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 mediation analysis in case-control studies, which can also be applied to cohort studies. Our estimator remains consistent in the presence of non-mediators and weak effect sizes in extensive simulations. Theoretical justification on consistency is provided under mild conditions and without requiring exact mediator selection. In a case-control substudy of the Women's Health Initiative involving 2150 individuals, we found that many metabolites were mediators with weak effects in the path from BMI to coronary heart disease, and an estimated 89% (95% CI: 73%–91%) of the variation explained by BMI in the underlying liability of coronary heart disease was estimated to be mediated by the measured metabolomics. The proposed estimation procedure is implemented in R package "r2MedCausal", available on GitHub.

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

*Z. Kang and L. Chen contributed equally. C. Li and T. Yang jointly supervised the project.

†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 Statistics, Iowa State University, IA 50011.

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 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: (i) Handling many weak mediators with cumulatively non-negligible effect sizes, as suggested by our preliminary analysis in Section 5.1; (ii) Accounting for ascertainment bias from the case-control design. Both challenges have received little attention in the mediation literature (Clark-Boucher et al., 2023; Rijnhart et al., 2021), highlighting the need for a more robust framework for high-dimensional mediation analysis with weak mediators and binary outcomes.

Quantifying the total mediation effect is nontrivial for high-dimensional settings. Most of the existing measures are variants of the mean-based product-of-coefficient (POC) measures, which are prone to cancellation in the presence of bidirectional mediation effects, a phenomenon that frequently occurs with omics mediators (Yang et al., 2021). In addition, many high-dimensional approaches rely on strong sparsity assumptions, which are not suitable for handling mediators with small and non-sparse effect sizes (Zhang et al., 2016; Gao et al., 2019). Some methods define the mediation effect based on transformed mediators, relaxing sparsity requirements on the observed scale but reducing interpretability as a trade-off (Huang and Pan, 2015).

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. (2021) extended the variance-based R^2 measure originally proposed by Fairchild et al. (2009) to accommodate multiple and high-dimensional mediators under a mixed model framework. This method captures the non-zero total mediation effect by quantifying the variance in the continuous 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 a likelihood-based McFadden's R^2 for its stability and relative independence from disease prevalence, compared to other R^2 measures under investigation (Chi, 2024). However, these R^2 measures (Yang et al., 2021; Fairchild et al., 2009; Chi, 2024) only capture the associative relationships rather than causal effects. In addition, the McFadden's R^2 defined for binary outcomes appears to be less sensitive to disease prevalence but remains to vary with it. Additionally, its interpretability is limited, as it is more closely linked to measuring goodness-of-fit than to explaining variance. In another line of work, Song et al. (2019) proposed a L_2 norm measure of the mediation effect while simply treating binary outcomes as continuous. However, both the use of the L_2 norm and the treatment of binary outcomes as continuous lack theoretical support. 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. The novel R^2 mediation effect measure is defined on a latent variable scale, i.e., liability, which is invariant to disease prevalence and remains its connection with the classic R^2 defined for continuous outcomes. However, the estimation of R^2 on the liability scale is not robust to ascertainment bias in a case-control study. To this end, 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 modify a heritability estimation method, phenotype-correlation genotype-correlation (PCGC) regression, a generalization of Haseman-Elston regression, to estimate consistent variance components under case-control studies (Golan et al., 2014). 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 methods. Furthermore, under mild conditions and without requiring exact mediator selection, we establish the consistency of the estimators. Applied to the motivating WHI case-control dataset, our method effectively captures 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 $R^2_{\text{med;causal}}$ 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 summarize the contributions and limitations of our work and outline future research directions. The theoretical proof of the statistical properties of the proposed measure in a high-dimensional context is provided in the Appendix. Finally, an R package, "r2MedCausal", has been developed to implement the proposed estimation procedure.

2 Methods

Our goal is to develop a variance-based total mediation effect measure that provides a causal interpretation for binary outcomes, particularly for a study with a case-control design. A foundational framework in mediation analysis is the approach by Baron and Kenny (1986), who proposed the following structural equations for continuous outcomes, capturing the relationships among an exposure X, a p-dimensional mediator $\boldsymbol{M}$, and an outcome Y:

$$\boldsymbol{M} = \boldsymbol{\alpha}^{(c)} X + \boldsymbol{\Psi}^{(c)} \boldsymbol{C} + \boldsymbol{\xi}, \tag{1}$$

$$Y = \gamma^{(c)} X + \boldsymbol{\beta}^{(c)\top} \boldsymbol{M} + \boldsymbol{\theta}^{(c)\top} \boldsymbol{C} + \epsilon, \tag{2}$$

where $\boldsymbol{C}$ is a $q \times 1$ vector representing covariates. The superscript $^{(c)}$ indicates coefficients corresponding to the continuous outcome setting. The coefficient $\boldsymbol{\alpha}^{(c)} = (\alpha_1^{(c)}, \alpha_2^{(c)}, \ldots, \alpha_p^{(c)})^\top$ captures the relationship between the exposure and mediator after adjusting for the effect of covariates. Similarly, $\boldsymbol{\beta}^{(c)} = (\beta_1^{(c)}, \beta_2^{(c)}, \ldots, \beta_p^{(c)})^\top$ captures the relationship between the mediators and the continuous outcome after adjusting for both the exposure and covariates. The parameter $\gamma^{(c)}$ is the direct effect of the exposure variable, and $\boldsymbol{\Psi}^{(c)} : p \times q$ and $\boldsymbol{\theta}^{(c)} : q \times 1$ represent the coefficients of covariates. The variables ϵ, $\boldsymbol{\xi} = (\xi_1, \xi_2, \ldots, \xi_p)^\top$ are normally distributed errors.

When the outcome is binary, (2) is replaced with a logistic regression model (MacKinnon et al., 2007), which naturally accommodates the case-control study design. However, the preferred mediation effect scale varies by outcome prevalence, with odds ratios for rare outcomes and risk ratios for common ones (see discussion in Section 2.1). Instead, the liability model provides a unified framework for analyzing both rare and common diseases. In addition, the variance-based mediation effect has consistent interpretation for both continuous outcome and binary outcome under the liability model. Its usage has also been demonstrated in the analysis of microbiome data (Sohn et al., 2022). To this end, we utilize the liability-based mediation model, which is defined

by the following structural equations:

$$\boldsymbol{M} = \boldsymbol{\alpha} X + \boldsymbol{\Psi} \boldsymbol{C} + \boldsymbol{\xi}, \tag{3}$$

$$l = \gamma X + \boldsymbol{\beta}^\top \boldsymbol{M} + \boldsymbol{\theta}^\top \boldsymbol{C} + \epsilon, \tag{4}$$

$$Y = \mathbf{1}(l > t), \tag{5}$$

where l is the underlying latent variable, Y is the observed binary outcome. The function $\mathbf{1}(\cdot)$ is an indicator function, with threshold t determining K, the prevalence of the trait in the population. The coefficients $\boldsymbol{\alpha}, \boldsymbol{\beta}, \gamma, \boldsymbol{\Psi}$, and $\boldsymbol{\theta}$ retain the same interpretations as in the continuous outcome model, now defined in relation to the latent liability. Similarly, the error terms ϵ and $\boldsymbol{\xi}$ remain normally distributed.

2.1 Existing mediation effect measures

In the mediation model defined by (1)–(2), the total effect of the exposure X on the continuous outcome Y is decomposed into a direct effect and a mediation effect. The total effect is typically estimated by regressing Y on X, after adjusting for covariates. The mediation effect can be estimated using either the difference-in-coefficients method, which subtracts the direct effect from the total effect, or the POC approach, which computes the inner product of $\boldsymbol{\alpha}^{(c)}$ and $\boldsymbol{\beta}^{(c)}$. For binary outcomes modeled with logistic regression, the difference-in-coefficients method may encounter the noncollapsibility problem (Greenland et al., 1999), making the POC approach preferable (VanderWeele, 2016). The POC measure is also used in high-dimensional mediation analysis methods, such as HIMA and HDMA (Zhang et al., 2016; Gao et al., 2019). Furthermore, to quantify the fraction of the total effect of an exposure on the outcome that is transmitted through mediators, the traditional proportion-mediated (PM) measure is given by $\boldsymbol{\alpha}^\top \boldsymbol{\beta} / (\gamma + \boldsymbol{\alpha}^\top \boldsymbol{\beta})$ (MacKinnon et al., 2007). These mean-based measures have lots of merits. However, they can be misleading in high-dimensional settings where numerous bidirectional component-wise mediation effects are present, potentially causing cancellation in the POC and PM measures.

