


RESEARCH ARTICLE

The impact of adjusting for pure predictors of exposure, mediator, and outcome on the variance of natural direct and indirect effect estimators

Awa Diop^{1,2} | Geneviève Lefebvre³  | Caroline S. Duchaine^{1,2,4} |
Danielle Laurin^{2,4,5} | Denis Talbot^{1,2} 

¹Département de Médecine Sociale et Préventive, Université Laval, Québec City, Québec, Canada

²Axe santé des Populations et Pratiques Optimales en Santé, Centre de Recherche du CHU de Québec - Université Laval, Québec City, Québec, Canada

³Département de Mathématiques, Université du Québec à Montréal, Montréal, Québec, Canada

⁴Centre de Recherche sur Les Soins et Les Services de Première Ligne de l'Université Laval, Québec City, Québec, Canada

⁵Faculté de Pharmacie, Université Laval, Québec City, Québec, Canada

Correspondence

Denis Talbot, Département de médecine sociale et préventive, Université Laval, Québec City, QC G1V 0A6, Canada.
Email: denis.talbot@fmed.ulaval.ca

Funding information

Fondation du CHU de Québec,
Grant/Award Number: 2710; Fonds de Recherche du Québec - Santé,
Grant/Award Number: 265385; Natural Sciences and Engineering Research Council of Canada, Grant/Award Numbers: 2016-06295, 2020-05473

It is now well established that adjusting for pure predictors of the outcome, in addition to confounders, allows unbiased estimation of the total exposure effect on an outcome with generally reduced standard errors (SEs). However, no analogous results have been derived for mediation analysis. Considering the simplest linear regression setting and the ordinary least square estimator, we obtained theoretical results showing that adjusting for pure predictors of the outcome, in addition to confounders, allows unbiased estimation of the natural indirect effect (NIE) and the natural direct effect (NDE) on the difference scale with reduced SEs. Adjusting for pure predictors of the mediator increases the SE of the NDE's estimator, but may increase or decrease the variance of the NIE's estimator. Adjusting for pure predictors of the exposure increases the variance of estimators of the NIE and NDE. Simulation studies were used to confirm and extend these results to the case where the mediator or the outcome is binary. Additional simulations were conducted to explore scenarios featuring an exposure-mediator interaction as well as the relative risk and odds ratio scales for the case of binary mediator and outcome. Both a regression approach and an inverse probability weighting approach were considered in the simulation study. A real-data illustration employing data from the Canadian Study of Health and Aging is provided. This analysis is concerned with the mediating effect of vitamin D in the effect of physical activity on dementia and its results are overall consistent with the theoretical and empirical findings.

KEYWORDS

causal inference, direct and indirect effects, mediation analysis, variable selection

1 | INTRODUCTION

Estimating exposure effects using observational data generally requires adjustment for confounding variables to obtain unbiased estimators. There exists a rich literature concerning the identification and selection of potentially confounding

Abbreviations: NDE, natural direct effect; NIE, Natural indirect effect; SE, standard error; IPW, inverse probability weighting

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2021 The Authors. *Statistics in Medicine* published by John Wiley & Sons Ltd.

variables when the goal is to estimate the effect of an exposure measured at one point in time on an outcome. Notably, approaches that employ causal graphs to help identifying confounders based on experts' knowledge have been proposed.¹⁻³ In recent years, various data-driven procedures have also been devised to supplement scientific knowledge with information from the data to help selecting adjustment covariates (see Witte and Didelez⁴ for an overview). Many of these methods aim not only to reduce confounding bias, but also to improve the precision of estimates. Indeed, theoretical and simulation results have revealed that unbiased estimators with reduced standard errors (SEs) can be obtained when adjusting for true confounders, pure predictors of the outcome and avoiding adjustment for pure predictors of the exposure.⁵⁻⁸

Another common goal of etiologic studies is to determine what portion of the exposure effect on the outcome is attributable to the effect of the exposure on a given putative intermediate or mediator variable. Analyses aimed at investigating how an exposure effect decomposes into a direct effect and an indirect effect through one or multiple mediators are referred to as mediation analyses. As for (total) exposure effect, identification of mediation effects from observational data requires adjustment for confounding variables.

It is now well established that unbiased estimation of natural direct and indirect effects requires adjustment for confounders of three relationships: exposure-outcome, mediator-outcome and exposure-mediator (see, eg, VanderWeele⁹). However, very little is currently known concerning how different types of variables affect the precision of mediation analysis estimators. The goal of the present paper is to address and fill this gap of knowledge. More specifically, the objective is to study the impact of adjusting for pure predictors of the exposure, mediator, and outcome on the variance of popular natural direct and indirect effect estimators used by practitioners. This is accomplished both from a conventional regression and a propensity-score perspectives, so to mimic previous investigations done on this issue for the estimation of total exposure effects.

The outline of the remainder of the paper is as follows. In Section 2, we introduce some background and set-up the notation. In Section 3, we provide theoretical insights regarding how adjusting for the aforementioned types of variables affects the variance of natural mediation effect estimators, on a difference scale, in the simplest regression setting with a continuous outcome and a continuous mediator. In Section 4, we present results from a simulation study designed to confirm these theoretical results and provide additional understanding when the mediator or the outcome is binary. Additional exploratory simulation results are also presented to investigate the extension in presence of an exposure-mediator interaction in the continuous mediator—continuous outcome case, as well as the relative risk and odds ratio scales in the binary mediator—binary outcome case. Section 5 illustrates how adjusting for hypothesized pure predictors affects results in a real data analysis concerning the mediating effect of vitamin D in the relationship between physical activity and dementia. We conclude the article with a summary of our results and a brief discussion on future perspectives.

2 | BACKGROUND

Let A be the exposure, M the mediator, Y the outcome and \mathbf{L} a set of measured covariates. We assume that a sample of independent observations $\{\mathbf{L}_i, A_i, M_i, Y_i\}$, $i = 1, \dots, n$, is randomly drawn from a given population. Further, let \mathcal{M} and \mathcal{L} be the sets of possible values of M and \mathbf{L} , respectively.

To define the causal quantities of interest, we consider Rubin's causal model.¹⁰ We denote by Y^a the outcome that would have been observed if $A = a$, $Y^{a,m}$ the outcome that would have been observed if $A = a$ and $M = m$, M^a the value that the mediator would have taken if $A = a$, and $Y^{a,M^{a'}}$ the outcome that would have been observed if $A = a$ and $M = M^{a'}$. The latter quantities are termed counterfactuals.

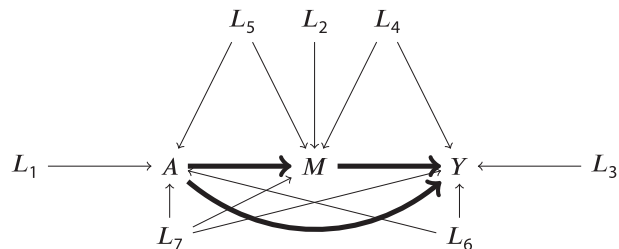
Using this notation, the total effect on the outcome of exposure level a versus level a' can be defined as $TE = \mathbb{E}(Y^a) - \mathbb{E}(Y^{a'})$. Multiple decompositions of TE in an indirect effect, due to the effect of the exposure on the mediator, and a direct effect, not due to the effect of the exposure on the mediator, have been proposed.¹¹⁻¹⁸ In this paper, we focus on decomposing TE in a natural indirect effect (NIE) and a natural direct effect (NDE), which are defined on a difference scale as

$$\begin{aligned} \text{NIE} &= \mathbb{E}(Y^{a,M^a}) - \mathbb{E}(Y^{a,M^{a'}}), \\ \text{NDE} &= \mathbb{E}(Y^{a,M^{a'}}) - \mathbb{E}(Y^{a',M^{a'}}). \end{aligned} \quad (1)$$

This natural effect decomposition is one of the most commonly employed in the applied literature.

