024 September 01.

Published in final edited form as:

*Epidemiology*. 2023 September 01; 34(5): 661–672. doi:10.1097/EDE.0000000000001643.

## Effect Measure Modification by Covariates in Mediation: Extending Regression-Based Causal Mediation Analysis

Yi Li<sup>1,2</sup>, Maya B. Mathur<sup>3</sup>, Daniel H. Solomon<sup>2,4,5</sup>, Paul M. Ridker<sup>4,6,7</sup>, Robert J. Glynn<sup>4,5,6,8</sup>, Kazuki Yoshida<sup>2,4,9</sup>

<sup>1</sup>Department of Epidemiology, Biostatistics and Occupational Health, School of Population and Global Health, Faculty of Medicine, McGill University, Montreal, QC, Canada.

<sup>2</sup>Division of Rheumatology, Inflammation, and Immunity, Department of Medicine, Brigham and Women's Hospital, Boston, MA, USA.

<sup>3</sup>Quantitative Science Unit, Department of Medicine, Stanford University, Palo Alto, CA, USA.

<sup>4</sup>Department of Medicine, Harvard Medical School, Boston, MA, USA.

<sup>5</sup>Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital, Boston, MA, USA.

<sup>6</sup>Division of Preventive Medicine, Department of Medicine, Brigham and Women's Hospital, Boston, MA, USA.

<sup>7</sup>Division of Cardiovascular Medicine, Department of Medicine, Brigham and Women's Hospital, Boston, MA, USA.

<sup>8</sup>Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, USA.

<sup>9</sup>OM1, Inc. MA, USA

### Abstract

Existing methods for regression-based mediation analysis assume that the exposure–mediator effect, exposure–outcome effect, and mediator–outcome effect are constant across levels of the baseline characteristics of patients. However, investigators often have insight into how these underlying effects may be modified by baseline characteristics and are interested in how the resulting mediation effects, such as the natural direct effect (NDE), the natural indirect effect

---

**CORRESPONDENCE:** Yi Li, MS. Department of Epidemiology, Biostatistics and Occupational Health, School of Population and Global Health, Faculty of Medicine, McGill University, 2001 McGill College Avenue, Montreal, QC H3A 1G1 (yi.li10@mail.mcgill.ca).

#### DISCLOSURES:

YL has no disclosures.

MBM has no disclosures.

DHS has received research support from Abbvie, Amgen, CorEvitas, and Genentech for unrelated work. As well, he receives royalties from UpToDate on unrelated work.

PMR has received research grant support from Novartis, Kowa, Amarin, Pfizer, and the National Heart, Lung, and Blood Institute; served as a consultant to Corvidia, Novartis, Novo Nordisk, Flame, Agepha, Inflazome, AstraZeneca, Janssen, Civi Biopharm, SOCAR, and Omeicos; and has a financial interest in Upton.

RJG has received funding from grants to the Brigham and Women's Hospital from AstraZeneca, Kowa, Novartis Pharmaceuticals Corporation, and Pfizer.

KY has received consulting fees from OM1, Inc. for unrelated work.

(NIE), and the proportion mediated, are modified by these baseline characteristics. Motivated by an empirical example of anti-interleukin-1 therapy's benefit on incident anemia reduction and its mediation by an early change in an inflammatory biomarker, we extended the closed-form regression-based causal mediation analysis with effect measure modification (EMM). Using a simulated numerical example, we demonstrated that naive analysis without considering EMM can give biased estimates of NDE and NIE and visually illustrate how baseline characteristics affect the presence and magnitude of EMM of NDE and NIE. We then applied the extended method to the empirical example informed by pathophysiologic insights into potential EMM by age, diabetes, and baseline inflammation. We found that the proportion modified through the early post-treatment inflammatory biomarker was greater for younger, non-diabetic patients with lower baseline level of inflammation, suggesting differential usefulness of the early post-treatment inflammatory biomarker in monitoring patients depending on baseline characteristics. To facilitate the adoption of EMM considerations in causal mediation analysis by the wider clinical and epidemiologic research communities, we developed a free and open-source R package, *regmedint*.

### Keywords

causal mediation; effect measure modification; interaction; direct effect; indirect effect

## INTRODUCTION

Causal mediation analysis has been gaining popularity in the last decade. Several estimation methods have been proposed to date<sup>1-6</sup> that share a foundation in the counterfactual potential outcome framework, though these methods differ in their estimation approaches and resulting interpretations. One of the most popular approaches is closed-form regression-based causal mediation analysis, introduced in a landmark paper by Valeri and VanderWeele with accompanying SAS and SPSS macros<sup>5,7</sup>. This approach involves fitting simple parametric models for the mediator and for the outcome. Closed-form variance estimation by delta method is computationally straightforward, compared to other major estimation approaches that require simulation or weighted data construction<sup>1,4,8</sup>.

In this article, we develop generalized methods for mediation analysis that allow the inclusion of effect measure modification (EMM). Existing methods and software for regression-based mediation analysis assume that the exposure–mediator effect, the exposure–outcome effect, and the mediator–outcome effect are constant across levels of the controlled covariates. This is a limitation because researchers may be interested in investigating how mediation varies across these covariates. For example, in the Canakinumab Anti-Inflammatory Thrombosis Outcome Study (CANTOS) trial, canakinumab (anti-interleukin-1 [IL-1] medication) demonstrated beneficial estimated effects on cardiovascular disease as well as other outcomes<sup>9,10</sup>, including incident anemia<sup>11</sup>. Causal mediation analysis has been conducted for some of the outcomes<sup>11,12</sup> to examine the mediating role of early post-treatment change in high-sensitivity C-reactive protein (hsCRP) biomarker, which relates to the activity of the IL-1 biologic pathway targeted by canakinumab. Moderate to high levels of mediation were observed confirming the relevance of the early post-treatment changes in the IL-1 pathway and the usefulness of the

hsCRP biomarker in assessing canakinumab benefits on subsequent clinical outcomes<sup>11,12</sup>. The total benefit of canakinumab on incident anemia appeared somewhat greater among older individuals, diabetic individuals, and individuals with higher baseline hsCRP<sup>11</sup>. However, to our knowledge the corresponding EMM of mediation by these variables was not previously investigated. The framework and software developed in this paper allow for formal investigation of potentially varying magnitudes of mediation through the early suppression of the IL-1 pathway across different baseline patient profiles.

The rest of the article is structured as follows. First, we extend the regression-based causal mediation approach to allow for the inclusion of EMM terms. Second, we develop closed-form formulas for the natural direct effect (NDE) and natural indirect effect (NIE) under four model specifications: (1) linear mediator model, linear outcome model; (2) logistic mediator model, linear outcome model; (3) linear mediator model, non-linear outcome model; (4) logistic mediator model, non-linear outcome model. We also describe settings when these conditional estimates: (a) coincide with marginal estimates, such that they can be interpreted as population average effects; and (b) are constant across levels of the covariates. Third, we use numerical examples to illustrate the EMM of the NDE and NIE. Fourth, we investigate the consequences of model misspecification in which the mediator and outcome models omit product terms when in fact EMM or causal interaction are present. Last, we revisit the above empirical example from the CANTOS trial to examine how EMM can impact NDE and NIE, and proportion mediated.

## DEFINITIONS, NOTATIONS, AND ASSUMPTIONS

We use the same notational convention as in Valeri and VanderWeele (2013)<sup>5</sup>. We denote the exposure as  $A$ , the mediator as  $M$ , the outcome as  $Y$ , and the vector of covariates as  $C$ . We assume the causal structure represented in the directed acyclic graph (DAG) in Figure 1. The effects of interest are the total effect (TE), NDE and NIE, as defined below<sup>4,5,7,10</sup>:

$$TE(\mathbf{c}) = g(E[Y_a | C = \mathbf{c}]) - g(E[Y_{a^*} | C = \mathbf{c}]),$$

$$NDE(\mathbf{c}) = g(E[Y_{aM_{a^*}} | C = \mathbf{c}]) - g(E[Y_{a^*M_{a^*}} | C = \mathbf{c}]),$$

$$NIE(\mathbf{c}) = g(E[Y_{aM_a} | C = \mathbf{c}]) - g(E[Y_{aM_{a^*}} | C = \mathbf{c}]),$$

where  $g$  is a general link function. We assume the same four identification assumptions as in Valeri and VanderWeele (2013)<sup>5</sup>. Detailed definitions and assumptions are in Section 1 of the eSupplement.

### EMM-Extended Closed-Form Regression-Based Causal Mediation Analysis Approach

To accommodate EMM by  $C$  in both the mediator and outcome models, we specify the following extended parametric models that include product terms for EMM (underlined):

$$M1. g_M[E(M | A = a, C = c)] = \beta_0 + \beta_1 a + \beta_2 c + \beta_3 ac,$$

$$M2. g_Y[E(Y | A = a, M = m, C = c)] = \theta_0 + \theta_1 a + \theta_2 m + \theta_3 am + \theta_4 c + \theta_5 ac + \theta_6 mc,$$

where  $g_M$  and  $g_Y$  are the link functions and  $\beta_2$ ,  $\beta_3$ ,  $\theta_4$ ,  $\theta_5$  and  $\theta_6$  are row vectors. In this article, we focus on identity and logit link functions, because the resulting formulas for these types of models can accommodate a wide range of outcome models, including survival outcomes.<sup>5-7</sup> Note that logistic and Cox proportional hazards model require rare outcome<sup>25</sup>. Details are in the Supplement (Section 2.1).