In recent years, causal mediation analysis within the counterfactual framework has garnered significant interest, as it generalizes to a wide range of mediation models (Pearl, 2001; VanderWeele and Vansteelandt, 2010). For a continuous outcome Y, the natural indirect effect (NIE) and the natural direct effect (NDE) for an exposure contrast between x and x' are defined as:

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

where $Y(x, M(x))$ represents the potential outcome when both exposure and mediator take their natural values under $X = x$, $Y(x, M(x'))$ represents the outcome when the exposure is set to x, but the mediator is fixed at the level it would take under $X = x'$, and $Y(x', M(x'))$ represents the outcome when both exposure and mediator take their natural values under $X = x'$. Both effects are conditioned on covariates $\boldsymbol{C}$. When there is no exposure–mediator interaction and the relationships are linear, NIE $= \boldsymbol{\alpha}^{(c)\top} \boldsymbol{\beta}^{(c)} (x - x')$. Without loss of generality, we set $x = 1$ and $x' = 0$. Therefore, NIE coincides with the POC measure NIE = POC $= \boldsymbol{\alpha}^{(c)\top} \boldsymbol{\beta}^{(c)}$. For the mediation and direct effect estimates to have a causal interpretation, the following confounding assumptions must be met (VanderWeele, 2016).

Assumption 1 (No Unmeasured Confounding)**.** The following assumptions are required for valid causal mediation analysis:

A1 No unmeasured confounding of the exposure-mediator effect.

A2 No unmeasured confounding of the mediator-outcome effect.

A3 No unmeasured confounding of the exposure-outcome effect.

A4 No mediator-outcome confounder that is itself affected by the exposure.

For a binary outcome, when Assumption 1 is satisfied and the outcome is rare, logistic regression can be used to estimate the causal mediation effect on an odds ratio scale as: $\mathrm{OR}^{\mathrm{NIE}} = \exp(\boldsymbol{\alpha}^\top \boldsymbol{\beta})$ (VanderWeele and Vansteelandt, 2010). However, when the outcome is not rare, odds ratios suffer from noncollapsibility, meaning that $\exp(\boldsymbol{\alpha}^\top \boldsymbol{\beta})$ does not provide an interpretable mediation effect estimate, as it depends on covariates adjustment. In such cases, log-binomial models are preferred to estimate the causal mediation effect on a risk ratio (RR) scale, given by $\mathrm{RR}^{\mathrm{NIE}} = \exp(\boldsymbol{\alpha}^\top \boldsymbol{\beta})$ (VanderWeele, 2016). Additionally, Gaynor et al. (2019) proposed a closed-form estimator for the NIE on the odds ratio scale using a probit approximation to the logit, relaxing the rare disease assumption when using logistic regression. However, regardless of the mediation framework or regression models used, this estimation of the mediation effect is fundamentally based on the POC calculation, thus suffering from the same limitation as the POC approach for a continuous outcome.

Recognizing the potential drawbacks of the POC measure, BAMA (Song et al., 2019) proposes an alternative L_2 norm for summarizing the component-wise mediation effects: $\mathrm{IE}_{\mathrm{BAMA}} = \sum_{i=1}^{p} (\alpha_i \beta_i)^2$. While this modification no longer suffers from cancellation of bidirectional component-wise mediation effects, it lacks rigorous theoretical support for the causal interpretation. An alternative mediation effect measure is the R^2 measure, which was first proposed in a single-mediator model for continuous outcomes by Fairchild et al. (2009). Yang et al. (2021) extended this measurement to high-dimensional mediation analysis for high-dimensional omics mediator $\boldsymbol{M}$. The measure is expressed as:

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

where $R^2_{Y,X} = [\mathrm{var}(Y) - \mathrm{var}(Y|X)]/\mathrm{var}(Y)$, $R^2_{Y,\boldsymbol{M}} = [\mathrm{var}(Y) - \mathrm{var}(Y|\boldsymbol{M})]/\mathrm{var}(Y)$, and $R^2_{Y,\boldsymbol{M}X} = [\mathrm{var}(Y) - \mathrm{var}(Y|X, \boldsymbol{M})]/\mathrm{var}(Y)$. The proposed method provides a robust, interpretable second-moment-based measure that has low bias and variance for continuous outcomes. However, it does not have a clear causal interpretation.

2.2 A novel R^2 measure with causal interpretation

Aiming to achieve the robustness of rare disease assumption and ease of interpretation, we define a new R^2 measure for high-dimensional mediators under the liability scale for binary outcomes. Following Yang et al. (2021), under the structural equation model specified by (3)–(5), the R^2 measure would be defined as:

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

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 $\boldsymbol{M}$ would not change the formulas, aside from notation differences, despite mediation analysis explicitly assuming directed causal pathways from

X to Y and from X to $\boldsymbol{M}$ to Y. To resolve this issue and explicitly incorporate the causal direction, we introduce the do-operators in our R^2 measure within Pearl's Structural Causal Framework (Pearl, 2010). Building on our group's previous work for continuous traits (Li et al., 2025), we formulate the new R^2 measure for binary outcomes as:

$$R^2_{\text{med;causal}} = R^2_{l,do(X)} + R^2_{l,do(\boldsymbol{M})} - R^2_{l,do(X,\boldsymbol{M})}. \tag{6}$$

We define

$$R^2_{l,do(\boldsymbol{M})} = \frac{\text{var}\left(\mathbb{E}\left(l \mid do(\boldsymbol{M} \sim \boldsymbol{M}')\right)\right)}{\text{var}(l)}, \tag{7}$$

where $\boldsymbol{M}'$ is identically distributed as $\boldsymbol{M}$ but independent of all other variables in the model and the do-operator is a stochastic intervention of $\boldsymbol{M}$. 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, in contrast to $R^2_{l,\boldsymbol{M}}$, which may be confounded by X. We similarly define

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

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

where X', X'', and $\boldsymbol{M}''$ are independent copies of X and $\boldsymbol{M}$, respectively, drawn from the same distribution but independent of other variables in the model. Furthermore, we assume that while the mediators may be correlated, there are no directed causal relationships among them. Instead, their correlation is driven by the exposure X and unobserved/unknown latent factors $\boldsymbol{U}$. This assumption is formalized as follows.

Assumption 2 (Conditional Parallel Mediator Assumption)**.** There exist random variables $\boldsymbol{U}$ that have a global influence on $\boldsymbol{M}$, where $M_k \perp\!\!\!\perp M_j | X, \boldsymbol{U}$ for any $k, j \in \{1, \ldots, p\}$.

We note that Assumption 2 is commonly used in methods and applications to high-dimensional omics mediation analysis (Huang and Pan, 2015; Chi, 2024; Perera et al., 2022), as usually the omics data were measured at a single time point rather than sequentially, exemplified by our motivating dataset, WHI. Given Assumptions 1 and 2, the estimated $R^2_{\text{med;causal}}$ can be interpreted as the proportion of variation in the underlying liability for the outcome that is *causally* explained by the exposure through the mediators. According to (3)–(5), we have

$$R^2_{\text{med;causal}} = \left(\sum_{k=1}^{p} \alpha_k^2 \beta_k^2\right) \frac{\text{var}(X)}{\text{var}(l)}. \tag{8}$$

As discussed in (MacKinnon et al., 2007), the PM measure in traditional mediation analysis is given by $\boldsymbol{\alpha}^\top \boldsymbol{\beta}/(\gamma + \boldsymbol{\alpha}^\top \boldsymbol{\beta})$. In the causal mediation analysis framework, assuming a rare binary outcome, the PM measure can be expressed as $\text{OR}^{\text{NDE}}(\text{OR}^{\text{NIE}} - 1)/(\text{OR}^{\text{NDE}}\text{OR}^{\text{NIE}} - 1)$, where $\text{OR}^{\text{NDE}} = \exp(\gamma)$ and $\text{OR}^{\text{NIE}} = \exp(\boldsymbol{\alpha}^\top \boldsymbol{\beta})$ denote the NDE and NIE on the odds ratio scale (VanderWeele, 2016). However, neither measure 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. As an alternative to the PM measure, we propose a new relative Q^2_{med} measure, which is bounded between 0 and 1:

$$Q^2_{\text{med}} = \frac{R^2_{\text{med;causal}}}{R^2_{\text{med;causal}} + \gamma^2\sigma_x^2/\text{var}(l)} = \frac{\sum_{k=1}^p \alpha_k^2\beta_k^2}{\sum_{k=1}^p \alpha_k^2\beta_k^2 + \gamma^2}. \tag{9}$$

This quantity is the proportion of variation of liability causally explained by X indirectly through the mediators over the sum of the proportion of the variation explained indirectly and directly by X. It can be interpreted as the proportion of the exposure explained variance on the liability scale of the outcome that is accounted for by the mediator.