Estimating NIE and NDE from the observed data can be done under the following causal assumptions: 1) $Y^{a,m} \perp\!\!\!\perp A | \mathbf{L}$; 2) $Y^{a,m} \perp\!\!\!\perp M | A, \mathbf{L}$; 3) $M^a \perp\!\!\!\perp A | \mathbf{L}$; 4) $Y^{a,m} \perp\!\!\!\perp M^{a'} | \mathbf{L}$; 5) $f(A | \mathbf{L}) > 0$ for $A \in \{a, a'\}$ and $\mathbf{L} \in \mathcal{L}$; 6) $f(m | A, \mathbf{L}) > 0$ for all $m \in \mathcal{M}$,

FIGURE 1 Causal graph depicting different types of variables that could be considered for estimating natural direct and indirect effects



$A \in \{a, a'\}$ and $\mathbf{L} \in \mathcal{L}$; 7) $Y^{a, M^a} = Y^a = Y$ if $A = a$; and 8) $M^a = M$ if $A = a$. Assumptions 1), 2) and 3) are conditional exchangeability assumptions that entail there are no unmeasured confounders of the relationships $A \rightarrow Y$, $M \rightarrow Y$ and $A \rightarrow M$. Assumption 4) is often informally interpreted as meaning that no confounders of the relationship $M \rightarrow Y$ is affected by A .^{19,20} Andrews and Didelez (2020)²¹ examine more formally the interpretation of this assumption and provide examples where it can be violated beyond post-exposure confounding. Note that Assumption 4) is unverifiable, even in randomized experiments. Assumptions 5) and 6) indicate there is a positive probability for all units to receive each exposure level of interest and each possible mediator values. Finally, assumptions 7) and 8) are consistency assumptions that stipulate that when a unit has exposure level a , its observed outcome and mediator values are the same as the counterfactual outcome and mediator values under exposure level a . The counterfactual expectations are then nonparametrically identified from the observed data as

$$\mathbb{E} \left(Y^{a, M^a} \right) = \int_{\mathbf{l} \in \mathcal{L}} \int_{m \in \mathcal{M}} \mathbb{E} (Y | A = a, M = m, \mathbf{L} = \mathbf{l}) f (m | A = a', \mathbf{L} = \mathbf{l}) f (\mathbf{l}) \mu (M) \mu (\mathbf{L}), \tag{2}$$

where $f (m | A = a', \mathbf{L} = \mathbf{l})$ and $f (\mathbf{l})$ are the densities of M and \mathbf{L} relative to appropriate measures $\mu (M)$ and $\mu (\mathbf{L})$, respectively (for example, a counting measure for discrete variables or the Lebesgue measure for continuous variables).

Due to the curse of dimensionality, NIE and NDE are generally estimated using parametric or semiparametric estimators.²² For example, Imai et al (2010)²³ have proposed estimators based on parametric outcome and mediator regression models, whereas VanderWeele (2009)²⁴ and Lange et al (2012)²⁵ have proposed inverse probability weighting estimators based on mediator and exposure regression models.

Various sets of covariates \mathbf{L} could meet assumptions 1)-3) stated previously. For instance, consider the causal graph depicted in Figure 1. In this graph, L_4, L_5, L_6 and L_7 are confounders of the $A \rightarrow Y$, $M \rightarrow Y$ and $A \rightarrow M$ relationships, whereas L_1, L_2 and L_3 are pure predictors of the exposure, mediator, and outcome, respectively. Adjusting for the confounders in a given mediation analysis is required to unbiasedly estimate NIE and NDE. While additionally adjusting for the pure predictors could be done without introducing bias, the variance of the estimators is expected to be impacted by their inclusion. In the sequel, we investigate the impact of the composition of adjustment set \mathbf{L} satisfying assumptions 1)-3) on the variance of NIE and NDE estimators, assuming that assumptions 4)-8) are also met.

3 | THEORETICAL RESULTS

We first consider a simple situation wherein analytical variance formulas for NIE and NDE estimators are available. In this case, the contribution of the pure predictors of the exposure, mediator and outcome on the variances can be conveniently isolated. To obtain these results, we rely on causal graph theory (for a brief overview, see the appendix of VanderWeele and Shpitser²).

We suppose that the causal graph depicted in Figure 1 represents the causal relationships between the variables and that both M and Y are continuous. Moreover, we assume that $\mathbb{E} (Y | A, M, \mathbf{L}) = \theta_0 + \theta_1 A + \theta_2 M + \mathbf{L} \theta_3$ and $\mathbb{E} (M | A, \mathbf{L}) = \beta_0 + \beta_1 A + \mathbf{L} \beta_2$. To simplify the presentation and without loss of generality, the exposure levels of interest we consider are $a = 1$ and $a' = 0$. Under these outcome and mediator models, it can be verified that

$$\begin{aligned} \text{NIE} &= \theta_2 \times \beta_1 \\ \text{NDE} &= \theta_1. \end{aligned} \tag{3}$$

Estimators of the NIE and NDE can be obtained by plugging in least square estimators of θ_1, θ_2 and β_1 in Equation (3). The asymptotic variance (delta-method) of the NIE estimator is $\text{Var} (\widehat{\text{NIE}}) = \theta_2^2 \text{Var} (\hat{\beta}_1) + \beta_1^2 \text{Var} (\hat{\theta}_2)$,²⁶

while $Var(\widehat{NDE}) = Var(\hat{\theta}_1)$. We can write²⁷

$$Var(\widehat{NIE}) = \theta_2^2 \frac{\sigma_{M|A,L}^2}{n(1 - R_{A|L}^2) Var(A)} + \beta_1^2 \frac{\sigma_{Y|A,M,L}^2}{n(1 - R_{M|A,L}^2) Var(M)} \tag{4}$$

$$Var(\widehat{NDE}) = \frac{\sigma_{Y|A,M,L}^2}{n(1 - R_{A|M,L}^2) Var(A)}, \tag{5}$$

where $\sigma_{M|A,L}^2$ is the population residual variance of the linear regression of M on A and L (similarly for $\sigma_{Y|A,M,L}^2$), and $R_{A|L}^2$ is the population coefficient of determination of the linear regression of A on L (similarly for $R_{M|A,L}^2$ and $R_{A|M,L}^2$). Note that in finite samples, the number of regression parameters is subtracted from the sample size n in the above formulas. To determine the impact of including pure predictors of the exposure, mediator and outcome in L , we assume that L includes all true confounders (L_4, L_5, L_6 and L_7), as required for unbiased estimation. Inspecting Equations (4) and (5), we notice that the choice of L affects $Var(\widehat{NIE})$ and $Var(\widehat{NDE})$ only through $\sigma_{M|A,L}^2, \sigma_{Y|A,M,L}^2, R_{A|L}^2, R_{M|A,L}^2$ and $R_{A|M,L}^2$. The discussion below focuses on the asymptotic behavior of the estimators' variances.

First, notice that L_1 and M are d-separated by $\{A, L_4, L_5, L_6, L_7\}$ in Figure 1. As such, the mediator is independent of the pure predictors of the exposure conditional on the exposure and confounders ($M \perp\!\!\!\perp L_1 | \{A, L_4, L_5, L_6, L_7\}$). It follows that including pure predictors of the exposure (L_1) in L does not affect $\sigma_{M|A,L}^2$ or $R_{M|A,L}^2$ (see Appendix for details). Similarly, remark that Y is d-separated from L_1 by $\{A, M, L_4, L_5, L_6, L_7\}$. Employing the same reasoning as before, we conclude that including pure predictors of the exposure in L does not affect $\sigma_{Y|A,M,L}^2$. Finally, including pure predictors of the exposure increases $R_{A|M,L}^2$ and $R_{A|L}^2$. As a consequence, including pure predictors of the exposure in L increases the variance of both \widehat{NIE} and \widehat{NDE} .

Analogously, it can be verified that including pure predictors of the outcome decreases $\sigma_{Y|A,M,L}^2$ and has no impact on $\sigma_{M|A,L}^2, R_{A|L}^2, R_{M|A,L}^2$ or $R_{A|M,L}^2$, therefore reducing the variance of both \widehat{NIE} and \widehat{NDE} .

Including pure predictors of the mediator decreases $\sigma_{M|A,L}^2$ and increases $R_{M|A,L}^2$. In addition, this increases $R_{A|M,L}^2$. Indeed, as M is a collider on the path $L_2 \rightarrow M \leftarrow A$, L_2 and A are associated after conditioning on M . On the other hand, the quantities $R_{A|L}^2$ and $\sigma_{Y|A,M,L}^2$ are unaffected by the inclusion of pure predictors of the mediator, because A and L_2 are d-separated by $\{L_4, L_5, L_6, L_7\}$, and Y and L_2 are d-separated by $\{A, M, L_4, L_5, L_6, L_7\}$. As a consequence, including pure predictors of the mediator increases $Var(\widehat{NDE})$. Because $\sigma_{M|A,L}^2$ appears in the numerator of the first term of $Var(\widehat{NIE})$ in Equation (4) and $R_{M|A,L}^2$ appears in the denominator of the second term, including pure predictors of the mediator may either decrease or increase $Var(\widehat{NIE})$. The exact consequence notably depends on the relative size of the effect of the mediator on the outcome (θ_2) and of the exposure on the mediator (β_1). It can be expected that when θ_2^2 is large relative to β_1^2 , the inclusion of pure predictors of the mediator generally decreases $Var(\widehat{NIE})$. Conversely, when θ_2^2 is small relative to β_1^2 , the inclusion of pure predictors of the mediator generally increases $Var(\widehat{NIE})$.

4 | SIMULATION STUDY

4.1 | Scenarios

We considered three main scenarios for each of the following cases (i) continuous mediator and outcome, (ii) binary mediator and continuous outcome, (iii) continuous mediator and binary outcome, and (iv) binary mediator and outcome. Additional exploratory simulations are described in Section 4.3.

