Biometrics 72, 402–413
June 2016

DOI: 10.1111/biom.12421

Hypothesis Test of Mediation Effect in Causal Mediation Model with High-Dimensional Continuous Mediators

Yen-Tsung Huang[1,*] and Wen-Chi Pan[2]

[1]Departments of Epidemiology and Biostatistics, Brown University, 121 South Main Street, Providence, Rhode Island 02912, U.S.A.

[2]Institute of Environmental and Occupational Health Sciences, National Yang-Ming University, No. 155, Section 2 Linong Street, Beitou District, Taipei City 112, Taiwan

email: Yen-Tsung_Huang@brown.edu

Summary. Causal mediation modeling has become a popular approach for studying the effect of an exposure on an outcome through a mediator. However, current methods are not applicable to the setting with a large number of mediators. We propose a testing procedure for mediation effects of high-dimensional continuous mediators. We characterize the marginal mediation effect, the multivariate component-wise mediation effects, and the $L2$ norm of the component-wise effects, and develop a Monte-Carlo procedure for evaluating their statistical significance. To accommodate the setting with a large number of mediators and a small sample size, we further propose a transformation model using the spectral decomposition. Under the transformation model, mediation effects can be estimated using a series of regression models with a univariate transformed mediator, and examined by our proposed testing procedure. Extensive simulation studies are conducted to assess the performance of our methods for continuous and dichotomous outcomes. We apply the methods to analyze genomic data investigating the effect of microRNA miR-223 on a dichotomous survival status of patients with glioblastoma multiforme (GBM). We identify nine gene ontology sets with expression values that significantly mediate the effect of miR-223 on GBM survival.

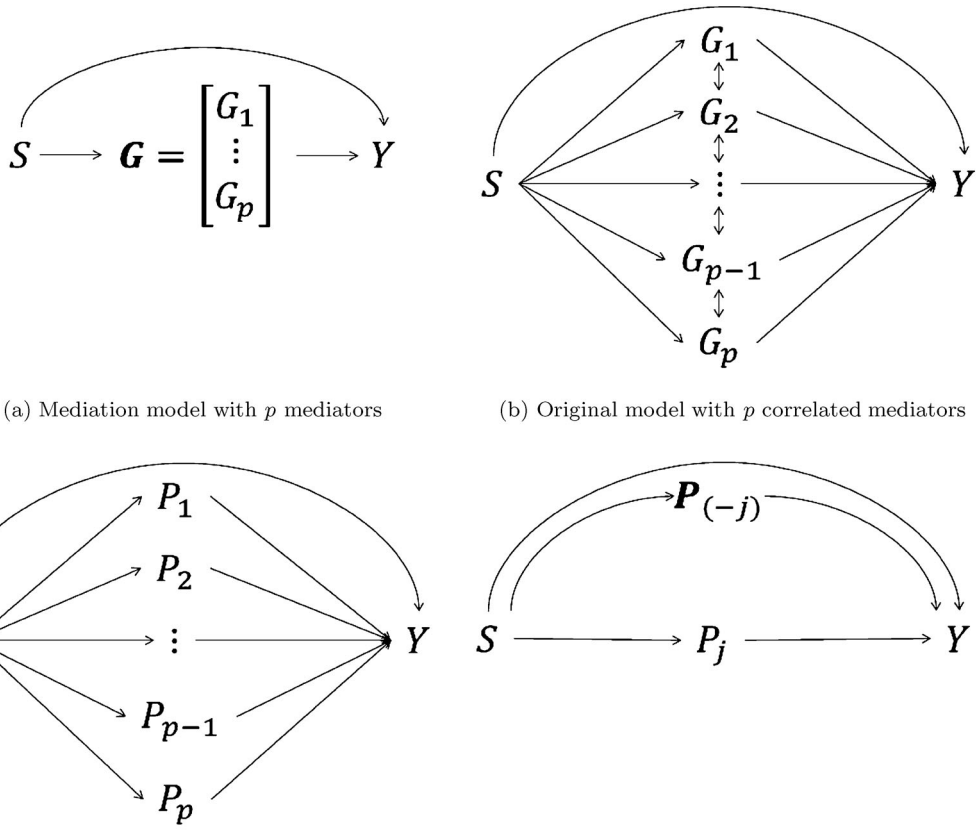
Key words: Causal mediation model; High-dimensional statistics; Hypothesis test; Integrative genomics; Multiple mediation; Natural indirect effect

1. Introduction

Given the wide availability of genomic data, studies based on high-throughput genomic data such as microarray or sequencing data have become popular in biomedical research. Successful examples include differentially expressed genes identified from gene expression microarray that classify disease subtypes; disease susceptibility genetic loci discovered from genome-wide association studies; and epigenetic markers identified from DNA methylation arrays that predict cancer prognosis. Despite these successes, of greater scientific interest may be to investigate by what mechanism the genes with differential expression, distinct genotypes or various epigenetic markers affect the outcome or phenotype. Such a mechanistic process is what a mediation analysis aims to characterize.

This article is motivated by a genomic study of glioblastoma multiforme (GBM) (Huang et al., 2015). The expression level of microRNA miR-223 has shown to be associated with GBM mortality, and the effect of miR-223 expression on mortality may be through its regulation on expression of numerous genes. Given the existing evidence, a natural question to ask is whether miR-223 S exerts its effect on cancer prognosis Y through altering the expression profile of a set of genes $\mathbf{G}$ with the same biological function (Figure 1a). A challenge is that the number of genes with the same function is usually large, compared to the sample size of the study, $n = 522$. For example, there are 187 genes in the gene set of leukocyte mediated immunity (Gene Ontology ID: GO0002443), and 983 genes in defense response (GO0006952). Given the large number of potential mediators p, the existing methods based on the model that includes all mediators as covariates are not applicable, i.e., large-p and small-n, and require additional development.

Mediation analysis was first proposed in social science literature (Baron and Kenny, 1986; MacKinnon, 2008). Recent advances in causal inference further extend the mediation model to incorporate nonlinearity and interactions (Robins and Greenland, 1992; Pearl, 2001; Imai et al., 2010; VanderWeele and Vansteelandt, 2010). Studies have attempted to develop mediation analyses for multiple mediators. Path-specific effects (PSEs) have been proposed to characterize the mediation effect of a pathway through multiple mediators (Avin et al., 2005; Daniel et al., 2015), for example, studying the effect of the path $S \to G_1 \to G_2 \to Y$ where G_1 and G_2 are the mediators 1 and 2, respectively. An approach based on regression or inverse probability weighting has been proposed to estimate PSEs (VanderWeele and Vansteelandt, 2013). Additionally, a generalized causal mediation model that incorporates mixed variable types has been proposed (Albert and Nelson, 2011). We have also developed a testing procedure to assess PSEs under the setting with high-dimensional exposures and two mediators (Huang, 2015). The methods focusing on PSEs explicitly model the causal structure among



(a) Mediation model with p mediators

(b) Original model with p correlated mediators

(c) Transformation model with p independent mediators, conditional on S

(d) Transformation model with a single mediator j

Figure 1. Causal diagram of the mediation model under the original ((a) and (b)) and transformation models ((c) and (d)). $\mathbf{P}_{(-j)}$ indicates all elements in $\mathbf{P}$ except P_j. Our proposed transformation is from (b) (or equivalently, (a)) to (c) and then to (d) (see details in Section 3).

mediators, but its generalization to high-dimensional mediators is challenging. For models with more than two mediators, estimation of the total and decomposed mediation effects has been developed based on the closed form expressions of the effects (Wang et al., 2013), bootstrap resampling (Preacher and Hayes, 2008) and estimating equations (Zhao et al., 2014). As they all require fitting a joint model of the outcome with all mediators $[Y|S, \mathbf{G}]$ where $\mathbf{G} = (G_1, ..., G_p)^T$ is the p mediators, p may still need to be small relative to the sample size to ensure model convergence. Therefore, despite the increasing interest in the multi-mediator model, existing methods can not be readily applicable to our data example with high-dimensional mediators, which is what this article aims to approach.