Remark 1 (Relationship with causal entropy)**.** The proposed causal coefficient of determination is connected to the existing concept of causal entropy (Simoes et al., 2024). It can be shown that for a single mediator,

$$R^2_{l,do(\boldsymbol{M})} = \mathbb{E}_{m_0\sim P_{\boldsymbol{M}'}}\left[\frac{\text{var}(l) - \text{var}(l|do(\boldsymbol{M}) = m_0)}{\text{var}(l)}\right], \tag{10}$$

where $do(\boldsymbol{M}) = m_0$ implies that we intervene $\boldsymbol{M}$ 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 (7) and (10) are equivalent for a single mediator. The connection between our definition and causal entropy can be seen directly through (10).

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 $\boldsymbol{M}_{\mathcal{T}}$, and three types of non-mediators ($\boldsymbol{M}_{\mathcal{I}_1}, \boldsymbol{M}_{\mathcal{I}_2}, \boldsymbol{M}_{\mathcal{I}_3}$), as illustrated in Figure 1:

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

We denote the proportions of $\boldsymbol{M}_{\mathcal{T}}$, $\boldsymbol{M}_{\mathcal{I}_1}$, $\boldsymbol{M}_{\mathcal{I}_2}$, and $\boldsymbol{M}_{\mathcal{I}_3}$ 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

$$\alpha_k \sim \pi_\alpha N(0, \sigma_\alpha^2) + (1 - \pi_\alpha)\delta_0, \tag{11}$$

$$\beta_k \sim \pi_{11} N(0, \frac{\sigma_{11}^2}{p\pi_{11}}) + \pi_{01} N(0, \frac{\sigma_{01}^2}{p\pi_{01}}) + (1 - \pi_\beta)\delta_0, \quad k = 1, \ldots, p, \tag{12}$$

where δ_0 denotes a point mass at zero. Under this working model and the binary mediation model defined by (3)–(5), it can be shown that

$$R^2_{\text{med;causal}} = \sigma_\alpha^2\sigma_x^2\sigma_{11}^2/\text{var}(l). \tag{13}$$

For the relative Q^2_{med} measure, we have

$$Q^2_{\text{med}} = \frac{\sigma_\alpha^2 \sigma_{11}^2}{\sigma_\alpha^2 \sigma_{11}^2 + \gamma^2}. \tag{14}$$

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

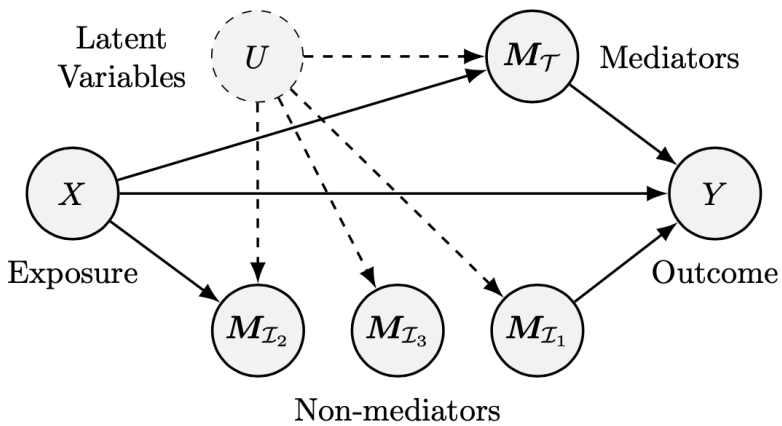


Figure 1: Graph representation of a mediation model.

3 Estimation procedure

To estimate the proposed R^2 measure, we adapted the cross-fitting framework as described by Xu et al. (2024), to improve the robustness of mediation effect estimation. Cross-fitting mitigates bias introduced by the winner's curse during the variable screening step (Göring et al., 2001). The approach involves a two-stage procedure. First, the sample is randomly split into two subsamples $\mathcal{D}^{(1)}$ and $\mathcal{D}^{(2)}$. In Step 1, mediation selection and variance estimation of $\boldsymbol{\alpha}$ are performed on subsample $\mathcal{D}^{(1)}$ to get the filtered mediator index set $\hat{\mathcal{T}}^{(1)}$ and variance estimate $\hat{\sigma}_\alpha^{(1)2}$. In Step 2, the estimation of other parameters is conducted on subsample $\mathcal{D}^{(2)}$ based on the selected mediators $\hat{\mathcal{T}}^{(1)}$. 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 $\hat{R}^{(1)2}_{\text{med;causal}}$ and $\hat{R}^{(2)2}_{\text{med;causal}}$. The procedure is visualized in Figure 2, and the mediator selection and parameter estimation process are detailed in Subsections 3.1-3.3.

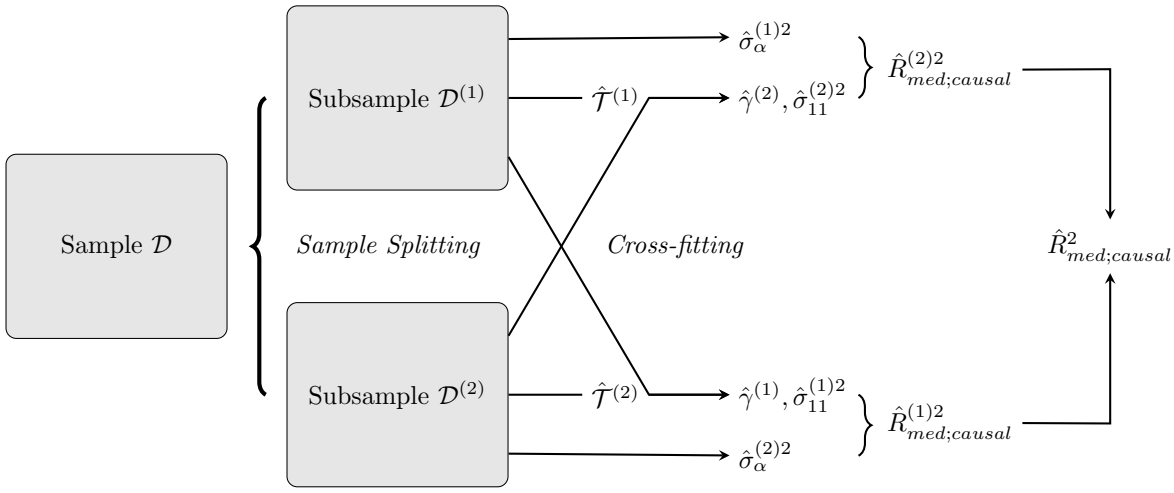


Figure 2: Cross-fitted estimation procedure of $R^2_{\text{med;causal}}$.

3.1 Estimation of σ_α^2 and mediator selection

The estimation of $\boldsymbol{\alpha}$ in our mediation analysis framework has been extensively studied as a secondary outcome analysis problem in case-control studies (Schifano, 2019). To estimate σ_α^2, we used the inverse probability weighting (IPW) method introduced by Horvitz and Thompson (1952). In this approach, each individual's contribution to the estimating equations is weighted by the reciprocal of their selection probability, effectively constructing a pseudo-sample that is more representative of the target population than the original case-control sample.

Let S_i denote the inclusion indicator for the i-th individual, and define the inverse weight as $w_i = 1/P(S_i = 1|Y_i, M_i, X_i)$. The IPW estimating equation for estimating each element of $\boldsymbol{\alpha}$ is given by (Sofer et al., 2017):

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

where $h(X_i)$ is a user-specified function ensuring that $\mathbb{E}(\partial U_{\text{IPW},i}/\partial \alpha_k)$ is invertible. Under the full-ascertainment assumption (Stene, 1977), $w_i = 1$ for case and $\frac{P(1-K)}{K(1-P)}$ for control, where K represents the population-level trait prevalence, and P denotes the case-control ratio in the study, both assumed to be known. Setting $h(X_i) = X_i$ simplifies the estimating equation into a weighted least squares regression problem.