For each scenario, 1000 independent datasets of size $n = 1000$ were generated according to equations compatible with the causal graph depicted in Figure 1. In all scenarios $L_1, \dots, L_7 \stackrel{iid}{\sim} N(0, 1)$ and $A \sim \text{Bernoulli}\{p = \text{expit}[L_1 + 0.2 \times (L_5 + L_6 + L_7)]\}$, where $\text{expit}(x) = [1 + \exp(-x)]^{-1}$. The linear predictor for generating M was $lp_M = \beta_1 A + L_2 + 0.2 \times (L_4 + L_5 + L_7)$ and the one for generating Y was $lp_Y = \theta_1 A + \theta_2 M + L_3 + 0.2 \times (L_4 + L_6 + L_7)$. Then, M was generated as $N(lp_M, 1)$ for the continuous case and as $\text{Bernoulli}[p = \text{expit}(lp_M)]$ in the binary case. Similarly, Y was either generated as $N(lp_Y, 1)$ or $\text{Bernoulli}[p = \text{expit}(lp_Y)]$ when Y was continuous or binary, respectively. In all scenarios, $\theta_1 = 0.4$, whereas (θ_2, β_1) were $(0.6, 0.6)$ in Scenario 1, $(0.3, 1.2)$ in Scenario 2, and $(1.2, 0.3)$ in Scenario 3.

The values of the simulation parameters were chosen so to exhibit clearly the impact of the pure predictors of the exposure, mediator and outcome on the variance of the estimators. In particular, the values of the coefficients associated with the pure predictors were chosen to be quite large, since the impact of these variables is expected to be proportional to the size of their associated coefficients. Moreover, we chose a relatively large sample size ($n = 1000$) so that the simulations are in adequation with the theoretical presentation which focused on the asymptotic variances of estimators.

4.2 | Analysis

For all main scenarios and types of outcomes, the parameters of interest were the NIE and NDE defined on an additive scale, as in Equation (1). For each generated dataset, the NIE and NDE were estimated using the model-based regression approach (hereafter called *regression*) presented in Imai et al²³ and implemented in the R package `mediation`,²⁸ as well as using the inverse probability weighting (*IPW*) approach presented in Lange et al.²⁵

Briefly, the *regression* approach requires modeling both the outcome and the mediator for estimating $\mathbb{E}(Y|A = a, M = m, \mathbf{L} = \mathbf{l})$ and $f(m|A = a', \mathbf{L} = \mathbf{l})$, respectively. The estimates are then plugged in an empirical version of Equation (2) to estimate the counterfactual expectations, yielding \widehat{NIE}_{reg} and \widehat{NDE}_{reg} . In our simulations, we employed main term linear regressions with normally distributed errors to model continuous variables and main term logistic regressions for binary variables.

The *IPW* estimates of the NIE and NDE were obtained by following the steps below (also implemented in the R package `medflex`²⁹):

1. Fit models for the mediator and exposure. Additionally, estimate the marginal probability of exposure, $P(A = 1)$.
2. Construct a new dataset where each original observation appears twice. Create a new variable $A^* = A$ for one copy of each original observation and $A^* = 1 - A$ for the other copy.
3. Using the output of the regression models of step 1, compute for each observation i of the dataset in step 2

$$w_i = \frac{\widehat{P}(A = a_i)}{\widehat{P}(A = a_i | \mathbf{L} = \mathbf{l}_i)} \frac{\widehat{P}(M = m_i | A = a_i^*, \mathbf{L} = \mathbf{l}_i)}{\widehat{P}(M = m_i | A = a_i, \mathbf{L} = \mathbf{l}_i)}.$$

4. Fit the regression $\mathbb{E}[Y|A, A^*] = \gamma_0 + \gamma_1 A + \gamma_2 A^*$ on the dataset of step 2 weighted according to w_i .
5. $\widehat{NIE}_{IPW} = \widehat{\gamma}_2$ and $\widehat{NDE}_{IPW} = \widehat{\gamma}_1$.

The same types of models were used for the *IPW* approach as for the *regression* approach, except in step 4 where a generalized linear model with a binomial distribution and identity link was used when the outcome was binary. Note that when the outcome is binary, $\mathbb{E}[Y|A, A^*] = P(Y = 1|A, A^*)$. Eight different adjustment sets were considered: 1) $\mathbf{L}_C = \{L_4, L_5, L_6, L_7\}$, 2) $\mathbf{L}_{C,A} = \{L_1, L_4, L_5, L_6, L_7\}$, 3) $\mathbf{L}_{C,M} = \{L_2, L_4, L_5, L_6, L_7\}$, 4) $\mathbf{L}_{C,Y} = \{L_3, L_4, L_5, L_6, L_7\}$, 5) $\mathbf{L}_{C,A,M} = \{L_1, L_2, L_4, L_5, L_6, L_7\}$, 6) $\mathbf{L}_{C,A,Y} = \{L_1, L_3, L_4, L_5, L_6, L_7\}$, 7) $\mathbf{L}_{C,M,Y} = \{L_2, L_3, L_4, L_5, L_6, L_7\}$, and 8) $\mathbf{L}_{C,A,M,Y} = \{L_1, L_2, L_3, L_4, L_5, L_6, L_7\}$.

Except in the continuous mediator–continuous outcome case, the true values of NIE and NDE were estimated using a Monte Carlo simulation. Briefly, we first generated $n = 2\,000\,000$ observations of $\{\mathbf{L}, M^1, M^0, Y^{1,M^1}, Y^{1,M^0}, Y^{0,M^1}, Y^{0,M^0}\}$ using the data generating equations described in Section 4.1. The true NIE and NDE were then estimated using the simulated counterfactual outcomes.

For each combination of case, scenario, estimator, and adjustment set, we computed the bias of the NIE and NDE estimators as the difference between the average of the estimates and the true value. We also estimated the true SE by computing the Monte Carlo standard deviation of the estimates. The relative SE was obtained by dividing the SE obtained using each adjustment set by the SE obtained with \mathbf{L}_C . Finally, we computed the proportion of the 95% confidence intervals that included the true effect, where confidence intervals were obtained with the nonparametric bootstrap with 1000 replicates using the percentile method.

4.3 | Additional simulations

We considered additional simulations to explore how our results further extend when considering the NIE and NDE on the risk ratio and odds ratio scales, as well as in situations featuring an exposure–mediator interaction. To reduce the

TABLE 1 Natural direct and indirect effect estimators from *regression* and *IPW* approaches by adjustment set (*L*) and scenario; continuous mediator–continuous outcome case

Scenarios	L	Regression approach				IPW approach			
		Bias		Relative SE		Bias		Relative SE	
		NIE	NDE	NIE	NDE	NIE	NDE	NIE	NDE
Scenario 1 ($\beta_1 = 0.6, \theta_2 = 0.6$)	L_C	< 0.01	< 0.01	1.00	1.00	< 0.01	< 0.01	1.00	1.00
	$L_{C,A}$	< 0.01	< 0.01	1.11	1.09	< 0.01	< 0.01	1.15	1.20
	$L_{C,M}$	< 0.01	< 0.01	0.81	1.02	< 0.01	< 0.01	0.87	1.04
	$L_{C,Y}$	< 0.01	< 0.01	0.97	0.73	< 0.01	< 0.01	0.97	0.74
	$L_{C,A,M}$	< 0.01	< 0.01	0.88	1.11	−0.01	< 0.01	1.02	1.26
	$L_{C,A,Y}$	< 0.01	< 0.01	1.09	0.79	< 0.01	< 0.01	1.13	0.95
	$L_{C,M,Y}$	< 0.01	< 0.01	0.74	0.75	< 0.01	< 0.01	0.81	0.78
	$L_{C,A,M,Y}$	< 0.01	< 0.01	0.82	0.80	< 0.01	< 0.01	0.98	1.01
Scenario 2 ($\beta_1 = 1.2, \theta_2 = 0.3$)	L_C	< 0.01	< 0.01	1.00	1.00	< 0.01	0.01	1.00	1.00
	$L_{C,A}$	< 0.01	< 0.01	1.04	1.08	−0.01	0.01	1.26	1.20
	$L_{C,M}$	< 0.01	< 0.01	1.20	1.07	< 0.01	< 0.01	1.60	1.22
	$L_{C,Y}$	< 0.01	< 0.01	0.81	0.73	< 0.01	< 0.01	0.87	0.77
	$L_{C,A,M}$	< 0.01	< 0.01	1.22	1.15	< 0.01	< 0.01	1.94	1.47
	$L_{C,A,Y}$	< 0.01	< 0.01	0.87	0.78	< 0.01	< 0.01	1.18	0.99
	$L_{C,M,Y}$	< 0.01	< 0.01	0.90	0.79	< 0.01	< 0.01	1.46	1.02
	$L_{C,A,M,Y}$	< 0.01	< 0.01	0.92	0.83	< 0.01	< 0.01	1.84	1.29
Scenario 3 ($\beta_1 = 0.3, \theta_2 = 1.2$)	L_C	< 0.01	< 0.01	1.00	1.00	< 0.01	< 0.01	1.00	1.00
	$L_{C,A}$	< 0.01	< 0.01	1.12	1.09	< 0.01	< 0.01	1.12	1.28
	$L_{C,M}$	< 0.01	< 0.01	0.71	1.01	< 0.01	< 0.01	0.71	1.01
	$L_{C,Y}$	< 0.01	< 0.01	1.00	0.73	< 0.01	< 0.01	1.00	0.73
	$L_{C,A,M}$	< 0.01	< 0.01	0.79	1.10	−0.01	< 0.01	0.80	1.29
	$L_{C,A,Y}$	< 0.01	< 0.01	1.12	0.79	< 0.01	< 0.01	1.12	1.05
	$L_{C,M,Y}$	< 0.01	< 0.01	0.70	0.74	< 0.01	< 0.01	0.71	0.75
	$L_{C,A,M,Y}$	< 0.01	< 0.01	0.79	0.79	−0.01	< 0.01	0.80	1.06