The rest of the article is organized as follows. In Section 2, we derive mediation effects for continuous and binary outcomes under a causal mediation model with high-dimensional continuous mediators and discuss the identifiability and model assumptions. In Section 3, utilizing the spectral decomposition, we transform mediators to be uncorrelated given the exposure such that we are able to estimate mediation effects using a series of low-dimensional regression models. In Section 4, we propose null hypotheses for the marginal mediation effect, the multivariate component-wise mediation effects and the $L2$ norm of the component-wise effects, and develop Monte-Carlo testing procedures for these effects. In Section 5, we conduct extensive simulation studies to evaluate the performance of different methods. In Section 6, we apply our methods to analyze the genomic data of GBM. We conclude with discussion in Section 7.

2. The Causal Mediation Model

We propose two regression models to represent the causal mediation model in Figure 1a. For subject i, the outcome Y_i is determined by q covariates $\mathbf{X}_i$ with the first element being 1 for the intercept, one exposure S_i, p mediators $\mathbf{G}_i$ and p possible interactions of S_i and $\mathbf{G}_i$:

$$Y_i = \mathbf{X}_i^T \boldsymbol{\beta}_X + S_i \beta_S + \mathbf{G}_i^T \boldsymbol{\beta}_G + S_i \mathbf{G}_i^T \boldsymbol{\beta}_C + \epsilon_{Yi}, \qquad (1)$$

where $\epsilon_{Yi} \sim N(0, \sigma^2)$, $\boldsymbol{\beta}_G = (\beta_{G1}, ..., \beta_{Gp})^T$, and $\boldsymbol{\beta}_C = (\beta_{C1}, ..., \beta_{Cp})^T$. And, the jth mediator can be further determined by the q covariates $\mathbf{X}_i$ and the exposure S_i: $G_{ji} = \mathbf{X}_i^T \boldsymbol{\alpha}_{Xj} + S_i \alpha_{Sj} + \epsilon_{Gji}$, or equivalently, one can specify a

multivariate model for the p mediators jointly:

$$\mathbf{G}_i = \mathbf{A}\mathbf{X}_i + S_i\boldsymbol{\alpha}_S + \boldsymbol{\epsilon}_{Gi}, \tag{2}$$

where $\boldsymbol{\epsilon}_{Gi} = (\epsilon_{G1i}, ..., \epsilon_{Gpi})^T \sim N_p(\mathbf{0}, \boldsymbol{\Sigma})$, $\mathbf{A}^T = (\boldsymbol{\alpha}_{X1}, ..., \boldsymbol{\alpha}_{Xp})$, $\boldsymbol{\alpha}_S^T = (\alpha_{S1}, ..., \alpha_{Sp})$. ϵ_{Yi} and $\boldsymbol{\epsilon}_{Gi}$ are assumed independent of S_i, $\mathbf{X}_i$, and each other.

The mediation effect (ME) can be defined with counterfactual notation as follows:

$$\mathrm{E}[Y(s_1, \mathbf{G}(s_1))|\mathbf{X}] - \mathrm{E}[Y(s_1, \mathbf{G}(s_0))|\mathbf{X}]. \tag{3}$$

$Y(s, \mathbf{g})$ denotes the potential outcome that Y would have attained had S and $\mathbf{G}$, respectively, been set to s and $\mathbf{g}$, and $\mathbf{G}(s)$ denotes the counterfactual gene expression values that would have been observed had S been set to s. Although the natural indirect effect has been defined and widely used in the causal inference literature, here we use the term interchangeably with mediation effect throughout the article, to better reflect its scientific interpretation in the motivating example. With assumptions that will be discussed in Section 2.1, the counterfactual mean can be expressed as an integral of model (1) with respect to (2):

$$\begin{aligned}&\mathrm{E}[Y(s_a, \mathbf{G}(s_b))|\mathbf{X}]\\ &= \int \mathrm{E}(Y|\mathbf{X}, S = s_a, \mathbf{G} = \mathbf{g})dF(\mathbf{G} = \mathbf{g}|\mathbf{X}, S = s_b).\end{aligned}$$

It follows that the ME can be expressed with the parameters in models (1) and (2):

$$\begin{aligned}&\mathrm{E}[Y(s_1, \mathbf{G}(s_1))|\mathbf{X}] - \mathrm{E}[Y(s_1, \mathbf{G}(s_0))|\mathbf{X}]\\ &= (s_1 - s_0)(\boldsymbol{\alpha}_S^T\boldsymbol{\beta}_G + s_1\boldsymbol{\alpha}_S^T\boldsymbol{\beta}_C).\end{aligned} \tag{4}$$

One can obtain similar results for the mediation model with a dichotomous outcome. For subject i, we propose a logistic regression model for the binary outcome:

$$\begin{aligned}&\mathrm{logit}[P(Y_i = 1|S_i, \mathbf{G}_i, \mathbf{X}_i)]\\ &= \mathbf{X}_i^T\boldsymbol{\beta}_X + S_i\beta_S + \mathbf{G}_i^T\boldsymbol{\beta}_G + S_i\mathbf{G}_i^T\boldsymbol{\beta}_C.\end{aligned} \tag{5}$$

With the assumptions in Section 2.1, one can express the following counterfactual outcomes with models (2) and (5):

$$\begin{aligned}&\mathrm{logit}[P(Y(s_a, \mathbf{G}(s_b)) = 1|\mathbf{X})]\\ &\approx c^*\{\mathbf{X}^T(\boldsymbol{\beta}_X + \mathbf{A}^T\boldsymbol{\beta}_G) + s_a\beta_S\\ &\quad + s_b\boldsymbol{\alpha}_S^T\boldsymbol{\beta}_G + s_a\mathbf{X}^T\mathbf{A}^T\boldsymbol{\beta}_C + s_as_b\boldsymbol{\alpha}_S^T\boldsymbol{\beta}_C\},\end{aligned}$$

where $c^* = (1 + 0.35 \times (\boldsymbol{\beta}_G + s_a\boldsymbol{\beta}_C)^T\boldsymbol{\Sigma}(\boldsymbol{\beta}_G + s_a\boldsymbol{\beta}_C))^{-1/2}$, by approximating logistic distribution with normal distribution for nonrare outcome (Zeger et al., 1988). The right-hand side is $\mathbf{X}^T(\boldsymbol{\beta}_X + \mathbf{A}^T\boldsymbol{\beta}_G) + s_a\beta_S + s_b\boldsymbol{\alpha}_S^T\boldsymbol{\beta}_G + s_a\mathbf{X}^T\mathbf{A}^T\boldsymbol{\beta}_C + s_as_b\boldsymbol{\alpha}_S^T\boldsymbol{\beta}_C + \frac{1}{2}\{\boldsymbol{\beta}_G + s_a\boldsymbol{\beta}_C\}^T\boldsymbol{\Sigma}\{\boldsymbol{\beta}_G + s_a\boldsymbol{\beta}_C\}$ if the outcome is rare (VanderWeele and Vansteelandt, 2010). We define the ME in the scale of log odds ratio, and express with the parameters in models (2) and (5):

$$\begin{aligned}\log[\mathrm{OR}^{ME}_{s_1s_0}(s_1)] &\equiv \mathrm{logit}[P(Y(s_1, \mathbf{G}(s_1)) = 1|\mathbf{X})]\\ &\quad -\mathrm{logit}[P(Y(s_1, \mathbf{G}(s_0)) = 1|\mathbf{X})]\\ &\approx c(s_1 - s_0)(\boldsymbol{\alpha}_S^T\boldsymbol{\beta}_G + s_1\boldsymbol{\alpha}_S^T\boldsymbol{\beta}_C),\end{aligned} \tag{6}$$

where c is 1 for rare outcomes and c^* with s_a being s_1 for nonrare outcomes.

