Estimation of total mediation effect for a binary trait in a case-control study for high-dimensional omics mediators[*][†]

Zhiyu Kang[‡] Li Chen[§] Peng Wei[¶] Zhichao Xu[‖] Chunlin Li[**]
Tianzhong Yang[††]

Abstract

Mediation analysis helps uncover how exposures impact outcomes through intermediate variables. Traditional mean-based total mediation effect measures can suffer from the cancellation of opposite component-wise effects and existing methods often lack the power to capture weak effects in high-dimensional mediators. Additionally, most existing work has focused on continuous outcomes, with limited attention to binary outcomes, particularly in case-control studies. To fill in this gap, we propose an R^2 total mediation effect measure under the liability framework, providing a causal interpretation and applicable to various high-dimensional mediation models. We develop a cross-fitted, modified Haseman-Elston regression-based estimation procedure tailored for case-control studies, which can also be applied to cohort studies with reduced efficiency. Our estimator remains consistent with non-mediators and weak effect sizes in extensive simulations. Theoretical justification on consistency is provided under mild conditions. In the Women's Health Initiative of 2150 individuals, we found that 89% (CI: 73% - 91%) of the variation in the underlying liability for coronary heart disease associated with BMI can be explained by metabolomics.

1 Introduction

Mediation analysis is a statistical tool used to explore how an exposure affects an outcome through one or more intermediate variables, known as mediators. With recent advances in high-throughput technologies, high-dimensional mediation analysis has emerged as an area of scientific interest. In this paper, we focus on estimating the total mediation effect, which provides critical insights for downstream analyses. Specifically, our method is motivated by the Women's

[*]Z. Kang and L. Chen have contributed equally.

[†]Correspondence: C. Li (chunlin@iastate.edu) and T. Yang (yang3704@umn.edu).
[‡]Division of Biostatistics and Health Data Science, University of Minnesota, MN 55455.
[§]School of Statistics, University of Minnesota, MN 55455.
[¶]Department of Biostatistics, University of Texas MD Anderson Cancer Center, TX 77030.
[‖]Department of Biostatistics, University of Texas MD Anderson Cancer Center, TX 77030.
[**]Department of Statistics, Iowa State University, IA 50011.
[††]Division of Biostatistics and Health Data Science, University of Minnesota, MN 55455.

Health Initiative (WHI) nested case-control substudy, where we are interested in examining the relationship between body mass index (BMI) and the risk of incident coronary heart disease (CHD) through metabolomic profiles in postmenopausal women. The mediation analysis of WHI dataset poses two main statistical challenges: (1) the preliminary analysis suggested the existence of many weak mediators with small effect sizes (Figure 4), which cannot be identified individually and (2) the necessity to adjust ascertainment bias from the case-control design, an issue that has received less attention in the mediation literature [1, 2]. These challenges highlight the need for a more robust framework for high-dimensional mediation analysis with weak mediators and binary outcomes.

The traditional total mediation measures struggle to effectively handle high-dimensional mediators, a challenge that arises for both continuous and binary outcomes. Most of these measures are variants of the mean-based product-of-cofficient (POC) estimand, suffering from cancellation in the presence of bidirectional mediation effects, a frequent scenario in high-dimensional settings [3]. In addition, many high-dimensional approaches rely on strong sparsity assumptions, which cannot detect mediators with small and non-sparse effect sizes [4, 5]. Some methods define mediation effect based on latent variables, relaxing sparsity requirements on the observed scale but reducing interpretability as a trade-off. [6].

There is growing literature on estimating the total mediation effect under high-dimensional settings in the presence of bidirectional component-wise mediation effects. Yang et al. extended the R^2 measure originally proposed by Fairchild et al. to accommodate multiple and high-dimensional mediators under a mixed model framework [3, 7]. This method captures the non-zero total mediation effect by quantifying the variance in outcome that is attributable to both the exposure and the mediators, attempting to address bidirectional effects issues in high-dimensional mediator settings. Subsequently, this measure was extended to binary outcomes using McFadden's R^2 for its stability and independence from disease prevalence [8]. However, these R^2 measures only capture the associative relationships rather than causality. In another line of work, Song et al. proposed a L_2 norm measure of the mediation effect while naively treating binary outcomes as continuous [9]. However, both the use of the L_2 norm and the treatment of binary outcomes as continuous lack theoretical support, limiting the interpretability of the method. Similarly, this L_2 norm measure does not have causal interpretation.

In response to these challenges, we propose a novel R^2 mediation effect measure that provides a causal interpretation for binary outcomes. Specifically, we propose a two-stage estimation procedure to account for ascertainment bias in case-control studies and to address non-mediators and weak effects. We further modify an heritability estimation method, phenotype-correlation genotype-correlation (PCGC) regression, to estimate consistent variance-components under case-control studies [10]. Although our procedure is designed for case-control studies, it remains applicable to cohort studies. Extensive simulation studies showcased the good performance of our proposed method under varying disease prevalence, weak effects, and model misspecification, consistently outperforming existing high-dimensional mediation models. Furthermore, we derived the asymptotic properties of the estimation in presence of non-mediators and noises. Applied to the motivating WHI case-control dataset, our method effectively capture a good amount of mediation effect of metabolites from BMI to CHD.

The rest of the paper is organized as follows. In Section 2, we review the existing mediation effect measures and highlight their limitations. We then introduce our proposed $R^2_{med;causal}$ mediation measure and discuss its causal interpretation. Section 3 details the estimation procedure and addresses implementation considerations specific to case-control designs. In Section 4, we

evaluate the performance of our method through extensive simulation studies, comparing it with existing approaches. Section 5 presents the application to WHI. In Section 6, we discuss the broader implications of our work and outline future research directions. The theoretical properties of the proposed measure in a high-dimensional context are provided in the Supplementary Materials.

2 Methods

A liability-based mediation model from a single exposure to a single binary outcome consists of the following structural equations,

$$\boldsymbol{M} = \boldsymbol{\alpha} X + \boldsymbol{\Phi} \boldsymbol{C} + \boldsymbol{\xi} \tag{1}$$

$$l = \gamma X + \boldsymbol{\beta}^\top \boldsymbol{M} + \boldsymbol{\theta}^\top \boldsymbol{C} + \varepsilon \tag{2}$$

$$Y = \mathbf{1}(l > t), \tag{3}$$

where $\boldsymbol{M}$ is a $p \times 1$ vector representing mediator, X is an exposure variable, $\boldsymbol{C}$ is a $q \times 1$ vector representing covariates, l is the underlying latent variable, Y is the observed binary outcome, $\mathbf{1}(\cdot)$ is the indicator function, t is the threshold that control K, the prevalence of trait in the population. The coefficient $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_p)^\top$ captures the relationship between the exposure and mediator, $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_p)^\top$ captures the relationship between mediators and outcome after conditioning on the exposure variable, γ is the direct effect of exposure variable, and $\boldsymbol{\Phi} : p \times q$ and $\boldsymbol{\theta} : q \times 1$ are the coefficients of covariates in equation 1 and 2. ε, $\boldsymbol{\xi} = (\xi_1, \xi_2, \ldots, \xi_p)^\top$ are normally distributed residuals.

2.1 Existing Mediation Effect Measures

In a traditional mediation model, the total effect of an exposure on an outcome is decomposed into a direct effect and an mediation effect mediated through one or more variables based on the structural equations first proposed in 1986[11]. The mediation effect can be estimated either by subtracting the direct effect from the total effect or by computing the inner product of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ (i.e., POC) when both the outcome and mediators are continuous. For binary outcomes, the left-hand side of equation 2 is usually replaced with a logit link function [12]. In such cases, the difference-in-coefficients method may encounter the non-collapsibility problem, making the POC approach preferable [13, 14]. The POC measure is also used in the high-dimensional mediation analysis methods, such as HIMA and HDMA [4, 5]. Furthermore, to quantify the fraction of the total effect of an exposure on the outcome that is transmitted through mediators, the traditional proportion-mediated measure is given by $\boldsymbol{\alpha}^\top \boldsymbol{\beta} / (\gamma + \boldsymbol{\alpha}^\top \boldsymbol{\beta})$[12]. These mean-based measures have lots of merits. However, it can be misleading at times because bidirectional component-wise mediation effect may exist under high-dimensional setting, thus leading to cancellation in the POC and proportion measures.

In recent years, causal mediation analysis in the counterfactual framework has been of great interest, which generalizes to any mediation model [15, 16, 17]. For a continuous outcome Y, the natural indirect effect (NIE) and the natural direct effect (NDE) are defined as:

$$\text{NIE} = E[Y(x, M(x))] - E[Y(x, M(x'))],$$
$$\text{NDE} = E[Y(x, M(x'))] - E[Y(x', M(x'))],$$

where $E[Y(x,M(x))]$ represents the expected outcome when both exposure and mediator take their natural values under $X = x$, $E[Y(x,M(x'))]$ represents the expected outcome when the exposure is set to x, but the mediator is fixed at the level it would take under $X = x'$, and $E[Y(x',M(x'))]$ represents the expected outcome when both exposure and mediator take their natural values under $X = x'$. For continuous outcomes, NIE coincide with the POC measure NIE = POC = $\boldsymbol{\alpha}^\top\boldsymbol{\beta}$ if interaction is absent. For the mediation and direct effects to have a causal interpretation, the following assumptions must be met[18]:

1. No unmeasured confounding of the exposure-mediator effect.
2. No unmeasured confounding of the mediator-outcome effect.
3. No unmeasured confounding of the exposure-outcome effect.
4. No mediator-outcome confounder that is itself affected by the exposure.

For a binary outcome, when the outcome is rare and these assumptions are satisfied, logistic regression can be used to estimate the causal mediation effect on an odds ratio scale as: $\text{OR}^{\text{NIE}} = \exp(\boldsymbol{\alpha}^\top\boldsymbol{\beta})$[16]. When the outcome is not rare, log-binomial model is fitted to estimate the causal mediation effect on a risk ratio (RR) scale $\text{RR}^{\text{NIE}} = \exp(\boldsymbol{\alpha}^\top\boldsymbol{\beta})$ [18]. Additionally, Gaynor, Schewartz, and Lin proposed a new estimator that utilizes a probit approximation to the logit, relaxing the rare disease assumption when using logistic regression [19]. However, regardless of the mediation framework and specific regression models applied, this estimation of the mediation effect is fundamentally based on the POC calculation, thus suffering from cancellation of the opposite component-wise effects.

To address these limitations, BAMA extends causal mediation analysis to high-dimensional scenarios by incorporating continuous shrinkage priors within a Bayesian framework [9]. Recognizing the potential drawbacks of POC measure, BAMA proposes an alternative L_2 norm for summarizing the component-wise mediation effects: $\text{NIE}_{BAMA} = \sum_{i=1}^{p}(\alpha_i\beta_i)^2$. While this modification helps reduce cancellation, it lacks rigorous theoretical support for the causal interpretation.

Another alternative mediation effect measure is the R^2 measure, which was first proposed in a single-mediator model under linear settings by Fairchild[7]. Yang et al. extended this measurement to high-dimensional mediation analysis for a continuous outcome Y and high-dimensional omics mediator $\boldsymbol{M}$ [3]. The measure is expressed as:

$$R^2_{med} = R^2_{Y,X} + R^2_{Y,\boldsymbol{M}} - R^2_{Y,\boldsymbol{M}X},$$

where $R^2_{Y,X} = [Var(Y) - Var(Y|X)]/Var(Y)$, $R^2_{Y,\boldsymbol{M}} = [Var(Y) - Var(Y|\boldsymbol{M})]/Var(Y)$, and $R^2_{Y,\boldsymbol{M}X} = [Var(Y) - Var(Y|X,\boldsymbol{M})]/Var(Y)$. The proposed method provides a robust, interpretable second-moment based measure that has low bias and variance for continuous outcomes.