Abbreviations: $L_C - L_{C,A,M,Y}$, adjustment sets, where *C* denotes confounders, *A* the pure predictor of the exposure, *M* the pure predictor of the mediator, and *Y* the pure predictor of the outcome. Relative SE, estimated SE of the estimator with a given adjustment set divided by the estimated SE with L_C , NDE, natural direct effect; NIE, natural indirect effect; β_1 is the exposure coefficient in the mediator model; θ_2 is the mediator coefficient in the outcome model.

computation burden, we did not compute confidence intervals in these additional scenarios. Moreover, we only considered the continuous mediator—continuous outcome case for the scenarios with an exposure–mediator interaction, and only the binary mediator—binary outcome case for exploring the risk ratio and odds ratio scales. Finally, we only employed the *IPW* approach for estimating the NIE and NDE on the risk ratio and odds ratio scales since the *mediation* package we used to implement the *regression* approach yields estimates on the difference scale exclusively. Details concerning these additional simulations are provided in Data S1.

4.4 | Results

In Tables 1 to 4, corresponding to the four cases of mediator and outcome types, we present the bias and relative SE of $(\widehat{NIE}_{reg}, \widehat{NDE}_{reg})$ and $(\widehat{NIE}_{IPW}, \widehat{NDE}_{IPW})$ on the difference scale for each adjustment set in Scenarios 1, 2, and 3. All

TABLE 2 Natural direct and indirect effect estimators from *regression* and *IPW* approaches by adjustment set (*L*) and scenario; binary mediator—continuous outcome case

Scenarios	L	Regression approach				IPW approach			
		Bias		Relative SE		Bias		Relative SE	
		NIE	NDE	NIE	NDE	NIE	NDE	NIE	NDE
Scenario 1 ($\beta_1 = 0.6, \theta_2 = 0.6$)	L_C	< 0.01	< 0.01	1.00	1.00	< 0.01	< 0.01	1.00	1.00
	$L_{C,A}$	< 0.01	< 0.01	1.05	1.08	< 0.01	< 0.01	1.13	1.15
	$L_{C,M}$	< 0.01	< 0.01	0.94	1.00	< 0.01	< 0.01	0.98	1.00
	$L_{C,Y}$	< 0.01	< 0.01	0.92	0.69	< 0.01	< 0.01	0.94	0.69
	$L_{C,A,M}$	< 0.01	< 0.01	0.99	1.08	< 0.01	< 0.01	1.12	1.15
	$L_{C,A,Y}$	< 0.01	< 0.01	0.99	0.75	< 0.01	< 0.01	1.09	0.85
	$L_{C,M,Y}$	< 0.01	< 0.01	0.89	0.69	< 0.01	< 0.01	0.90	0.70
	$L_{C,A,M,Y}$	< 0.01	< 0.01	0.93	0.75	< 0.01	< 0.01	1.06	0.86
Scenario 2 ($\beta_1 = 1.2, \theta_2 = 0.3$)	L_C	< 0.01	< 0.01	1.00	1.00	< 0.01	< 0.01	1.00	1.00
	$L_{C,A}$	< 0.01	< 0.01	0.99	1.08	< 0.01	< 0.01	1.19	1.15
	$L_{C,M}$	< 0.01	< 0.01	1.06	1.01	< 0.01	< 0.01	1.10	1.01
	$L_{C,Y}$	< 0.01	< 0.01	0.77	0.69	< 0.01	< 0.01	0.78	0.69
	$L_{C,A,M}$	< 0.01	< 0.01	1.05	1.08	< 0.01	< 0.01	1.31	1.16
	$L_{C,A,Y}$	< 0.01	< 0.01	0.78	0.74	< 0.01	< 0.01	1.01	0.85
	$L_{C,M,Y}$	< 0.01	< 0.01	0.83	0.69	< 0.01	< 0.01	0.87	0.70
	$L_{C,A,M,Y}$	< 0.01	< 0.01	0.82	0.75	< 0.01	< 0.01	1.12	0.87
Scenario 3 ($\beta_1 = 0.3, \theta_2 = 1.2$)	L_C	< 0.01	< 0.01	1.00	1.00	< 0.01	< 0.01	1.00	1.00
	$L_{C,A}$	< 0.01	< 0.01	1.08	1.08	< 0.01	< 0.01	1.10	1.15
	$L_{C,M}$	< 0.01	< 0.01	0.89	1.00	< 0.01	< 0.01	0.91	1.00
	$L_{C,Y}$	< 0.01	< 0.01	0.97	0.69	< 0.01	< 0.01	1.00	0.69
	$L_{C,A,M}$	< 0.01	< 0.01	0.97	1.08	< 0.01	< 0.01	1.01	1.16
	$L_{C,A,Y}$	< 0.01	< 0.01	1.03	0.75	< 0.01	< 0.01	1.10	0.86
	$L_{C,M,Y}$	< 0.01	< 0.01	0.88	0.69	< 0.01	< 0.01	0.90	0.69
	$L_{C,A,M,Y}$	< 0.01	< 0.01	0.94	0.75	< 0.01	< 0.01	1.00	0.87

Abbreviations: $L_C - L_{C,A,M,Y}$, adjustment sets, where *C* denotes confounders, *A* the pure predictor of the exposure, *M* the pure predictor of the mediator, and *Y* the pure predictor of the outcome. Relative SE, estimated SE of the estimator with a given adjustment set divided by the estimated SE with L_C ; NDE, natural direct effect; NIE, natural indirect effect, β_1 is the exposure coefficient in the mediator model, θ_2 is the mediator coefficient in the outcome model.

confidence interval coverages were found close to 95% (between 92.9% and 98.3%). A slight overcoverage, due to too large confidence intervals produced by the percentile bootstrap, was observed for the *regression* approach for estimating the NIE in the cases where the mediator was binary. The detailed coverage results are omitted from the tables. The results for the additional simulations are reported in Tables S1 to S3. Box plots of the estimates for all simulations are depicted in Figures S1 to S14.

As expected, the results indicate that the NIE and the NDE were estimated with little or negligible bias by both the *regression* and the *IPW* approaches, regardless of the set of covariates that was selected for adjustment.

The main simulation results for the continuous mediator–continuous outcome case (Table 1) are aligned with the theoretical results of Section 3. Overall, regardless of the approach employed and the strength of either θ_2 and β_1 , the SE of \widehat{NIE} and \widehat{NDE} increased when including the pure predictor of the exposure. Oppositely, the SE decreased or remained similar when including the pure predictor of the outcome. For both the *regression* and *IPW* approaches the SE of \widehat{NDE} increased or remained similar when adjusting for the pure predictor of the mediator regardless of the strength of θ_2 and

TABLE 3 Natural direct and indirect effect estimators from *regression* and *IPW* approaches by adjustment set (*L*) and scenario; continuous mediator–binary outcome case