2.1. *Identifiability and Model Assumptions*

Denoting A independent of B conditional on C as $A \perp B|C$, we list four sufficient assumptions for identifying ME: with adjustment for covariates $\mathbf{X}$, 1) $Y(s) \perp S|\mathbf{X}$, i.e., no confounding for the relationship of the exposure S and the outcome Y, 2) $Y(s, \mathbf{g}) \perp \mathbf{G}|\mathbf{X}, S$, i.e., no confounding for the relationship of the mediators $\mathbf{G}$ and the outcome Y, conditional on the exposure S, 3) $\mathbf{G}(s) \perp S|\mathbf{X}$, i.e., no confounding for the relationship of the exposure S and mediators $\mathbf{G}$, and 4) $Y(s, \mathbf{g}) \perp \mathbf{G}(s^*)|\mathbf{X}$, i.e., no confounder for the mediators-outcome ($\mathbf{G}$-Y) relationship that is affected by the exposure S.

These assumptions can be viewed as a multivariate extension of those for the single-mediator model (VanderWeele and Vansteelandt, 2010). In general, such an extension makes stronger assumptions for 2) and 3), but does not affect assumption 1). There has been discussion that incorporating multiple mediators actually renders assumption 4) more plausible (VanderWeele and Vansteelandt, 2013). For the single-mediator model $S \to G_1 \to Y$, all the other G_j, $j = 2, ..., p$ can potentially violate assumption 4) provided all gene expressions are closely related. Similarly, one may argue that accounting for multiple mediators simultaneously may reduce the possibility of confounding for 2), e.g., G_2 can be a common cause of G_1 and Y and thus a confounder in $S - G_1 - Y$ model but not in $S - (G_1, G_2) - Y$ model. In our motivating example, the extension is from a single gene expression to a set of gene expressions with similar biological functions. Based on the biological characteristics, we argue that the $Y - \mathbf{G}$ and $Y - G_j$ associations share most if not all confounders, and similarly, for the $\mathbf{G} - S$ and $G_j - S$ associations. Therefore, in our data example, the multivariate extension makes slightly stronger assumptions 2) and 3) but weakens 4) (and maybe 2)), compared to the single-mediator model. See more discussion about 2) and 4) in Supplementary Materials.

An implicit model assumption of our method is that mediators are treated as an entity with correlation among their members, but no directed causal paths are assumed among the mediators. If causal mechanisms among mediators are of interest, one should resort to the approach focusing on path-specific effects. For the two-mediator model, the PSE approach is able to identify the effect of the path $S \to G_1 \to Y$, $S \to G_2 \to Y$ and $S \to G_1 \to G_2 \to Y$ with additional assumptions (Albert and Nelson, 2011; Daniel et al., 2015); and our method focuses on $S \to (G_1, G_2) \to Y$. Under certain causal structures of $\mathbf{G}$, our model assumption can be violated (see discussion of disjunctive effect in Supplement). As the mediators in our example $\mathbf{G}$ are gene expression values measured simultaneously as a snapshot rather than a cascade, the undirected correlation among expression values may be more plausible than a directed causal structure. Our model (1) or (5) allows exposure-by-mediator cross-product

interactions but not mediator-by-mediator interactions due to nonidentifiability for mediator-by-mediator cross-product coefficients under the transformed univariate model to be discussed next. However, as we focus on hypothesis testing, failure to incorporate potential mediator-by-mediator interactions does not invalidate the test but simply affects the power. The power gain if any may be very limited given additional $p(p-1)/2$ parameters to be estimated even under the model with only two-way interactions. Alternatively, one can incorporate more flexible kernels of $\mathbf{G}$ to capture nonlinear interactions (Wu et al., 2011).

3. Transformation of Mediators

We illustrate the transformation with diagrams in Figure 1. We transform the correlated mediators in the original model (Figure 1b) to be uncorrelated and study the ME of transformed mediators (Figure 1c). Moreover, we exploit such transformation for dimensional reduction (Figure 1d). For the $p \times p$ covariance matrix $\mathbf{\Sigma}$, there exists an orthogonal $p \times p$ matrix $\mathbf{u}$ such that $\mathbf{u\Sigma u}^T = \text{diag}(\sigma_1^2, ..., \sigma_p^2)$ where $\sigma_1^2 > ... > \sigma_p^2$ (Harville, 2001). It follows that

$$\mathbf{P}_i = \mathbf{A}^* \mathbf{X}_i + S_i \boldsymbol{\alpha}_S^* + \boldsymbol{\epsilon}_{Pi}, \tag{7}$$

where the vector of transformed mediators $\mathbf{P}_i = (P_{1i}, ..., P_{pi})^T = \mathbf{u}\mathbf{G}_i$, $\boldsymbol{\epsilon}_{Pi} = \mathbf{u}\boldsymbol{\epsilon}_{Gi} \sim N_p(\mathbf{0}, \text{diag}(\sigma_1^2, ..., \sigma_p^2))$, $\mathbf{A}^* = \mathbf{uA}$ with jth row being $\mathbf{A}_j^*$, and $\boldsymbol{\alpha}_S^* = \mathbf{u}\boldsymbol{\alpha}_S = (\alpha_{S1}^*, ..., \alpha_{Sp}^*)^T$. Models (1) and (5) can be rewritten as:

$$\mathcal{G}(\text{E}[Y_i|S_i, \mathbf{G}_i, \mathbf{X}_i)]) = \mathbf{X}_i^T \boldsymbol{\beta}_X + S_i \beta_S + \mathbf{P}_i^T \boldsymbol{\beta}_G^* + S_i \mathbf{P}_i^T \boldsymbol{\beta}_C^* \tag{8}$$

where $\boldsymbol{\beta}_G^* = \mathbf{u}\boldsymbol{\beta}_G = (\beta_{G1}^*, ..., \beta_{Gp}^*)^T$, $\boldsymbol{\beta}_C^* = \mathbf{u}\boldsymbol{\beta}_C = (\beta_{C1}^*, ..., \beta_{Cp}^*)^T$, and $\mathcal{G}(\cdot)$ denotes the link function: identity and logit links for (1) and (5), respectively. The ME can be expressed as

$$\text{ME} = c(s_1 - s_0)(\boldsymbol{\alpha}_S^{*T} \boldsymbol{\beta}_G^* + s_1 \boldsymbol{\alpha}_S^{*T} \boldsymbol{\beta}_C^*), \tag{9}$$

where $c = 1$ for continuous or rare dichotomous outcomes, and $c = (1 + 0.35 \times (\boldsymbol{\beta}_G + s_1\boldsymbol{\beta}_C)^T \mathbf{\Sigma} (\boldsymbol{\beta}_G + s_1 \boldsymbol{\beta}_C))^{-1/2}$ for nonrare dichotomous outcomes. We further denote (4), (6) and (9) as marginal MEs because they refer to the overall effect mediated by $\mathbf{G}$ regardless through which element of $\mathbf{G}$. A consistent estimator for $\mathbf{\Sigma}$ is the sample covariance of the residuals in model (2). After transformation, the P's are uncorrelated; and under the assumption that G's are Gaussian, the P's are also independent, which is needed for deriving the following model (10) in the Appendix.