2.2 A novel R^2 measurement with causal interpretation

The R^2 measure for high-dimensional mediators can be extended from a continuous outcome to a binary outcome under the liability scale. Under the structural equation model specified by equation 1, 2 and 3, the measure is defined as:

$$R^2_{med} = R^2_{l,X} + R^2_{l,\boldsymbol{M}} - R^2_{l,\boldsymbol{M}X}.$$

However, this definition of R^2_{med} does not inherently take into account the directional relationship between X and $\boldsymbol{M}$. Specifically, exchanging X and $\boldsymbol{M}$ will not change the formulas except for notation difference, even though mediation analysis explicitly assumes a causal pathway from X to Y and from X to $\boldsymbol{M}$ to Y. To resolve this issue and explicitly incorporate the casual direction, we introduce a do-operator in our R^2 measure within Pearl's Structural Causal Framework [20]. The new R^2 measure is formulated as:

$$R^2_{med;causal} = R^2_{l,do(X)} + R^2_{l,do(\boldsymbol{M})} - R^2_{l,do(\boldsymbol{M}X)}. \tag{4}$$

We define

$$R^2_{l,do(\boldsymbol{M})} = var\left[\frac{E(l|do(M \sim M'))}{Var(l)}\right], \tag{5}$$

where $\boldsymbol{M}'$ is identically distributed as M but independent of all other variables in the model and the do-operator is a stochastic intervention of M. This definition isolates the causal effect from $\boldsymbol{M}$ to l by removing the indirect contribution of X via the path that involves $\boldsymbol{M}$. This ensures that $R^2_{l,do(\boldsymbol{M})}$ reflects the direct causal effect of $\boldsymbol{M}$ on l, unlike $R^2_{l,\boldsymbol{M}}$, which may confound this with X's influence. Under no unmeasured confounder assumption, the $R^2_{med;causal}$ can be interpreted as the variance of outcome *causally* explained by exposure X through the mediators $\boldsymbol{M}$.

We similarly define

$$R^2_{l,do(X)} = \frac{\mathrm{var}\left(\mathbb{E}\left(l \mid do(X \sim X')\right)\right)}{\mathrm{var}(l)},$$
$$R^2_{l,do(X,M)} = \frac{\mathrm{var}\left(\mathbb{E}\left(l \mid do((X,M_k) \sim (X'',M''))\right)\right)}{\mathrm{var}(l)},$$

Assumption 1 (Conditional parallel mediator assumption)**.** There exist random variables $\boldsymbol{U}$ that has a global influence on $\boldsymbol{M}$, where $M_k \perp\!\!\!\perp M_j | X, \boldsymbol{U}$ for any $k, j \in \{1, \ldots, p\}$.

Theorem 1. *In Eq.* (1) - (3) *, if the identifiability conditions and assumption are satisfied, then we have*

$$R^2_{\text{med,causal}} = \left(\sum_{k=1}^{p} \alpha_k^2 \beta_k^2\right) \frac{\mathrm{var}(X)}{\mathrm{var}(l)}. \tag{6}$$

As an alternative to the proportion-mediated measure, we propose a new relative Q^2_{med} measure, which is bounded between 0 and 1:

$$Q^2_{med} = \frac{R^2_{med;causal}}{R^2_{med;causal} + \gamma^2 \sigma_x^2 / Var(l)}. \tag{7}$$

This approach offers a more interpretable, variance-based metric for assessing the relative importance of mediated effects while circumventing the cancellation problem.

2.2.1 Relationship with causal entropy

It can be shown that for a single mediator,

$$R^2_{l,do(\boldsymbol{M})} = E_{m_0 \sim P_{\boldsymbol{M}'}}\left[\frac{Var(l) - Var(l|do(\boldsymbol{M}) = m_0)}{Var(l)}\right], \tag{8}$$

where $do(M) = m_0$ implies that we intervene M and set it to be m_0, ignoring the dependencies between M and its parent variables. This notation explicitly accounts for causality (direction) in our formula. it can be shown that Eq. (8) and Eq. (5) are equivalent for a single mediator. The connection between our definition and causal entropy [21] can be seen directly through Eq. (8).

2.3 Modeling with non- and weak mediators

In real-world data, the identities of true mediators are usually unknown, necessitating the inclusion of non-mediators in our mediation model. Potential mediators $\boldsymbol{M}$ are partitioned into true mediators $\boldsymbol{M}_{\mathscr{T}}$, and three types of non-mediators $(\boldsymbol{M}_{\mathscr{I}_1}, \boldsymbol{M}_{\mathscr{I}_2}, \boldsymbol{M}_{\mathscr{I}_3})$ and illustrated in Figure 1:

- $\boldsymbol{M}_{\mathscr{T}} = \{M_j | \alpha_j \neq 0 \text{ and } \beta_j \neq 0 \text{ for } j \in \mathscr{T},\}$
- $\boldsymbol{M}_{\mathscr{I}_1} = \{M_j | \alpha_j = 0 \text{ and } \beta_j \neq 0 \text{ for } j \in \mathscr{I}_1\}$
- $\boldsymbol{M}_{\mathscr{I}_2} = \{M_j | \alpha_j \neq 0 \text{ and } \beta_j = 0 \text{ for } j \in \mathscr{I}_2\}$
- $\boldsymbol{M}_{\mathscr{I}_3} = \{M_j | \alpha_j = 0 \text{ and } \beta_j = 0 \text{ for } j \in \mathscr{I}_3\}$

We denote the proportions of $\boldsymbol{M}_{\mathscr{T}}$,, $\boldsymbol{M}_{\mathscr{I}_1}$, $\boldsymbol{M}_{\mathscr{I}_2}$, and $\boldsymbol{M}_{\mathscr{I}_3}$ as π_{11}, π_{01}, π_{10} and π_{00} respectively. The overall proportions of non-zero elements in $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are then given by $\pi_\alpha = \pi_{11} + \pi_{10}$ and $\pi_\beta = \pi_{11} + \pi_{01}$. We develop a working model under the assumptions that

$$\alpha_k \sim \pi_\alpha N(0, \sigma_\alpha^2) + (1 - \pi_\alpha)\delta_0,$$
$$\beta_k \sim \pi_{11} N(0, \frac{\sigma_{11}^2}{p\pi_{11}}) + \pi_{01} N(0, \frac{\sigma_{01}^2}{p\pi_{01}}) + (1 - \pi_\beta)\delta_0,$$

where δ_0 denotes a point mass at zero. Furthermore, we assume the correlation between mediators is undirected [6] and is driven by the existence of latent factors $\boldsymbol{U}$ following Assumption 1.

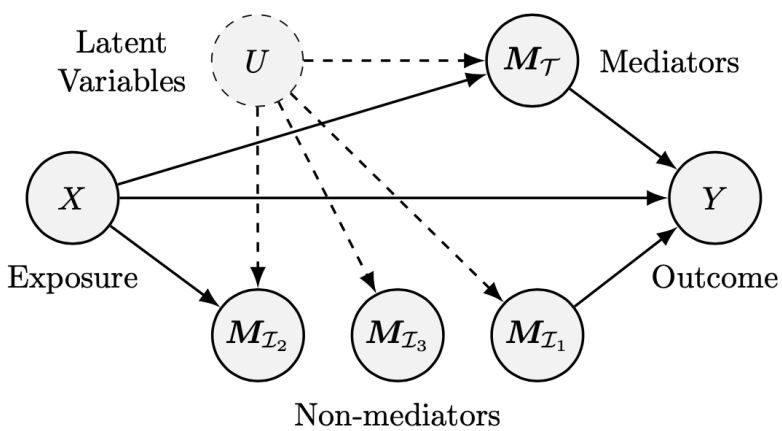


Figure 1: Graph representation of a mediation model.

Under the above assumptions and binary mediation model, it can be shown that

$$R^2_{med;causal} = \sigma^2_\alpha \sigma^2_x \sigma^2_{11} / Var(l). \tag{9}$$

To quantify the fraction of the total effect of an exposure on an outcome that is explained through the mediators, the traditional proportion-mediated measure is given by $\boldsymbol{\alpha}^\top \boldsymbol{\beta}/(\gamma + \boldsymbol{\alpha}^\top \boldsymbol{\beta})$[12]. In the causal mediation analysis framework, assuming a rare binary outcome, the proportion-mediated measure can be expressed as $\mathrm{OR}^{\mathrm{NDE}}(\mathrm{OR}^{\mathrm{NIE}} - 1)/(\mathrm{OR}^{\mathrm{NDE}}\mathrm{OR}^{\mathrm{NIE}} - 1)$, where $\mathrm{OR}^{\mathrm{NDE}} = \exp(\gamma)$ and $\mathrm{OR}^{\mathrm{NIE}} = \exp(\boldsymbol{\alpha}^\top \boldsymbol{\beta})$ represents the NDE and NIE on an odds ratio scale [18]. However, neither of the measures is guaranteed to range from 0 to 1. Additionally, both rely on the sum of component-wise effects, suffering from the cancellation problem when effects are in different directions. For the relative Q^2_{med} measure,

$$Q^2_{med} = \frac{\sigma^2_\alpha \sigma^2_{11}}{\sigma^2_\alpha \sigma^2_{11} + \gamma^2}. \tag{10}$$

It can be seen that Q^2_{med} is guaranteed to range from 0 and 1 (0 when no mediation effect and 1 when completely mediated) and do not have cancellation problems when effects are in different directions.

3 Estimation procedure

To estimate the proposed R^2 measure, we adapted the cross-fitting framework as described by [22], to improve the robustness of mediation effect estimation. Cross-fitting mitigates bias introduced by the winner's curse during the variable screening step. The approach involves a two-stage procedure where the sample is randomly split into two subsample $\mathscr{D}^{(1)}$ and $\mathscr{D}^{(2)}$. Mediation selection and variance estimation of $\boldsymbol{\alpha}$ is first performed on one subsample $\mathscr{D}^{(1)}$ to get the filtered mediator index set $\hat{\mathscr{T}}^{(1)}$ and variance estimate $\hat{\sigma_\alpha}^{(1)2}$. Then estimation of other parameters is conducted on subsampe $\mathscr{D}^{(2)}$ based on the selected mediators $\hat{\mathscr{T}}^{(1)}$. The roles of the two subsamples are then reversed, allowing the whole sample to be used in the estimation process. Finally, the R^2 estimate is given by the mean of cross-fitted R^2 estimates $\hat{R}^{(1)2}_{med;causal}$ and $\hat{R}^{(2)2}_{med;causal}$. The procedure is visualized in Figure 2 and the mediator selection and parameter estimation process are detailed in the Subsections 3.1-3.3.