The extended models include several product terms. Product terms can be used to model either causal interaction or EMM<sup>14</sup>, but there are critical differences in the interpretation and assumptions. Causal interaction entails a joint intervention on the two variables, while EMM entails heterogeneity in the causal effect of one variable across subgroups of the other<sup>14</sup>. Stronger assumptions regarding control of confounding will be required to identify causal interaction, compared to EMM<sup>14</sup>. Under the assumptions required for mediation analysis (A1-A4), the product term between the exposure and mediator represents causal interaction, but additional assumptions would be required for exposure-covariate or mediator-covariate product terms to represent causal interactions rather than mere EMM<sup>5,7,8,13</sup>. Throughout this paper, we use the same identification assumptions as are required for mediation analysis<sup>5,7,8,13</sup>, and so we describe product terms involving covariates as representing EMM rather than causal interaction.

## EMM OF NATURAL DIRECT AND INDIRECT EFFECTS BY COVARIATES

Based on Models M1 and M2, we derive closed-form formulas for NDE and NIE (Table 1). When  $\beta_3 = \theta_5 = \theta_6 = 0$  (*i.e.*, no EMM product terms are included in either the mediator or the outcome models), the formulas reduce to the original formulas in VV2013<sup>5</sup>. The extended formulas in Table 1 are implemented in the R *regmedint* package (version 1.0.0 and later)<sup>15-17</sup>.

Proportion mediated, which most studies are interested in reporting, is defined as  $\frac{NIE}{NDE + NIE}$  when the outcome model is linear and  $\frac{\exp(NDE) \cdot [\exp(NIE) - 1]}{\exp(NDE) \cdot \exp(NIE) - 1}$  when the outcome model is non-linear. Derivations of point estimates and standard errors by the delta method are given in the eSupplement Section 3.

Due to the way that mediator and outcome models are constructed (eSupplement Section 2.1), these formulas provide conditional effect estimates of NDE and NIE. That is, these natural effect formulas contain the covariate vector  $\mathbf{C}$  in their expressions, indicating that these natural effects are potentially modified by covariates. It is tempting to interpret the conditional effect estimates evaluated at the mean covariate level as the corresponding population average (marginal) effect estimates. However, this marginal interpretation does not generally hold. Most of the time, numerical integration is necessary to obtain marginal

effects from the conditional mediator and outcome models<sup>1–3,8</sup>. With the extended formulas in Table 1, we first examine the model specifications under which conditional natural effect estimates evaluated at the mean covariate levels coincide with the marginal effect estimates. Further, we discuss model specifications under which conditional effect estimates are constant.

### **When do conditional NDE and NIE estimates at the mean covariate levels coincide with the marginal estimates?**

Conditional NDE and NIE estimates evaluated at the mean covariate levels coincide with corresponding marginal estimates when their effect formulas (Table 1) are linear in  $C$ . From Table 1, we see that the outcome model must be linear, because a non-linear outcome model cannot give NDE or NIE estimates that are linear in  $C$  due to the exponentiation, regardless of the type of the mediator model.

Table 2 summarizes, for different specifications of the outcome and mediator models, the conditions under which the NDE or NIE evaluated at the mean covariate level will have a marginal interpretation. If the outcome model is linear, then the NDE estimate evaluated at the mean covariate level will have a marginal interpretation if the mediator model is also linear. However, if the mediator model is logistic rather than linear, then the conditional NDE estimate will not have a marginal interpretation unless  $\theta_3 = 0$  (no exposure–mediator causal interaction  $[A \times M]$ ). Again for a linear outcome model, the NIE estimate evaluated at the mean covariate level may not have a marginal interpretation unless (1) the mediator model is also linear and either  $\beta_3 = 0$  (no EMM of the exposure effect by covariates  $[A \times C]$  on mediator) or  $\theta_6 = 0$  (no modification of the mediator effect by covariates  $[M \times C]$  on outcome); or (2) the mediator model is logistic and  $\beta_2 = \beta_3 = 0$  (no covariate effects in the mediator model). These conditions are sufficient, but not necessary, conditions because some perfect cancellations can also eliminate covariates from the formulas in Table 1.

### **When are conditional NDE and NIE estimates constant?**

Table 3 shows the model specifications for which the conditional NDE and NIE are constant with respect to covariates, where the conditional NDE and NIE can be interpreted without reference to covariates. The conditions given are again sufficient, but not necessary, conditions because some perfect cancellations can also eliminate covariates from the formulas in Table 1.

## **NUMERICAL ILLUSTRATIONS**

To provide intuition, we use numerical examples to illustrate how the conditional NDE and NIE change when conditioning on different choices of covariate levels and when using different link functions for the mediator and outcome models. We outline the data generating, examine the true conditional NDE and NIE, and illustrate the bias in estimates of NDE and NIE when EMM terms are omitted.

## Data Generating Process

We randomly generated data for C, A, M, and Y in the following order. We generated one continuous covariate:

$$C \sim \text{Normal}(0, 2^2).$$

We generated a binary exposure:

$$A \sim \text{Bernoulli}(\text{expit}(C + C^2)).$$

For the link functions for mediator and outcome models, we consider the following four model specifications:

1. Linear mediator model, linear outcome model
2. Logistic mediator model, linear outcome model
3. Linear mediator model, logistic outcome model
4. Logistic mediator model, logistic outcome model

We generated continuous or binary mediators and outcomes as follows:

$$\begin{aligned}
 M_{\text{cont}} &= 0.2 + 0.4A + 0.5C + \beta_3 AC + \varepsilon, \text{ where } \varepsilon \sim \text{Normal}(0, 0.5^2), \\
 &\text{or } M_{\text{bin}} \sim \text{Bernoulli}(\text{expit}(0.2 + 0.4A + 0.5C + \beta_3 AC)), \\
 Y_{\text{cont}} &= 0.5 + 0.3A + 0.2M + \theta_3 AM + 0.1C + \theta_5 AC + \theta_6 MC + \varepsilon, \text{ where } \varepsilon \sim \text{Normal}(0, 0.5^2), \\
 Y_{\text{bin}} &\sim \text{Bernoulli}(\text{expit}(-5 + 0.3A + 0.2M + \theta_3 AM + 0.1C + \theta_5 AC + \theta_6 MC)), \text{ for scenarios 1 - 4,} \\
 &\text{model specifications 1, 2 and 4.} \\
 Y_{\text{bin}} &\sim \text{Bernoulli}(\text{expit}(-10 + 0.3A + 0.2M + \theta_3 AM + 0.1C + \theta_5 AC + \theta_6 MC)), \text{ for scenarios 1 - 4,} \\
 &\text{model specification 3.}
 \end{aligned}$$

We varied the magnitude of  $\beta_3$  (*i.e.*, EMM by the covariate of the exposure's effect on the mediator),  $\theta_3$  (*i.e.*, causal interaction between the exposure and the mediator on the outcome),  $\theta_5$  (*i.e.*, EMM by the covariate of the exposure's effect on the outcome), and  $\theta_6$  (*i.e.*, EMM by the covariate of the mediator's effect on the outcome). We had a total of 48 settings representing all possible combinations of the 4 scenarios (*i.e.*, inclusion of product terms), the 4 link function specifications, and the 3 parameter variations (Table 4). We then examined how the conditional NDE and NIE depend on the choice of conditioned covariate levels by evaluating them at levels of the covariate that ranged from  $-2$  to  $+2$  with a step size of  $0.1$ .

## Behavior of True Conditional NDE and NIE

In this section, we consider the true conditional NDE and NIE given the data generating process outlined above. Regarding the inclusion of product terms in the mediator and outcome models, we considered the following four possible model specifications, all of

which were correctly specified. In Scenario 1, the model included an  $A \times M$  product term but not any other product terms. In Scenarios 2–4, other possible product terms were nonzero, corresponding to  $A \times C$  in the mediator model,  $A \times C$  in the outcome model, and/or  $M \times C$  in the outcome model (Table 4). The true NDE, NIE and total effects (TE) were calculated by directly plugging in the true coefficients ( $\beta$ 's and  $\theta$ 's) in the mediator and the outcome models from the data generating process in the previous section.

Figure 2 shows a special case of Scenario 1 ( $\theta_3 = 0$ ) in which neither regression model contained any of the possible product terms ( $\theta_3 = \beta_3 = \theta_5 = \theta_6 = 0$ ).

In this case, the conditional NDE, NIE, and TE are all constant when the mediator model is linear. However, the conditional NIE varies slightly by the covariate level when the mediator model is logistic, despite having no explicit EMM terms in either mediator or outcome model.