With $P_j \perp P_k | S$ for all j, k, and the rare disease assumption for the dichotomous outcome, we show in the Appendix that the outcome model can be re-expressed based on P_j:

$$\begin{aligned}\mathcal{G}(\text{E}[Y_i|\mathbf{X}_i, S_i, P_{ji}]) = \mathbf{X}_i^T \boldsymbol{\gamma}_X + S_i \gamma_S + S_i \mathbf{X}_i^T \boldsymbol{\gamma}_{SX} + S_i^2 \gamma_{S2} \\ + P_{ji} \beta_{Gj}^* + S_i P_{ji} \beta_{Cj}^*,\end{aligned} \tag{10}$$

where $\boldsymbol{\gamma}_X = \boldsymbol{\beta}_X + \sum_{k \neq j} \beta_{Gk}^* \mathbf{A}_k^*$, $\gamma_S = \beta_S + \sum_{k \neq j} \beta_{Gk}^* \alpha_{Sk}^*$, $\boldsymbol{\gamma}_{SX} = \sum_{k \neq j} \beta_{Ck}^* \mathbf{A}_k^*$, $\gamma_{S2} = \sum_{k \neq j} \beta_{Ck}^* \alpha_{Sk}^*$. In the model $[Y|\mathbf{X}, S, P_j]$, β_{Gj}^* and β_{Cj}^*, are the regression coefficients for P_j and SP_j, adjusting for the covariates $\mathbf{X}$, the exposure S, its quadratic term S^2 and the exposure-by-covariate interactions $S\mathbf{X}$. Note the S^2 and $S\mathbf{X}$ terms exist provided that S-by-P_k interaction, β_{Ck}^* is assumed ($k \neq j$) nonzero. Since P_k is a linear function of S, to account for the S-by-P_k cross-product effect requires a quadratic term of S. The regression coefficients for P_j and SP_j are identical to those in model (8). Therefore, rather than fitting the large joint model $[Y|\mathbf{X}, S, \mathbf{P}]$ as (8), one can fit p low-dimensional models $[Y|\mathbf{X}, S, P_j]$ as (10), to estimate $\boldsymbol{\beta}_G^*$ and $\boldsymbol{\beta}_C^*$, the two parameters involved in ME in (9). Under the model with a single mediator P_j, other independent mediators $\mathbf{P}_{(-j)}$ constitute a part of direct effect not through P_j (Figure 1d). In the absence of exposure-by-mediator interactions, $\beta_{Cj}^* = \beta_{Ck}^* = 0$ for j and all $k(\neq j)$, model (10) can be simplified to the regression model in Baron and Kenny (1986): $\mathcal{G}(\text{E}[Y_i|\mathbf{X}_i, S_i, P_{ji}]) = \mathbf{X}_i^T \boldsymbol{\gamma}_X + S_i \gamma_S + P_{ji} \beta_{Gj}^*$.

4. Hypothesis Test of Mediation Effect

4.1. *Null Hypotheses*

We focus on hypothesis testing because it serves as a critical step to scan through high-dimensional data such as genomic data. The proposed procedure can be adapted for estimation. For both continuous and dichotomous outcomes, mediation effects have similar expressions (4) and (6) based on the original mediators $\mathbf{G}$, and (9) based on the transformed mediators $\mathbf{P}$. Marginal ME comparing any arbitrary s_1 versus. s_0 is proportional to Δ or Δ^*, and the null hypothesis of no marginal ME can be expressed as:

$$H_0 : \Delta \equiv \boldsymbol{\alpha}_S{}^T (\boldsymbol{\beta}_G + \boldsymbol{\beta}_C) = 0, \tag{11}$$

$$H_0 : \Delta^* \equiv \boldsymbol{\alpha}_S^{*T} (\boldsymbol{\beta}_G^* + \boldsymbol{\beta}_C^*) = 0, \tag{12}$$

where we assume $s_1 = 1$. We use separate notations Δ and Δ^* to denote two estimands from the original and transformation models, respectively, but it should be noted $\Delta = \Delta^*$, and they can be tested under the two different models. It can be shown that the method proposed by Zhao et al. (2014) tests the marginal mediation effect Δ as (11) (see Section 2 of the Supplement). Denote the component-wise mediation effects of Δ and Δ^*, respectively, as $\boldsymbol{\delta} = (\delta_1, ..., \delta_p)^T$ and $\boldsymbol{\delta}^* = (\delta_1^*, ..., \delta_p^*)^T$, where $\delta_j \equiv \alpha_{Sj}(\beta_{Gj} + \beta_{Cj})$ and $\delta_j^* \equiv \alpha_{Sj}^*(\beta_{Gj}^* + \beta_{Cj}^*)$, one can specify the following null hypotheses for no ME with respect to each element of Δ or Δ^*:

$$H_0 : \boldsymbol{\delta} = \mathbf{0}, \quad \text{or} \quad H_0 : \boldsymbol{\delta}^* = \mathbf{0}. \tag{13}$$

Null hypothesis (11) represents a broader class of null, compared to (13). For example, under (11), one may have a half of δ_j's equal to the other half but with opposite directions, which does not belong to null (13). The discrepancy between (11) (or (12)) and (13) comes from the fact that the component-wise mediation effects δ_j may cancel each other and result in a zero marginal ME Δ. Since tests for a marginal effect and component-wise effects may have different implications, one has to consider the scientific context to choose a proper null. In our data example, the tests for component-wise effects are of greater scientific interest. For example, if a very strong positive ME through G_1 and a strong negative effect through G_2 are discovered, one may construe the finding as a

significant ME and investigate the mechanisms of the two opposite effects, rather than focusing only on the zero marginal effect.

Although the null (13) is of scientific interest, to perform a multivariate test for $\boldsymbol{\delta}$ or $\boldsymbol{\delta}^*$ has a severe limitation. Compared to the test for the marginal ME, tests for $\boldsymbol{\delta}$ or $\boldsymbol{\delta}^*$ have p degrees of freedom (DF). Conventional multivariate tests such as likelihood ratio tests with large DF have been shown to have very low power (Huang et al., 2014). To overcome the difficulty, we propose to test the $L2$ norm of the component-wise mediation effects:

$$H_0 : \tau \equiv \|\boldsymbol{\delta}\|^2 = 0, \quad \text{or} \quad H_0 : \tau_p \equiv \|\boldsymbol{\delta}^*\|^2 = 0. \tag{14}$$

Note the null (13) is identical to (14). We highlight the two different estimands that the two sets of hypotheses focus on. (14) is equivalent to the null for variance component tests (Lin, 1997) or the empirical Bayes approach that assumes δ_j follow an arbitrary distribution with mean 0 and variance τ or τ_p and tests for the variance being zero (Cai et al., 2012).

We propose three null hypotheses under the models with mediators $\mathbf{G}$: a null hypothesis for the marginal ME (11), a null hypothesis for the p component-wise MEs (13) and a null hypothesis for $L2$ norm of the p component-wise MEs (14), and we also propose corresponding three null hypotheses under the transformation model with mediators $\mathbf{P}$. To test null hypotheses with mediators $\mathbf{G}$ requires fitting a large joint model $[Y|\mathbf{X}, S, \mathbf{G}]$ (1) or (5). When the number of mediators p is large relative to the sample size n, there would be a serious numerical obstacle. In contrast, as shown in Section 3, the parameters in the null (12), (13), and (14) can be estimated from a series of low-dimensional models $[Y|\mathbf{X}, S, P_j]$ (10), even under the setting with a large p and a small n.

4.2. *Testing Procedure*

To test the null hypotheses (11)–(14), we propose a Monte-Carlo procedure for approximating the distribution of the marginal ME Δ and Δ^*, the component-wise MEs $\boldsymbol{\delta}$ and $\boldsymbol{\delta}^*$ and their $L2$ norms τ and τ_p. Denote $\boldsymbol{\theta} = (\boldsymbol{\alpha}_S^T, \boldsymbol{\beta}_G^T, \boldsymbol{\beta}_C^T)^T$, and $\hat{\boldsymbol{\theta}}$ being the maximum likelihood estimator (MLE) of $\boldsymbol{\theta}$ under the parametric models assumed in (1) (or (5)) and (2). The distribution of $\hat{\boldsymbol{\theta}}$ can be approximated by a multivariate normal distribution,

$$\hat{\boldsymbol{\theta}} \dot{\sim} N_{3p}(\hat{\boldsymbol{\theta}}, \hat{\text{Cov}}(\hat{\boldsymbol{\theta}})), \tag{15}$$

where $\hat{\text{Cov}}(\hat{\boldsymbol{\theta}})$ is a $3p \times 3p$ block diagonal matrix with the upper $p \times p$ block diagonal being the covariance matrix of $\hat{\boldsymbol{\alpha}}_S$ and the lower $2p \times 2p$ being the covariance of $(\hat{\boldsymbol{\beta}}_G{}^T, \hat{\boldsymbol{\beta}}_C{}^T)^T$, provided that $\boldsymbol{\epsilon}_Y$ and $\boldsymbol{\epsilon}_G$ are independent. Alternatively, one can approximate the distribution of $\hat{\boldsymbol{\theta}}$ by bootstrapping. We will show in simulation that the bootstrap-based and normality-based methods have very similar performance, but normality-based method is computationally much more efficient.