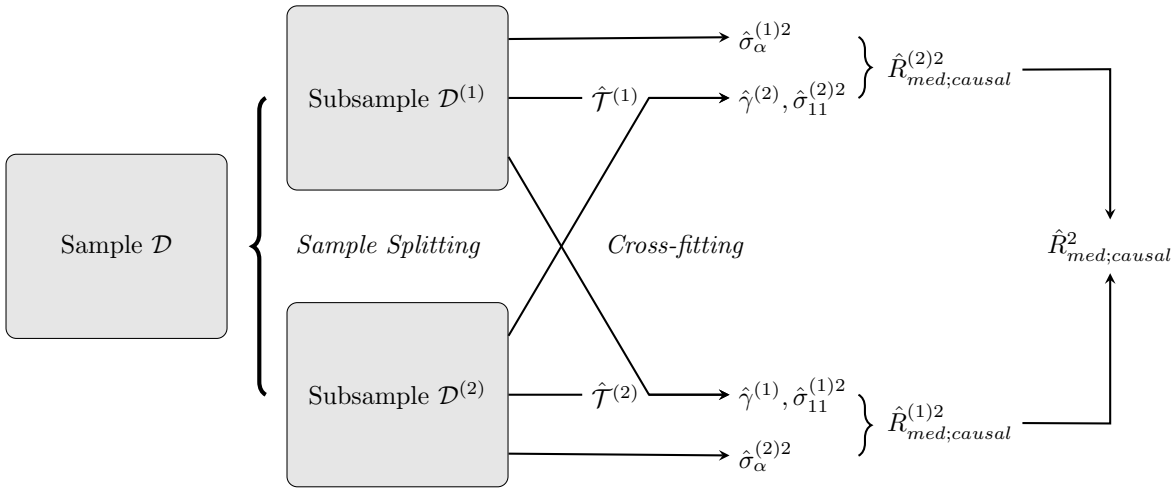


Figure 2: Cross-fitted estimation procedure of $R^2_{med;causal}$

3.1 Estimation of σ_α^2 and mediator selection

Estimation of $\boldsymbol{\alpha}$ in our mediation analysis framework has been intensively studied as a secondary outcome analysis problem in case-control studies. We used Inverse Probability Weighting (IPW) method to estimate σ_α^2. IPW method involves weighting each individual's contribution to the score or estimating equations from the secondary outcome model by the reciprocal of the their selection probability [23]. A univariate IPW estimation problem for estimating each element of $\boldsymbol{\alpha}$ is described:

Let S_i denote the inclusion indicator for the i-th individual. $w_i = 1/P(S_i = 1|Y_i, M_i, X_i)$ is the inverse weighting. The IPW estimating equation for this problem is

$$\sum_{i=1}^{n} U_{IPW,i}(\alpha_k) = \sum_{i=1}^{n} S_i w_i h(X_i)[\boldsymbol{M}_{i,k} - \alpha_k X_i] = 0,$$

$k = 1, \ldots, p$, $h(X_i)$ is specified by the user such that $E(\partial U_{IPW,i}/\partial \alpha_k)$ is invertible. Under the full-ascertainment assumption, w_i is 1 for case and $\frac{P(1-K)}{K(1-P)}$ for control[10]. Setting $h(X_i) = X_i$ simplifies the estimating equation problem into a weighted least square regression problem.

Simultaneously, multiple testing is conducted to identify non-mediators $\boldsymbol{M}_{\mathscr{I}_1}(\alpha = 0, \beta \neq 0)$ and noise mediators $\boldsymbol{M}_{\mathscr{I}_3}(\alpha = 0, \beta = 0)$. Given that $Var(\alpha) = \pi_\alpha \sigma_\alpha^2$, we can derive the estimator as $\hat{\sigma}_\alpha^2 = \widehat{Var(\alpha)}/\hat{\pi}_\alpha$, where $Var(\alpha)$ is estimated empirically from the IPW results and π_α is estimated by a false discovery rate (FDR) control method with FDR threshold set to be less than 0.01. The filtered mediator set is denoted by $\boldsymbol{M}_{\hat{\mathscr{T}}}$ for later use in cross-fitted variance component estimation.

3.2 Estimation of γ

The IPW approach can also be applied to estimate γ parameter. By choosing a proper estimating equation, the IPW problem becomes a weighted probit regression problem with the same weight as described in section 2.1. However, directly regressing outcome Y on exposure variable X will

lead to biased result because it ignores the mediated pathway through which X influences Y via $\boldsymbol{M}$. To address this, we include the top principal components (PCs) of $\boldsymbol{M}$, $\boldsymbol{W}$, as the additional covariates in the model. By incorporating these components, we can effectively control for the influence of $\boldsymbol{M}$ that may introduce bias in the estimation of γ.

3.3 Estimation of σ_{11}^2

For the estimation of σ_{11}^2, we propose to use the PCGC regression [10]. PCGC regression is a method of moments that estimates genetic variance proportion on the liability scale while accounting for the oversampling in case-control designs. This method generalizes Haseman–Elston regression and models the correlation between phenotypic similarity and genetic similarity, assuming independence between genetic and environmental effects, as well as among environmental effects of individuals. This approach implicitly assumes that genetic factors cannot be strongly correlated, an assumption that likely to be violated by mediators due to their relationship with the exposure (equation 1). Therefore, we propose to modify PCGC and estimate σ_{11}^2 following the steps described below:

1. Identify the non-mediators with $\alpha = 0$, i.e., $\boldsymbol{M}_{\mathscr{I}_1}$ and $\boldsymbol{M}_{\mathscr{I}_3}$ with FDR < 0.01 as described in section 3.1. The remaining mediator set is denoted by $\hat{\mathscr{T}}$.

2. Estimate standardized residual $\hat{\boldsymbol{\xi}}$ by regressing scaled $\boldsymbol{M}$ on X, $\boldsymbol{C}$, $\boldsymbol{W}$ using IPW approach. Retain only the filtered mediator set $\hat{\mathscr{T}}$ identified in the other subsample.

3. Apply PCGC regression with $\hat{\boldsymbol{\xi}}_{\hat{\mathscr{T}}}$ as the standardized genotypic matrix and X, $\boldsymbol{C}$ and $\boldsymbol{W}$ as covariates to estimate $\hat{\sigma}_{11}^2$.

In presence of covariates, PCGC approach consists of a two-step procedure:

1. Estimate $\hat{P}_i = P(Y_i = 1|X_i, \boldsymbol{C}, \boldsymbol{W}_i)$ among samples using a logistic regression model by regressing Y on X, $\boldsymbol{C}$ and $\boldsymbol{W}$. We can then have

$$\hat{K}_i = \frac{\frac{K(1-P)}{P(1-K)}\hat{P}_i}{1 + \frac{K(1-P)}{P(1-K)}\hat{P}_i - \hat{P}_i}$$
$$\hat{t}_i = \Phi^{-1}(1 - \hat{K}_i),$$

   where K_i and t_i represent trait prevalence and threshold conditioned on covariates; K is the prevalence of trait in the population, which is typically treated as known.

2. Correct the effect of the ascertainment conditioned on covariates and then regress corrected phenotypic correlations on corrected genotypic correlations.

   (a) The refined phenotypic correlation can be expressed as

$$Z_{ij} = \frac{(y_i - P_i)(y_j - P_j)}{\sqrt{P_i(1-P_i)}\sqrt{P_j(1-P_j)}}.$$

   (b) Genetic correlation matrix among individuals can be calculated by standardized residual: $\hat{G} = \hat{\boldsymbol{\xi}}_{\hat{\mathscr{T}}}\hat{\boldsymbol{\xi}}_{\hat{\mathscr{T}}}^\top / p$.

(c) By regressing Z_{ij} on

$$\frac{\varphi(\hat{t}_i)\varphi(\hat{t}_j)\left[1-(\hat{P}_i+\hat{P}_j)\left(\frac{P-K}{P(1-K)}\right)+\hat{P}_i\hat{P}_j\left(\frac{P-K}{P(1-K)}\right)^2\right]\hat{G}_{ij}}{\sqrt{\hat{P}_i(1-\hat{P}_i)}\sqrt{\hat{P}_j(1-\hat{P}_j)}\left(\hat{K}_i+(1-\hat{K}_i)\frac{K(1-P)}{P(1-K)}\right)\left(\hat{K}_j+(1-\hat{K}_j)\frac{K(1-P)}{P(1-K)}\right)}$$

yields an estimator of $\hat{\sigma}_{11}^2$.

3.4 Theoretical results on consistency

In this section, we present the theoretical properties of the proposed method. Theorem 2 established the consistency of our proposed total mediation effect estimator $\widehat{R}^2_{med;causal}$ and relative measurement estimator $\widehat{Q}^2_{med}$ under mild regularity conditions.

Theorem 2. *Under Assumptions 1-9, and assume there exists $c_0 > 0$ that $\frac{\text{TP}}{p} > c_0$, then for any $\varepsilon > 0$, there exists $\delta > 0$, such that if $\frac{\text{TP}}{\text{TP+FP}} > 1-\delta$ for p,n large enough, we have*

$$\lim_{p\to\infty,n\to\infty}\mathbb{P}(|\widehat{R}^2_{med;causal}-R^2_{med;causal}|<\varepsilon)=1.$$

and

$$\lim_{p\to\infty,n\to\infty}\mathbb{P}(|\widehat{Q}^2_{med}-Q^2_{med}|<\varepsilon)=1.$$

Let $A=\{\alpha\neq 0,\hat{\alpha}\neq 0\}$ and $B=\{\alpha=0,\hat{\alpha}\neq 0\}$. Herein, TP and FP correspond to $|A|$ and $|B|$ in the initial filtering step of α (Section 3.1). A key assumption underlying this result is that the proportion of true mediators and non-mediators type $\boldsymbol{M}_{\mathscr{I}_2}$ among all potential mediators remains above a positive threshold as $p\to\infty$. This assumption ensures a sufficiently large number of mediators contribute to the total mediation effect. Another important assumption is the control of FDR when identifying non-mediators type $\boldsymbol{M}_{\mathscr{I}_1}$ and noise mediators $\boldsymbol{M}_{\mathscr{I}_3}$. We assume that the FDR for selecting variables impacted by exposure is stringent enough to minimize the inclusion of spurious mediators. In practice, our simulations suggest that a reasonable choice, such as FDR ≤ 0.01, yields robust performance across various sample sizes and mediator sparsity levels (Section 4). Notably, the proposed estimator is shown to be consistent without requiring mediator selection consistency on β and α but only sufficiently low FDR in testing p marginal association (i.e., $\alpha=0$). This is especially desirable in high-dimensional situations, as exact selection is rarely achieved.

4 Simulation studies

4.1 Simulation Settings

In this section we evaluated the performance of $R^2_{med;causal}$ and Q^2_{med} estimators in comparison with existing high-dimensional mediation models: HIMA, HDMA and BAMA. Simulation was conducted with different combination of non-mediators to evaluate our proposed measure and estimation procedure under the following 4 scenarios:

1. Scenario 1 (S1): 180 $\boldsymbol{M}_{\tau}$, 420 $\boldsymbol{M}_{\mathscr{I}_1}$, 420 $\boldsymbol{M}_{\mathscr{I}_2}$, 980 $\boldsymbol{M}_{\mathscr{I}_3}$;
2. Scenario 2 (S2): 360 $\boldsymbol{M}_{\tau}$, 240 $\boldsymbol{M}_{\mathscr{I}_1}$, 840 $\boldsymbol{M}_{\mathscr{I}_2}$, 560 $\boldsymbol{M}_{\mathscr{I}_3}$;
3. Scenario 3 (S3): 360 $\boldsymbol{M}_{\tau}$, 840 $\boldsymbol{M}_{\mathscr{I}_1}$, 240 $\boldsymbol{M}_{\mathscr{I}_2}$, 560 $\boldsymbol{M}_{\mathscr{I}_3}$;
4. Scenario 4 (S4): 720 $\boldsymbol{M}_{\tau}$, 480 $\boldsymbol{M}_{\mathscr{I}_1}$, 480 $\boldsymbol{M}_{\mathscr{I}_2}$, 320 $\boldsymbol{M}_{\mathscr{I}_3}$.

The population prevalence of disease K was set at 0.05 and disease prevalence in the study P was set at 0.5. Mediators were generated with dimension fixed at $p = 2000$, sample size N varying from 500 to 2000 with a step size of 500. We set the single exposure variable $X \sim N(0,1)$ and its direct effect parameter $\gamma = 0.5$. The beta variance of true mediators was set to be $\sigma_{11}^2 = 0.5$, and that of type-1 non-mediators was set to be $\sigma_{01}^2 = 0.3$. The error term was then generated from $\varepsilon \sim N(0, 1-\sigma_{11}^2-\sigma_{01}^2)$. To ensure Assumptions 1 and 2 hold, we let $\xi_k = \alpha_0 U + N(0,1)$ where $U \sim N(0,1)$ and $\alpha_0 \sim N(0, 0.3)$. For each scenario, we run 100 replications across different sample size N. The mean absolute bias, standard deviation and mean square error were reported for both measures.