To identify non-mediators and noise variables, Wald tests are conducted to detect $\boldsymbol{M}_{\mathcal{I}_1} = \{M_j|\alpha_j = 0 \text{ and } \beta_j \neq 0 \text{ for } j \in \mathcal{I}_1\}$ and $\boldsymbol{M}_{\mathcal{I}_3} = \{M_j|\alpha_j = 0 \text{ and } \beta_j = 0 \text{ for } j \in \mathcal{I}_3\}$. Wald test p-values are adjusted using a multiple testing correction procedure to control the false discovery rate (FDR) at a pre-specified threshold. The resulting filtered set of mediators, denoted by $\boldsymbol{M}_{\hat{\mathcal{T}}}$, is subsequently used in cross-fitted variance component estimation. Given that $\text{var}(\alpha) = \pi_\alpha \sigma_\alpha^2$, the variance estimator is obtained as $\hat{\sigma}_\alpha^2 = \widehat{\text{var}(\alpha)}/\hat{\pi}_\alpha$, where $\text{var}(\alpha)$ is empirically estimated from the IPW results, and π_α is estimated based on the ratio of the number of selected mediators and the total number of potential mediators.

3.2 Estimation of γ

The IPW approach can also be applied to estimate the parameter γ. By selecting an appropriate estimating equation, the IPW problem becomes a weighted probit regression problem with the same weight as described in Section 3.1. However, directly regressing the outcome Y on the exposure variable X may lead to biased estimates, as it does not account for $\boldsymbol{M}$. To mitigate this bias, we include the top principal components (PCs) of the mediator matrix across individuals, denoted as $\boldsymbol{W}$, as 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 (Lemma 3, Appendix).

3.3 Estimation of σ_{11}^2

For the estimation of σ_{11}^2, we propose to use PCGC regression (Golan et al., 2014). PCGC regression is a method of moments that estimates variance components under a liability model while accounting for the oversampling in case-control designs. As a generalization of the Haseman–Elston regression method (Haseman and Elston, 1972), PCGC regression models the expected correlation of the observed phenotype as a function of variance components and the correlation of its corresponding design matrix, which, in our context, corresponds to the mediator matrix across individuals. Unlike the restricted maximum likelihood (REML) method (Lee et al., 2011), which tends to underestimate variance components in case-control studies due to ascertainment bias, PCGC explicitly accounts for this bias by modeling the inclusion probability of individuals in the study. However, this approach implicitly assumes weak correlation among rows of the design matrix, an assumption that is likely violated by mediators due to their relationship with the exposure (Equation (3)). To account for this, we modify the PCGC regression framework and estimate σ_{11}^2 following the cross-fitting procedure as described in Section **??**:

Step 1 (First subsample): Identify the mediators with $\alpha \neq 0$, i.e., $\boldsymbol{M}_{\mathcal{T}}$ and $\boldsymbol{M}_{\mathcal{I}_2}$ following Section 3.1 and denote the selected mediator set as $\hat{\mathcal{T}}$. Estimate standardized residual $\hat{\boldsymbol{\xi}}_{\hat{\mathcal{T}}}$ by regressing scaled $\boldsymbol{M}$ on X, $\boldsymbol{C}$, $\boldsymbol{W}$ using the IPW approach.

Step 2 (Second subsample): Apply PCGC regression with $\hat{\boldsymbol{\xi}}_{\hat{\mathcal{T}}}$ as the standardized design matrix and include X, $\boldsymbol{C}$ and $\boldsymbol{W}$ as covariates to estimate $\hat{\sigma}_{11}^2$.

After adjusting for covariates that drive the correlation between mediators, the resulting residuals $\hat{\boldsymbol{\xi}}_{\hat{\mathcal{T}}}$ in Step 1 would satisfy the weak correlation assumption in PCGC regression. We note that adjusting for $\boldsymbol{W}$ in the analysis not only provides us a consistent estimate of direct effect under high-dimensional setting (Section 3.2), it also potentially controls for unknown confounding in the same spirit as Xu et al. (2024). Additionally, ascertainment bias is appropriately addressed in both steps through the IPW approach for residual estimation and PCGC regression for variance estimation.

In the presence of covariates, PCGC regression follows a two-step procedure to estimate σ_{11}^2. First, the conditional probability of the trait for the i-th individual $\hat{P}_i = P(Y_i = 1|X_i, \boldsymbol{C}_i, \boldsymbol{W}_i)$ is estimated using a logistic regression model by regressing Y on X, $\boldsymbol{C}$ and $\boldsymbol{W}$. Based on this estimate, we compute the adjusted trait prevalence $\hat{K}_i$ and corresponding liability threshold $\hat{t}_i$ for

the i-th individual as:

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

where $\Phi(\cdot)$ denotes the cumulative distribution function (CDF) of the standard normal distribution.

Second, PCGC regression adjusts for ascertainment bias by examining the relationship between phenotypic correlation and correlations among rows of the design matrix, conditioned on the effect of covariates. The phenotypic correlation is expressed as

$$Z_{ij} = \frac{(Y_i - \hat{P}_i)(Y_j - \hat{P}_j)}{\sqrt{\hat{P}_i(1-\hat{P}_i)}\sqrt{\hat{P}_j(1-\hat{P}_j)}},$$

where Y_i and Y_j are the observed binary outcomes for the i-th and j-th individuals. The correlation matrix, which captures the correlation between individuals based on residuals, is computed as $\hat{G} = \hat{\boldsymbol{\xi}}_{\hat{\mathcal{T}}}\hat{\boldsymbol{\xi}}_{\hat{\mathcal{T}}}^{\top}/p$. Finally, we obtain an estimator for σ_{11}^2 by regressing the phenotypic correlation Z_{ij} on the modified correlation:

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

3.4 Theoretical results on consistency

In this section, we present the theoretical properties of the proposed method. Theorem 1 established the consistency of our proposed total mediation effect estimator $\widehat{R}^2_{\text{med;causal}}$ and relative measurement estimator $\widehat{Q}^2_{\text{med}}$ under mild regularity conditions.

Theorem 1. *Under Assumptions 1-10 in the Appendix, and assume there exists $c_0 > 0$ that $\frac{\text{TP}}{p} > c_0$, then for any $\epsilon > 0$, there exists $\delta > 0$, such that if $\frac{\text{TP}}{\text{TP+FP}} > 1 - \delta$ for p, n large enough, we have*

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

and

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

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 $\boldsymbol{M}_{\mathcal{I}_2}$ 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 $\boldsymbol{M}_{\mathcal{I}_1}$ and noise mediators $\boldsymbol{M}_{\mathcal{I}_3}$. We assume that the

FDR for selecting variables impacted by exposure is stringent enough to minimize the inclusion of non-mediators type $\boldsymbol{M}_{\mathcal{I}_1}$ and noise mediators $\boldsymbol{M}_{\mathcal{I}_3}$. In practice, our simulations suggest that a reasonable choice, such as FDR ≤ 0.01 using Benjamini-Hochberg (BH) procedure, 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 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

We conducted simulation studies to evaluate the consistency of the proposed estimators of $R^2_{\text{med;causal}}$ and Q^2_{med}. Four scenarios were considered, each varying the composition of mediator types while fixing the total number of mediators at $p = 2000$:

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

The binary outcome was generated under the liability threshold model, with population disease prevalence $K = 0.05$ and study disease prevalence $P = 0.5$ to mimic a case-control design with a 1:1 case-control ratio. Sample size n varies from 500 to 2000 with an increment of 500. We set a single exposure variable $X \sim N(0, 1)$ and its direct effect parameter $\gamma = 0.5$. We specified $\sigma^2_\alpha = 0.3$, the cumulative variance of the true mediators $\boldsymbol{M}_{\mathcal{T}}$ as $\sigma^2_{11} = 0.5$, and that of the non-mediators $\boldsymbol{M}_{\mathcal{I}_1}$ as $\sigma^2_{01} = 0.3$. The coefficients $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ were generated from the working model (11). To ensure Assumptions 1 and 2 hold, we generated $\xi_k = \alpha_0 U + \nu_k$ for $k = 1, \ldots, p$, where $U \sim N(0, 1)$, $\nu_k \sim N(0, 1)$ and $\alpha_0 \sim N(0, 0.3)$, with U, ν_k and α_0 mutually independent. The vector of $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_p)^\top$ was then used to generate mediators as $\boldsymbol{M} = \boldsymbol{\alpha} X + \boldsymbol{\xi}$. The error term for the outcome model was generated from $\epsilon \sim N(0, 1 - \sigma^2_{11} - \sigma^2_{01})$. For each scenario, we ran 100 replications across all sample sizes. The mean absolute bias, standard deviation and mean square error were reported for both measures.