By generating independent $\tilde{\boldsymbol{\theta}}$ repeatedly from $N_{3p}(\hat{\boldsymbol{\theta}}, \hat{\text{Cov}}(\hat{\boldsymbol{\theta}}))$, the realization for the null distributions of $\hat{\Delta} = \Delta(\hat{\boldsymbol{\theta}})$ and $\hat{\boldsymbol{\delta}} = \boldsymbol{\delta}(\hat{\boldsymbol{\theta}})$ can be approximated, respectively by $\{\Delta(\tilde{\boldsymbol{\theta}}, 0)^{(b)} \equiv \Delta(\tilde{\boldsymbol{\theta}})^{(b)} - B^{-1}\sum_b \Delta(\tilde{\boldsymbol{\theta}})^{(b)}\}$ and $\{\boldsymbol{\delta}(\tilde{\boldsymbol{\theta}}, 0)^{(b)} \equiv \boldsymbol{\delta}(\tilde{\boldsymbol{\theta}})^{(b)} - B^{-1}\sum_b \boldsymbol{\delta}(\tilde{\boldsymbol{\theta}})^{(b)}\}$, where $b = 1, ..., B$, the number of resampling draws. The justification for approximating the null distribution by centering at zero is that the estimates of component-wise mediation effects have been shown to follow the product of two normal distributions (MacKinnon et al., 2004), and the mean of the distribution is the product of the means for the two normal variables plus their correlation (Aroian, 1947), which is zero under the null and no-unmeasured confounding assumptions. See more discussion in Supplementary Materials. The p-value for $\hat{\Delta}$ can be approximated as the tail probability by comparing $\hat{\Delta}$ with $\{\Delta(\tilde{\boldsymbol{\theta}}, 0)^{(b)}\}$. To test for (13), we calculate $Q_{\boldsymbol{\delta}} = \hat{\boldsymbol{\delta}}^T \hat{\text{Cov}}^{-1}(\hat{\boldsymbol{\delta}})\hat{\boldsymbol{\delta}}$ and the p-value by comparing $Q_{\boldsymbol{\delta}}$ with p-DF χ^2 distribution or $\{Q_{\boldsymbol{\delta}}(\tilde{\boldsymbol{\theta}}, 0)^{(b)} \equiv \boldsymbol{\delta}(\tilde{\boldsymbol{\theta}}, 0)^{(b)T} \hat{\text{Cov}}^{-1}(\hat{\boldsymbol{\delta}})\boldsymbol{\delta}(\tilde{\boldsymbol{\theta}}, 0)^{(b)}\}$ where $\hat{\text{Cov}}(\hat{\boldsymbol{\delta}})$ is the sample covariance of the B resampling draws $\{\boldsymbol{\delta}(\tilde{\boldsymbol{\theta}}, 0)^{(b)}\}$. We emphasize that although we use $\boldsymbol{\delta}$ to motivate tests for the $L2$ norm of the component-wise mediation effects, we do not advocate the p-DF tests for $\boldsymbol{\delta}$ due to its low power.

4.2.1. Testing procedure under the transformation model. Similarly, the distribution of $\hat{\boldsymbol{\theta}}^* = (\hat{\boldsymbol{\alpha}}_S^{*T}, \hat{\boldsymbol{\beta}}_G^{*T}, \hat{\boldsymbol{\beta}}_C^{*T})^T$ can be approximated by independent draws $\tilde{\boldsymbol{\theta}}^*$ from $N_{3p}(\hat{\boldsymbol{\theta}}^*, \hat{\text{Cov}}(\hat{\boldsymbol{\theta}}^*))$, and the null distribution of $\Delta^*(\hat{\boldsymbol{\theta}}^*)$ and $\boldsymbol{\delta}^*(\hat{\boldsymbol{\theta}}^*)$ can be approximated accordingly with $\{\Delta^*(\tilde{\boldsymbol{\theta}}^*, 0)^{(b)} \equiv \Delta^*(\tilde{\boldsymbol{\theta}}^*)^{(b)} - B^{-1}\sum_b \Delta^*(\tilde{\boldsymbol{\theta}}^*)^{(b)}\}$ and $\{\boldsymbol{\delta}^*(\tilde{\boldsymbol{\theta}}^*, 0)^{(b)} \equiv \boldsymbol{\delta}^*(\tilde{\boldsymbol{\theta}}^*)^{(b)} - B^{-1}\sum_b \boldsymbol{\delta}^*(\tilde{\boldsymbol{\theta}}^*)^{(b)}\}$. We have already shown that $\boldsymbol{\theta}^*$ can be estimated by a series of low-dimensional regression models (7) and (10). $\hat{\text{Cov}}(\hat{\boldsymbol{\theta}}^*)$ is a $3p \times 3p$ block diagonal matrix with the upper $p \times p$ block diagonal being the covariance matrix of $\hat{\boldsymbol{\alpha}}_S^*$ and the lower $2p \times 2p$ matrix $\mathbf{J}_{2p}^{-1}$. $\mathbf{J}_{2p}^{-1}$ is the lower $2p \times 2p$ matrix for the inverse of $\mathbf{J}$, where

$$\mathbf{J} = \sum_i w_i \mathbf{Z}_i \mathbf{Z}_i^T,$$

$\mathbf{Z}_i^T = (\mathbf{X}_i^T, S_i, \mathbf{P}_i^T, S_i\mathbf{P}_i^T)$, and $w_i = \sigma^{-2}$ and $\frac{e^{\eta_i}}{(1+e^{\eta_i})^2}$ for continuous and dichotomous outcomes, respectively. With $n > p$, w_i can be estimated by fitting model (1), (5), or (8). With $p > n$, we propose estimators $\hat{\sigma}^{*2}$ to estimate σ^2 from fitting model (10) on P_1, the transformed mediator corresponding to the largest variance estimate $\hat{\sigma}_1^2$, and $\hat{\eta}_i^* = \mathbf{X}_i^T\hat{\boldsymbol{\beta}}_X^* + S_i\hat{\beta}_S^* + \hat{\boldsymbol{\beta}}_G^{*T}\mathbf{P}_i + \hat{\boldsymbol{\beta}}_C^{*T}S_i\mathbf{P}_i$ for η where $\hat{\boldsymbol{\beta}}_G^* = (\hat{\beta}_{G1}^*, ..., \hat{\beta}_{Gp}^*)^T$ and $\hat{\boldsymbol{\beta}}_C^* = (\hat{\beta}_{C1}^*, ..., \hat{\beta}_{Cp}^*)^T$ are MLEs in a series of (10) for $j = 1, ..., p$, $\hat{\boldsymbol{\beta}}_X^* = \hat{\boldsymbol{\gamma}}_X - \sum_{k\neq j}\hat{\beta}_{Gk}^*\hat{\mathbf{A}}_k^*$, $\hat{\beta}_S^* = \hat{\gamma}_S - \sum_{k\neq j}\hat{\beta}_{Gk}^*\hat{\alpha}_{Sk}^*$, and $\hat{\boldsymbol{\gamma}}_X$ and $\hat{\gamma}_S$ are MLEs for (10) with $j = 1$. We discuss in Supplementary Materials that $\hat{\sigma}^{*2}$ is a consistent estimator for σ^2 under the null and overestimates σ^2 under the alternative.