For scenario 1, we also compared our methods to other competing methods: HIMA, HDMA and BAMA in terms of total mediation effect estimate and proportion-mediated estimate. Prevalence of disease K varied from 0.5 (cohort study) to 0.05 (case control study) and sample size varied from 500 to 2000. For the competing methods, we calculated our proposed measurement ($R^2_{med;causal}$ and Q^2_{med}) empirically from the component-wise estimates $\hat{\boldsymbol{\alpha}}$, $\hat{\boldsymbol{\beta}}$ and $\hat{\gamma}$. We also calculated measurements of mediation effect and proportion-mediated as outlined in their paper. For method HIMA and HDMA, their IE measure is the inner product of $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\beta}}$. Their proportion mediated, following the definition in traditional analysis, is $\hat{\boldsymbol{\alpha}}^\top\hat{\boldsymbol{\beta}}/(\hat{\gamma}+\hat{\boldsymbol{\alpha}}^\top\hat{\boldsymbol{\beta}})$. While for BAMA, the mediation effect measure is given by $\text{NIE}_{BAMA} = \sum_{i=1}^{p}(\hat{\alpha}_i\hat{\beta}_i)^2$. The results are summarized in the following section. Relative bias ($\frac{\text{Bias}}{\text{True Measure}}$) were calculated to better compare different types of measurements.

4.2 Simulation Results

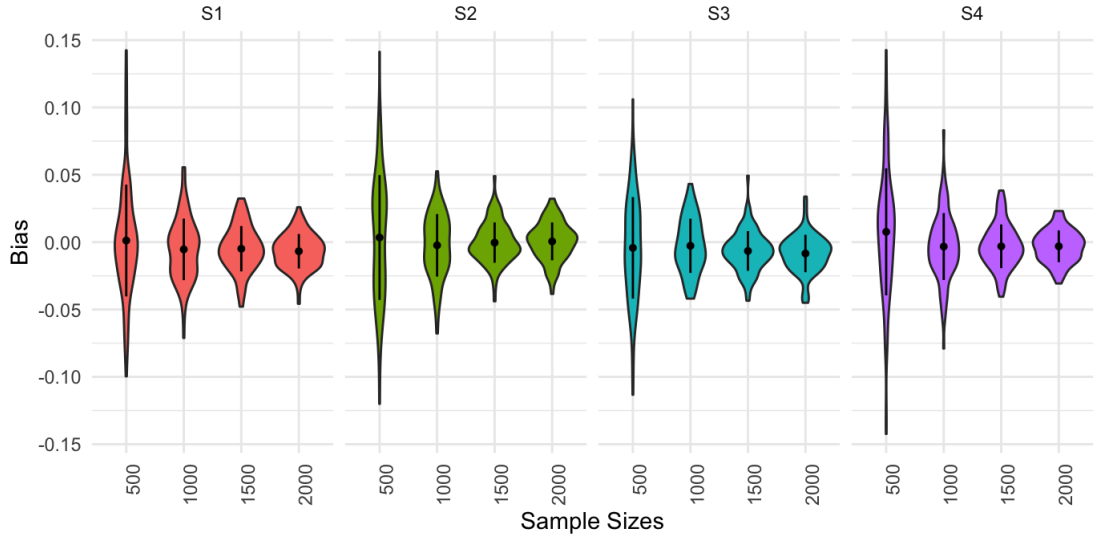
In this section, we discuss the performance of our proposed measures under varying sample sizes, number of different types of non-mediators, and disease prevalence scenarios, and compare them with competing methods, i.e., HIMA, HDMA and BAMA.

Figure 3 and Table 1 summarize the estimation performance of two proposed measurements in terms of absolute bias, standard deviation, and mean square error. Overall, the proposed measurements performed well across all scenarios, with minimal bias and error observed across different sample sizes and number of non-mediators. As expected, the performance metrics improved with larger sample sizes. Furthermore, the violin plots indicate the consistency of the proposed estimation methods as well as their robustness under various conditions.

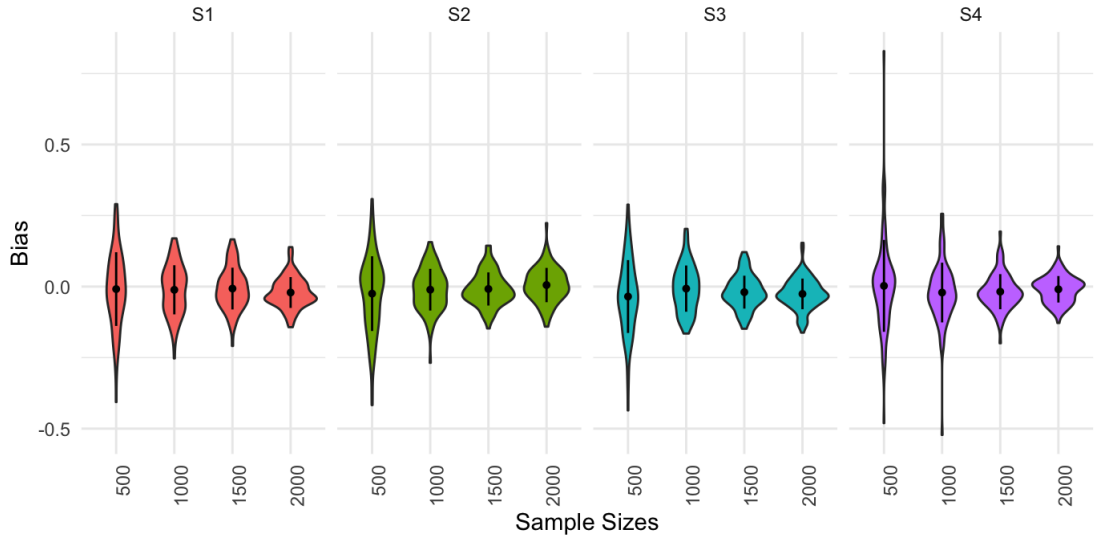
Tables 2 and 3 present a comparison of our method with HIMA, HDMA, and BAMA in terms of absolute bias, standard deviation, and mean square error for estimating IE and proportion-mediated, respectively. Our proposed method consistently outperformed competing methods across all simulation scenarios. HIMA and HDMA frequently identified only a small number of mediators (typically fewer than 10), which indicated their limited ability to capture mediators with small effect sizes, resulting in higher bias and less reliable estimates. BAMA, on the other hand, explicitly assumes that mediator effects follow a normal distribution with either large or

small variance, which may explain its improved performance over HIMA and HDMA in some simulation scenarios. However, the amount of bias was still higher than desired. As the prevalence of disease K increased, our method showed reduced efficiency in small sample sizes, as reflected by the higher estimation variance. However, with larger sample sizes, the method's performance became more stable, showing comparable results across different K values.

Table 4 summarizes their computation speed in both case-control and cohort study. No significant difference in running time was observed between cohort ($K = 0.5$) and case-control ($K = 0.05$) studies across all methods. However, the growth of running time with increasing sample size for other methods was notable. Under a Bayesian sampling estimation framework, the running time for BAMA exhibited approximately linear growth as N increased, making it the slowest approach. While HDMA was faster than BAMA, it remained slower than HIMA and the proposed method. Overall, the proposed method and HIMA demonstrated the most efficient performance, with the slowest growth with running time.



(a) Bias for $R^2_{med;causal}$ estimation



(b) Bias for Q^2_{med} estimation

Figure 3: Violin plots of bias in a case-control study ($K = 0.05$) under varying sample sizes and proportions of non-mediators across four scenarios. The black dots represent the average bias, while the vertical lines correspond to the range of one standard deviation centered at the mean bias.

Table 1: Performance of proposed estimators under different scenarios. |Bias| represents average absolute bias. SD represents standard deviation of estimators based on 100 repetitions. MSE stands for mean square error of estimators.

		R^2_{Med}			Q^2_{med}		
Scenario	N	|Bias|	SD	MSE	|Bias|	SD	MSE
S1	500	0.0297	0.0415	0.0017	0.1007	0.1297	0.0167
	1000	0.0183	0.0229	0.0005	0.0714	0.0864	0.0075
	1500	0.0139	0.0168	0.0003	0.0592	0.0735	0.0054
	2000	0.0115	0.0128	0.0002	0.0469	0.0543	0.0034
S2	500	0.0387	0.0464	0.0021	0.1028	0.1315	0.0177
	1000	0.0189	0.0232	0.0005	0.0592	0.0734	0.0055
	1500	0.0116	0.015	0.0002	0.0469	0.0583	0.0034
	2000	0.0111	0.014	0.0002	0.047	0.0603	0.0036
S3	500	0.0302	0.0377	0.0014	0.1035	0.1282	0.0175
	1000	0.0168	0.0201	0.0004	0.0666	0.0813	0.0066
	1500	0.0125	0.0147	0.0003	0.0492	0.0579	0.0037
	2000	0.0124	0.0139	0.0003	0.0469	0.054	0.0035
S4	500	0.0359	0.0471	0.0023	0.1079	0.1617	0.0259
	1000	0.0188	0.0248	0.0006	0.0771	0.1051	0.0114
	1500	0.013	0.0162	0.0003	0.0505	0.0618	0.0041
	2000	0.0098	0.0118	0.0001	0.0371	0.0466	0.0022

Table 2: Comparison of relative performance of IE estimators. Each cell reports the average absolute relative bias, with the standard deviation of the relative bias shown in parentheses.

		R^2 Measure				Original mean-based Measure		
K	N	Proposed	HIMA	HDMA	BAMA	HIMA	HDMA	BAMA
0.05	500	0.08	1.45	1.74	0.91	1.17	2.85	0.92
		(0.09)	(1.74)	(1.71)	(0.01)	(0.61)	(2.34)	(0.01)
	1000	0.14	1	1.77	0.92	1.36	3.67	0.93
		(0.18)	(1.12)	(1.14)	(0.01)	(0.53)	(2.27)	(0.01)
	1500	0.1	0.89	1.77	0.92	1.38	3.58	0.92
		(0.12)	(0.99)	(0.8)	(0.01)	(0.4)	(2.38)	(0.01)
	2000	0.08	0.68	1.76	0.92	1.49	3.88	0.92
		(0.09)	(0.77)	(0.7)	(0.01)	(0.37)	(2.33)	(0.01)
0.2	500	0.32	1.17	1.34	0.9	0.71	1.84	0.92
		(0.39)	(1.32)	(1.45)	(0.01)	(0.58)	(2.06)	(0.01)
	1000	0.17	0.83	1.18	0.92	0.78	2.65	0.93
		(0.22)	(0.98)	(0.79)	(0.01)	(0.41)	(2.18)	(0.01)
	1500	0.12	0.65	1.07	0.92	0.78	2.48	0.93
		(0.15)	(0.77)	(0.75)	(0.01)	(0.34)	(1.89)	(0.01)
	2000	0.1	0.54	1.12	0.92	0.78	2.67	0.93
		(0.12)	(0.66)	(0.59)	(0.01)	(0.28)	(1.74)	(0.01)
0.5	500	0.4	1.14	1.31	0.9	0.54	1.67	0.92
		(0.49)	(1.27)	(1.44)	(0.01)	(0.51)	(1.89)	(0.01)
	1000	0.18	0.78	1.03	0.92	0.57	2.22	0.93
		(0.23)	(0.95)	(0.85)	(0.01)	(0.38)	(1.91)	(0.01)
	1500	0.14	0.65	0.96	0.92	0.52	2.16	0.93
		(0.17)	(0.79)	(0.69)	(0.01)	(0.31)	(1.73)	(0.01)
	2000	0.11	0.45	0.83	0.93	0.59	2.17	0.93
		(0.13)	(0.58)	(0.59)	(0.01)	(0.29)	(1.97)	(0.01)

Table 3: Comparison of relative performance of proportion-mediated estimators. Each cell reports the average absolute relative bias, with the standard deviation of the relative bias shown in parentheses.