Figure 3 shows 16 selected settings with  $\theta_3 \neq 0$  ( $A \times M$  causal interaction is present) and with the middle coefficient values from Table 4. Each row represents one choice of mediator and outcome link functions. The remaining settings are in the eSupplement (Sections 4.1–4.4, the three columns on the left).

The true conditional NDE is dependent on covariate levels in all settings when  $\theta_3 \neq 0$ . In the “linear, linear” case (*i.e.*, a linear mediator model and a linear outcome model), the conditional NDE is always linearly dependent on covariate levels. As a result, as noted above, the conditional NDE evaluated at the mean covariate level coincides with the marginal NDE in this case. This marginal interpretation at the mean covariate level does not generally hold for other choices of link functions (*i.e.*, cases other than the “linear, linear” case). In the “logistic, linear” case, the marginal interpretation does not hold in general, but may hold approximately if the effect of the covariate on the mediator is very weak (1(b), 2(b), 3(b) and 4(b) in Figure 3). The dependence on covariates of NDE and NIE stratified by the combination of mediator and outcome models' link function and the inclusion of product term is summarized in Table 5.

The true conditional NIE is constant in the “linear, linear” and “linear, logistic” cases (*i.e.*, linear mediator model) when no EMM terms are involved (1(a) and 1(c) in Figure 3). However, if there are EMM terms, the conditional NIE is also modified by covariates even in the “linear, linear” and “linear, logistic” cases (2(a), 3(a), 4(a), 2(c), 3(c) and 4(c)). In the “linear, linear” case, the conditional NIE is linearly dependent on covariates, allowing for the conditional NIE at the mean covariate level to coincide with the marginal NIE, unless all three forms of EMM are present. Otherwise, such a marginal interpretation generally does not hold. In the “logistic, linear” case, approximate linearity and marginal interpretation holds if there is very weak covariate effect on the mediator.

### Estimation of conditional NDE and NIE with misspecified regression models

Here we demonstrate the influence of misspecifying the regression models by omitting EMM terms (when they in fact exist in the data generating process) on the estimation of conditional NDE and NIE. Importantly, such misspecification could arise if one uses



existing statistical methods and software tools for regression-based mediation analysis, which do not accommodate product terms involving the covariates, if in fact such EMM is present. We used the same data generating process in the previous section that included EMM terms. One large ( $n = 5,000$ ) dataset was simulated for each one of the 48 settings. In Scenario 1, we omitted the  $A \times M$  causal interaction in the fitted outcome model despite its presence in the data-generating process. In Scenarios 2–4, we included the  $A \times M$  causal interaction, but omitted all  $A \times C$  and  $M \times C$  terms in the fitted models, despite their presence in the data-generating process.

Figure 4 shows 16 selected settings, corresponding to Figure 3 (true conditional NDE, NIE and TE). The remaining settings are in the eSupplement (Sections 4.1–4, the three columns on the right). Figure 5 shows how much misspecified models biases the estimated NDE and NIE away from the true effects, comparing the 16 selected settings in Figures 3 and 4. In Figure 5, we see that omitting EMMs severely biases the conditional effects at most covariate levels, especially in “linear, logistic” cases when at least exposure–covariate product term is in both mediator and outcome models (3(c), 3(d), 4(c), 4(d)). The magnitude of bias is substantive, compared to the range of true effects.

Not only the magnitude of effects, but also their direction, can be estimated incorrectly. For instance, in Scenario 1 in the “linear, linear” case, the true conditional NDE at covariate level = 2 is negative (1(a) in Figure 3), while the estimated NDE when omitting  $A \times M$  causal interaction is positive (1(a) in Figure 4). In Scenario 3 in the “logistic, linear” case, the true conditional NDE at covariate level =  $-2$  is negative (3(b) in Figure 3), while the estimated NDE when omitting  $A \times C$  term in both mediator and outcome models is positive. Although the bias in Scenario 1 (*i.e.*, from omitting the exposure–mediator interaction) can be resolved using existing methods for mediation analysis that accommodate exposure–mediator interaction<sup>6</sup>, it is critical to note that the bias in Scenarios 2–4 (*i.e.*, from omitting product terms involving the covariates) can only be resolved using the proposed extended method.

Additionally, omitting  $A \times M$  causal interaction and EMM terms can also lead one to: (1) *Falsely conclude constant conditional effects*; (2) *Falsely conclude conditional effects coincide with marginal effects*. For (1), in Scenario 1 (omitting  $A \times M$  causal interaction), for all mediator and outcome models, the conditional NDEs in Figure 4 are all falsely constant. For “linear, linear” case, Scenarios 2–4, the NIEs in Figure 4 are falsely constant. In both cases, the natural effect estimates should vary with  $C$ , as seen in the corresponding panels in Figure 3. For (2), when all three EMM terms are present, the conditional NIEs at the mean covariate level from the correctly specified model do not correspond to the marginal counterparts under “linear, linear” and “linear, logistic” cases. Omitting them gives the false impression of a marginal interpretation.

## EMPIRICAL EXAMPLE DEMONSTRATION

The previously published analysis from the CANTOS trial found that the observed reduction in anemia associated with use of canakinumab (anti-interleukin-1 [IL-1] medication) was strongly mediated through the change in hsCRP in the first three months among 8,683



patients without anemia at trial entry<sup>11</sup>. Here we examined whether this previously reported mediation is modified by clinically plausible modifiers of the underlying mechanisms.

Baseline hsCRP, age, and diabetes are three effect modifiers of interest<sup>11</sup> (Figure 6). Baseline hsCRP is a potential effect modifier of canakinumab's impact on the change in hsCRP (mediator) because those with a higher baseline inflammation is expected to experience a greater decline in hsCRP due to the biological impact of IL-1 blockade as well as regression to the mean. As aging and diabetes both have detrimental influence on hematopoietic stem cell functions<sup>18–22</sup>, we hypothesized that the beneficial impact of early reduction in hsCRP (mediator) on incident anemia is decreased by a higher age and diabetes.

Informed by the above clinical reasoning, the regression models were constructed as follows. The exposure was binary canakinumab use (treatment vs. placebo; treatment group: all doses combined), the continuous mediator was the change in log hsCRP (a negative value means improvement) from baseline to the 3<sup>rd</sup> month, and the outcome was time to incident anemia. We adjusted for the same baseline covariates as the previous paper (age, sex, log hsCRP, heart failure status, diabetes status, and hypertension status)<sup>11</sup>. We fit a linear mediator model and a Weibull accelerated failure time outcome model. We included baseline log hsCRP, age, diabetes as effect measure modifiers, where log hsCRP modifies the treatment's effect on the mediator, and age and diabetes modify the mediator's effect on the outcome. We evaluated the extent of mediation with the “lower anemia risk” profile, *i.e.* those who are younger (within 25<sup>th</sup> percentile value of age), having lower baseline log hsCRP (within 25<sup>th</sup> percentile value of baseline log hsCRP), and having no diabetes, to the “higher anemia risk” profile, *i.e.* those who are older (within 75<sup>th</sup> percentile value of age), having higher log hsCRP (within 75<sup>th</sup> percentile value of log hsCRP), and having diabetes. The R code is available at the Github repository <https://github.com/einsley1993/emm-ext-med-cantos> and the eSupplement Section 5.

We analyzed data from 8,302 subjects without anemia at baseline after excluding subjects with missing values. The trial protocol was approved by institutional review boards or ethics committees in the 39 countries involved in the study. All participants provided written informed consent to participate, and the trial underwent surveillance by an independent data and safety monitoring committee. Table 6 shows the causal mediation results. For the “lower anemia risk” profile, 64.0% of the canakinumab effect on the incident anemia outcome was mediated by the change of log hsCRP in the first 3 months. For the “higher anemia risk” profile, the proportion mediated was 48.7%. Although the total effect estimates were similar, we observed a greater proportion mediated for the lower anemia risk profile compared to the higher anemia risk profile.

Table 7 shows the coefficients from the underlying mediator and outcome models, which can give further insights into the lower proportion mediated in the “higher anemia risk” profile. The coefficient of the Canakinumab×baseline log(hsCRP) in the mediator model was negative, although the magnitude of coefficients was small, confirming a greater canakinumab-driven reduction in hsCRP among individuals with a higher baseline hsCRP. The coefficient of change in log(hsCRP) on the survival time was negative, indicating that the effect of reduction in log(hsCRP) was protective. The coefficients of product terms  $\Delta$

$\log(\text{hsCRP}) \times \text{Age}$  and  $\Delta \log(\text{CRP}) \times \text{Diabetes}$  were both positive, suggesting that the benefit of early reduction in hsCRP (mediator) on incidence anemia weakens with increasing age or diabetes, although the magnitude of coefficients was small.