The $L2$ norm of the component-wise MEs τ and τ_p can also be denoted as functions of $\boldsymbol{\theta}$ or $\boldsymbol{\theta}^*$: $\tau(\boldsymbol{\theta})$ and $\tau_p(\boldsymbol{\theta}^*)$. Their null distributions can be approximated by $\{\tau(\tilde{\boldsymbol{\theta}}, 0)^{(b)} \equiv \|\boldsymbol{\delta}(\tilde{\boldsymbol{\theta}}, 0)^{(b)}\|^2\}$ and $\{\tau_p(\tilde{\boldsymbol{\theta}}^*, 0)^{(b)} \equiv \|\boldsymbol{\delta}^*(\tilde{\boldsymbol{\theta}}^*, 0)^{(b)}\|^2\}$, which can be obtained from $\{\boldsymbol{\delta}(\tilde{\boldsymbol{\theta}}, 0)^{(b)}\}$ and $\{\boldsymbol{\delta}^*(\tilde{\boldsymbol{\theta}}^*, 0)^{(b)}\}$ by calculating $L2$ norm of p component-wise effects within the same draw. Of note, our testing procedure assumes normality for $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\theta}}^*$, but not for $\hat{\Delta}$, $\hat{\boldsymbol{\delta}}$ and $\hat{\tau}$ (or, $\hat{\Delta}^*$, $\hat{\boldsymbol{\delta}}^*$ and $\hat{\tau}_p$). In fact, $\hat{\boldsymbol{\delta}}$ follows the normal product distribution. Therefore, normal approximation using the Delta method may fail to capture its vari-

Table 1

The empirical size of the proposed testing procedures. The size is estimated from 5000 simulations across α_S from 0 to 15 and with $\boldsymbol{\beta}_G = \boldsymbol{\beta}_C = \mathbf{0}$. Δ^*: the test for the marginal mediation effect under the transformation model; $\boldsymbol{\delta}^*$: the p-DF test for the component-wise mediation effects under the transformation model; τ_p: the test for L2 norm of $\boldsymbol{\delta}^*$; Δ: the test for the marginal mediation effect under the original model; $\boldsymbol{\delta}$: the p-DF test for the component-wise mediation effects under the original model; τ: the test for L2 norm of $\boldsymbol{\delta}$.

	Δ^*	$\boldsymbol{\delta}^*$	τ_p	Δ	$\boldsymbol{\delta}$	τ
			Continuous outcome			
$n = 500,\ p = 50$	5.40	2.62	4.56	5.76	8.66	5.52
$n = 500,\ p = 50$, non-normal	4.40	0.00	4.00	5.00	0.80	3.40
$n = 50,\ p = 50$	3.98	1.93	3.84			
			Dichotomous outcome			
$n = 2000,\ p = 50$	4.38	0.12	4.48	5.12	4.02	6.08
$n = 400,\ p = 50$	2.74	0.00	2.94			

ability because $\hat{\boldsymbol{\delta}}$ is not normally distributed when the effect or sample size is small (MacKinnon et al., 2004). It should be stressed that tests for τ_p as well as Δ^* and $\boldsymbol{\delta}^*$ under the transformation model based on $\mathbf{P}$ can still be conducted even with a large p, while those for τ, Δ and $\boldsymbol{\delta}$ under the original model based on $\mathbf{G}$ can be constructed only if the sample size is large enough for fitting model (1) or (5).

5. Simulation Studies

We conduct extensive simulation studies for both continuous and dichotomous outcomes under large and small studies with $p = 50$. For the continuous outcome, the sample size n is 500 in the large study and 50 in the small study. The exposure S has values 0, 1, and 2 randomly assigned to approximately an equal number of subjects. With S, we generate p mediators $\mathbf{G}$ using model (2) with $\alpha_{S1} = ... = \alpha_{S50} = \alpha$, and $\boldsymbol{\epsilon}_G$ follow a multivariate normal with mean $\mathbf{0}$ and covariance $\boldsymbol{\Sigma}$ with the diagonal being 1 and the off-diagonal being ρ. The magnitude of α ranges from 0 to 0.15 for $n = 500$ and from 0 to 0.4 for $n = 50$, as indicated in power plots. We generate Y according to model (1) with ϵ_Y following the standard normal $N(0, 1)$ and β's indicated in the plots. For the dichotomous outcome, we simulate $n = 2000$ for the large study and $n = 400$ for the small study, with S and $\mathbf{G}$ generated similar to those for the continuous outcome. The magnitude of α ranges from 0 to 0.24 for $n = 2000$ and from 0 to 0.15 for $n = 400$. We generate the outcome Y according to model (5), with the intercept being -0.35 and β's indicated in power plots. For simulation studies with a large sample size, we evaluate the performance of tests for all hypotheses (11)–(14). With a small sample size, only tests based on the transformation models can be conducted. Significance level is set as p-value < 0.05. Type I error rate and power are estimated as the proportion of p-value < 0.05 via 5000 and 2000 simulation studies, respectively. Different configurations of β's as indicated in the power plots are explored, including 1) nonsparse mediation effects without cancelation, 2) nonsparse effects with cancelation, 3) sparse effects without cancelation, and 4) sparse effects with cancelation, where cancelation means some component-wise mediation effects are positive and some are negative. We compare our methods with that by Zhao et al. (2014).

Under the null, type I error is protected except for $\boldsymbol{\delta}$ (Table 1). The tests for smaller studies tend to be conservative, probably due to inadequacy of the normal approximation. The tests for p component-wise MEs $\boldsymbol{\delta}$ or $\boldsymbol{\delta}^*$ have conservative or inflated type I errors, which may result from relatively large DF ($p = 50$) given the sample size. For illustration, we focus here on results of statistical power for the continuous outcome, and present the results for the dichotomous outcome in Supplementary Materials. In the setting without effect cancelation, different tests have similar performance (Figures 2a and c). With mixed positive and negative effects, tests for the $L2$ norm have better power (Figures 2b and d). Performance of the test by Zhao et al. (2014) is very similar to that for $H_0 : \Delta = 0$ across various settings since both tests evaluate the marginal mediation effect. With a small n and a large p, the test for τ_p has better power in the presence of cancelation while the test for the marginal effect Δ^* has slightly better power if no cancelation (Figures 2e and f). The results in Figure 2 are based on $\rho = 0.3$ and those with $\rho = 0.7$ are presented in Figure S2 of Supplementary Materials. With the high correlation among mediators, the difference in power diminishes.

Our proposed methods are based on resampling from a multivariate normal (15). To examine sensitivity to the normality assumption, we conduct another set of simulation studies with $\boldsymbol{\epsilon}_G \sim N_p(\mathbf{0}, \boldsymbol{\Sigma}) + \mathbf{U}$ and $\epsilon_Y \sim N(0, 1) +$ Uniform$(-0.5, 0.5)$ where $\mathbf{U}$ is a p-by-p matrix with each element following independent Uniform$(-1, 1)$. Type I error is protected under the null (Table 1). As discussed in Section 4.2, we can obtain $\{\Delta^*(\tilde{\boldsymbol{\theta}}^*, 0)^{(b)}\}$, $\{\Delta(\tilde{\boldsymbol{\theta}}, 0)^{(b)}\}$, $\{\tau_p(\tilde{\boldsymbol{\theta}}^*, 0)^{(b)}\}$ and $\{\tau(\tilde{\boldsymbol{\theta}}, 0)^{(b)}\}$ from (15) assuming normality or via bootstrap without normality assumption. Under the non-normal random errors, we compare the two approaches. In the setting the same as Figure 2a and $\alpha_{Sj} = 0.1$, the power of normality-based and bootstrap-based methods, respectively are 51.6 and 53.8% for Δ, 43.2 and 42.0% for Δ^*, 50.8 and 49.6% for τ, and 34.0 and 34.6% for τ_p. Individual p-values are also very similar to each other under the null (Figure 3) and alternative (Figure S3 in Supplementary Materials) with correlation > 0.9. Despite their similar performance, we note a significant difference in the computation cost. For one run with 2.90 GHz