		Q^2_{med} Measure				Proportion mediated	
K	N	Proposed	HIMA	HDMA	BAMA	HIMA	HDMA
0.05	500	0.11	0.82	0.9	0.27	0.27	1.17
		(0.13)	(0.76)	(0.98)	(0.16)	(0.33)	(1.19)
	1000	0.16	0.5	0.94	0.24	0.25	1.48
		(0.2)	(0.56)	(0.78)	(0.13)	(0.27)	(1.18)
	1500	0.12	0.47	0.86	0.26	0.24	1.41
		(0.14)	(0.51)	(0.68)	(0.09)	(0.23)	(1.19)
	2000	0.11	0.37	0.96	0.27	0.27	1.55
		(0.13)	(0.4)	(0.69)	(0.09)	(0.21)	(1.16)
0.2	500	0.3	0.82	0.81	0.2	0.29	1.05
		(0.35)	(0.65)	(0.92)	(0.2)	(0.37)	(1.3)
	1000	0.18	0.52	0.8	0.24	0.21	1.37
		(0.23)	(0.54)	(0.73)	(0.14)	(0.26)	(1.31)
	1500	0.15	0.46	0.78	0.27	0.19	1.24
		(0.18)	(0.46)	(0.74)	(0.11)	(0.22)	(1.17)
	2000	0.12	0.48	0.86	0.28	0.17	1.34
		(0.14)	(0.41)	(0.71)	(0.09)	(0.19)	(1.09)
0.5	500	0.33	0.81	0.84	0.2	0.29	1.04
		(0.42)	(0.65)	(0.97)	(0.2)	(0.35)	(1.3)
	1000	0.21	0.61	0.76	0.24	0.22	1.25
		(0.25)	(0.57)	(0.78)	(0.16)	(0.28)	(1.28)
	1500	0.14	0.52	0.79	0.28	0.17	1.21
		(0.16)	(0.48)	(0.74)	(0.1)	(0.21)	(1.18)
	2000	0.13	0.48	0.75	0.29	0.17	1.24
		(0.15)	(0.38)	(0.74)	(0.1)	(0.21)	(1.3)

Table 4: Computation speed for different approaches (in seconds). Each cell presents the average running time, with the standard deviation of the running time provided in parentheses.

	Proposed		HIMA		HDMA		BAMA	
N, K	0.05	0.5	0.05	0.5	0.05	0.5	0.05	0.5
500	12.92	25.74	2.44	2.64	4.78	5.02	19.03	18.92
	(0.21)	(0.74)	(0.12)	(0.35)	(0.1)	(0.56)	(2.72)	(3.15)
1000	15.79	16.8	3.46	4.58	9.67	12.28	37.44	39.29
	(0.25)	(0.67)	(0.17)	(1.48)	(0.23)	(4.2)	(6.62)	(7.35)
1500	19.62	21	4.75	5.23	17.73	18.79	57.06	58.56
	(0.28)	(0.75)	(0.31)	(0.63)	(0.38)	(1.26)	(10.25)	(10.54)
2000	23.85	25.59	6.33	8.19	30.87	36.82	78.71	81.34
	(0.27)	(0.72)	(0.49)	(2.55)	(0.71)	(8.77)	(13.65)	(14.07)

5 Application to the Women's Health Initiative Study

This study utilizes data from the WHI nested case-control substudy (dbGaP accession number: phs001334), which examines the relationship between metabolomic profiles and incident CHD in postmenopausal women. The substudy is consisted of subjects from both the WHI Observational Study (WHI-OS) and the WHI Hormone Therapy Trials (WHI-HT). A total of 2,306 samples (681 case-control pairs from WHI-HT and 472 case-control pairs from WHI-OS) were analyzed for metabolomic profiling, representing 2,151 unique consented participants in the analysis. Baseline targeted metabolomic profiles were measured using liquid chromatography–mass spectrometry, yielding 371 metabolites after initial quality control. Cases were defined as postmenopausal women who experienced incident myocardial infarction (MI) or CHD death during follow-up. In WHI-HT, cases were those who developed MI or CHD death during the hormone therapy intervention period. In WHI-OS, cases were selected from previously identified incident CHD cases in the prior biomarker study (BAA 22), excluding those who experienced a stroke before MI or CHD death. Controls were matched by age, race/ethnicity, enrollment period, and hysterectomy status, excluding those with prior cardiovascular diseases.

BMI is a well-established risk factor for CHD [24], however, how BMI impacts CHD through metabolites is less studied. Metabolites with more than 50% missing values were excluded from the analysis. For the remaining metabolites, missing values were imputed by assigning half of the lowest observed value [25]. Metabolomic profiles were log-transformed. We further remove an outlier by inspecting a plot of the top two principal components of metabolomic profiles. Smoking status, alcohol consumption, and education level were included in the model as covariates, with missing values imputed using multiple imputation by chained equations[26]. After these steps, the final dataset contained 2,150 samples with 366 metabolite variables. The prevalence of disease in the population was estimated as a weighted sum of disease prevalence across ethnic groups, using data from the American Heart Association's Heart Disease and Stroke Statistics report [27].

Table 5: Total mediation effect of metabolites from BMI to CHD in WHI nested case-control study. 95% confidence interval is present in parenthesis.

	$\hat{R}^2_{\text{med;causal}}$	$\hat{\boldsymbol{\alpha}}^\top\hat{\boldsymbol{\beta}}$	$\sum_i(\hat{\alpha}_i\hat{\beta}_i)^2$	$\hat{Q}^2_{med}$	$\frac{\hat{\boldsymbol{\alpha}}^\top\hat{\boldsymbol{\beta}}}{\hat{\boldsymbol{\alpha}}^\top\hat{\boldsymbol{\beta}}+\hat{\gamma}}$
Proposed	0.0072			0.8935	
	(0.065, 0.0093)			(0.7335, 0.9114)	
HIMA	0.0033	0.09	0.0031	0.8962	0.821
	(0.0021, 0.0041)	(0.0677, 0.0914)	(0.003, 0.0043)	(0.4909, 0.9352)	(0.5472, 0.8469)
HDMA	0.0172	0.0088	0.016	0.4449	0.0558
	(0.0098, 0.021)	(-0.068, 0.095)	(0.0085, 0.0221)	(0.2315, 0.9883)	(-0.3326, 0.8911)
BAMA	0.0002	0.0231	0.0002	0.4954	0.7361
	(0.0001, 0.0002)	(0.0122, 0.0344)	(0.0001, 0.0002)	(0.0873, 0.9947)	(0.02538, 2.0473)

Figure 4: Heatmap of Counts Across P-value Intervals for $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$

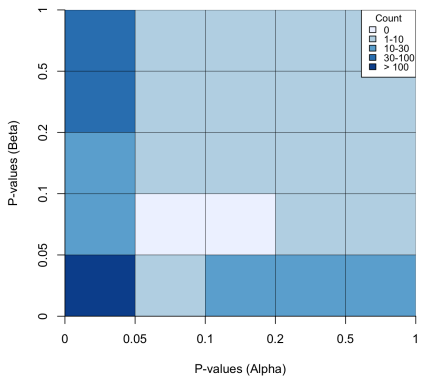


Table 5 summarizes the total mediation estimates of BMI on CHD through metabolites obtained from different methods. For HIMA, HDMA, and BAMA, empirical $R^2_{med;causal}$ values were calculated based on the component-wise mediation estimates. The confidence intervals (CIs) for the estimates were calculated using the jackknife method, where each iteration involved omitting one individual from the dataset and recomputing the estimates. The CI bounds were determined by the 2.5th and 97.5th percentiles of the distribution of all jackknife estimates.

Figure 4 visualizes the distribution of mediator counts based on the marginal P-values of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. While over 100 metabolites exhibit small p values for their marginal effects in both the mediator and outcome models, only a few were identified by HIMA and HDMA, showing the prevalence of potential weak mediators. In addition, evidence of opposing component-wise mediation effect estimates was observed across methods. For example, HDMA method produced the smallest traditional mediation effect estimate $\hat{\boldsymbol{\alpha}}^\top\hat{\boldsymbol{\beta}}$. Among the 13 significant mediators identified by HDMA, 7 showed positive effect, while 6 showed negative effect. In contrast, the HIMA method identified 9 significant mediators, 8 of which had positive effects, resulting in a similar total mediation effect estimate despite a much smaller R^2 value. Our proposed method produced the second largest R^2 estimate, demonstrating its capability to capture the total mediation effect that other methods may miss. While HDMA had the largest R^2 estimate by including three mediators filtered out by our screening step. In other words, HDMA inadvertently included non-mediators due to inflated type I error in estimating α, which arose from not accounting for selection effects.

6 Discussion

We have developed a novel mediation analysis framework tailored for case-control studies to estimate the total mediation effect under high-dimensional settings. The key contributions of our method include a new R^2 measure designed for binary outcomes, which addresses the challenge of opposite direction effects and provides a clear causal interpretation. Furthermore, we introduced an efficient estimation procedure that accounts for ascertainment bias, which is computationally scalable, and effectively handles non-mediators while capturing mediators with weak effects. We further prove that under mild conditions, our estimator is consistent (Appendix). Through simulation studies, we demonstrated its robustness in estimating mediation effects under varying scenarios, and its applicability was further validated in the WHI study.

7 Acknowledgement

The WHI program is funded by the National Heart, Lung, and Blood Institute, National Institutes of Health, U.S. Department of Health and Human Services through contracts HHSN268201600018C, HHSN268201600001C, HHSN268201600002C, HHSN268201600003C, and HHSN268201600004C. Metabolomic analyses were funded by NHLBI contract HHSN268201300008C. This manuscript was not prepared in collaboration with investigators of the WHI, has not been reviewed and/or approved by WHI, and does not necessarily reflect the opinions of the WHI investigators or the NHLBI. The WHI dataset used for the analyses described in this manuscript was obtained from dbGaP at https://www.ncbi.nlm.nih.gov/gap/ through accession number phs001334. PW was partially supported by National Institutes of Health (NIH) grant R21HL170213. TY was supported by St. Baldrick Career Award and nih grant R01AG074858.

A Proof of consistency of $R^2_{med;causal}$ and Q^2_{med} estimators

In this section, we prove the theoretical result in the main text.

A.1 Technical Conditions for the consistency of $\widehat{R}^2_{med;causal}$ and $\widehat{Q}^2_{med}$

Assumption 2. Let $\widetilde{\mathbf{X}}_i = (I_p \otimes X_i, I_p \otimes \boldsymbol{C}_i^\top)$, there exists a positive constant R such that

$$0 < \frac{1}{R} < \lambda_{\min}(\frac{1}{n}\sum_{i=1}^{n}\widetilde{\mathbf{X}}_i^\top\widetilde{\mathbf{X}}_i) \leq \lambda_{\max}(\frac{1}{n}\sum_{i=1}^{n}\widetilde{\mathbf{X}}_i^\top\widetilde{\mathbf{X}}_i) < R < \infty$$

and

$$0 < \frac{1}{R} < \pi_\alpha\sigma_\alpha^2 < R < \infty$$

and for all i,

$$0 < \frac{1}{R} < w_i < R < \infty$$

Assumption 3. We have $\mathbf{E}[\xi_{ij}^4] < \infty$ and $\sup_{jk}\mathbf{E}[(\xi_{ij}\xi_{ik})^4] < \infty$ for $i = 1,\ldots,n$ and $j,k = 1,\ldots,p$.

Assumptions 2 and 3 are standard conditions in case-control study, see Conditions (C2),(C3) and (C9) in [28] for an example.

Assumption 4. The variance of X explained by $\boldsymbol{M}$ is strictly less than 1.

Assumption 5. Let $\boldsymbol{c}_i = (X_i, \boldsymbol{C}_i, \boldsymbol{W}_i)^\top$ and $\widehat{\boldsymbol{c}}_i = (X_i, \boldsymbol{C}_i, \widehat{\boldsymbol{W}}_i)^\top$. Assume that there exists a constant $C > 0$, such that $\max(\|\boldsymbol{c}_i\|_2, \|\widehat{\boldsymbol{c}}_i\|_2) < C$ and $\frac{1}{n}\sum_{i=1}^n \|\widehat{\boldsymbol{c}}_i - \boldsymbol{c}_i\|_2 = o_p(1)$.

In the factor model literature, Assumption 2 corresponds to the standard eigenvalue condition and Assumption 5 is very mild, with Theorem 4 in [29] as an example.