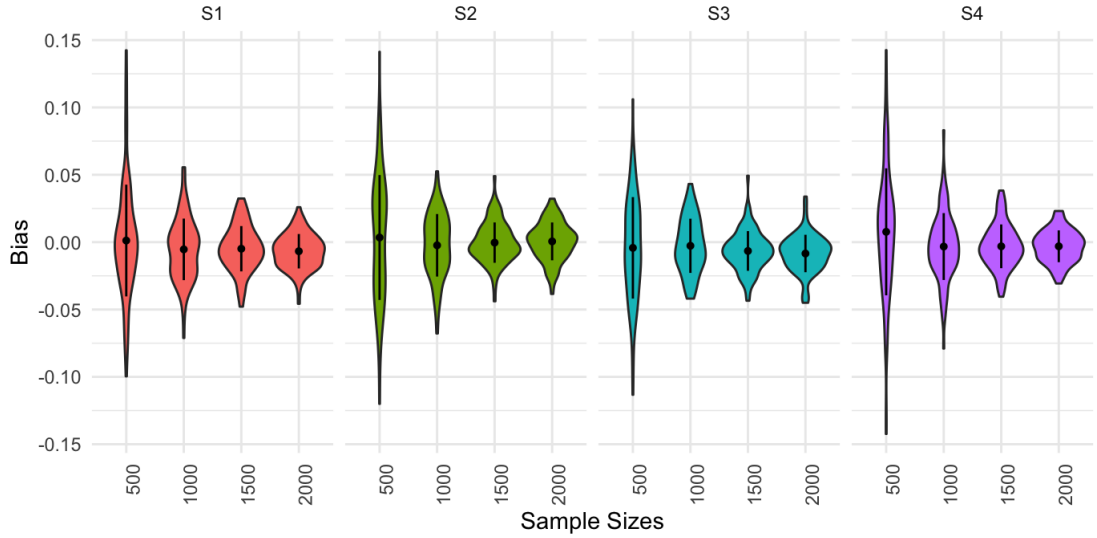
Under S1, we further compared our methods to competing methods: HIMA (Zhang et al., 2016), HDMA (Gao et al., 2019) and BAMA (Song et al., 2019) in terms of total mediation effect estimate and PM estimate. We varied the prevalence of disease K from 0.5 (cohort study) to 0.05 (case-control study) and the sample size from 500 to 2000. For the competing methods, we calculated mean-based mediation effects as recommended in their paper. For methods 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 $\text{IE}_{\text{BAMA}} = \sum_{i=1}^p (\hat{\alpha}_i \hat{\beta}_i)^2$. The results are summarized in the following section. Furthermore, we calculated the corresponding $R^2_{\text{med;causal}}$ and Q^2_{med} based on their component-wise estimates $\hat{\boldsymbol{\alpha}}$, $\hat{\boldsymbol{\beta}}$ and $\hat{\gamma}$. Relative bias was also reported, defined as Bias/True Value.

4.2 Simulation results

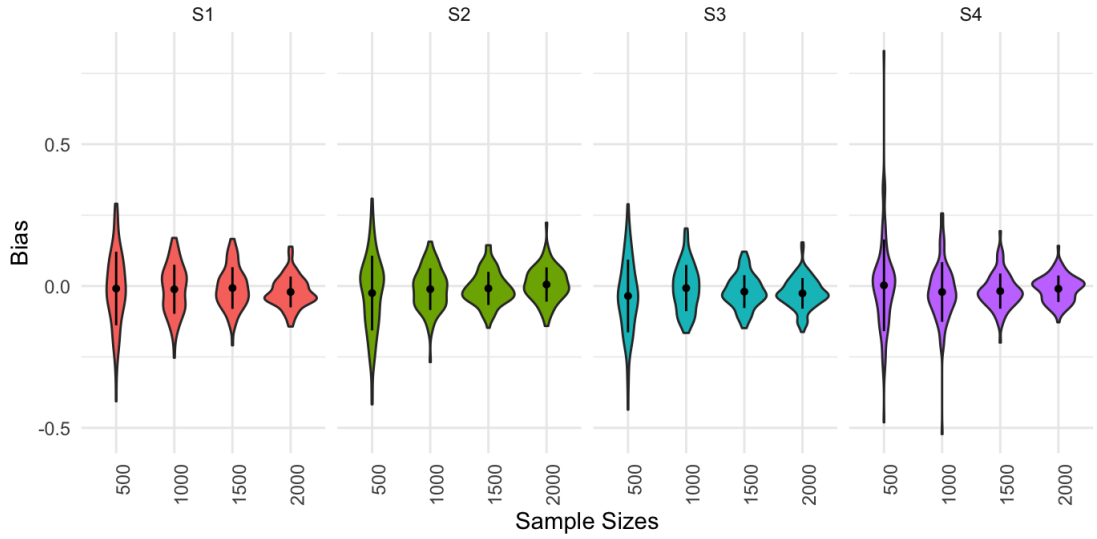
Figure 3 and Table 1 summarize the estimation performance of two proposed measures in terms of absolute bias, standard deviation, and mean square error. Overall, the proposed measures had minimal bias and error observed across different sample sizes and number/types of non-mediators. As expected, the performance metrics improved as the sample size increased. Furthermore, the violin plots (Figure 3) 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 PM under case-control designs ($K = 0.05$). Our proposed method consistently outperformed competing methods across all simulation scenarios. HIMA and HDMA often 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. Moreover, neither HIMA nor HDMA accounts for ascertainment bias when estimating the exposure–mediator associations $\boldsymbol{\alpha}$, which likely contributes to their biased effect estimates under case-control designs. 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. Simulation results with $K = 0.2$ and 0.5 are presented in Tables 6 and 7 in the Appendix. As the prevalence of disease K increased, our method showed reduced efficiency in small sample sizes, indicated by the higher estimation variance. However, with larger sample sizes, the method's performance became more stable, yielding comparable results across different K values.

Computation times in both case-control and cohort studies are summarized in Table 4. 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 running time of other methods increased noticeably with larger sample sizes. Overall, HIMA, HDMA, and BAMA exhibited approximately linear growth as n increased, while the proposed method has the slowest growth with running time. Particularly, BAMA, under a Bayesian sampling estimation framework, is the slowest approach. While HDMA was faster than BAMA, it remained slower than HIMA and the proposed method.



(a) Bias for $R^2_{\text{med;causal}}$ 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 Scenarios S1-S4. 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.

		$R^2_{\text{med;causal}}$			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 the relative performance of IE estimators in a case-control study ($K = 0.05$). Results of the proposed $R^2_{\text{med;causal}}$ measure and the original method-specific measure, i.e., POE for HIMA and HDMA, $\sum_{i=1}^{p}(\hat{\alpha}_i\hat{\beta}_i)^2$ for BAMA, are reported. Each cell contains the average absolute relative bias, with the standard deviation of the relative bias shown in parentheses.

	$R^2_{\text{med;causal}}$ Measure				Original Method-Specific Measure		
n	Proposed	HIMA	HDMA	BAMA	HIMA	HDMA	BAMA
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)

Table 3: Comparison of the relative performance of PM estimators in a case-control study ($K = 0.05$). Results of the proposed $R^2_{\text{med;causal}}$ measure and the original method-specific measure, i.e., $\hat{\boldsymbol{\alpha}}^\top\hat{\boldsymbol{\beta}}/(\hat{\gamma} + \hat{\boldsymbol{\alpha}}^\top\hat{\boldsymbol{\beta}})$ for HIMA and HDMA, $\sum_{i=1}^{p}(\hat{\alpha}_i\hat{\beta}_i)^2/(\sum_{i=1}^{p}(\hat{\alpha}_i\hat{\beta}_i)^2 + \hat{\gamma}^2)$ for BAMA, are reported. 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	
n	Proposed	HIMA	HDMA	BAMA	HIMA	HDMA
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)

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 (WHI) study

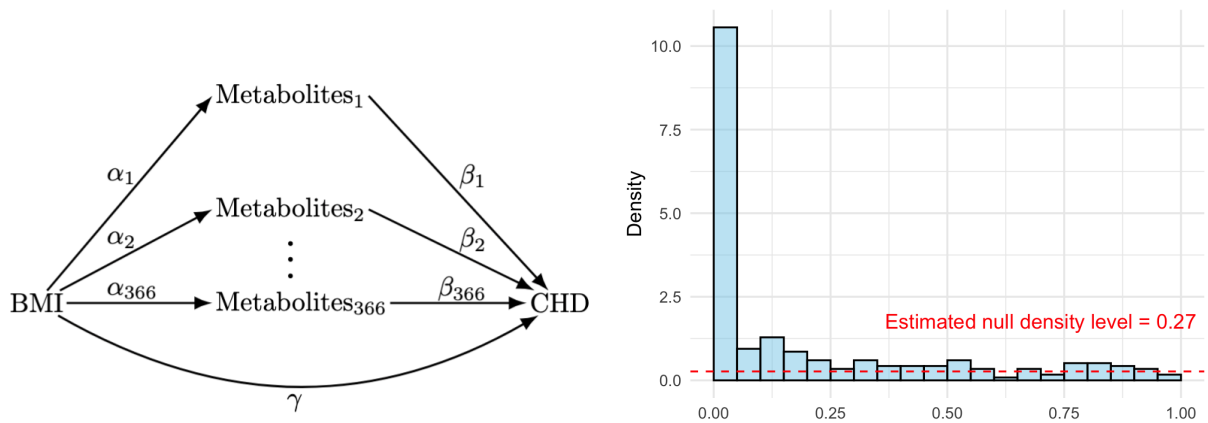
The WHI nested case-control substudy (dbGaP accession number: phs001334) examines the relationship between metabolomic profiles and incident CHD in postmenopausal women. The substudy consists of subjects from both the WHI Observational Study (WHI-OS) and the WHI Hormone Therapy Trials (WHI-HT). A total 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 (Almuwaqqat et al., 2024), but its impact on CHD through metabolites remains 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 (Paynter et al., 2018). Metabolomic profiles were log-transformed. We further removed 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 (van Buuren and Groothuis-Oudshoorn, 2011). After these steps, the final dataset contained 2150 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 (Tsao et al., 2022).