## DISCUSSION

Motivated by an empirical example on the causal mediation of canakinumab's beneficial effect on incident anemia through early reduction in an inflammatory marker hsCRP, we extended the regression-based causal mediation approach to allow for EMM. EMM is a well-studied topic in the context of total effect estimation, but has not gained much attention in the context of causal mediation. Our methodologic results show that in many cases, existing methods will provide biased estimates of the NIE or NDE if there is exposure-covariate or mediator-covariate EMM by the covariates in the exposure-mediator effect, the exposure-outcome effect, or the mediator-outcome effect. The empirical example illustrates how substantive knowledge can inform the modeling process of the EMM in causal mediation. In the empirical analysis, we found that the level of mediation through the reduction in the hsCRP in the early post-treatment was higher for "low anemia risk" profile than for "higher anemia risk" profile. This finding may indicate that the usefulness of early post-treatment hsCRP as a therapeutic monitoring biomarker is greater among individuals with the "low anemia risk" baseline patient profile, providing a more nuanced understanding of hsCRP monitoring adding to the previous mediation investigation without EMM assessment<sup>11</sup>. To facilitate the adoption of EMM considerations for applied researchers, we provided the free and open-source R package, *regmedint*.

Our EMM-extended regression-based causal mediation generalizes Valeri and VanderWeele's closed-form regression-based approach<sup>5</sup>. However, it is beneficial to make a brief comparison to other causal mediation analysis approaches, including Imai *et al.*<sup>1-3,8</sup> and Lange *et al.*,<sup>4</sup> and discuss how they address EMM. Imai *et al.*'s approach relies on the same set of mediator and outcome regression models, but supports the evaluation of EMM (discussed as "moderated mediation", the common terminology used in the social science literature). However, their method gives NDE and NIE estimates only on the difference scale, whereas our proposed methods can provide such estimates on either the difference or the ratio scale. Additionally, our approach that uses closed-form formulas can be more computationally efficient than Imai *et al.*'s simulation-based approach<sup>26</sup>. Lange *et al.*'s approach is a modification of the inverse probability weighted estimation of marginal structural model<sup>4,23</sup>. This method accommodates EMM of the conditional NDE and NIE by including two product terms, one between the covariate and the reference exposure level, and the other between the covariate and the new exposure level, in the weighted counterfactual outcome model. The advantages of this approach are that users can read off NDE and NIE and how much they are modified by covariates directly from the coefficients of the model. Nevertheless, a downside of this approach is that inverse probability of weighting may give less-stable estimates for continuous exposures and mediators than the regression-based approach.

There are nonetheless a few limitations in our extended models. First, we consider only the cases when the additional product terms represent EMMs. To interpret the exposure-

covariate and mediator–covariate product terms as causal interactions, we need further identification assumptions. Indeed, a covariate can be an effect modifier and causally interacts with exposure or mediator simultaneously<sup>14</sup>. Our intention is to extend the original closed-form mediation approach, so we extend the underlying mediator and outcome models without introducing additional assumptions. Hence, the limitations of Valeri and VanderWeele (2013)<sup>5</sup> such as untestable cross-world independency assumptions still remain in our extended models. To accommodate the violation of the assumption of no exposure-induced mediator–outcome confounding<sup>24,25</sup>, the randomized interventional analogue approach has been developed where the mediator values are sampled from a distribution, instead of fixed to the natural value it would take when treated or untreated. However, the natural direct and indirect effects redefined under a randomized interventional analogue are different from the classical NDE and NIE<sup>10</sup>. Although there is continued interest in developing mediation methods in the natural effects framework in recent years<sup>27–35</sup>, some have also criticized the use of natural effects because the cross-world independency assumptions do not have corresponding real-world implications<sup>36–42</sup>, though these views are not universal<sup>7,43</sup>. A new approach called “separable treatment effect approach” (or “interventionist approach”) has been proposed that could mitigate the concerns around implausible cross-world assumptions by decomposing the treatment to two components to make the natural direct and indirect paths have an empirical meaning from the interventional perspective in the real world<sup>36–42</sup>. If the investigators can build a story around treatment decomposition, the identification formulas provided in our paper also apply to the separable treatment effect approach.

In conclusion, we extended the regression-based causal mediation method<sup>5–7</sup> with EMM terms in the underlying mediator and outcome regressions models. Given the influence of covariates on conditional NDE and NIE, investigators are encouraged to consider EMM terms in the underlying regression models if substantive knowledge justifies these as exemplified by the CANTOS trial analysis. Our free and open-source R package *regmedint* is a convenient tool to conduct the EMM-extended closed-form regression-based causal mediation analysis and investigate the nature of covariate dependence of NDE and NIE.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## FUNDING:

The research was supported by the Shore/Brigham and Women’s Hospital Department of Medicine Fellowship Award to KY.

MBM received support from National Institutes of Health (NIH) grants R01CA222147, R01LM013866, UL1TR003142, P30CA124435, and P30DK116074.

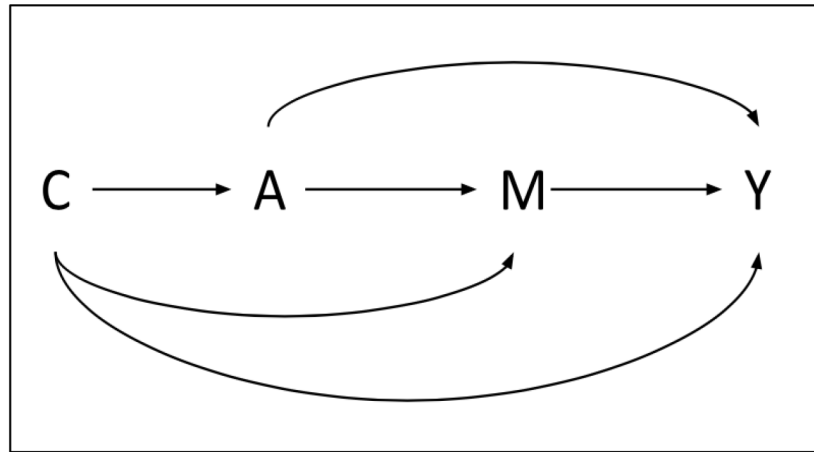
KY received additional support from the Rheumatology Research Foundation K Bridge Award and NIAMS K23 AR076453 and R03AR081309.

## REFERENCES

1. Imai K, Keele L, Tingley D. A general approach to causal mediation analysis. *Psychological Methods*. 2010;15(4):309–334. doi:10.1037/a0020761 [PubMed: 20954780]
2. Imai K, Tingley D, Yamamoto T. Experimental designs for identifying causal mechanisms: Experimental Designs for Identifying Causal Mechanisms. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*. 2013;176(1):5–51. doi:10.1111/j.1467-985X.2012.01032.x
3. Imai K, Keele L, Tingley D, Yamamoto T. Unpacking the Black Box of Causality: Learning about Causal Mechanisms from Experimental and Observational Studies. *Am Polit Sci Rev*. 2011;105(4):765–789. doi:10.1017/S0003055411000414
4. Lange T, Vansteelandt S, Bekaert M. A Simple Unified Approach for Estimating Natural Direct and Indirect Effects. *American Journal of Epidemiology*. 2012;176(3):190–195. doi:10.1093/aje/kwr525 [PubMed: 22781427]
5. Valeri L, VanderWeele TJ. Mediation analysis allowing for exposure–mediator interactions and causal interpretation: Theoretical assumptions and implementation with SAS and SPSS macros. *Psychological Methods*. 2013;18(2):137–150. doi:10.1037/a0031034 [PubMed: 23379553]
6. Valeri L, VanderWeele TJ. SAS Macro for Causal Mediation Analysis with Survival Data: *Epidemiology*. 2015;26(2):e23–e24. doi:10.1097/EDE.0000000000000253 [PubMed: 25643116]
7. VanderWeele TJ. (2015). *Explanation in causal inference: methods for mediation and interaction*. Oxford University Press.
8. Imai K, Keele L, Yamamoto T. Identification, Inference and Sensitivity Analysis for Causal Mediation Effects. *Statist Sci*. 2010;25(1). doi:10.1214/10-STS321
9. Ridker PM, Everett BM, Thuren T, et al. Antiinflammatory Therapy with Canakinumab for Atherosclerotic Disease. *N Engl J Med*. 2017;377(12):1119–1131. doi:10.1056/NEJMoa1707914 [PubMed: 28845751]
10. Solomon DH, Glynn RJ, MacFadyen JG, et al. Relationship of Interleukin-1 $\beta$  Blockade With Incident Gout and Serum Uric Acid Levels: Exploratory Analysis of a Randomized Controlled Trial. *Ann Intern Med*. 2018;169(8):535. doi:10.7326/M18-1167 [PubMed: 30242335]
11. Vallurupalli M, MacFadyen JG, Glynn RJ, et al. Effects of Interleukin-1 $\beta$  Inhibition on Incident Anemia: Exploratory Analyses From a Randomized Trial. *Ann Intern Med*. 2020;172(8):523. doi:10.7326/M19-2945 [PubMed: 32203978]
12. Yoshida K, Glynn RJ, Choi HK, et al. Canakinumab’s Effect Against Subsequent Gout Flares and High-Sensitivity C-Reactive Protein Levels: A Causal Mediation Analysis. *Arthritis Care & Research* Published online November 26, 2022;acr.24832. doi:10.1002/acr.24832
13. Robins JM, Greenland S. Identifiability and Exchangeability for Direct and Indirect Effects: *Epidemiology*. 1992;3(2):143–155. doi:10.1097/00001648-199203000-00013 [PubMed: 1576220]
14. VanderWeele TJ. On the Distinction Between Interaction and Effect Modification. *Epidemiology*. 2009;20(6):863–871. doi:10.1097/EDE.0b013e3181ba333c [PubMed: 19806059]
15. Yoshida K, Li Y, Mathur M. regmedint: Regression-Based Causal Mediation Analysis with an Interaction Term Regression-Based Causal Mediation Analysis with an Interaction Term. R package version 1.0. 0. Published online 2022.
16. Li Y, Mathur MB, Yoshida K. R Package Regmedint: Extension of Regression-Based Causal Mediation Analysis with Effect Measure Modification by Covariates. *Open Science Framework*; 2022. doi:10.31219/osf.io/d4brv
17. Li Y, Yoshida K, Kaufman JS, Mathur MB. A brief primer on conducting regression-based causal mediation analysis. *Psychological Trauma: Theory, Research, Practice, and Policy*. Published online January 26, 2023. doi:10.1037/tra0001421
18. Pradhan AD. C-Reactive Protein, Interleukin 6, and Risk of Developing Type 2 Diabetes Mellitus. *JAMA*. 2001;286(3):327. doi:10.1001/jama.286.3.327 [PubMed: 11466099]
19. Thomas MC. Anemia in diabetes: marker or mediator of microvascular disease? *Nat Rev Nephrol*. 2007;3(1):20–30. doi:10.1038/ncpneph0378
20. Bousounis P, Bergo V, Trompouki E. Inflammation, Aging and Hematopoiesis: A Complex Relationship. *Cells*. 2021;10(6):1386. doi:10.3390/cells10061386 [PubMed: 34199874]