Following derivation of heritability estimation with known fixed effects in [10], we have

$$\mathbb{E}\left[Z_{ij} \mid \mathscr{S} = 1; G_{ij}, t_i, t_j\right] = \frac{A'\left(0; t_i, t_j\right)}{B\left(0; t_i, t_j\right)} G_{ij} + b_{ij}(G_{ij}; t_i, t_j) G_{ij}^2$$

where

$$\begin{aligned} A\left(G_{ij}; t_i, t_j\right) = & \frac{(1-P_i)\left(1-P_j\right)}{\sqrt{P_i(1-P_i)}\sqrt{P_j\left(1-P_j\right)}} \mathbb{P}\left(y_i = y_j = 1; G_{ij}, t_i, t_j\right) + \\ & \frac{K(1-P)}{P(1-K)} \frac{-P_i\left(1-P_j\right)}{\sqrt{P_i(1-P_i)}\sqrt{P_j\left(1-P_j\right)}} \mathbb{P}\left(y_i = 0, y_j = 1; G_{ij}, t_i, t_j\right) + \\ & \frac{K(1-P)}{P(1-K)} \frac{-P_j(1-P_i)}{\sqrt{P_i(1-P_i)}\sqrt{P_j\left(1-P_j\right)}} \mathbb{P}\left(y_i = 1, y_j = 0; G_{ij}, t_i, t_j\right) + \\ & \left(\frac{K(1-P)}{P(1-K)}\right)^2 \frac{P_i P_j}{\sqrt{P_i(1-P_i)}\sqrt{P_j\left(1-P_j\right)}} \mathbb{P}\left(y_i = 0, y_j = 0; G_{ij}, t_i, t_j\right). \end{aligned}$$

and

$$B\left(\rho; t_i, t_j\right) = \mathbb{P}(\mathscr{S} = 1; \rho, t_i, t_j)$$

Denote $\Delta_{ij} = \frac{A'\left(0;t_i,t_j\right)}{B\left(0;t_i,t_j\right)\sigma_{11}^2} = \frac{\varphi(t_i)\varphi(t_j)\left[1-(P_i+P_j)\left(\frac{P-K}{P(1-K)}\right)+P_iP_j\left(\frac{P-K}{P(1-K)}\right)^2\right]}{\sqrt{P_i(1-P_i)}\sqrt{P_j(1-P_j)}\left(K_i+(1-K_i)\frac{K(1-P)}{P(1-K)}\right)\left(K_j+(1-K_j)\frac{K(1-P)}{P(1-K)}\right)}$

Assumption 6. Let $e_{ij} = Z_{ij} - \mathbb{E}\left[Z_{ij} \mid \mathscr{S} = 1; G_{ij}, t_i, t_j\right]$, we assume that

$$\frac{\sum_{i<j} \Delta_{ij} b_{ij}(G_{ij}) G_{ij}^3}{\sum_{i<j} \Delta_{ij}^2 G_{ij}^2} \to 0$$

and

$$\frac{\sum_{i<j} \Delta_{ij} G_{ij} e_{ij}}{\sum_{i<j} \Delta_{ij}^2 G_{ij}^2} \to 0$$

as $n \to \infty$.

The consistency of the PCGC estimator without fixed effects when $\frac{p}{n}$ converges to a positive constant was established in [30]. Under their assumptions, the first part of Assumption 6 is implied by their Lemma 1 and Lemma 2, and the second part of Assumption 6 is implied by their Lemma 4.

Assumption 7. We have that

$$\max(\sum_{i<j}|\Delta_{ij}G_{ij}-\widehat{\Delta}_{ij}\widehat{G}_{ij}|,|\sum_{i<j}(\widehat{\Delta}_{ij}^2\widehat{G}_{ij}^2-\Delta_{ij}^2G_{ij}^2)|)=o_p(n(n-1))$$

Assumption 8. There exists a positive constant $0<r<0.5$ such that

$$0<r<P<1-r<1,0<r<K<1-r<1$$

and for all i,

$$0<r<P_i<1-r<1,0<r<K_i<1-r<1$$

For $\sigma>0$, let

$$L_n(\sigma)=\frac{1}{n(n-1)}\sum_{i<j}\left(Z_{ij}-\sigma^2\widehat{\Delta}_{ij}\widehat{G}_{ij}\right)^2$$

and

$$G_n(\sigma)=\frac{1}{n(n-1)}\sum_{i<j}\left(\widetilde{Z}_{ij}-\sigma^2\Delta_{ij}G_{ij}\right)^2$$

where $\widetilde{Z}_{ij}=Z_{ij}-\frac{\sum_{i<j}\Delta_{ij}b_{ij}(G_{ij})G_{ij}^3}{\sum_{i<j}\Delta_{ij}^2G_{ij}^2}\Delta_{ij}G_{ij}$.

Assumption 9. There exists a function $G(\sigma)$ possessing a unique minimizer such that $\forall\sigma>0$, $G_n(\sigma)\to G(\sigma)$ in probability as $n\to\infty$.

The Assumption 9 is an assumption of convenience and a relaxation could be employed by studying sub-sequences of $G_n(\cdot)$, see Theorem 10.8 and 10.9 in [31]. A similar condition was adopted in Condition (C1) of [32] to prove the consistency of estimator in logistic regression with measurement error.

A.2 Proof of Theorem 2

This is a direct result of Lemma 1 and the Bolzano-Weierstrass theorem.

A.3 Technical lemmas

Suppose the confusion matrix is given, the variance to be estimated by the PCGC model is the variance of random effect:

$$\text{TP1}\times\frac{\sigma_{11}^2}{p\pi_{11}}+\text{FP1}\times\frac{\sigma_{01}^2}{p\pi_{01}},$$

where TP1 and FP1 represent the number of $\boldsymbol{M}_\tau$ and $\boldsymbol{M}_{\mathscr{I}_1}$ that are not filtered out, respectively. Furthermore, we let FN1 and TN1 represent the numbers of $\boldsymbol{M}_\tau$ and $\boldsymbol{M}_{\mathscr{I}_1}$ that are filtered out, respectively.

Assumption 10. There exists constants $c_{\text{TP}},c_{\text{TN}}>0$ and $c_{\text{FP}},c_{\text{FN}}\geq 0$ such that $c_{\text{TP}}+c_{\text{FP}}+c_{\text{TN}}+c_{\text{FN}}=1$ and

$$\lim_{p\to\infty,n\to\infty}\text{TP}:\text{FP}:\text{TN}:\text{FN}=c_{\text{TP}}:c_{\text{FP}}:c_{\text{TN}}:c_{\text{FN}}$$

Lemma 1. *Under Assumptions in Lemma 2, Lemma 3 and Lemma 4, for any $\varepsilon > 0$, there exists $\delta > 0$, such that if $\frac{c_{\text{TP}}}{c_{\text{TP}}+c_{\text{FP}}} > 1-\delta$, we have*

$$\lim_{p\to\infty, n\to\infty} \mathbb{P}(|\widehat{R}^2_{med;causal} - R^2_{med;causal}| < \varepsilon) = 1.$$

and

$$\lim_{p\to\infty, n\to\infty} \mathbb{P}(|\widehat{Q}^2_{med} - Q^2_{med}| < \varepsilon) = 1.$$

Proof. By law of large number, under Assumption 10, we have

$$\lim_{p\to\infty, n\to\infty} \frac{\text{TP1}}{\text{TP1}+\text{FN1}} = \frac{c_{\text{TP}}}{c_{\text{TP}}+c_{\text{FN}}} \qquad \lim_{p\to\infty, n\to\infty} \frac{\text{FP1}}{\text{FP1}+\text{TN1}} = \frac{c_{\text{FP}}}{c_{\text{FP}}+c_{\text{TN}}}$$

Regarding the estimate of σ^2_α, we have $Var(\alpha) = \pi_\alpha \sigma^2_\alpha$, where $\pi_\alpha = c_{\text{TP}} + c_{\text{FN}}$. Given that $\hat{\pi}_\alpha = c_{\text{TP}} + c_{\text{FP}}$ we can have the estimate

$$\hat{\sigma}^2_\alpha = \frac{\widehat{Var(\alpha)}}{\hat{\pi}_\alpha} = \frac{\widehat{Var(\alpha)}}{c_{\text{TP}}+c_{\text{FP}}} = \frac{\widehat{Var(\alpha)}}{c_{\text{TP}}+c_{\text{FN}}} \frac{c_{\text{TP}}+c_{\text{FN}}}{c_{\text{TP}}+c_{\text{FP}}} = \sigma^2_\alpha \frac{\widehat{Var(\alpha)}}{Var(\alpha)} \frac{c_{\text{TP}}+c_{\text{FN}}}{c_{\text{TP}}+c_{\text{FP}}}$$

By Lemma 2 and Lemma 4, we have:

$$\begin{aligned}
\hat{\sigma}^2_\alpha * \hat{\sigma}^2_{11} &\to \sigma^2_\alpha \frac{c_{\text{TP}}+c_{\text{FN}}}{c_{\text{TP}}+c_{\text{FP}}} \times \left(\frac{c_{\text{TP}}}{c_{\text{TP}}+c_{\text{FN}}} \sigma^2_{11} + \frac{c_{\text{FP}}}{c_{\text{FP}}+c_{\text{TN}}} \sigma^2_{01} \right) \\
&= \sigma^2_\alpha \times \left(\frac{c_{\text{TP}}}{c_{\text{TP}}+c_{\text{FP}}} \sigma^2_{11} + \frac{c_{\text{FP}}}{c_{\text{TP}}+c_{\text{FP}}} \frac{\pi_\alpha}{1-\pi_\alpha} \sigma^2_{01} \right)
\end{aligned}$$

as $p, n \to \infty$. The above arguments along with Assumption 2 and Lemma 3 complete the proof. □

Lemma 2. *Under Assumption 1, 2 and 3, we have*

$$|\widehat{Var(\alpha)} - Var(\alpha)| = O_p(\frac{1}{\sqrt{p}} + \frac{1}{\sqrt{n}})$$

Proof. Let $\widetilde{\alpha} = (\text{Vec}(\alpha)^\top, \text{Vec}(\Phi^\top)^\top$ and $D(\boldsymbol{u}) = Q(\widetilde{\alpha}^* + \boldsymbol{u}) - Q(\widetilde{\alpha}^*) = J_1(\boldsymbol{u}) + J_2(\boldsymbol{u})$, where Q is the quadratic loss and

$$\begin{aligned}
J_1(\boldsymbol{u}) &= -2 \left(\sum_{i=1}^n w_i \boldsymbol{\xi}_i^\top \widetilde{\mathbf{X}}_i \right) \boldsymbol{u} \\
J_2(\boldsymbol{u}) &= \boldsymbol{u}^T \left(\sum_{i=1}^n w_i \widetilde{\mathbf{X}}_i^\top \widetilde{\mathbf{X}}_i \right) \boldsymbol{u}
\end{aligned}$$

Let $\widetilde{\mathbf{W}} = \text{diag}(w_1 1_p^\top, \ldots, w_n 1_p^\top)$, $\widetilde{\boldsymbol{\xi}} = (\boldsymbol{\xi}_1^\top, \ldots, \boldsymbol{\xi}_n^\top)^\top$ and $\widetilde{\mathbf{X}} = (\widetilde{\mathbf{X}}_1^\top, \ldots, \widetilde{\mathbf{X}}_n^\top)^\top$.