5.1 Preliminary analysis: Evidence of weak effects

In this section, we show that data exhibited evidence of mediation signals in the paths of interest (BMI-metabolites-CHD as in Figure 4), particularly from weak mediators. Figure 4(a) presents a conceptual diagram of the hypothesized mediation structure in the WHI case-control study, where BMI influences CHD risk through high-dimensional metabolomics profiles. As a preliminary analysis, we first conducted exposure-mediator association tests to identify a subset of mediator candidates, using the significance threshold of 0.05 for p-values. Figure 4(b) displays a histogram of p-values for testing the effect of the subset of mediators on the outcome, conditional on BMI. Both the exposure–mediator and mediator–outcome tests were marginal tests and adjusted for relevant covariates including smoking status, alcohol consumption, and education level. The red dashed line represents the estimated null density level, computed as the proportion of null mediators using Storey's method (Storey, 2002), multiplied by the expected uniform density under the null hypothesis. P-values above this level are more likely to originate from mediators with a non-zero effect on the outcome. The prevalence of such p-values, particularly for $p > 0.05$, suggests the presence of diffuse weak mediation signals. This pattern contrasts with the expected distribution under a sparse alternative, where only a few strong mediators produce a sharp concentration of p-values near zero, while the rest follow a uniform distribution under the null.

In addition, a total of 123 metabolites exhibited statistically significant associations in both marginal tests on exposure-mediator association and mediator-outcome association. However, HIMA and HDMA identified only 13 and 9 mediators, respectively, as they primarily capture mediators with strong effects. This contrast highlights the importance of capturing weaker mediation effects, which may collectively contribute to the mediation process for BMI and CHD, as well as other potential exposure-outcome relationships of interest, especially those with widespread impact on molecular phenotypes (Moore et al., 2014; Sebastiani et al., 2024) .



(a) Conceptual mediation diagram.

(b) Histogram of p-values for testing mediator-outcome associations.

Figure 4: (a) Conceptual diagram illustrating the hypothesized mediation structure between BMI, metabolomics mediators, and CHD in the WHI case-control study. (b) Histogram of p-values for testing $\beta_i = 0$ in the outcome model, where β_i represents the effect of the i-th metabolite on CHD, conditional on BMI. The tests shown are restricted to mediator candidates that passed a prior test for exposure-mediator associations (i.e., with α_i significantly different from 0). The red dashed line represents the estimated density of null p-values using Storey's method.

5.2 Comparison of total mediation effect estimates

Table 5 summarizes the total mediation estimates of BMI on CHD through metabolites obtained from different methods. For HIMA, HDMA, and BAMA, empirical $R^2_{\text{med;causal}}$ 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.

Notably, evidence of opposing component-wise mediation effect estimates was observed across methods. HDMA method produced the smallest POC mediation effect estimate $\hat{\boldsymbol{\alpha}}^\top\hat{\boldsymbol{\beta}}$. Among the 13 significant mediators identified by HDMA, 7 exhibited positive effects while 6 exhibited negative effects, leading to substantial cancellation in the total estimate. In contrast, HIMA identified 9 significant mediators, 8 of which had positive effects, resulting in a POC mediation effect estimate that was approximately ten times larger than that of HDMA, despite having a much smaller $R^2_{\text{med;causal}}$ estimate. While HDMA had the largest $R^2_{\text{med;causal}}$ estimate, it included three potential non-mediators with p-values of $\alpha > 0.05$, i.e., these three non-mediators may be included due to an inflated type I error in estimating α, which arose from not accounting for selection effects. Our proposed method produced the second largest $R^2_{\text{med;causal}}$ estimate, demonstrating its capability to capture the total mediation effect that other methods may miss.

The proposed method estimated the total mediation effect of BMI on CHD through the selected set of 366 metabolites as $R^2_{\text{med;causal}} = 0.0072$, with a 95% confidence interval of (0.0065, 0.0093), indicating that approximately 0.7% of the variance in the underlying liability for CHD was causally explained by the metabolomic mediation pathway. In addition, the proportion of the total exposure 's explained variance on the liability scale that is accounted for by metabolomic profileseffect mediated by metabolites, as quantified by Q^2_{med}, was estimated at 0.8935, with a 95% confidence interval of (0.7335, 0.9114), suggesting that a substantial portion of BMI's effect on CHD liability was mediated through these metabolomic profiles.

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

	$\hat{R}^2_{\text{med;causal}}$	$\hat{\boldsymbol{\alpha}}^\top\hat{\boldsymbol{\beta}}$	$\sum_i(\hat{\alpha}_i\hat{\beta}_i)^2$	$\hat{Q}^2_{\text{med}}$	$\frac{\hat{\boldsymbol{\alpha}}^\top\hat{\boldsymbol{\beta}}}{\hat{\boldsymbol{\alpha}}^\top\hat{\boldsymbol{\beta}}+\hat{\gamma}}$
Proposed	0.0072 (0.065, 0.0093)	-	-	0.8935 (0.7335, 0.9114)	-
HIMA	0.0033 (0.0021, 0.0041)	0.09 (0.0677, 0.0914)	0.0031 (0.003, 0.0043)	0.8962 (0.4909, 0.9352)	0.821 (0.5472, 0.8469)
HDMA	0.0174 (0.0098, 0.021)	0.0082 (-0.068, 0.095)	0.0162 (0.0085, 0.0221)	0.4443 (0.2315, 0.9883)	0.0517 (-0.3326, 0.8911)
BAMA	0.0002 (0.0001, 0.0002)	0.0231 (0.0122, 0.0344)	0.0002 (0.0001, 0.0002)	0.4954 (0.0873, 0.9947)	0.7361 (0.02538, 2.0473)

6 Discussion

We have developed a novel mediation analysis framework to estimate the total mediation effect under high-dimensional settings for binary outcomes, particularly tailored for case-control studies. Existing high-dimensional mediation methods for binary outcomes are primarily designed for

cohort studies, often overlooking ascertainment bias in estimating the parameter α. In addition, the choice of mediation measures often varies by disease prevalence. The key contributions of our method include a new R^2 measure designed for binary outcomes, which addresses the challenge of opposite-direction effects, invariant to disease prevalence, 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 many mediators with weak effects. To ensure theoretical validity, we demonstrate that under mild conditions, and without requiring exact mediator selection, which permits the inclusion of mediators with weak effects, our estimators converge reliably when the false discovery rate is appropriately controlled (see 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. The proposed method is implemented in the R package "r2MedCausal", available on GitHub at `https://github.com/zhiyu-kang/r2MedCausal`.

Despite its strengths, our proposed $R^2_{\text{med;causal}}$ estimation procedure has some limitations. First of all, our estimation method specifically targets to capture weak but non-negligible mediation effects. It does not take into account strong effects explicitly. In other words, when only a few sparse strong effects exist, our current estimation procedure may not be the most suitable choice. Therefore, in practice, we recommend performing preliminary analysis as in Section 5.1 before using our method. Employing a multiple-step procedure may solve this issue (Li et al., 2025); however, it is not directly assessed in this work. Secondly, while PCGC was proposed to account for ascertainment bias in case-control studies, it may be less efficient than likelihood-based methods for cohort studies. Thirdly, as a proportion estimator, Q^2_{med} becomes unstable when the denominator gets close to zero, i.e., when no mediation and direct effects are present. We suggest cautious use of this relative measure when the total effect is small, following the recommendations of Baron and Kenny (1986). In terms of further research directions, we are interested in extending our causal mediation framework to multi-omics or tensor data, where mediators originate from multiple sources on different scales.