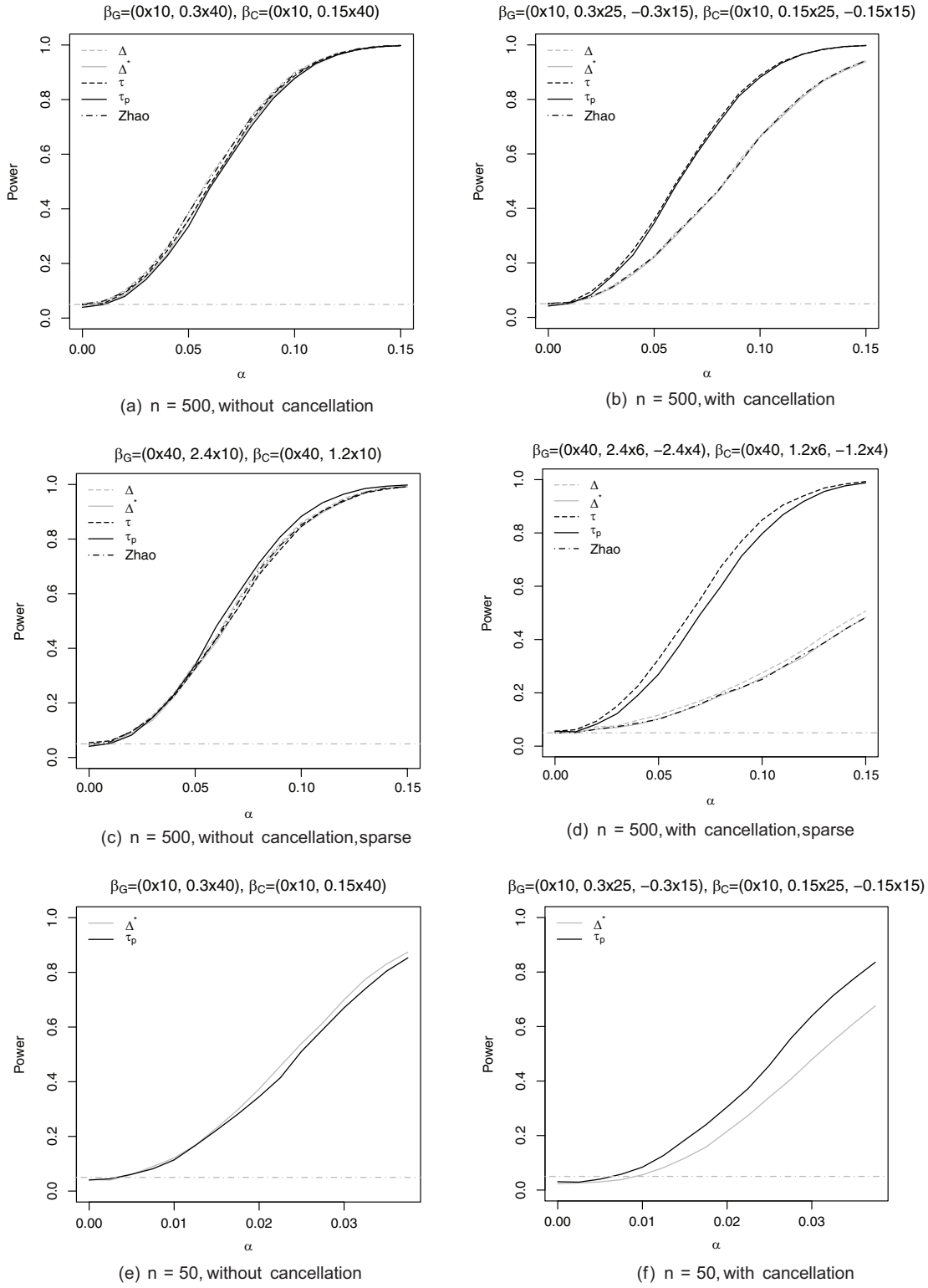


Figure 2. Empirical power for continuous outcome with $p = 50$ and $\rho = 0.3$. The label for parameter configurations has ordering, e.g., $(0 \times 10, 0.3 \times 40)$ indicates the first 10 coefficients are 0 and the other 40 are 0.3; $(0 \times 10, 0.3 \times 25, -0.3 \times 15)$ indicates the first 10 coefficients are 0, the next 25 are 0.3, and the last 15 are -0.3. The horizontal dash-dot line in gray indicates 0.05. In 3b and 3d, the two highest lines are tests for τ and τ_p.

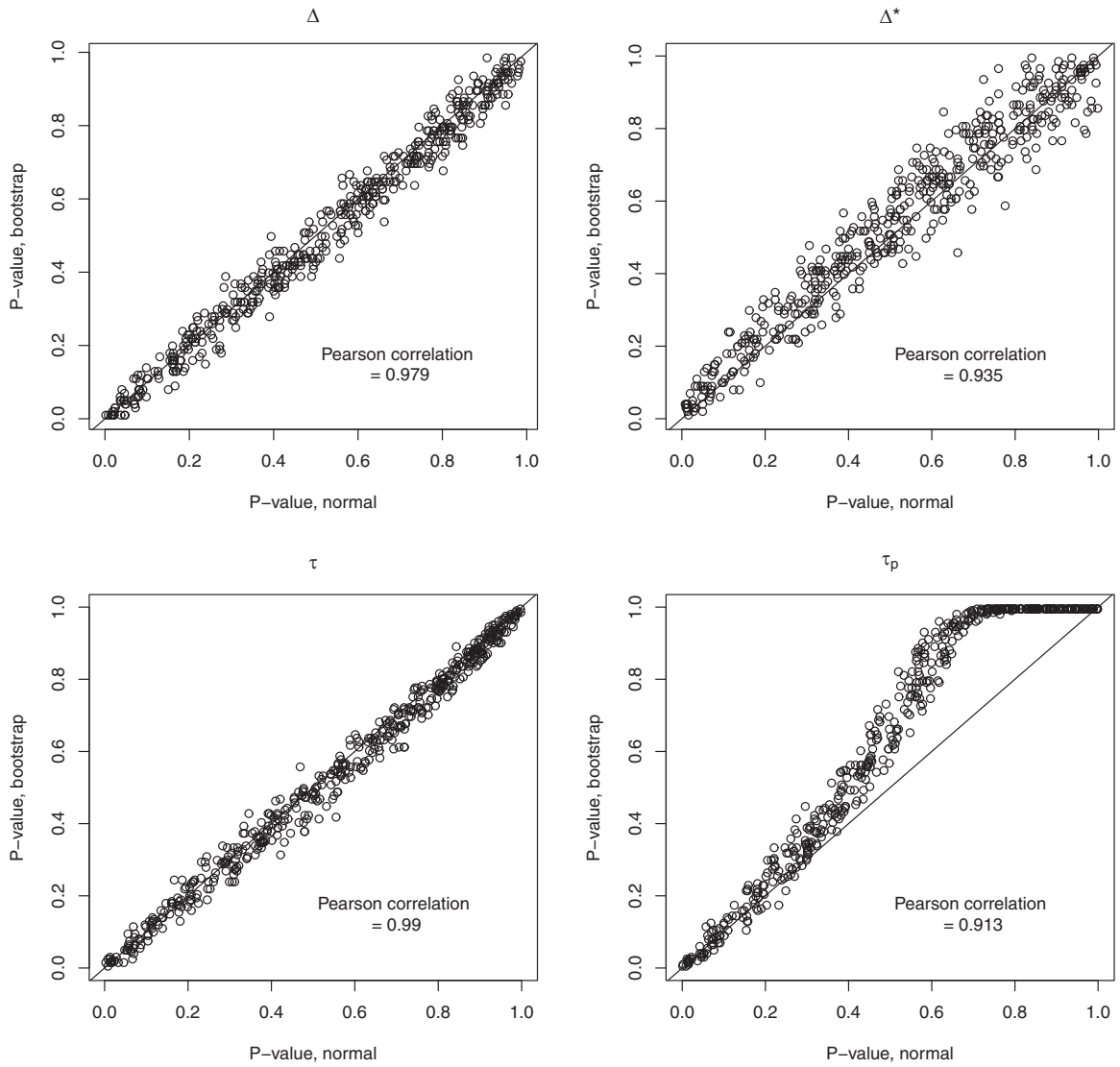


Figure 3. *P*-values between normality-based and bootstrap-based resampling. Pearson correlation is calculated based on logit-transformed *p*-values.