By Lemma 5, we have

$$|J_1(\boldsymbol{u})| = 2 \left\langle \widetilde{\mathbf{X}}^\top \widetilde{\mathbf{W}} \widetilde{\boldsymbol{\xi}}, \boldsymbol{u} \right\rangle \le 2 \|\widetilde{\mathbf{X}}^\top \widetilde{\mathbf{W}} \widetilde{\boldsymbol{\xi}}\| \|\boldsymbol{u}\| = 2 O_p(\sqrt{n} p) \|\boldsymbol{u}\|$$

By Assumption 2, we have

$$J_2(\boldsymbol{u}) \geq n\|\boldsymbol{u}\|^2 R^{-2}$$

Then for any $0 < c < 1$, we can choose sufficiently large M such that

$$\mathbb{P}(\inf_{\|\boldsymbol{u}\|=1} Q(\widetilde{\alpha}^* + Mp/\sqrt{n}\boldsymbol{u}) - Q(\widetilde{\alpha}^*) > 0) \geq 1 - c$$

for n sufficiently large. Thus $\|\widehat{\widetilde{\alpha}^*} - \widetilde{\alpha}^*\| = O_p(\frac{p}{\sqrt{n}})$. Then we have

$$|\frac{1}{p}\|\widehat{\boldsymbol{\alpha}}\|^2 - \frac{1}{p}\|\boldsymbol{\alpha}^*\|^2| = O_p(\frac{1}{\sqrt{n}})$$

By the fact that $|Var(\alpha) - \frac{1}{p}\|\boldsymbol{\alpha}^*\|^2| = O_p(\frac{1}{\sqrt{p}})$, we have

$$|\widehat{Var(\alpha)} - Var(\alpha)| = O_p(\frac{1}{\sqrt{p}} + \frac{1}{\sqrt{n}})$$

□

Lemma 3. *Under Assumption 1, 2, 4 and 5, we have $\|\widehat{\boldsymbol{\nu}} - \boldsymbol{\nu}_0\|_2 = o_p(1)$, where $\boldsymbol{\nu}$ corresponds to $(X_i, \boldsymbol{C}_i, \boldsymbol{W}_i)$.*

Proof. Let

$$L_n(\boldsymbol{\nu}) = \frac{1}{n}\sum_1^n w_i \left\{ y_i \log\Phi\left(\widehat{\boldsymbol{c}}_i^\top \boldsymbol{\nu}\right) + (1 - y_i)\log\Phi\left(-\widehat{\boldsymbol{c}}_i^\top \boldsymbol{\nu}\right)\right\}$$

and $\widehat{\boldsymbol{\nu}} = \arg\max_{\boldsymbol{\nu}} L_n(\boldsymbol{\nu})$. Let

$$G_n(\boldsymbol{\nu}) = \frac{1}{n}\sum_1^n w_i \left\{ \Phi\left(\boldsymbol{c}_i^\top \boldsymbol{\nu}_0\right)\log\Phi\left(\boldsymbol{c}_i^\top \boldsymbol{\nu}\right) + \Phi\left(-\boldsymbol{c}_i^\top \boldsymbol{\nu}_0\right)\log\Phi\left(-\boldsymbol{c}_i^\top \boldsymbol{\nu}\right)\right\}$$

Then

$$L_n(\boldsymbol{\nu}) - G_n(\boldsymbol{\nu}) = R_{n,1}(\boldsymbol{\nu}) + R_{n,2}(\boldsymbol{\nu})$$

where

$$R_{n,1}(\boldsymbol{\nu}) = \frac{1}{n}\sum_1^n w_i \left\{ \left(y_i - \Phi\left(\boldsymbol{c}_i^\top \boldsymbol{\nu}_0\right)\right)\log\Phi\left(\widehat{\boldsymbol{c}}_i^\top \boldsymbol{\nu}\right) + \left(1 - y_i - \Phi\left(-\boldsymbol{c}_i^\top \boldsymbol{\nu}_0\right)\right)\log\Phi\left(-\widehat{\boldsymbol{c}}_i^\top \boldsymbol{\nu}\right)\right\}$$

and

$$\begin{aligned} R_{n,2}(\boldsymbol{\nu}) = & \frac{1}{n}\sum_1^n w_i \{\Phi\left(\boldsymbol{c}_i^\top \boldsymbol{\nu}_0\right)\left(\log\Phi\left(\boldsymbol{c}_i^\top \boldsymbol{\nu}\right) - \log\Phi\left(\widehat{\boldsymbol{c}}_i^\top \boldsymbol{\nu}\right)\right) \\ & + \Phi\left(-\boldsymbol{c}_i^\top \boldsymbol{\nu}_0\right)\left(\log\Phi\left(-\boldsymbol{c}_i^\top \boldsymbol{\nu}\right) - \log\Phi\left(-\widehat{\boldsymbol{c}}_i^\top \boldsymbol{\nu}\right)\right)\} \end{aligned}$$

For any fixed $\boldsymbol{\nu}$, under Assumption 2 and 5, $R_{n,1}(\boldsymbol{\nu})$ has mean zero and asymptotically negligible variance, and

$$|R_{n,2}(\boldsymbol{\nu})| \leq 2R\frac{1}{n}\sum_{i=1}^n \frac{\|\widehat{\boldsymbol{c}}_i - \boldsymbol{c}_i\|_2 \|\boldsymbol{\nu}\|_2}{\sqrt{2\pi}\Phi(-C\|\boldsymbol{\nu}\|_2)} = o_p(1)$$

Note $G_n(\boldsymbol{\nu})$ converges in probability to $\mathbb{E}[G_n(\boldsymbol{\nu})]$ for any fixed $\boldsymbol{\nu}$ and $\mathbb{E}[G_n(\boldsymbol{\nu})]$ has a unique maximum at $\boldsymbol{\nu}_0$ under Assumption 4. We have that $L_n(\boldsymbol{\nu})$ converges in probability to $\mathbb{E}[G_n(\boldsymbol{\nu})]$ for any fixed $\boldsymbol{\nu}$. This completes the proof by Theorem 2.2 in [33]. □

We define $\sigma^2_{11(adj)} = \frac{c_{\text{TP}}}{c_{\text{TP}}+c_{\text{FN}}}\sigma^2_{11} + \frac{c_{\text{FP}}}{c_{\text{FP}}+c_{\text{TN}}}\sigma^2_{01}$ and introduce the following lemma.

Lemma 4. *Under Assumption 1, 6, 7, 8, 9 and 10*

$$|\hat{\sigma}^2_{11} - \sigma^2_{11(adj)}| = o_p(1)$$

Proof. Note that

$$\begin{aligned}
L_n(\sigma) - G_n(\sigma) =& \frac{2\sigma^2}{n(n-1)}\sum_{i<j}(\Delta_{ij}G_{ij} - \widehat{\Delta}_{ij}\widehat{G}_{ij})Z_{ij} \\
&+ \frac{\sigma^4}{n(n-1)}\sum_{i<j}(\widehat{\Delta}^2_{ij}\widehat{G}^2_{ij} - \Delta^2_{ij}G^2_{ij}) \\
&- \frac{2\sigma^2}{n(n-1)}\frac{\sum_{i<j}\Delta_{ij}b_{ij}(G_{ij})G^3_{ij}}{\sum_{i<j}\Delta^2_{ij}G^2_{ij}}\sum_{i<j}\Delta^2_{ij}G^2_{ij} \\
&+ \frac{1}{n(n-1)}\frac{\sum_{i<j}\Delta_{ij}b_{ij}(G_{ij})G^3_{ij}}{\sum_{i<j}\Delta^2_{ij}G^2_{ij}}\sum_{i<j}\left(2Z_{ij} - \frac{\sum_{i<j}\Delta_{ij}b_{ij}(G_{ij})G^3_{ij}}{\sum_{i<j}\Delta^2_{ij}G^2_{ij}}\Delta_{ij}G_{ij}\right)G_{ij}
\end{aligned}$$

By Assumption 8, there exists a constant $M_1 > 0$, such that $|Z_{ij}| < M_1$. Then for any fixed $\sigma > 0$, we have

$$\left|\frac{2\sigma^2}{n(n-1)}\sum_{i<j}(\Delta_{ij}G_{ij} - \widehat{\Delta}_{ij}\widehat{G}_{ij})Z_{ij}\right| \leq \frac{2M_1\sigma^2}{n(n-1)}\sum_{i<j}\left|\Delta_{ij}G_{ij} - \widehat{\Delta}_{ij}\widehat{G}_{ij}\right| = o_p(1)$$

and

$$\left|\frac{\sigma^4}{n(n-1)}\sum_{i<j}(\widehat{\Delta}^2_{ij}\widehat{G}^2_{ij} - \Delta^2_{ij}G^2_{ij})\right| = o_p(1)$$

by Assumption 7. There exists a constant $M_2 > 0$, such that $|\Delta_{ij}| < M_2$ by Assumption 8 and

$$\left|2Z_{ij} - \frac{\sum_{i<j}\Delta_{ij}b_{ij}(G_{ij})G^3_{ij}}{\sum_{i<j}\Delta^2_{ij}G^2_{ij}}\Delta_{ij}G_{ij}\right| = O_p(1)$$

by Assumption 6. Then for any fixed $\sigma > 0$, we have that

$$\left|\frac{2\sigma^2}{n(n-1)}\frac{\sum_{i<j}\Delta_{ij}b_{ij}(G_{ij})G^3_{ij}}{\sum_{i<j}\Delta^2_{ij}G^2_{ij}}\sum_{i<j}\Delta^2_{ij}G^2_{ij}\right| \leq 2M_2^2\sigma^2\left|\frac{\sum_{i<j}\Delta_{ij}b_{ij}(G_{ij})G^3_{ij}}{\sum_{i<j}\Delta^2_{ij}G^2_{ij}}\right| = o_p(1)$$

and

$$\left|\frac{1}{n(n-1)}\frac{\sum_{i<j}\Delta_{ij}b_{ij}(G_{ij})G^3_{ij}}{\sum_{i<j}\Delta^2_{ij}G^2_{ij}}\sum_{i<j}\left(2Z_{ij} - \frac{\sum_{i<j}\Delta_{ij}b_{ij}(G_{ij})G^3_{ij}}{\sum_{i<j}\Delta^2_{ij}G^2_{ij}}\Delta_{ij}G_{ij}\right)G_{ij}\right| = o_p(1)$$

Thus we have $|L_n(\sigma) - G_n(\sigma)| = o_p(1)$ for any fixed $\sigma > 0$.

Denote $\widehat{\sigma}_{L_n} = \arg\max_{\sigma>0} L_n(\sigma)$ and $\widehat{\sigma}_{G_n} = \arg\max_{\sigma>0} G_n(\sigma)$. By Assumption 9, both $L_n(\sigma)$ and $G_n(\sigma)$ converge to $G(\sigma)$ in probability for any $\sigma > 0$. Denote the unique minimizer as σ_*,

by Theorem 2.2 in [33], we have $|\widehat{\sigma}^2_{L_n} - \sigma^2_*| = o_p(1)$ and $|\widehat{\sigma}^2_{G_n} - \sigma^2_*| = o_p(1)$. Thus it remains to compute $\lim_{n\to\infty} \widehat{\sigma}^2_{G_n}$.

$$\begin{aligned}
\widehat{\sigma}^2_{G_n} =& \frac{\sum_{i<j} \widetilde{Z}_{ij}\Delta_{ij}G_{ij}}{\sum_{i<j}\Delta^2_{ij}G^2_{ij}} = \frac{\sum_{i<j} Z_{ij}\Delta_{ij}G_{ij}}{\sum_{i<j}\Delta^2_{ij}G^2_{ij}} - \frac{\sum_{i<j}\Delta_{ij}b_{ij}(G_{ij})G^3_{ij}}{\sum_{i<j}\Delta^2_{ij}G^2_{ij}} \\
=& \frac{\sum_{i<j}\mathbb{E}\left[Z_{ij} \mid \mathscr{S} = 1; G_{ij}, t_i, t_j\right]\Delta_{ij}G_{ij}}{\sum_{i<j}\Delta^2_{ij}G^2_{ij}} + \frac{\sum_{i<j} e_{ij}\Delta_{ij}G_{ij}}{\sum_{i<j}\Delta^2_{ij}G^2_{ij}} - \frac{\sum_{i<j}\Delta_{ij}b_{ij}(G_{ij})G^3_{ij}}{\sum_{i<j}\Delta^2_{ij}G^2_{ij}} \\
=& \sigma^2_{11(adj)} + \frac{\sum_{i<j} e_{ij}\Delta_{ij}G_{ij}}{\sum_{i<j}\Delta^2_{ij}G^2_{ij}} \to \sigma^2_{11(adj)}.
\end{aligned}$$