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. The authors acknowledge the Minnesota Supercomputing Institute (MSI) at the University of Minnesota for providing resources that contributed to the research results reported in this paper. The author also acknowledges the use of ChatGPT, a generative AI tool, to assist in refining sentence structure, improving clarity, and ensuring conciseness in the manuscript. PW was partially supported by National Institutes of Health (NIH) grant R21HL170213. TY was supported by St. Baldrick Career Award and NIH grant R01AG074858. TY also wants to thank Ms. Josey Sorenson in helping improving the interpretation of the proposed measures.

A Additional simulation results

Table 6: 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.

		$R^2_{\text{med;causal}}$ Measure				Method-Specific Original Measure		
K	n	Proposed	HIMA	HDMA	BAMA	HIMA	HDMA	BAMA
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 7: Comparison of relative performance of PM 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
0.2	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)

B Technical proofs

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

B.1 Technical conditions

Assumption 3. 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 4. 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 3 and 4 are standard conditions in case-control studies, see Conditions (C2),(C3) and (C9) in Sofer et al. (2017) for an example.

Assumption 5. There exists a constant $c > 0$, such that $\mathbb{E}[\mathrm{var}(X \mid \boldsymbol{M})] > c$.

Assumption 6. 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 3 corresponds to the standard eigenvalue condition and Assumption 6 is very mild, with Theorem 4 in Fan et al. (2013) as an example.

Following the derivation of heritability estimation with known fixed effects in Golan et al. (2014), we have

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

where

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

and

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

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 7. Let $e_{ij} = Z_{ij} - \mathbb{E}\left[Z_{ij} \mid \mathcal{S} = 1; G_{ij}, t_i, t_j\right]$, 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 Bonnet (2018). Under their assumptions, the first part of Assumption 7 is implied by their Lemma 1 and Lemma 2, and the second part of Assumption 7 is implied by their Lemma 4.

Assumption 8. We have that

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

Assumption 9. 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 $\widetilde{Z}_{ij} = Z_{ij} - \frac{\sum_{i<j} \Delta_{ij} b_{ij}(G_{ij}) G_{ij}^3}{\sum_{i<j} \Delta_{ij}^2 G_{ij}^2} \Delta_{ij} G_{ij}$.

Assumption 10. There exists a function $G(\sigma)$ possessing a unique minimizer such that $\forall \sigma > 0$, $G_n(\sigma) \to G(\sigma)$ in probability as $n \to \infty$.

Assumption 10 is an assumption of convenience, and relaxation could be employed by studying sub-sequences of $G_n(\cdot)$, see Theorem 10.8 and 10.9 in Rockafellar (1970). A similar condition was adopted in Condition (C1) of Stefanski and Carroll (1985) to prove the consistency of the estimator in logistic regression with measurement error.

B.2 Proof of Theorem 1

This is a direct result of Lemma 1 (see below) and the Bolzano-Weierstrass theorem.

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

Assumption 11. 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 $\epsilon > 0$, there exists $\delta > 0$, such that if $\frac{c_{\text{TP}}}{c_{\text{TP}}+c_{\text{FP}}} > 1 - \delta$, we have*

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

and

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

Proof. By law of large number, under Assumption 11, 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 $\text{var}(\alpha) = \pi_\alpha \sigma^2_\alpha$, 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

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

By Lemma 2 and Lemma 4, we have:

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

as $p, n \to \infty$. The above arguments along with Assumption 3 and Lemma 3 complete the proof. Specifically, δ can chosen to be

$$\min\left(\frac{\epsilon\, \text{var}(l)}{2\sigma^2_x \sigma^2_\alpha |\sigma^2_{01} \frac{\pi_\alpha}{1-\pi_\alpha} - \sigma^2_{11}|}, \frac{\epsilon(\sigma^2_\alpha \sigma^2_{11} + \gamma^2)^2}{4|\sigma^2_{01} \frac{\pi_\alpha}{1-\pi_\alpha} - \sigma^2_{11}|(2\gamma^2 + \epsilon(\sigma^2_\alpha \sigma^2_{11} + \gamma^2))}, 1\right),$$

if $\sigma^2_{11} \neq \sigma^2_{01} \times \frac{\pi_\alpha}{1-\pi_\alpha}$ or $\delta = 1$ if $\sigma^2_{11} = \sigma^2_{01} \times \frac{\pi_\alpha}{1-\pi_\alpha}$. □

Lemma 2. *Under Assumption 1, 2, 3 and 4, we have*

$$|\widehat{\text{var}(\alpha)} - \text{var}(\alpha)| = O_p\left(\frac{1}{\sqrt{p}} + \frac{1}{\sqrt{n}}\right).$$

Proof. Let $\widetilde{\alpha} = (\text{Vec}(\alpha)^\top, \text{Vec}(\Psi^\top)^\top$ 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

$$J_1(\boldsymbol{u}) = -2\left(\sum_{i=1}^n w_i \boldsymbol{\xi}_i^\top \widetilde{\mathbf{X}}_i\right) \boldsymbol{u},$$

$$J_2(\boldsymbol{u}) = \boldsymbol{u}^T \left(\sum_{i=1}^n w_i \widetilde{\mathbf{X}}_i^\top \widetilde{\mathbf{X}}_i\right) \boldsymbol{u}.$$

Let $\widetilde{\mathbf{W}} = \mathrm{diag}(w_1, \ldots, w_n) \otimes I_p$, $\widetilde{\boldsymbol{\xi}} = (\boldsymbol{\xi}_1^\top, \ldots, \boldsymbol{\xi}_n^\top)^\top$ and $\widetilde{\mathbf{X}} = (\widetilde{\mathbf{X}}_1^\top, \ldots, \widetilde{\mathbf{X}}_n^\top)^\top$, where $\mathrm{diag}(w_1, \ldots, w_n)$ is the $n \times n$ diagonal matrix with the $(w_1, \ldots, w_n)$ on its main diagonal and $\otimes$ denotes the Kronecker product.

By Lemma 5, we have

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

By Assumption 3, 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 $|\,\mathrm{var}(\alpha) - \frac{1}{p}\|\boldsymbol{\alpha}^*\|^2| = O_p(\frac{1}{\sqrt{p}})$, we have

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

□

Lemma 3. *Under Assumptions 1, 2, 3, 5 and 6, 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 3 and 6, $R_{n,1}(\boldsymbol{\nu})$ 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 5. 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 Andersen and Gill (1982). □

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

Lemma 4. *Under Assumption 1, 2, 7, 8, 9, 10 and 11*

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

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 9, 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 8. There exists a constant $M_2 > 0$, such that $|\Delta_{ij}| < M_2$ by Assumption 9 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 7. 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

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

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 10, 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 Andersen and Gill (1982), we have $|\widehat{\sigma}_{L_n}^2 - \sigma_*^2| = o_p(1)$ and $|\widehat{\sigma}_{G_n}^2 - \sigma_*^2| = o_p(1)$. Thus it remains to compute $\lim_{n\to\infty}\widehat{\sigma}_{G_n}^2$.

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

This completes the proof. □

Lemma 5. *Let Assumptions 1, 2, 3 and 4 hold. Then*

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

Proof. Note that

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

For any $j = 1, \ldots, (q+1)p$, by Assumption 3, 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

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

where the last inequality follows the same proof of Lemma 3 in Sofer et al. (2017) by setting $\delta = 2$. This completes the proof. □

References

Almuwaqqat, Z., Hui, Q., Liu, C., Zhou, J. J., Voight, B. F., Ho, Y.-L., Posner, D. C., Vassy, J. L., Gaziano, J. M., Cho, K., et al. (2024). Long-term body mass index variability and adverse cardiovascular outcomes. *JAMA Network Open* **7,** e243062.

Andersen, P. K. and Gill, R. D. (1982). Cox's regression model for counting processes: a large sample study. *The Annals of Statistics* pages 1100–1120.

Baron, R. and Kenny, D. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology* **51,** 1173–1182.

Bonnet, A. (2018). Heritability estimation in case-control studies. *Electronic Journal of Statistics* **12,** 1662 – 1716.

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

Clark-Boucher, D., Zhou, X., Du, J., Liu, Y., Needham, B. L., Smith, J. A., and Mukherjee, B. (2023). Methods for mediation analysis with high-dimensional dna methylation data: Possible choices and comparisons. *PLOS Genetics* **19,** 1–26.

Fairchild, A. J., Mackinnon, D. P., Taborga, M. P., and Taylor, A. B. (2009). R2 effect-size measures for mediation analysis. *Behavior Research Methods* **41,** 486—498.