21. Dimmeler S, Leri A. Aging and Disease as Modifiers of Efficacy of Cell Therapy. *Circulation Research*. 2008;102(11):1319–1330. doi:10.1161/CIRCRESAHA.108.175943 [PubMed: 18535269]
22. Kojima H, Kim J, Chan L. Emerging roles of hematopoietic cells in the pathobiology of diabetic complications. *Trends in Endocrinology & Metabolism*. 2014;25(4):178–187. doi:10.1016/j.tem.2014.01.002 [PubMed: 24507996]
23. Steen J, Loeys T, Moerkerke B, Vansteelandt S. medflex : An R Package for Flexible Mediation Analysis using Natural Effect Models. *J Stat Soft*. 2017;76(11). doi:10.18637/jss.v076.i11
24. Lin SH, Young J, Logan R, Tchetgen Tchetgen EJ, VanderWeele TJ. Parametric Mediation g-Formula Approach to Mediation Analysis with Time-varying Exposures, Mediators, and Confounders: *Epidemiology*. 2017;28(2):266–274. doi:10.1097/EDE.0000000000000609 [PubMed: 27984420]
25. Lin SH, Young JG, Logan R, VanderWeele TJ. Mediation analysis for a survival outcome with time-varying exposures, mediators, and confounders. *Statistics in Medicine*. 2017;36(26):4153–4166. doi:10.1002/sim.7426 [PubMed: 28809051]
26. Tingley D, Yamamoto T, Hirose K, Keele L, Imai K. mediation : R Package for Causal Mediation Analysis. *J Stat Soft*. 2014;59(5). doi:10.18637/jss.v059.i05
27. Cheng C, Spiegelman D, Li F. Estimating the natural indirect effect and the mediation proportion via the product method. *BMC Med Res Methodol*. 2021;21(1):253. doi:10.1186/s12874-021-01425-4 [PubMed: 34800985]
28. Hesser H Identifying causal mechanisms in psychotherapy: What can we learn from causal mediation analysis? *Clin Psychology and Psychoth*. 2022;29(3):1050–1058. doi:10.1002/cpp.2687
29. Celli V Causal mediation analysis in economics: Objectives, assumptions, models. *Journal of Economic Surveys*. 2022;36(1):214–234. doi:10.1111/joes.12452
30. Cheng L, Guo R, Liu H. Causal Mediation Analysis with Hidden Confounders Published online January 6, 2022. Accessed February 3, 2023. <http://arxiv.org/abs/2102.11724>
31. Cai X, Zhu Y, Huang Y, Ghosh D. High-dimensional causal mediation analysis based on partial linear structural equation models. *Computational Statistics & Data Analysis*. 2022;174:107501. doi:10.1016/j.csda.2022.107501
32. Tai AS, Lin SH, Chu YC, Yu T, Puhán MA, VanderWeele T. Causal Mediation Analysis with Multiple Time-varying Mediators. *Epidemiology*. 2023;34(1):8–19. doi:10.1097/EDE.0000000000001555 [PubMed: 36455244]
33. Weir IR, Rider JR, Trinquart L. Counterfactual mediation analysis in the multistate model framework for surrogate and clinical time-to-event outcomes in randomized controlled trials. *Pharmaceutical Statistics*. 2022;21(1):163–175. doi:10.1002/pst.2159 [PubMed: 34346173]
34. Linero AR, Zhang Q. Mediation analysis using Bayesian tree ensembles. *Psychological Methods*. Published online July 4, 2022. doi:10.1037/met0000504
35. Caubet M, Samoilenko M, Drouin S, et al. Bayesian joint modeling for causal mediation analysis with a binary outcome and a binary mediator: Exploring the role of obesity in the association between cranial radiation therapy for childhood acute lymphoblastic leukemia treatment and the long-term risk of insulin resistance. *Computational Statistics & Data Analysis*. 2023;177:107586. doi:10.1016/j.csda.2022.107586
36. Robins JM, Richardson TS. Alternative Graphical Causal Models and the Identification of Direct Effects. In: *Causality and Psychopathology*. Oxford University Press; 2011. doi:10.1093/oso/9780199754649.003.0011
37. Robins JM, Richardson TS, Shpitser I. An Interventionist Approach to Mediation Analysis. arXiv:200806019 [stat]. Published online August 28, 2021. Accessed April 27, 2022. <http://arxiv.org/abs/2008.06019>
38. Didelez V Defining causal mediation with a longitudinal mediator and a survival outcome. *Lifetime Data Anal*. 2019;25(4):593–610. doi:10.1007/s10985-018-9449-0 [PubMed: 30218418]
39. Stensrud MJ, Hernán MA, Tchetgen Tchetgen EJ, Robins JM, Didelez V, Young JG. A generalized theory of separable effects in competing event settings. *Lifetime Data Anal*. 2021;27(4):588–631. doi:10.1007/s10985-021-09530-8 [PubMed: 34468923]

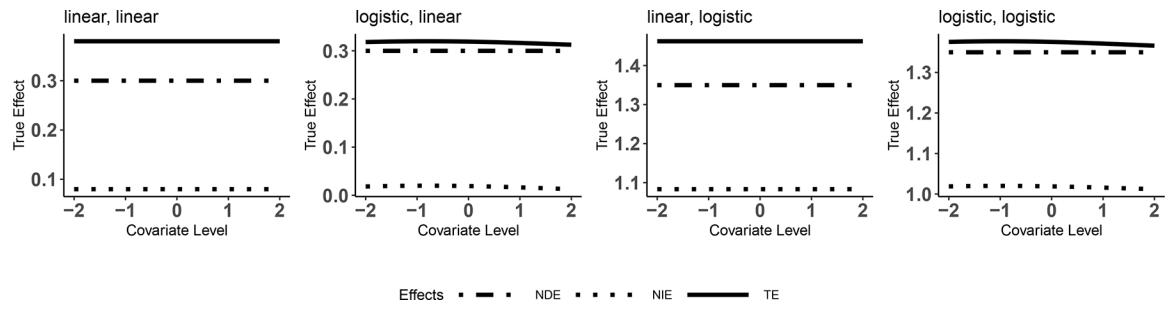
40. Stensrud MJ, Young JG, Didelez V, Robins JM, Hernán MA. Separable Effects for Causal Inference in the Presence of Competing Events. *Journal of the American Statistical Association*. 2022;117(537):175–183. doi:10.1080/01621459.2020.1765783
41. Stensrud MJ, Robins JM, Sarvet A, Tchetgen Tchetgen EJ, Young JG. Conditional Separable Effects. *Journal of the American Statistical Association*. Published online July 8, 2022:1–13. doi:10.1080/01621459.2022.2071276
42. VanderWeele TJ, Valeri L, Ananth CV. Counterpoint: Mediation Formulas With Binary Mediators and Outcomes and the “Rare Outcome Assumption.” *American Journal of Epidemiology*. 2019;188(7):1204–1205. doi:10.1093/aje/kwy281 [PubMed: 30590417]
43. VanderWeele TJ. Mediation Analysis: A Practitioner’s Guide. *Annu Rev Public Health*. 2016;37(1):17–32. doi:10.1146/annurev-publhealth-032315-021402 [PubMed: 26653405]