This completes the proof. □

Lemma 5. *Let Assumption 1, 2 and 3 holds. Then*

$$\|\frac{1}{n}\widetilde{\mathbf{X}}^\top\widetilde{\mathbf{W}}\widetilde{\boldsymbol{\xi}}\| = O_p(\frac{p}{\sqrt{n}})$$

Proof. Note that

$$\begin{aligned}
\|\frac{1}{n}\widetilde{\mathbf{X}}^\top\widetilde{\mathbf{W}}\widetilde{\boldsymbol{\xi}}\| \le& \left((q+1)p \sup_{j=1,\ldots,(q+1)p}\left(\frac{1}{n}\widetilde{\mathbf{X}}^\top_{\cdot,j}\widetilde{\mathbf{W}}\widetilde{\boldsymbol{\xi}}\right)^2\right)^{1/2} \\
\le& \sqrt{(q+1)p}\left(\sup_{j=1,\ldots,(q+1)p}\|\frac{1}{n}\widetilde{\mathbf{W}}\widetilde{\mathbf{X}}_{\cdot,j}\|^2 \sup_{\|\boldsymbol{u}\|=1}|\boldsymbol{u}^\top\widetilde{\boldsymbol{\xi}}|^2\right)^{1/2}
\end{aligned}$$

For any $j = 1,\ldots,(q+1)p$, by Assumption 2, we have

$$\|\frac{1}{n}\widetilde{\mathbf{W}}\widetilde{\mathbf{X}}_{\cdot,j}\|^2 \le \frac{1}{n}R^3$$

This leads to

$$\|\frac{1}{n}\widetilde{\mathbf{X}}^\top\widetilde{\mathbf{W}}\widetilde{\boldsymbol{\xi}}\| \le \sqrt{\frac{(q+1)p}{n}}R^{3/2}(\sup_{\|\boldsymbol{u}\|=1}|\boldsymbol{u}^\top\widetilde{\boldsymbol{\xi}}|^2)^{1/2}$$

For sequences l_n such that $l_n/\frac{(q+1)p}{\sqrt{n}} \to \infty$, there exists $K > 0$ such that

$$\begin{aligned}
\mathbb{P}(\|\frac{1}{n}\widetilde{\mathbf{X}}^\top\widetilde{\mathbf{W}}\widetilde{\boldsymbol{\xi}}\| > l_n) \le& \mathbb{P}([\sqrt{\frac{(q+1)p}{n}}R^{3/2}]^4(\sup_{\|\boldsymbol{u}\|=1}|\boldsymbol{u}^\top\widetilde{\boldsymbol{\xi}}|^2)^2 > l_n^4) \\
\le& \frac{(q+1)^2p^2R^6\mathbb{E}(\sup_{\|\boldsymbol{u}\|=1}|\boldsymbol{u}^\top\widetilde{\boldsymbol{\xi}}|^2)^2}{n^2l_n^4} \\
\le& KR^6\frac{(q+1)^4p^4}{n^2l_n^4} \to 0
\end{aligned}$$

where the last inequality follows the same proof of Lemma 3 in [28] by setting $\delta = 2$. This completes the proof. □

References

[1] Dylan Clark-Boucher, Xiang Zhou, Jiacong Du, Yongmei Liu, Belinda L. Needham, Jennifer A. Smith, and Bhramar Mukherjee. Methods for mediation analysis with high-dimensional dna methylation data: Possible choices and comparisons. *PLOS Genetics*, 19(11):1–26, 11 2023.

[2] Judith J.M. Rijnhart, Sophia J. Lamp, Matthew J. Valente, David P. MacKinnon, Jos W.R. Twisk, and Martijn W. Heymans. Mediation analysis methods used in observational research: a scoping review and recommendations. *BMC Medical Research Methodology*, 21(1), December 2021. Publisher Copyright: © 2021, The Author(s).

[3] Tianzhong Yang, Jingbo Niu, Han Chen, and Peng Wei. Estimation of total mediation effect for high-dimensional omics mediators. *BMC Bioinformatics*, 22(1):414, August 2021.

[4] Haixiang Zhang, Yinan Zheng, Zhou Zhang, Tao Gao, Brian Joyce, Grace Yoon, Wei Zhang, Joel Schwartz, Allan Just, Elena Colicino, Pantel Vokonas, Lihui Zhao, Jinchi Lv, Andrea Baccarelli, Lifang Hou, and Lei Liu. Estimating and testing high-dimensional mediation effects in epigenetic studies. *Bioinformatics*, 32(20):3150–3154, 06 2016.

[5] Yuzhao Gao, Haitao Yang, Ruiling Fang, Yanbo Zhang, Ellen L. Goode, and Yuehua Cui. Testing mediation effects in high-dimensional epigenetic studies. *Frontiers in Genetics*, 10, 2019.

[6] Yen-Tsung Huang and Wen-Chi Pan. Hypothesis Test of Mediation Effect in Causal Mediation Model With High-Dimensional Continuous Mediators. *Biometrics*, 72(2):402–413, 09 2015.

[7] Amanda J Fairchild, David P Mackinnon, Marcia P Taborga, and Aaron B Taylor. R2 effect-size measures for mediation analysis. *Behavior research methods*, 41(2):486—498, May 2009.

[8] Sunyi Chi. High-dimensional mediation analysis of multi-omics data. *PhD Dissertation*, 2024. Available at: `https://digitalcommons.library.tmc.edu/utgsbs_dissertations/1332/`.

[9] Yanyi Song, Xiang Zhou, Min Zhang, Wei Zhao, Yongmei Liu, Sharon L. R. Kardia, Ana V. Diez Roux, Belinda L. Needham, Jennifer A. Smith, and Bhramar Mukherjee. Bayesian Shrinkage Estimation of High Dimensional Causal Mediation Effects in Omics Studies. *Biometrics*, 76(3):700–710, 11 2019.

[10] David Golan, Eric S. Lander, and Saharon Rosset. Measuring missing heritability: Inferring the contribution of common variants. *Proceedings of the National Academy of Sciences*, 111(49):E5272–E5281, 2014.

[11] Reuben Baron and David Kenny. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51:1173–1182, 01 1986.

[12] D.P. MacKinnon, C.M. Lockwood, C.H. Brown, W. Wang, and J.M. Hoffman. The intermediate endpoint effect in logistic and probit regression. *Clinical Trials*, 4(5):499–513, 2007.

[13] Judith J. M. Rijnhart, Judith J.M. Rijnhart, Matthew J. Valente, Matthew J. Valente, Heather L. Smyth, Heather L. Smyth, David P. MacKinnon, and David P. MacKinnon. Statistical mediation analysis for models with a binary mediator and a binary outcome: the differences between causal and traditional mediation analysis. *Prevention Science*, 24:408–418, 2021.

[14] Judith J.M. Rijnhart, Matthew J. Valente, David P. MacKinnon, Jos W.R. Twisk, and Martijn W. Heymans. The use of traditional and causal estimators for mediation models with a binary outcome and exposure-mediator interaction. *Structural Equation Modeling: A Multidisciplinary Journal*, 28(3):345–355, 2021.

[15] Judea Pearl. Direct and indirect effects. January 2013.

[16] Tyler J. VanderWeele and Stijn Vansteelandt. Odds Ratios for Mediation Analysis for a Dichotomous Outcome. *American Journal of Epidemiology*, 172(12):1339–1348, 10 2010.

[17] Tyler VanderWeele and Stijn Vansteelandt. Mediation analysis with multiple mediators. *Epidemiologic Methods*, 2(1):95–115, 2014.

[18] Tyler J. VanderWeele. Mediation analysis: A practitioner's guide. *Annual Review of Public Health*, 37(Volume 37, 2016):17–32, 2016.

[19] Sheila M. Gaynor, Joel Schwartz, and Xihong Lin. Mediation analysis for common binary outcomes. *Statistics in Medicine*, 38(4):512–529, 2019.

[20] Judea Pearl. An Introduction to Causal Inference. *The International Journal of Biostatistics*, 6(2):7, February 2010.

[21] Francisco N. F. Q. Simoes, Mehdi Dastani, and Thijs van Ommen. Fundamental properties of causal entropy and information gain. *ArXiv*, abs/2402.01341, 2024.

[22] Zhichao Xu, Chunlin Li, Sunyi Chi, Tianzhong Yang, and Peng Wei. Speeding up interval estimation for R2-based mediation effect of high-dimensional mediators via cross-fitting. *Biostatistics*, page kxae037, 10 2024.

[23] Elizabeth D. Schifano. A review of analysis methods for secondary outcomes in case-control studies. *Communications for Statistical Applications and Methods*, 26(2):103–129, March 2019. Edition: 2019/03/31 Publisher: The Korean Statistical Society, and Korean International Statistical Society.

[24] Zakaria Almuwaqqat, Qin Hui, Chang Liu, Jin J Zhou, Benjamin F Voight, Yuk-Lam Ho, Daniel C Posner, Jason L Vassy, J Michael Gaziano, Kelly Cho, et al. Long-term body mass index variability and adverse cardiovascular outcomes. *JAMA Network Open*, 7(3):e243062–e243062, 2024.

[25] Nina P. Paynter, Raji Balasubramanian, Franco Giulianini, Dong D. Wang, Lesley F. Tinker, Shuba Gopal, Amy A. Deik, Kevin Bullock, Kerry A. Pierce, Justin Scott, Miguel A. Martínez-González, Ramon Estruch, JoAnn E. Manson, Nancy R. Cook, Christine M. Albert, Clary B. Clish, and Kathryn M. Rexrode. Metabolic predictors of incident coronary heart disease in women. *Circulation*, 137(8):841–853, 2018.

[26] Stef van Buuren and Karin Groothuis-Oudshoorn. mice: Multivariate imputation by chained equations in r. *Journal of Statistical Software*, 45(3):1–67, 2011.

[27] Connie W. Tsao, Aaron W. Aday, Zaid I. Almarzooq, Alvaro Alonso, Andrea Z. Beaton, Marcio S. Bittencourt, Amelia K. Boehme, Alfred E. Buxton, April P. Carson, Yvonne Commodore-Mensah, Mitchell S.V. Elkind, Kelly R. Evenson, Chete Eze-Nliam, Jane F. Ferguson, Giuliano Generoso, Jennifer E. Ho, Rizwan Kalani, Sadiya S. Khan, Brett M. Kissela, Kristen L. Knutson, Deborah A. Levine, Tené T. Lewis, Junxiu Liu, Matthew Shane Loop, Jun Ma, Michael E. Mussolino, Sankar D. Navaneethan, Amanda Marma Perak, Remy Poudel, Mary Rezk-Hanna, Gregory A. Roth, Emily B. Schroeder, Svati H. Shah, Evan L. Thacker, Lisa B. VanWagner, Salim S. Virani, Jenifer H. Voecks, Nae-Yuh Wang, Kristine Yaffe, Seth S. Martin, on behalf of the American Heart Association Council on Epidemiology, Prevention Statistics Committee, and Stroke Statistics Subcommittee. Heart disease and stroke statistics—2022 update: A report from the american heart association. *Circulation*, 145(8):e153–e639, 2022.

[28] Tamar Sofer, Elizabeth D Schifano, David C Christiani, and Xihong Lin. Weighted pseudolikelihood for snp set analysis with multiple secondary outcomes in case-control genetic association studies. *Biometrics*, 73(4):1210–1220, 2017.

[29] Jianqing Fan, Yuan Liao, and Martina Mincheva. Large covariance estimation by thresholding principal orthogonal complements. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 75(4):603–680, 2013.

[30] Anna Bonnet. Heritability estimation in case-control studies. *Electronic Journal of Statistics*, 12(1):1662 – 1716, 2018.

[31] R. Tyrrell Rockafellar. *Convex Analysis*. Princeton University Press, Princeton, N.J., 1970.

[32] Leonard A Stefanski and Raymond J Carroll. Covariate measurement error in logistic regression. *The annals of statistics*, pages 1335–1351, 1985.

[33] Per Kragh Andersen and Richard D Gill. Cox's regression model for counting processes: a large sample study. *The annals of statistics*, pages 1100–1120, 1982.