Fan, J., Liao, Y., and Mincheva, M. (2013). Large covariance estimation by thresholding principal orthogonal complements. *Journal of the Royal Statistical Society Series B: Statistical Methodology* **75,** 603–680.

Gao, Y., Yang, H., Fang, R., Zhang, Y., Goode, E. L., and Cui, Y. (2019). Testing mediation effects in high-dimensional epigenetic studies. *Frontiers in Genetics* **10,** 1195.

Gaynor, S. M., Schwartz, J., and Lin, X. (2019). Mediation analysis for common binary outcomes. *Statistics in Medicine* **38,** 512–529.

Golan, D., Lander, E. S., and Rosset, S. (2014). Measuring missing heritability: Inferring the contribution of common variants. *Proceedings of the National Academy of Sciences* **111,** E5272–E5281.

Greenland, S., Pearl, J., and Robins, J. M. (1999). Confounding and collapsibility in causal inference. *Statistical Science* **14,** 29 – 46.

Göring, H. H., Terwilliger, J. D., and Blangero, J. (2001). Large upward bias in estimation of locus-specific effects from genomewide scans. *The American Journal of Human Genetics* **69,** 1357–1369.

Haseman, J. and Elston, R. (1972). The investigation of linkage between a quantitative trait and a marker locus. *Behavior genetics* **2,** 3–19.

Horvitz, D. G. and Thompson, D. J. (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association* **47,** 663–685.

Huang, Y.-T. and Pan, W.-C. (2015). Hypothesis Test of Mediation Effect in Causal Mediation Model With High-Dimensional Continuous Mediators. *Biometrics* **72,** 402–413.

Lee, S. H., Wray, N. R., Goddard, M., and Visscher, P. M. (2011). Estimating missing heritability for disease from genome-wide association studies. *American journal of human genetics* **88,** 294–305.

Li, C., Chen, L., Kang, Z., Pankow, J. S., and Yang, T. (2025). Measuring weak effects in high dimensional mediation analysis. *medRxiv* pages 2025–01.

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

Moore, S. C., Matthews, C. E., Sampson, J. N., Stolzenberg-Solomon, R. Z., Zheng, W., Cai, Q., Tan, Y. T., Chow, W.-H., Ji, B.-T., Liu, D. K., Xiao, Q., Boca, S. M., Leitzmann, M. F., Yang, G., Xiang, Y. B., Sinha, R., Shu, X. O., and Cross, A. J. (2014). Human metabolic correlates of body mass index. *Metabolomics* **10,** 259–269.

Paynter, N. P., Balasubramanian, R., Giulianini, F., Wang, D. D., Tinker, L. F., Gopal, S., Deik, A. A., Bullock, K., Pierce, K. A., Scott, J., Martínez-González, M. A., Estruch, R., Manson, J. E., Cook, N. R., Albert, C. M., Clish, C. B., and Rexrode, K. M. (2018). Metabolic predictors of incident coronary heart disease in women. *Circulation* **137,** 841–853.

Pearl, J. (2001). Direct and indirect effects. *Proceedings of the 17th Conference in Uncertainty in Artificial Intelligence* **411,** 411–420.

Pearl, J. (2010). An Introduction to Causal Inference. *The International Journal of Biostatistics* **6,** 7.

Perera, C., Zhang, H., Zheng, Y., Hou, L., Qu, A., Zheng, C., Xie, K., and Liu, L. (2022). HIMA2: high-dimensional mediation analysis and its application in epigenome-wide DNA methylation data. *BMC Bioinformatics* **23,** 296.

Rijnhart, J., Lamp, S., Valente, M., MacKinnon, D., Twisk, J., and Heymans, M. (2021). Mediation analysis methods used in observational research: a scoping review and recommendations. *BMC Medical Research Methodology* **21,**.

Rockafellar, R. T. (1970). *Convex Analysis*. Princeton University Press, Princeton, N.J.

Schifano, E. D. (2019). A review of analysis methods for secondary outcomes in case-control studies. *Communications for Statistical Applications and Methods* **26,** 103–129.

Sebastiani, P., Monti, S., Lustgarten, M. S., Song, Z., Ellis, D., Tian, Q., Schwaiger-Haber, M., Stancliffe, E., Leshchyk, A., Short, M. I., Ardisson Korat, A. V., Gurinovich, A., Karagiannis, T., Li, M., Lords, H. J., Xiang, Q., Marron, M. M., Bae, H., Feitosa, M. F., Wojczynski, M. K., O'Connell, J. R., Montasser, M. E., Schupf, N., Arbeev, K., Yashin, A., Schork, N., Christensen, K., Andersen, S. L., Ferrucci, L., Rappaport, N., Perls, T. T., and Patti, G. J. (2024). Metabolite signatures of chronological age, aging, survival, and longevity. *Cell Reports* **43,** 114913.

Simoes, F. N. F. Q., Dastani, M., and van Ommen, T. (2024). Fundamental properties of causal entropy and information gain. *arXiv* **abs/2402.01341,**.

Sofer, T., Cornelis, M. C., Kraft, P., and Tchetgen, E. J. T. (2017). Control function assisted ipw estimation with a secondary outcome in case-control studies. *Statistica Sinica* **27,** 785–804.

Sofer, T., Schifano, E. D., Christiani, D. C., and Lin, X. (2017). Weighted pseudolikelihood for snp set analysis with multiple secondary outcomes in case-control genetic association studies. *Biometrics* **73,** 1210–1220.

Sohn, M. B., Lu, J., and Li, H. (2022). A compositional mediation model for a binary outcome: Application to microbiome studies. *Bioinformatics* **38,** 16–21.

Song, Y., Zhou, X., Zhang, M., Zhao, W., Liu, Y., Kardia, S. L. R., Diez Roux, A. V., Needham, B. L., Smith, J. A., and Mukherjee, B. (2019). Bayesian Shrinkage Estimation of High Dimensional Causal Mediation Effects in Omics Studies. *Biometrics* **76,** 700–710.

Stefanski, L. A. and Carroll, R. J. (1985). Covariate measurement error in logistic regression. *The Annals of Statistics* pages 1335–1351.

Stene, J. (1977). Assumptions for different ascertainment models in human genetics. *Biometrics* **33,** 523–527.

Storey, J. D. (2002). A direct approach to false discovery rates. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* **64,** 479–498.

Tsao, C. W., Aday, A. W., Almarzooq, Z. I., Alonso, A., Beaton, A. Z., Bittencourt, M. S., Boehme, A. K., Buxton, A. E., Carson, A. P., Commodore-Mensah, Y., Elkind, M. S., Evenson, K. R., Eze-Nliam, C., Ferguson, J. F., Generoso, G., Ho, J. E., Kalani, R., Khan, S. S., Kissela, B. M., Knutson, K. L., Levine, D. A., Lewis, T. T., Liu, J., Loop, M. S., Ma, J., Mussolino, M. E., Navaneethan, S. D., Perak, A. M., Poudel, R., Rezk-Hanna, M., Roth, G. A., Schroeder, E. B., Shah, S. H., Thacker, E. L., VanWagner, L. B., Virani, S. S., Voecks, J. H., Wang, N.-Y., Yaffe, K., Martin, S. S., on behalf of the American Heart Association Council on Epidemiology, Committee, P. S., and Subcommittee, S. S. (2022). Heart disease and stroke statistics—2022 update: A report from the american heart association. *Circulation* **145,** e153–e639.

van Buuren, S. and Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in r. *Journal of Statistical Software* **45,** 1–67.

VanderWeele, T. J. (2016). Mediation analysis: A practitioner's guide. *Annual Review of Public Health* **37,** 17–32.

VanderWeele, T. J. and Vansteelandt, S. (2010). Odds Ratios for Mediation Analysis for a Dichotomous Outcome. *American Journal of Epidemiology* **172,** 1339–1348.

Xu, Z., Li, C., Chi, S., Yang, T., and Wei, P. (2024). Speeding up interval estimation for R2-based mediation effect of high-dimensional mediators via cross-fitting. *Biostatistics* **26,** kxae037.

Yang, T., Niu, J., Chen, H., and Wei, P. (2021). Estimation of total mediation effect for high-dimensional omics mediators. *BMC Bioinformatics* **22,** 414.

Zhang, H., Zheng, Y., Zhang, Z., Gao, T., Joyce, B., Yoon, G., Zhang, W., Schwartz, J., Just, A., Colicino, E., Vokonas, P., Zhao, L., Lv, J., Baccarelli, A., Hou, L., and Liu, L. (2016). Estimating and testing high-dimensional mediation effects in epigenetic studies. *Bioinformatics* **32,** 3150–3154.