Scenarios	L	Regression approach				IPW approach			
		Bias		Relative SE		Bias		Relative SE	
		NIE	NDE	NIE	NDE	NIE	NDE	NIE	NDE
Scenario 1 ($\beta_1 = 0.6, \theta_2 = 0.6$)	L_C	< 0.01	< 0.01	1.00	1.00	< 0.01	< 0.01	1.00	1.00
	$L_{C,A}$	< 0.01	< 0.01	1.05	1.09	< 0.01	< 0.01	1.13	1.17
	$L_{C,M}$	< 0.01	< 0.01	0.98	1.02	< 0.01	< 0.01	1.01	1.02
	$L_{C,Y}$	< 0.01	< 0.01	0.98	0.90	< 0.01	< 0.01	0.99	0.90
	$L_{C,A,M}$	< 0.01	< 0.01	1.01	1.11	< 0.01	< 0.01	1.20	1.21
	$L_{C,A,Y}$	< 0.01	< 0.01	1.04	0.99	< 0.01	< 0.01	1.11	1.08
	$L_{C,M,Y}$	< 0.01	< 0.01	0.93	0.92	< 0.01	< 0.01	0.98	0.93
	$L_{C,A,M,Y}$	< 0.01	< 0.01	0.96	1.01	< 0.01	< 0.01	1.15	1.12
Scenario 2 ($\beta_1 = 1.2, \theta_2 = 0.3$)	L_C	< 0.01	< 0.01	1.00	1.00	< 0.01	< 0.01	1.00	1.00
	$L_{C,A}$	< 0.01	< 0.01	1.01	1.08	< 0.01	< 0.01	1.15	1.16
	$L_{C,M}$	< 0.01	< 0.01	1.33	1.07	< 0.01	< 0.01	1.59	1.18
	$L_{C,Y}$	< 0.01	< 0.01	0.91	0.89	< 0.01	< 0.01	0.94	0.90
	$L_{C,A,M}$	< 0.01	< 0.01	1.34	1.14	< 0.01	< 0.01	1.83	1.38
	$L_{C,A,Y}$	< 0.01	< 0.01	0.92	0.96	< 0.01	< 0.01	1.08	1.07
	$L_{C,M,Y}$	< 0.01	< 0.01	1.22	0.96	< 0.01	< 0.01	1.55	1.10
	$L_{C,A,M,Y}$	< 0.01	< 0.01	1.23	1.02	< 0.01	< 0.01	1.79	1.31
Scenario 3 ($\beta_1 = 0.3, \theta_2 = 1.2$)	L_C	< 0.01	< 0.01	1.00	1.00	< 0.01	< 0.01	1.00	1.00
	$L_{C,A}$	< 0.01	< 0.01	1.08	1.09	< 0.01	< 0.01	1.09	1.19
	$L_{C,M}$	< 0.01	< 0.01	0.74	1.01	< 0.01	< 0.01	0.74	1.00
	$L_{C,Y}$	< 0.01	< 0.01	1.00	0.92	< 0.01	< 0.01	1.00	0.92
	$L_{C,A,M}$	< 0.01	< 0.01	0.80	1.09	< 0.01	< 0.01	0.83	1.19
	$L_{C,A,Y}$	< 0.01	< 0.01	1.09	1.01	< 0.01	< 0.01	1.09	1.12
	$L_{C,M,Y}$	< 0.01	< 0.01	0.74	0.93	< 0.01	< 0.01	0.74	0.93
	$L_{C,A,M,Y}$	< 0.01	< 0.01	0.80	1.02	< 0.01	< 0.01	0.83	1.13

Abbreviations: $L_C - L_{C,A,M,Y}$, adjustment sets, where C denotes confounders, A the pure predictor of the exposure, M the pure predictor of the mediator, and Y the pure predictor of the outcome. Relative SE, estimated SE of the estimator with a given adjustment set divided by the estimated SE with L_C , NDE, natural direct effect; NIE, natural indirect effect, β_1 is the exposure coefficient in the mediator model, θ_2 is the mediator coefficient in the outcome model.

β_1 . However the impact on the SE of \widehat{NIE} varied according to the values of these parameters. In Scenario 1, where $\theta_2 = \beta_1$ and in Scenario 3, where $\theta_2 > \beta_1$, including the pure predictor of the mediator decreased the SE of \widehat{NIE} and, contrariwise, the SE increased in Scenario 2, where $\theta_2 < \beta_1$. Hence, the lowest SE for estimating the NDE was obtained when adjusting for the pure predictor of the outcome, in addition to confounders. For estimating the NIE, depending on the scenario, either including both the pure predictor of the outcome and the mediator, or only the pure predictor of the outcome, in addition to confounders, produced the estimator with the smallest SE.

Results of the main simulations were overall similar in the binary mediator–continuous outcome, the continuous mediator–binary outcome, the binary mediator–binary outcome cases. We thus only highlight the main differences as compared to the continuous mediator–continuous outcome case without interaction. In both binary outcome cases, the SE of \widehat{NIE}_{IPW} generally increased when adjusting for the pure predictor of the mediator in Scenario 1, instead of decreasing. In both binary mediator cases, the inclusion of the pure predictor of the mediator did not affect much the SE of \widehat{NDE} ,

TABLE 4 Natural direct and indirect effect estimators from *regression* and *IPW* approaches by adjustment set (*L*) and scenario; binary mediator–binary outcome case

Scenarios	L	Regression approach				IPW approach			
		Bias		Relative SE		Bias		Relative SE	
		NIE	NDE	NIE	NDE	NIE	NDE	NIE	NDE
Scenario 1 ($\beta_1 = 0.6, \theta_2 = 0.6$)	L_C	< 0.01	< 0.01	1.00	1.00	< 0.01	< 0.01	1.00	1.00
	$L_{C,A}$	< 0.01	< 0.01	1.02	1.10	< 0.01	< 0.01	1.15	1.16
	$L_{C,M}$	< 0.01	< 0.01	0.97	1.00	< 0.01	< 0.01	1.02	1.00
	$L_{C,Y}$	< 0.01	< 0.01	0.96	0.91	< 0.01	< 0.01	0.96	0.92
	$L_{C,A,M}$	< 0.01	< 0.01	1.03	1.10	< 0.01	< 0.01	1.17	1.16
	$L_{C,A,Y}$	< 0.01	< 0.01	0.98	1.00	< 0.01	< 0.01	1.13	1.08
	$L_{C,M,Y}$	< 0.01	< 0.01	0.95	0.91	< 0.01	< 0.01	0.98	0.92
	$L_{C,A,M,Y}$	< 0.01	< 0.01	0.98	1.00	< 0.01	< 0.01	1.15	1.08
Scenario 2, ($\beta_1 = 1.2, \theta_2 = 0.3$)	L_C	< 0.01	< 0.01	1.00	1.00	< 0.01	< 0.01	1.00	1.00
	$L_{C,A}$	< 0.01	< 0.01	1.00	1.08	< 0.01	< 0.01	1.12	1.14
	$L_{C,M}$	< 0.01	< 0.01	1.07	1.00	< 0.01	< 0.01	1.11	1.01
	$L_{C,Y}$	< 0.01	< 0.01	0.91	0.90	< 0.01	< 0.01	0.92	0.91
	$L_{C,A,M}$	< 0.01	< 0.01	1.09	1.08	< 0.01	< 0.01	1.25	1.14
	$L_{C,A,Y}$	< 0.01	< 0.01	0.90	0.98	< 0.01	< 0.01	1.06	1.05
	$L_{C,M,Y}$	< 0.01	< 0.01	0.99	0.91	< 0.01	< 0.01	1.02	0.92
	$L_{C,A,M,Y}$	< 0.01	< 0.01	0.98	0.98	< 0.01	< 0.01	1.18	1.05
Scenario 3, ($\beta_1 = 0.3, \theta_2 = 1.2$)	L_C	< 0.01	< 0.01	1.00	1.00	< 0.01	-0.01	1.00	1.00
	$L_{C,A}$	< 0.01	< 0.01	1.04	1.12	< 0.01	-0.01	1.11	1.18
	$L_{C,M}$	< 0.01	< 0.01	0.91	1.00	< 0.01	-0.01	0.92	1.00
	$L_{C,Y}$	< 0.01	< 0.01	0.99	0.92	< 0.01	-0.01	1.00	0.92
	$L_{C,A,M}$	< 0.01	< 0.01	0.96	1.12	< 0.01	-0.01	1.02	1.18
	$L_{C,A,Y}$	< 0.01	< 0.01	1.05	1.03	< 0.01	-0.01	1.11	1.11
	$L_{C,M,Y}$	< 0.01	< 0.01	0.92	0.92	< 0.01	-0.01	0.92	0.93
	$L_{C,A,M,Y}$	< 0.01	< 0.01	0.96	1.03	< 0.01	-0.01	1.02	1.11

Abbreviations: $L_C - L_{C,A,M,Y}$, adjustment sets, where *C* denotes confounders, *A* the pure predictor of the exposure, *M* the pure predictor of the mediator, and *Y* the pure predictor of the outcome. Relative SE, estimated SE of the estimator with a given adjustment set divided by the estimated SE with L_C , NDE, natural direct effect; NIE, natural indirect effect, β_1 is the exposure coefficient in the mediator model, θ_2 is the mediator coefficient in the outcome model.

instead of increasing the SE. In 67 replicates of the continuous mediator–binary outcome case of Scenario 2, the *IPW* approach failed to converge (no valid solution was found by the R function).

The results of the additional simulations were also aligned with the other findings. In particular, the relative SEs we observed on the risk ratio and odds ratio scales for the binary mediator–binary outcome case were virtually the same as when using the difference scale. In simulations with a “strong” exposure–mediator interaction in the continuous mediator–continuous outcome case, we observed that the inclusion of pure predictors of the mediator reduced the SE of \widehat{NIE} in Scenario 2, instead of increasing it as in the simulations without an interaction or with a “weak” interaction. We hypothesize that this difference may be caused by the fact that the addition of an exposure–mediator interaction term increased the effect of the mediator on the outcome. As such, the effect of the mediator on the outcome, when the interaction was added, was no longer “weak” as compared to the effect of the exposure on the mediator in Scenario 2.

Although the goal of the simulation was not to compare estimators, we note that the plots depicted in Figures S1 to S14 showcase that the *IPW* approach has a greater variance and is more susceptible to yield outlying estimates than the *regression* approach.