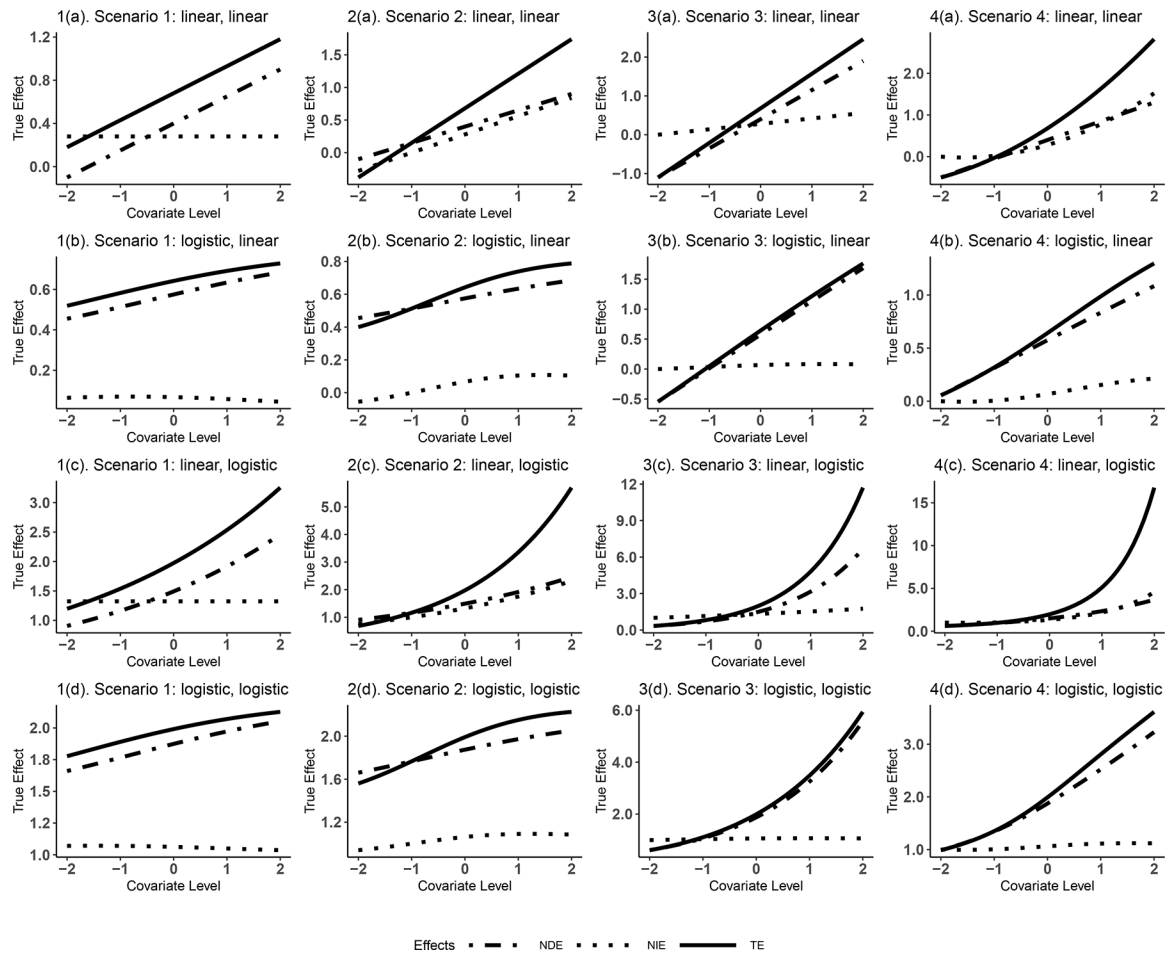
**Abbreviations.** A: exposure; M: mediator; Y: outcome; C: covariate (including effect measure modifier and confounders).

**Figure 1.**  
Assumed causal structure.

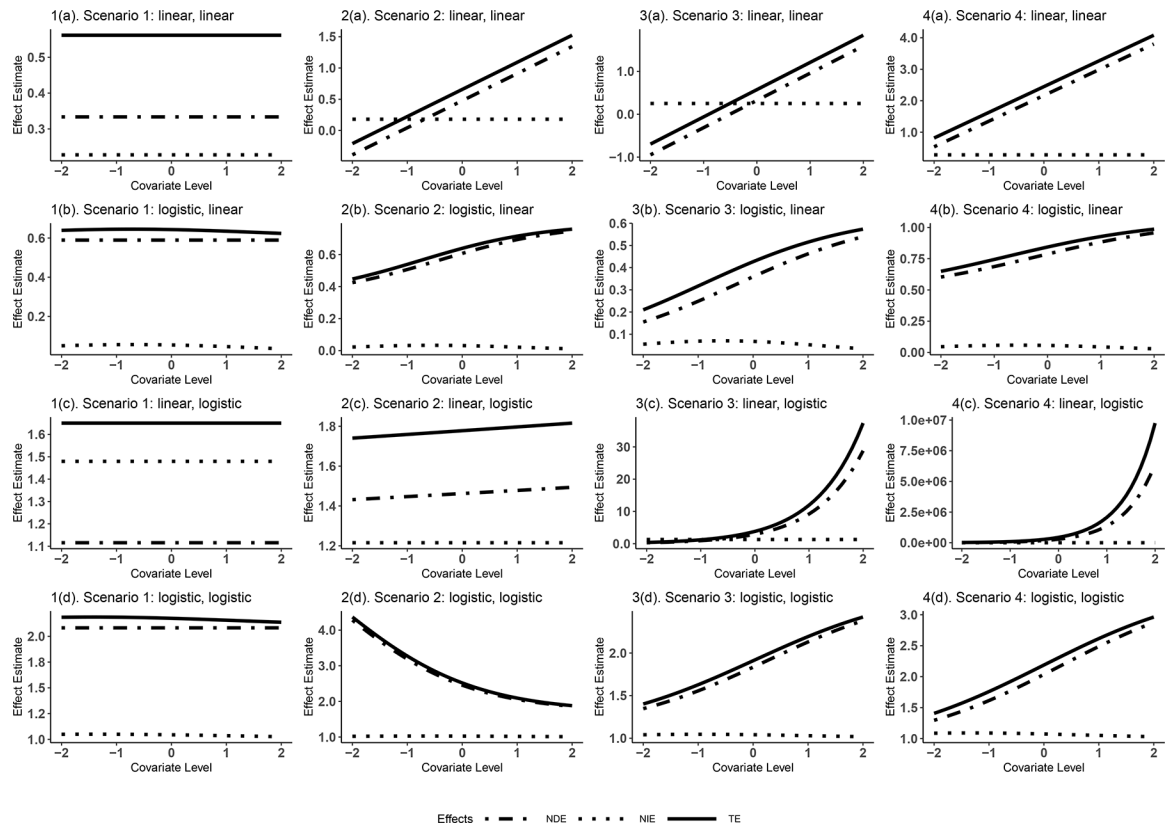




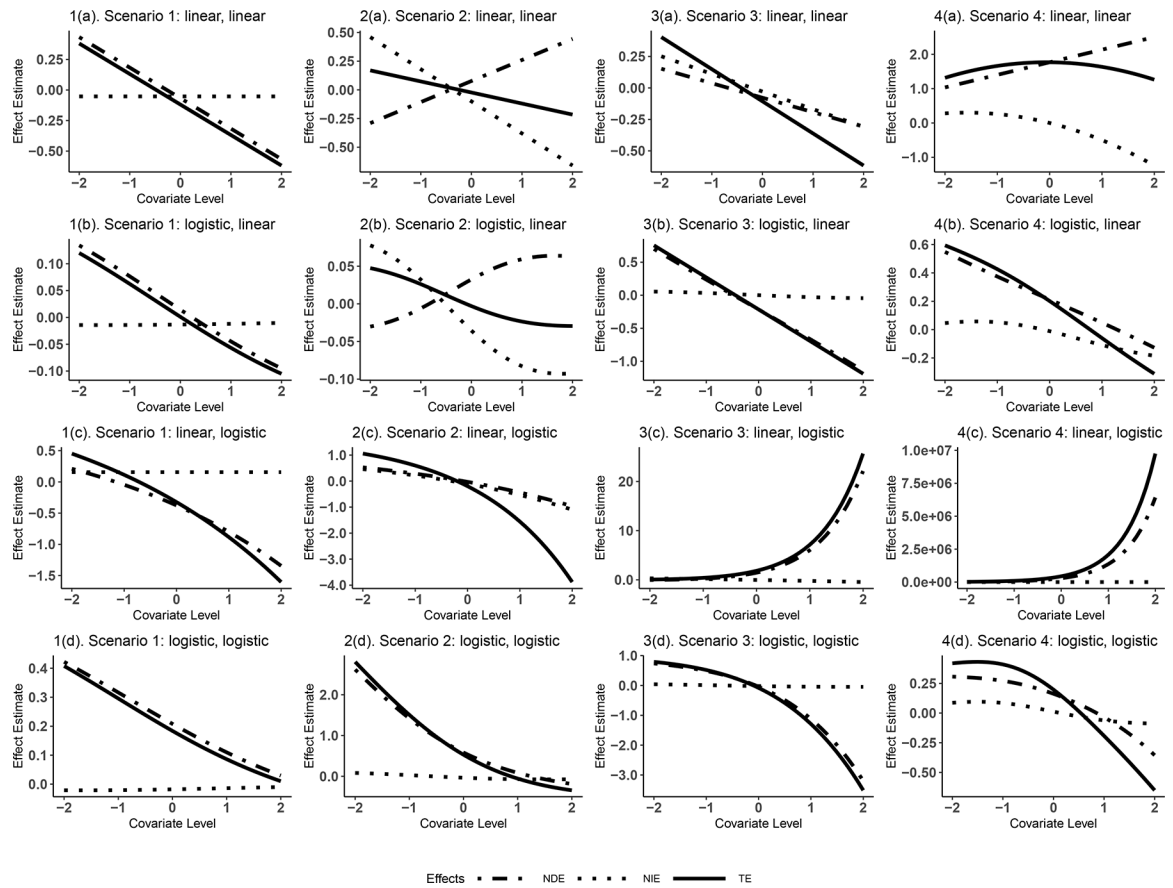
**Figure 2.** Effect measure modification (EMM) of conditional natural effects by covariates despite having no product terms in mediator or outcome regression model (*i.e.*, Scenario 1).



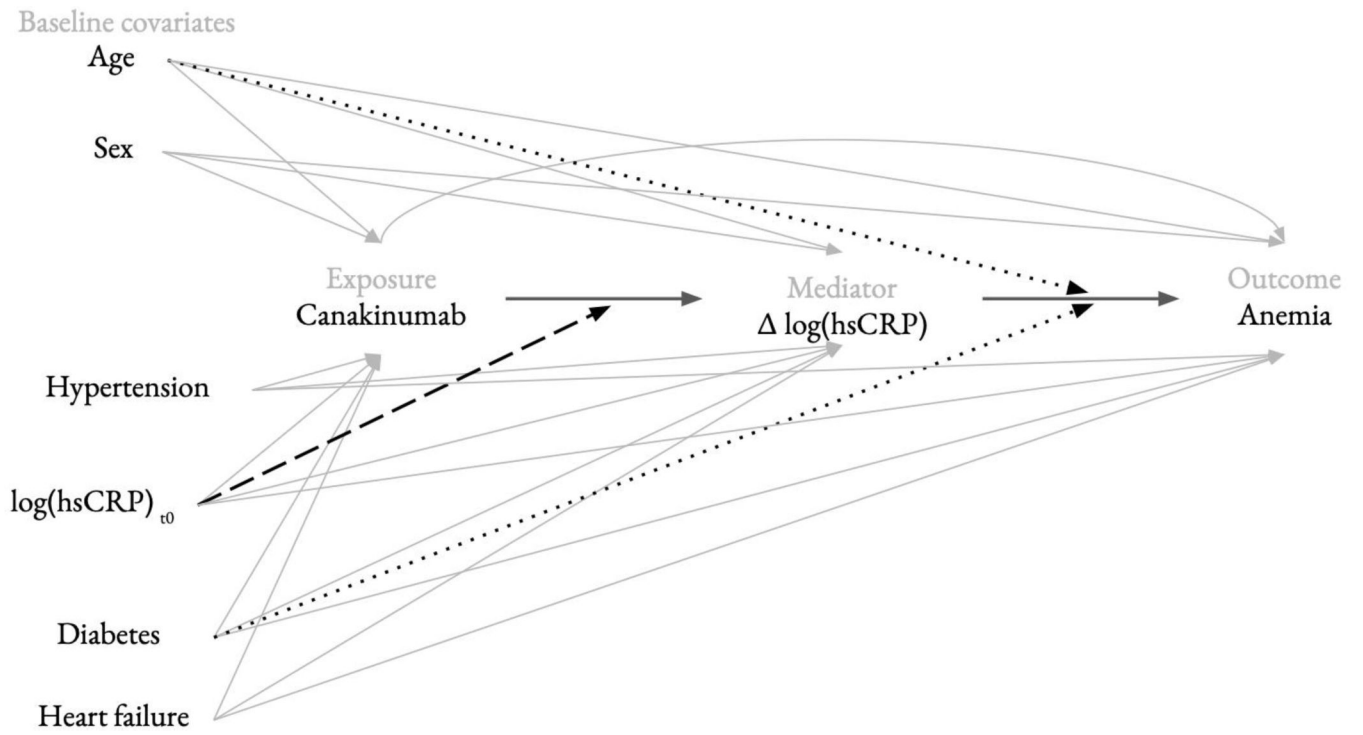
**Figure 3.** Selected plots across combinations of scenarios and link functions for the true NDE, NIE and TE effects.



**Figure 4.** Selected plots across combinations of scenarios and link functions when models are misspecified.



**Figure 5.** Selected plots showing bias, comparing true effects versus estimated effects from misspecified models.



**Figure 6.**  
 Effect measure modification graph for the CANTOS trial example  
 Long-dashed arrow: baseline  $\log(\text{CRP})$  modifies exposure–mediator relationship  
 Dotted arrow: Age and diabetes modify mediator–outcome relationship

**Table 1.**

Closed-form formulas for NDE and NIE with different mediator and outcome models.

<b>NDE</b>	
Linear mediator model, Linear outcome model	$(a - a^*)[\theta_1 + \theta_3(\beta_0 + \beta_1 a^* + \beta_2 c + \beta_3 a^* c) + \theta_5 c]$
Logistic mediator model, Linear outcome model	$(a - a^*)[\theta_1 + \theta_3 \cdot \text{expit}(\beta_0 + \beta_1 a^* + \beta_2 c + \beta_3 a^* c) + \theta_5 c]$
Linear mediator model, Non-linear outcome model	$\exp\left([a - a^*] \cdot \left[\theta_1 + \theta_5 c + \theta_3(\beta_0 + \beta_1 a^* + \beta_2 c + \beta_3 a^* c) + \frac{1}{2}\sigma^2\theta_3(2\theta_2 + \theta_3 a + \theta_3 a^* + 2\theta_6 c)\right]\right)$
Logistic mediator model, Non-linear outcome model	$\exp[(\theta_1 + \theta_5 c)(a - a^*)] \frac{1 + \exp(\beta_0 + \beta_1 a^* + \beta_2 c + \beta_3 a^* c + \theta_2 + \theta_3 a + \theta_6 c)}{1 + \exp(\beta_0 + \beta_1 a^* + \beta_2 c + \beta_3 a^* c + \theta_2 + \theta_3 a^* + \theta_6 c)}$
<b>NIE</b>	
Linear mediator model, Linear outcome model	$(\theta_2 + \theta_3 a + \theta_6 c)(\beta_1 + \beta_3 c)(a - a^*)$
Logistic mediator model, Linear outcome model	$(\theta_2 + \theta_3 a + \theta_6 c)[\text{expit}(\beta_0 + \beta_1 a + \beta_2 c + \beta_3 a c) - \text{expit}(\beta_0 + \beta_1 a^* + \beta_2 c + \beta_3 a^* c)]$
Linear mediator model, Non-linear outcome model	$\exp\{(\theta_2 + \theta_3 a + \theta_6 c)(\beta_1 + \beta_3 c)(a - a^*)\}$
Logistic mediator model, Non-linear outcome model	$\frac{1 + \exp(\beta_0 + \beta_1 a + \beta_2 c + \beta_3 a c + \theta_2 + \theta_3 a + \theta_6 c)}{1 + \exp(\beta_0 + \beta_1 a^* + \beta_2 c + \beta_3 a^* c + \theta_2 + \theta_3 a + \theta_6 c)} \cdot \frac{1 + \exp(\beta_0 + \beta_1 a^* + \beta_2 c + \beta_3 a^* c)}{1 + \exp(\beta_0 + \beta_1 a + \beta_2 c + \beta_3 a c)}$

The formulas for non-linear outcome model accommodate logistic, log-linear, Poisson, negative binomial, accelerated failure time model and a Cox proportional hazards model, with appropriate link functions and modeling assumptions. Note that logistic and Cox proportional hazards model require the outcome to be rare. Details are in the Supplement (Section 2.1). NDE=natural direct effect; NIE=natural indirect effect.

**Table 2.**

Sufficient conditions for conditional natural direct effect (NDE) and natural indirect effect (NIE) estimates at the mean covariate level to coincide with the marginal NDE and NIE estimates.