The Monte Carlo SE for estimating the bias was below 0.005, and the one for estimating the coverage of confidence intervals was at most 0.8%.³⁰

Summary of theoretical and empirical results

- Adjusting for pure predictors of the exposure tends to increase the SE of estimators of the natural direct and indirect effects.
- Adjusting for pure predictors of the outcome tends to decrease the SE of estimators of the natural direct and indirect effects.
- Adjusting for pure predictors of the mediator tends to increase the SE of estimators of the natural direct effect.
- When the effect of the mediator on the outcome is large relative to the effect of the exposure on the mediator, adjusting for pure predictors of the mediator generally decreases the SE of estimators of the natural indirect effect.
- When the effect of the mediator on the outcome is small relative to the effect of the exposure on the mediator, adjusting for pure predictors of the mediator generally increases the SE of estimators of the natural indirect effect.

5 | ILLUSTRATION

5.1 | Context

Identification of protective factors for dementia is a priority for public health.³¹ There is accumulating evidence concerning the beneficial effect of physical activity on the risk of dementia.^{32,33} Vitamin D could also present a protective effect on the risk of dementia because of its neuroprotective, anti-inflammatory and antioxidative properties,³⁴ but the evidence is still inconsistent.^{35,36} Furthermore, increased physical activity has been associated with higher concentration of blood vitamin D partly because of a greater sun exposure, the major source of vitamin D.³⁷ The aim of this analysis was to evaluate the mediating effect of plasma vitamin D concentration in the association between physical activity and dementia risk.

5.2 | Method

This mediation analysis was conducted using data from the Canadian Study of Health and Aging.³⁸ This is a cohort study of individuals, representative of the Canadian population aged 65 years or older, with three measurement times at 5-yearly intervals over 10 years (T1, T2, and T3). Participants without dementia at T1 and T2, living in the community, and with frozen blood sample collected at T2 were considered for the current analysis.³⁹ Practice of regular physical activity (yes or no) was assessed in a self-reported risk factor questionnaire at T1 with two other questions related to the physical activity frequency and intensity in the past two weeks.⁴⁰ Vitamin D plasma concentration at T2 (in nmol/L) was evaluated according to plasma 25(OH)D,⁴¹ its main circulating biomarker. The physician and the neuropsychologist made independent diagnoses of dementia at T3 according to published criteria,⁴² followed by a case conference for a consensus diagnosis. The choice to consider physical activity at T1, vitamin D at T2 and dementia at T3 was made to ensure a temporal ordering between the exposure, mediator, and outcome. Based on subject-matter knowledge, a causal diagram for this mediation analysis problem was drawn (see Figure 2). Age, sex, and education were chosen as main confounders. The season of the physical activity measurement, the season of the vitamin D measurement and the presence of the allele e4 on apolipoprotein E gene (ApoE4) were identified as pure predictors of the exposure, mediator and outcome, respectively.

In the sample of participants with a measure of 25(OH)D, 16% had missing data on physical activity, but all the other variables had no missing data. Missing data on physical activity were imputed using chained equations. A single imputation was performed to simplify the illustration. The NIE and NDE were estimated on an additive scale, as in Equation (1), using both the *IPW* and *regression* approaches described in Section 4.2. Physical activity was modeled using a logistic

FIGURE 2 Causal graph depicting the hypothesized mediational relationship between physical activity, vitamin D and dementia

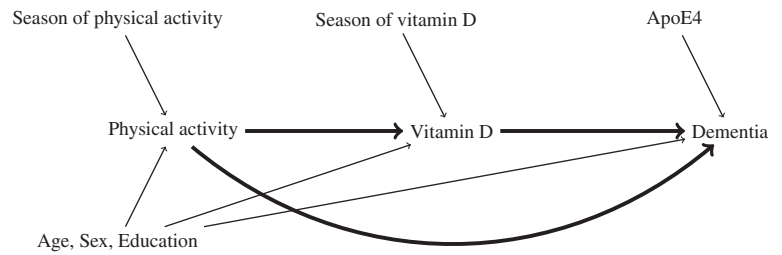


TABLE 5 Descriptive statistics of the subsample of the Canadian Study of Health and Aging employed for the mediation analysis ($n = 461$)

Characteristics	<i>n</i> (%) or mean (SD)
Woman sex, <i>n</i> (%)	265 (57.5)
Years of education, mean (sd)	10.13 (4.03)
Age (years), mean (sd)	80.97 (6.02)
Physical activity, <i>n</i> (%)	312 (67.7)
Log(vitamin D) in nmol/L, mean (sd)	3.73 (0.58)
Dementia, <i>n</i> (%)	94 (20.4)
ApoE4, <i>n</i> (%)	98 (21.3)
Season of physical activity, <i>n</i> (%)	
Winter	80 (17.4)
Spring	135 (29.3)
Summer	107 (23.2)
Fall	139 (30.2)
Season of vitamin D, <i>n</i> (%)	
Winter	53 (11.5)
Spring	165 (35.8)
Summer	155 (33.6)
Fall	88 (19.1)

regression model and the log of vitamin D was modeled using a linear regression model. This log-transformation was applied because the distribution of vitamin D was positively skewed. Dementia was modeled using a logistic regression model for the *regression* approach and a generalized linear model with a binomial distribution and identity link for the *IPW* approach. As in the simulation study, the analyses were either adjusted for confounders only (L_C), or for confounders and pure predictors of the exposure ($L_{C,A}$), the mediator ($L_{C,M}$), the outcome ($L_{C,Y}$), the exposure and mediator ($L_{C,A,M}$), the exposure and outcome ($L_{C,A,Y}$), the mediator and outcome ($L_{C,M,Y}$), or the exposure, mediator and outcome ($L_{C,A,M,Y}$). SEs were estimated using 10 000 nonparametric bootstrap replicates to reduce the Monte Carlo error in the estimation.

5.3 | Results

In our sample of 461 participants, a majority of them were female (57.5%), the mean age at baseline was 81 years, 67.7% practiced regular physical activity and 20.4% had developed dementia by T3 (Table 5). The results for NDE and NIE estimates are consistent across approaches and adjustment sets (Table 6). Regular physical activity is associated with a direct risk difference of dementia of approximately -16%, but the indirect association is close to 0. In this illustration, the SE estimates do not differ much according to the adjustment set considered, but the differences are overall consistent with what is expected based on our theoretical and empirical findings. That is, the estimated SE is generally lower when adjusting for the pure predictor of the outcome, and generally greater when adjusting for the pure predictor of the exposure or the mediator. Tables S4 to S6 provide results for the regressions of physical activity, vitamin D, and dementia.

TABLE 6 Natural indirect (NIE) and direct (NDE) effect estimate on the difference scale and SE for the relationship between regular physical activity (yes/no) and dementia through vitamin D levels in the Canadian Study of Health and Aging ($n = 461$)

Adjustment	Regression approach				IPW approach			
	NIE	SE NIE	NDE	SE NDE	NIE	SE NIE	NDE	SE NDE
L_C	0.0003	0.0032	-0.1658	0.0422	0.0007	0.0034	-0.1552	0.0437
$L_{C,A}$	0.0001	0.0033	-0.1655	0.0440	0.0007	0.0034	-0.1602	0.0438
$L_{C,M}$	0.0012	0.0035	-0.1744	0.0429	0.0017	0.0040	-0.1649	0.0435
$L_{C,Y}$	0.0004	0.0032	-0.1603	0.0422	0.0004	0.0032	-0.1505	0.0435
$L_{C,A,M}$	0.0010	0.0036	-0.1728	0.0430	0.0015	0.0040	-0.1671	0.0449
$L_{C,A,Y}$	0.0002	0.0033	-0.1610	0.0420	0.0004	0.0034	-0.1551	0.0448
$L_{C,M,Y}$	0.0013	0.0036	-0.1692	0.0430	0.0012	0.0040	-0.1587	0.0442
$L_{C,A,M,Y}$	0.0011	0.0036	-0.1685	0.0429	0.0011	0.0040	-0.1610	0.0442

Abbreviations: $L_C - L_{C,A,M,Y}$, adjustment sets, where C denotes confounders, A the pure predictor of the exposure, M the pure predictor of the mediator, and Y the pure predictor of the outcome.

5.4 | Discussion of substantive results

The result of the application showed robust associations between regular physical activity and reduction of dementia risk at follow-up, which is in line with several previous longitudinal studies.^{32,40} Increasing physical activity is one lifestyle recommendation for the prevention of dementia among older adults.³¹ However, no indirect association was observed for the plasma vitamin D concentration. To our knowledge no previous study has evaluated the mediating effect of vitamin D in the association between physical activity and dementia risk. The direct association of physical activity on vitamin D was observed in some longitudinal studies.^{43,44} This association was also observed in our analysis (Table S5). This confirms the hypothesis that higher physical activity levels could result in a greater sun exposure, which in turn, increases blood vitamin D concentration.³⁷ The protective effect of vitamin D on the risk of dementia was not confirmed in our study. As reported in a previous paper, methodological issues have to be taken into account in the interpretation of the result.³⁹ First, a survival bias might have occurred. As participants were aged on average of 81 years at baseline, those with lower vitamin D concentrations who did not die or develop dementia before the beginning of the study might have shown some other protective factors for dementia. This selection bias is frequent in studies on older adults.⁴⁵ The selection of subjects without dementia at T2 and alive until the end of the study may also induce a selection bias. This potential bias has been previously addressed in detail,³⁹ and could have been controlled to some extent using inverse probability of censoring weights, for example. We did not apply such a correction here in order to simplify the illustration and because the `mediation` package we used to implement the *regression* approach does not allow weighting of observations. In addition, the association could have been confounded by vitamin D supplementation. Participants with more comorbidities, showing vitamin D deficiency before the beginning of the study could have been prescribed vitamin D supplements for the treatment of osteoporosis for example, which could be reflected in the 25(OH)D measurement.⁴⁶ Sensitivity analysis adjusted for vitamin D supplementation gave results in the same direction, but our self-reported measure for vitamin supplement was subject to misclassification bias and under-reporting. There are likely other unmeasured covariates, such as frailty, dietary intake and neighborhood walkability that may induce residual confounding bias. Finally, effect modification by sex could be present. Several sex and gender differences are present for cognitive function and dementia risk,⁴⁷ and women also tend to be more subject to vitamin D deficiency.⁴⁶ All these issues should be taken into account in future studies.

6 | DISCUSSION

As far as we know, there has been no previous work regarding the impact of adjusting for pure predictors of the outcome, exposure or mediator in mediation analysis models. The purpose of this paper was to address this gap in knowledge. The most basic mediation analysis uses main term linear regressions for modeling the mediator and the outcome given the mediator, and corresponding natural direct and indirect effects are simple functions of the regression coefficients of

these models. In this simple context, we established, using standard ordinary least squares results, that adjusting for pure predictors of the outcome increases the precision of NIE and NDE estimators, whereas adjusting for pure predictors of the exposure decreases the precision of these estimators. We have also shown that adjusting for pure predictors of the mediator decreases the precision of the NDE estimator, but has a variable impact on the precision of the NIE estimator, notably depending on the strength of the exposure–mediator and mediator–outcome associations. A simulation study was then conducted to further examine the impact of pure predictor adjustment for two causal approaches implemented in the R software for the estimation of natural mediation effects. The empirical results we obtained were in line with our analytical results, suggesting that the theoretical insight we developed extends beyond the simple case we considered to more complex estimation approaches and types of outcomes and mediators. These results are additionally consistent with those of the current literature regarding the effect of adjusting for pure predictors of the outcome and exposure when estimating a total exposure effect (see References 5,6,8). Finally, we illustrated those results in a real-data analysis concerning the intermediate role of vitamin D in the relationship between physical activity and dementia. Hypothesized pure predictors of the exposure, mediator and outcome, in addition to confounders, were identified based on substantive knowledge. In this illustration, the SE estimates did not vary much according to which variables were adjusted for, but the differences were overall consistent with our expectations. However, it is important to note that the observed differences in the SE estimates may partly reflect sampling variability instead of true differences, especially because of the small sample size. More fundamentally, our real data analysis has revealed that the impact of the adjustment set on the precision of mediation analysis estimators can be limited if pure predictors are weaker than those considered in our simulation studies.

Our findings suggest that analysts should adjust for pure predictors of the outcome, in addition to true confounders, and avoid adjustment for pure predictors of the exposure when performing mediation analyses. Since adjustment for pure predictors of the mediator increases the variance of the NDE estimator and may either increase or decrease the variance of the NIE estimator, it may seem advisable to avoid adjusting for such variables. In situations where it is expected that adjusting for pure predictors of the mediator should improve the precision of the NIE estimator (eg, the anticipated mediator–outcome association is much stronger than the exposure–mediator association), an alternative strategy may be to estimate the NDE and the NIE separately. When estimating the NDE, pure predictors of the mediators would be avoided, whereas these variables would be included when estimating the NIE. It is important to note that the sum of the estimated NIE and NDE may not be exactly equal to the total effect estimate when employing such a strategy, either because of random fluctuations or to model misspecifications. If it is unknown a priori whether the pure predictors of the mediator would increase or decrease the variance of the NIE estimator, it may be tempting to consider both an estimate adjusted and not adjusted for these variables and choose as a final estimate the one with the lowest estimated variance. However, we advise against doing this because the observed difference in the estimated variances may be due to sampling variations and not reflect a real difference in the (unknown) true variances. Choosing the estimator with the lowest estimated variance may thus lead to underestimating the true variance and yield invalid inferences.

Some limitations to consider concerning the present paper are that the theoretical results were derived from considering a simple setting and the simulation scenarios were based on a single DAG. While it could be of interest to derive theoretical results for more general settings, corresponding closed-form variance formulas are generally not available or it is difficult to determine the impact of adjusting for specific variables in these formulas. In this aspect, our work shares the same limitation as previous works concerning the impact of adjusting for pure predictors of the exposure or outcome when estimating a total effect. Additional simulation studies may be useful to determine if there are situations where our recommendations do not apply. In particular, in small sample or high-dimensional settings (where the number of covariates is close to the sample size), it is unclear if adjusting for pure predictors of the outcome, in addition to true confounders, would reduce the variance of estimators. Simulation studies inspired by real data settings would also be helpful to better inform on the potential practical benefits of covariate selection in mediation analysis. In this regard, the plasmode simulation framework, which combines real data with synthetic data, may prove particularly interesting.⁴⁸ An important consideration when conducting such simulations is that realistic scenarios are likely to vary from one field of application to the other. There are multiple directions in which our work could be extended either through theoretical derivations or additional simulation results. Notably, it would be interesting to consider other estimands than the NIE and NDE, cases with multiple mediators and time-to-event outcomes. A more thorough investigation of scenarios with interactions between exposure, mediator and confounders would also be warranted.

Lastly, we point out that the natural effect decomposition has been criticized by some authors because the NIE and NDE cannot be identified without making unverifiable assumptions, even in a randomized trial (see, eg, Naimi et al²⁰ and references therein). In addition, the natural decomposition does not inform on the effect of any public health or clinical intervention.²⁰ However, others have argued that natural effects may nonetheless be of interest, notably when the goal is

to describe natural mechanisms.^{49,50} Moreover, in the simplest linear regression setting that we considered for deriving theoretical results, the estimators of the natural decomposition also correspond to less controversial estimands, such as the controlled direct effect or interventional effects.^{17,19} This correspondence leads us to hypothesize that our conclusions are likely to extend to other mediation estimands more generally. In conclusion, we believe our results are an important step to better understand variable selection in mediation analysis and will prove helpful to guide practitioners performing such analyses.

ACKNOWLEDGEMENTS

This research was supported by Fondation du CHU de Québec [#2710 to DT], the Natural Sciences and Engineering Research Council of Canada [#2016-06295 to DT, #2020-05473 to GL], the Fonds de recherche du Québec—Santé [#265385 to DT]. DT and GL are Fonds de recherche du Québec—Santé Chercheurs-Boursiers.

DATA AVAILABILITY STATEMENT

The R code used for performing the simulation study is available as Data S1.