<i>NDE</i>	Linear outcome model	Non-linear outcome model*
Linear mediator model	Always	
Logistic mediator model	$\theta_3 = 0$	Never
<i>NIE</i>	Linear outcome model	Non-linear outcome model*
Linear mediator model	$\beta_3 = 0$ or $\theta_6 = 0$	
Logistic mediator model	$\beta_2 = \beta_3 = 0$	Never

**Abbreviations.** EMM: effect measure modification; A: exposure; C: covariates; M: mediator.

**Coefficients.**

$\theta_5 = 0$ : no EMM of the exposure effect by covariates [A×C] on outcome

$\theta_3 = 0$ : no A×M interaction

$\beta_2 = 0$ : no confounder in mediator outcome

$\beta_3 = 0$ : no EMM of the exposure effect by covariates [A×C] on mediator

$\theta_6 = 0$ : no EMM of the mediator effect by covariates [M×C] on outcome

\* : The non-linear outcome model accommodates logistic, log-linear, Poisson, negative binomial, accelerated failure time model and a Cox proportional hazards model, with appropriate link functions and modeling assumptions. Note that logistic and Cox proportional hazards model require the outcome to be rare. Details are in the Supplement (Section 2.1).



**Table 3.**

Sufficient conditions for constant conditional natural direct effect (NDE) and natural indirect effect (NIE) estimates with respect to covariates.

<i>NDE</i>	Linear outcome model	Non-linear outcome model*
Linear mediator model	$\theta_5 = 0$ AND $\{\theta_3 = 0$ OR $\beta_2 = \beta_3 = \theta_6 = 0\}$	
Logistic mediator model	$\theta_5 = 0$ AND $\beta_2 = \beta_3 = \theta_6 = 0$	
<i>NIE</i>	Linear outcome model	Non-linear outcome model*
Linear mediator model	$\beta_3 = \theta_6 = 0$	
Logistic mediator model	$\beta_2 = \beta_3 = \theta_6 = 0$	

**Abbreviations.** EMM: effect measure modification; A: exposure; C: covariates; M: mediator.

**Coefficients.**

$\theta_5 = 0$ : no EMM of the exposure effect by covariates [A×C] on outcome

$\theta_3 = 0$ : no A×M interaction

$\beta_2 = 0$ : no confounder in mediator outcome

$\beta_3 = 0$ : no EMM of the exposure effect by covariates [A×C] on mediator

$\theta_6 = 0$ : no EMM of the mediator effect by covariates [M×C] on outcome

\* : The non-linear outcome model accommodates logistic, log-linear, Poisson, negative binomial, accelerated failure time model and a Cox proportional hazards model, with appropriate link functions and modeling assumptions. Note that logistic and Cox proportional hazards model require the outcome to be rare. Details are in the Supplement (Section 2.1).

**Table 4.**

Parameter settings under four simulation scenarios.

Scenario	Product terms in mediator and outcome models	$\theta_3$	$\beta_3$	$\theta_5$	$\theta_6$
1	A×M in outcome model	{0, 0.5, 0.8}	0	0	0
2	A×M in outcome model, A×C in mediator model	0.5	{0.1, 0.4, 0.7}	0	0
3	A×M in outcome model, A×C in mediator model, A×C in outcome model	0.5	0.2	{0.2, 0.5, 0.8}	0
4	A×M in outcome model, A×C in mediator model, A×C in outcome model, M×C in outcome model	0.5	0.2	0.3	{0.3, 0.6, 0.9}

**Coefficients.** $\theta_3$ : A×M causal interaction in outcome model $\beta_3$ : A×C effect measure modification in mediator model $\theta_5$ : A×C effect measure modification in outcome model $\theta_6$ : M×C effect measure modification in outcome model

**Table 5.**

Summary of effect modification by covariates, stratified by combinations of scenarios and link functions.

<b>Model 1: linear mediator model, linear outcome model</b>				
Simulation Scenario	Coefficient	Inclusion of product terms	Dependence on covariates	
			<u>NDE</u>	<u>NIE</u>
1	$\beta_3 = \theta_5 = \theta_6 = 0$	only A×M	linearly <sup>a</sup>	constant
2	$\theta_5 = \theta_6 = 0$	adding A×C in mediator model	linearly	linearly
3	$\theta_6 = 0$	adding A×C in outcome model	linearly	linearly
4		adding M×C in outcome model	linearly	non-linearly
<b>Model 2: logistic mediator model, linear outcome model</b>				
Simulation Scenario	Coefficient	Inclusion of product terms	Dependence on covariates	
			<u>NDE</u>	<u>NIE</u>
1	$\beta_3 = \theta_5 = \theta_6 = 0$	only A×M	non-linearly <sup>a</sup>	non-linearly
2	$\theta_5 = \theta_6 = 0$	adding A×C in mediator model	non-linearly	non-linearly
3	$\theta_6 = 0$	adding A×C in outcome model	non-linearly	non-linearly
4		adding M×C in outcome model	non-linearly	non-linearly
<b>Model 3: linear mediator model, non-linear outcome model<sup>b</sup></b>				
Simulation Scenario	Coefficient	Inclusion of product terms	Dependence of covariates	
			<u>NDE</u>	<u>NIE</u>
1	$\beta_3 = \theta_5 = \theta_6 = 0$	only A×M	non-linearly <sup>a</sup>	constant
2	$\theta_5 = \theta_6 = 0$	adding A×C in mediator model	non-linearly	non-linearly
3	$\theta_6 = 0$	adding A×C in outcome model	non-linearly	non-linearly
4		adding M×C in outcome model	non-linearly	non-linearly
<b>Model 4: logistic mediator model, non-linear outcome model<sup>b</sup></b>				
Simulation Scenario	Coefficient	Inclusion of product terms	Dependence of covariates	
			<u>NDE</u>	<u>NIE</u>
1	$\beta_3 = \theta_5 = \theta_6 = 0$	only A×M	non-linearly <sup>a</sup>	non-linearly
2	$\theta_5 = \theta_6 = 0$	adding A×C in mediator model	non-linearly	non-linearly
3	$\theta_6 = 0$	adding A×C in outcome model	non-linearly	non-linearly
4		adding M×C in outcome model	non-linearly	non-linearly

<sup>a</sup>: NDE is constant when  $\theta_3 = 0$ . This is corresponding to the special case shown in Figure 2.

<sup>b</sup>: The non-linear outcome model accommodates logistic, log-linear, Poisson, negative binomial, accelerated failure time model and a Cox proportional hazards model, with appropriate link functions and modeling assumptions. Note that logistic and Cox proportional hazards model require the outcome to be rare. Details are in the Supplement (Section 2.1).

**Abbreviations.** NDE: natural direct effect; NIE: natural indirect effect.

**Coefficients.**

$\theta_3$ : A×M causal interaction in outcome model

$\beta_3$ : A×C effect measure modification in mediator model

$\theta_5$ : A×C effect measure modification in outcome model

$\theta_6$ : M×C effect measure modification in outcome model

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 6.**

Results of mediation analyses for the CANTOS trial example.

Subgroup	Effect	Point Estimate	95% CI
Lower anemia risk profile <sup>a</sup>	NDE	1.149	(0.951, 1.388)
	NIE	1.230	(1.137, 1.330)
	TE	1.413	(1.182, 1.689)
	PM	63.96%	(30.05%, 97.87%)
Higher anemia risk profile <sup>b</sup>	NDE	1.210	(1.012, 1.447)
	NIE	1.164	(1.065, 1.273)
	TE	1.409	(1.162, 1.708)
	PM	48.65%	(21.06%, 76.23%)

**Abbreviations.** EMM: effect measure modification; NDE: natural direct effect; NIE: natural indirect effect; PM: proportion mediated; CI: confidence interval.

The effects are on the mean time ratio scale, a value greater than 1.0 indicates a protective effect of canakinumab on the incident anemia outcome.

<sup>a</sup>: Conditioning on 25<sup>th</sup> percentile value of age, 25<sup>th</sup> percentile value of baseline logarithm of high-sensitivity C-reactive protein, and having no diabetes.

<sup>b</sup>: Conditioning on 75<sup>th</sup> percentile value of age, 75<sup>th</sup> percentile value of logarithm of high-sensitivity C-reactive protein, and having diabetes.

**Table 7.**

Coefficients in the fitted mediator and outcome models in the CANTOS trial example

<i>Mediator model</i>		
	<b>Estimate</b>	<b>SE</b>
(Intercept)	0.423	0.069
Canakinumab	-0.604	0.047
Age	-0.001	0.001
Sex (female)	0.013	0.021
Baseline log(hsCRP)	-1.437	0.023
Heart failure	0.093	0.019
Hypertension	0.016	0.023
Diabetes	0.043	0.023
Canakinumab×Baseline log(hsCRP)	-0.031	0.028
<i>Outcome model</i>		
	<b>Estimate</b>	<b>SE</b>
(Intercept)	7.401	0.536
Canakinumab	0.082	0.149
log(hsCRP) <sup>a</sup>	-0.484	0.205
Age	-0.048	0.008
Sex (female)	-0.009	0.090
Baseline log(hsCRP)	-0.734	0.089
Heart failure	-0.639	0.082
Hypertension	-0.067	0.106
Diabetes	-0.214	0.170
Canakinumab× log(hsCRP)	-0.045	0.062
log(hsCRP)×Age	0.004	0.003
log(hsCRP)×Diabetes	0.047	0.064

<sup>a</sup>:  $\log(\text{hsCRP}) = \log(\text{hsCRP})$  at month 3 -  $\log(\text{hsCRP})$  at month 0