ORCID

Geneviève Lefebvre  <https://orcid.org/0000-0002-7155-8218>

Denis Talbot  <https://orcid.org/0000-0003-0431-3314>

REFERENCES

- Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology*. 1999;10(1):37-48.
- VanderWeele TJ, Shpitser I. A new criterion for confounder selection. *Biometrics*. 2011;67(4):1406-1413.
- VanderWeele TJ. Principles of confounder selection. *Eur J Epidemiol*. 2019;34:211-219.
- Witte J, Didelez V. Covariate selection strategies for causal inference: classification and comparison. *Biom J*. 2019;61(5):1270-1289.
- Pearl J. Invited commentary: understanding bias amplification. *Am J Epidemiol*. 2011;174(11):1223-1227.
- Lefebvre G, Delaney JA, Platt RW. Impact of mis-specification of the treatment model on estimates from a marginal structural model. *Stat Med*. 2008;27(18):3629-3642.
- De Luna X, Waernbaum I, Richardson TS. Covariate selection for the nonparametric estimation of an average treatment effect. *Biometrika*. 2011;98(4):861-875.
- Shortreed SM, Ertefaie A. Outcome-adaptive lasso: variable selection for causal inference. *Biometrics*. 2017;73(4):1111-1122.
- VanderWeele T. *Explanation in Causal Inference: Methods for Mediation and Interaction*. New York, NY: Oxford University Press; 2015.
- Rubin DB. Causal inference using potential outcomes: design, modeling decisions. *J Am Stat Assoc*. 2005;100(469):322-331.
- VanderWeele TJ, Vansteelandt S. Conceptual issues concerning mediation, interventions and composition. *Stat Interface*. 2009;2(4):457-468.
- VanderWeele TJ, Vansteelandt S. Odds ratios for mediation analysis for a dichotomous outcome. *Am J Epidemiol*. 2010;172(12):1339-1348.
- Robins JM, Richardson TS. Alternative graphical causal models and the identification of direct effects. *Finding the Determinants of Disorders and Their Cures*. Oxford, UK: Oxford University Press; 2010:103-158.
- Vansteelandt S, VanderWeele TJ. Natural direct and indirect effects on the exposed: effect decomposition under weaker assumptions. *Biometrics*. 2012;68(4):1019-1027.
- Lok JJ. Defining and estimating causal direct and indirect effects when setting the mediator to specific values is not feasible. *Stat Med*. 2016;35(22):4008-4020.
- VanderWeele TJ, Tchetgen EJT. Mediation analysis with time varying exposures and mediators. *J R Stat Soc Ser B Stat Methodol*. 2017;79(3):917-938.
- Vansteelandt S, Daniel RM. Interventional effects for mediation analysis with multiple mediators. *Epidemiol*. 2017;28(2):258-265.
- Stensrud MJ, Young JG, Didelez V, Robins JM, Hernán MA. Separable effects for causal inference in the presence of competing events. *J Am Stat Assoc*. 2020;1-9.
- Valeri L, VanderWeele TJ. Mediation analysis allowing for exposure–mediator interactions and causal interpretation: theoretical assumptions and implementation with SAS and SPSS macros. *Psychol Methods*. 2013;18(2):137-150.
- Naimi AI, Kaufman JS, MacLehose RF. Mediation misgivings: ambiguous clinical and public health interpretations of natural direct and indirect effects. *Int J Epidemiol*. 2014;43(5):1656-1661.
- Andrews RM, Didelez V. Insights into the ϵ -cross-world independence assumption of causal mediation analysis; 2020. arXiv preprint arXiv:2003.10341.
- Lange T, Hansen KW, Sørensen R, Galatius S. Applied mediation analyses: a review and tutorial. *Epidemiol Health*. 2017;39:e2017035-e2017030.
- Imai K, Keele L, Tingley D. A general approach to causal mediation analysis. *Psychol Methods*. 2010;15(4):309-334.
- VanderWeele TJ. Marginal structural models for the estimation of direct and indirect effects. *Epidemiology*. 2009;20(1):18-26.

25. Lange T, Vansteelandt S, Bekaert M. A simple unified approach for estimating natural direct and indirect effects. *Am J Epidemiol*. 2012;176(3):190-195.
26. Sobel ME. Asymptotic confidence intervals for indirect effects in structural equation models. *Sociol Methodol*. 1982;13:290-312.
27. O'Brien RM. A caution regarding rules of thumb for variance inflation factors. *Qual Quant*. 2007;41(5):673-690.
28. Tingley D, Yamamoto T, Hirose K, Keele L, Imai K. Mediation: r package for causal mediation analysis. *J Stat Softw*. 2014;59(5):1-38.
29. Steen J, Loeys T, Moerkerke B, Vansteelandt S. medflex: an R package for flexible mediation analysis using natural effect models. *J Stat Softw*. 2017;76(11):1-46.
30. Morris TP, White IR, Crowther MJ. Using simulation studies to evaluate statistical methods. *Stat Med*. 2019;38(11):2074-2102.
31. Livingston G, Sommerlad A, Orgeta V, et al. Dementia prevention, intervention, and care. *Lancet*. 2017;390(10113):2673-2734.
32. Lee J. The relationship between physical activity and dementia: a systematic review and meta-analysis of prospective cohort studies. *J Gerontol Nurs*. 2018;44(10):22-29.
33. Najjar J, Östling S, Gudmundsson P, et al. Cognitive and physical activity and dementia: a 44-year longitudinal population study of women. *Neurology*. 2019;92(12):e1322-e1330.
34. Aspell N, Lawlor B, O'Sullivan M. Is there a role for vitamin D in supporting cognitive function as we age? *Proc Nutr Soc*. 2018;77(2):124-134.
35. Chai B, Gao F, Wu R, et al. Vitamin D deficiency as a risk factor for dementia and Alzheimer's disease: an updated meta-analysis. *BMC Neurol*. 2019;19(284):1-11.
36. Yang K, Chen J, Li X, Zhou Y. Vitamin D concentration and risk of Alzheimer disease: a meta-analysis of prospective cohort studies. *Medicine (Baltimore)*. 2019;98(35):1-6.
37. Fernandes MR, Barreto WdR Jr. Association between physical activity and vitamin D: a narrative literature review. *Rev Assoc Med Bras*. 2017;63(6):550-556.
38. Lindsay J, Sykes E, McDowell I, Verreault R, Laurin D. More than the epidemiology of Alzheimer's disease: contributions of the Canadian study of health and aging. *Can J Psychiatry*. 2004;49(2):83-91.
39. Duchaine CS, Talbot D, Nafti M, et al. Vitamin D status, cognitive decline and incident dementia: the Canadian study of health and aging. *Can J Public Health*. 2020;111:312-321.
40. Laurin D, Verreault R, Lindsay J, MacPherson K, Rockwood K. Physical activity and risk of cognitive impairment and dementia in elderly persons. *Arch Neurol*. 2001;58(3):498-504.
41. Jones G. Metabolism and biomarkers of vitamin D. *Scand J Clin Lab Invest*. 2012;72(sup243):7-13.
42. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 4th ed. Washington, DC: American Psychiatric Association; 1994.
43. Scott D, Blizzard L, Fell J, Ding C, Winzenberg T, Jones G. A prospective study of the associations between 25-hydroxy-vitamin D, sarcopenia progression and physical activity in older adults. *Clin Endocrinol*. 2010;73(5):581-587.
44. McKibben RA, Zhao D, Lutsey PL, et al. Factors associated with change in 25-hydroxyvitamin D levels over longitudinal follow-up in the ARIC study. *J Clin Endocrinol*. 2016;101(1):33-43.
45. Mayeda ER, Tchetgen Tchetgen EJ, Power MC, et al. A simulation platform for quantifying survival bias: an application to research on determinants of cognitive decline. *Am J Epidemiol*. 2016;184(5):378-387.
46. Greene-Finestone LS, Berger D, De Groh M, et al. 25-Hydroxyvitamin D in Canadian adults: biological, environmental, and behavioral correlates. *Osteoporos Int*. 2011;22(5):1389-1399.
47. Malpetti M, Ballarini T, Presotto L, et al. Gender differences in healthy aging and Alzheimer's dementia: a 18F-FDG-PET study of brain and cognitive reserve. *Hum Brain Mapp*. 2017;38(8):4212-4227.
48. Gadbury GL, Xiang Q, Yang L, Barnes S, Page GP, Allison DB. Evaluating statistical methods using plasmide data sets in the age of massive public databases: an illustration using false discovery rates. *PLoS Genet*. 2008;4(6):e1000098.
49. Pearl J. Direct and indirect effects. Paper presented at: Proceedings of the 17th Conference on Uncertainty in Artificial Intelligence; 2001Morgan Kaufmann, San Francisco, CA.
50. Schwartz S, Hafeman D, Campbell U, Gattton N. Authors' response to: commentary gilding the black box. *Int J Epidemiol*. 2010;39:1399-1401.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Diop A, Lefebvre G, Duchaine CS, Laurin D, Talbot D. The impact of adjusting for pure predictors of exposure, mediator, and outcome on the variance of natural direct and indirect effect estimators. *Statistics in Medicine*. 2021;40:2339-2354. <https://doi.org/10.1002/sim.8906>

APPENDIX

We provide details to explain why conditional independence $M \perp\!\!\!\perp L_1 | \{A, L_4, L_5, L_6, L_7\}$ implies that including L_1 in \mathbf{L} has no impact on the population parameters $\sigma_{M|A,\mathbf{L}}^2$ or $R_{M|A,\mathbf{L}}^2$. First, we can write the regression model $\mathbb{E}[M|A, L_1, L_4, L_5, L_6, L_7] = \beta'_0 + \beta'_1 A + (L_1, L_4, L_5, L_6, L_7) \boldsymbol{\beta}'_2$. Because $M \perp\!\!\!\perp L_1 | \{A, L_4, L_5, L_6, L_7\}$, $\mathbb{E}[M|A, L_1, L_4, L_5, L_6, L_7] = \mathbb{E}[M|A, L_4, L_5, L_6, L_7] = \beta_0 + \beta_1 A + (L_4, L_5, L_6, L_7) \boldsymbol{\beta}_2$. Since the preceding equality must hold for all values of $\{A, L_1, L_4, L_5, L_6, L_7\}$, it follows that the true coefficient associated with L_1 in $\mathbb{E}[M|A, L_1, L_4, L_5, L_6, L_7]$ must be zero. In other words, the two (true) regression models are identical. As such, their population residual variance and their population coefficient of determination must coincide. We conclude that including L_1 in \mathbf{L} does not affect $\sigma_{M|A,\mathbf{L}}^2$ or $R_{M|A,\mathbf{L}}^2$.