CPU and 8.00 GB RAM, the bootstrap-based and normality-based methods with 1000 resampling take 190.95 and 0.29 seconds, respectively.

Finally, we investigate a special case, termed disjunctive effect where the mediator j is linked to either the exposure ($\alpha_{Sj} \neq 0$) or the outcome ($\beta_{Gj} \neq 0$ or $\beta_{Cj} \neq 0$), but not both (Figure S4). Specifically, we set $\alpha_{S1} = ... = \alpha_{S25} = 0.1$, $\alpha_{S26} = ... = \alpha_{S50} = 0$, $\beta_{G1} = ... = \beta_{G25} = 0$, $\beta_{G26} = ... = \beta_{G50} = 0.3$, $\beta_{Cj} = \beta_{Gj}/2$, and $\rho = 0.3$. As these satisfy the null hypothesis for Δ and τ, we observe 4.80 and 2.80% of p-values < 0.05 respectively for the two hypotheses. Under the transformation model, a linear combination of all p G_j's may bring the nonzero α_{Sj} and β_{Gj} (or β_{Cj}) together for the transformed mediator P_j. The power of τ_p is 14.8%. The disjunctive effect may or may not represent a true mediation effect depending on the causal structure. See more discussion in Supplementary Materials.

6. Data Application

Glioblastoma multiforme (GBM) is the most common malignant brain tumor that is rapidly fatal with median survival time of 15 months (Stupp et al., 2005). Our previous study has shown that microRNA miR-223 may be a prognostic marker by regulating expression levels of multiple genes (Huang et al., 2015). We found that 306 out of 518 genes regulated by miR-223 showed significant univariate MEs (false discovery rate (FDR) (Storey, 2002) < 0.01) on GBM survival. We hypothesize that miR-223 coordinates with its downstream gene expressions to affect the GBM survival, particularly on the basis of a gene set with similar biological functions. We use the Gene Ontology (Ashburner et al., 2000) to classify genes associated with miR-223 into 120 gene sets containing gene numbers ranging from 17 to 10 157. We apply the proposed testing procedures to investigate whether expression values of gene sets with similar biological functions are mediators for

Table 2
P-values of the top nine gene sets with expression levels that mediate the effect of miR-223 on 3-month survival in GBM patients.

GO ID	Gene Function	No. of genes	*p*-value (FDR)		
			Δ^*	$\boldsymbol{\delta}^*$	τ_p
GO0009607	Response to biotic stimulus	543	0.022 (0.041)	1.000 (1.00)	0.235 (0.173)
GO0051707	Response to other organism	521	0.028 (0.041)	1.000 (1.00)	0.263 (0.173)
GO0016064	Immunoglobulin mediated immune response	90	0.032 (0.041)	0.832 (1.00)	0.085 (0.173)
GO0019724	B cell mediated immunity	93	0.032 (0.041)	0.824 (1.00)	0.081 (0.173)
GO0045088	Regulation of innate immune response	212	0.033 (0.041)	0.997 (1.00)	0.132 (0.173)
GO0002443	Leukocyte mediated immunity	187	0.041 (0.041)	0.984 (1.00)	0.122 (0.173)
GO0002285	Lymphocyte activation involved immune response	88	0.046 (0.041)	0.996 (1.00)	0.066 (0.173)
GO0002764	Immune response regulating signaling pathway	222	0.048 (0.041)	0.993 (1.00)	0.245 (0.173)
GO0006952	Defense response	983	0.049 (0.041)	1.000 (1.00)	0.295 (0.178)

the prognostic effect of miR-223. Expression data of microRNAs and genes in the tumor genome, survival information as well as demographics of 522 GBM patients are extracted from The Cancer Genome Atlas (TCGA). The outcome of interest is dichotomous 3-month survival status.

We have discussed the plausibility of identifiability assumptions in Section 2.1. The analyses adjust for age at diagnosis, sex, ethnicity as well as the center or batch effect using Combat (Johnson et al., 2007), and are limited to GBM, the grade 4 astrocytoma. Because the number of genes within gene sets is large, we can only conduct tests for Δ^*, $\boldsymbol{\delta}^*$, and τ_p based on the transformation models. Among the 120 gene sets, tests for the marginal ME Δ^* identify 9, 44, and 79 gene sets with *p*-values less than 0.05, 0.1, and 0.2, respectively, compared to 0, 9, and 44 gene sets using tests for $L2$ norm of component-wise MEs τ_p. No gene set is identified using tests for *p*-DF component-wise effects. We tabulate the 9 gene sets with *p*-value for the marginal ME less than 0.05 (all with FDR = 0.041) in Table 2. Seven of the nine gene sets have immune-related functions.

No unmeasured confounding assumptions by demographic or design characteristics should be satisfied with the above covariate adjustment and restriction. Due to the biological complexity, expression levels of the gene sets are not necessarily independent. The correlation between gene sets may still violate the identifiability assumptions if the other gene sets also have prognostic effect on survival. To address the potential confounding, one can evaluate the influence by sensitivity analysis, which has been developed for the setting with a single mediator (Imai et al., 2010) and requires additional development to incorporate high-dimensional mediators.

In Figure 4, we show directionality of univariate MEs estimated with univariate mediation analyses and their corresponding *p*-values for gene sets shown in Table 2. The univariate MEs in the nine gene sets are dominated by negative effects, which explains the better performance of tests for the marginal ME than those for the $L2$ norm of component-wise MEs. We note the effect estimates may be biased in univariate mediation analysis due to confounding by correlated genes and simply use it to characterize the directionality and distribution of the association. Taken together, our findings suggest that miR-223 expression has a significant effect on the 3-month GBM survival through altering the expression profile of nine gene sets involved in immune functions or responses to external stimulus.

7. Discussion

We study in this article the mediation effect of a causal mediation model with high-dimensional continuous mediators. We show that the mediation effect can be decomposed into component-wise mediation effects, which can be estimated by a series of low-dimensional regression models with a transformed mediator using spectral decomposition. Although we focus on hypothesis testing in this article, the proposed Monte-Carlo procedure can be adapted to make inference on estimation. For example, $\{\Delta(\tilde{\boldsymbol{\theta}})^{(b)}\}$ and $\{\boldsymbol{\delta}(\tilde{\boldsymbol{\theta}})^{(b)}\}$ can be used to construct confidence interval for $\hat{\Delta}$ and $\hat{\boldsymbol{\delta}}$, respectively.

Our development assumes the mediators to be normally distributed, which can be relaxed to accommodate arbitrary continuous distribution. One can model arbitrary distributions via a nonparametric transformation: $h(\mathbf{G}_i) = \mathbf{A}\mathbf{X}_i + S_i\boldsymbol{\alpha}_S + \boldsymbol{\epsilon}_{Gi}$, where $\boldsymbol{\epsilon}_{Gi}$ is normally distributed, and $h(\cdot)$ is an unspecified monotone transformation function. A nonparametric maximum likelihood estimator or a rank-based estimator can be constructed to perform statistical inference for the semiparametric model (Zeng and Lin, 2007; Xue and Zou, 2012).

Although our transformation approach is able to estimate the high-dimensional mediation effect through a series of low-dimensional regression models, the corresponding component-wise mediation effect δ_j^* does not necessarily have an intuitive interpretation. δ_j^* indicates the mediation effect of the exposure S on the outcome Y mediated through the transformed mediator P_j, a linear combination of the original p mediators $\mathbf{G}$. Our approach aims to evaluate the mediation through the p mediators *en bloc* rather than emphasizing the contribution from individual mediators. It is noteworthy that the component-wise mediation effect δ_j corresponds to the mediation through G_j if G_j is independent of other G's, according to our result (10). However, under the setting with correlation among G's, δ_j does not necessarily correspond to the effect of G_j due to the nonidentifiability of certain paths (VanderWeele

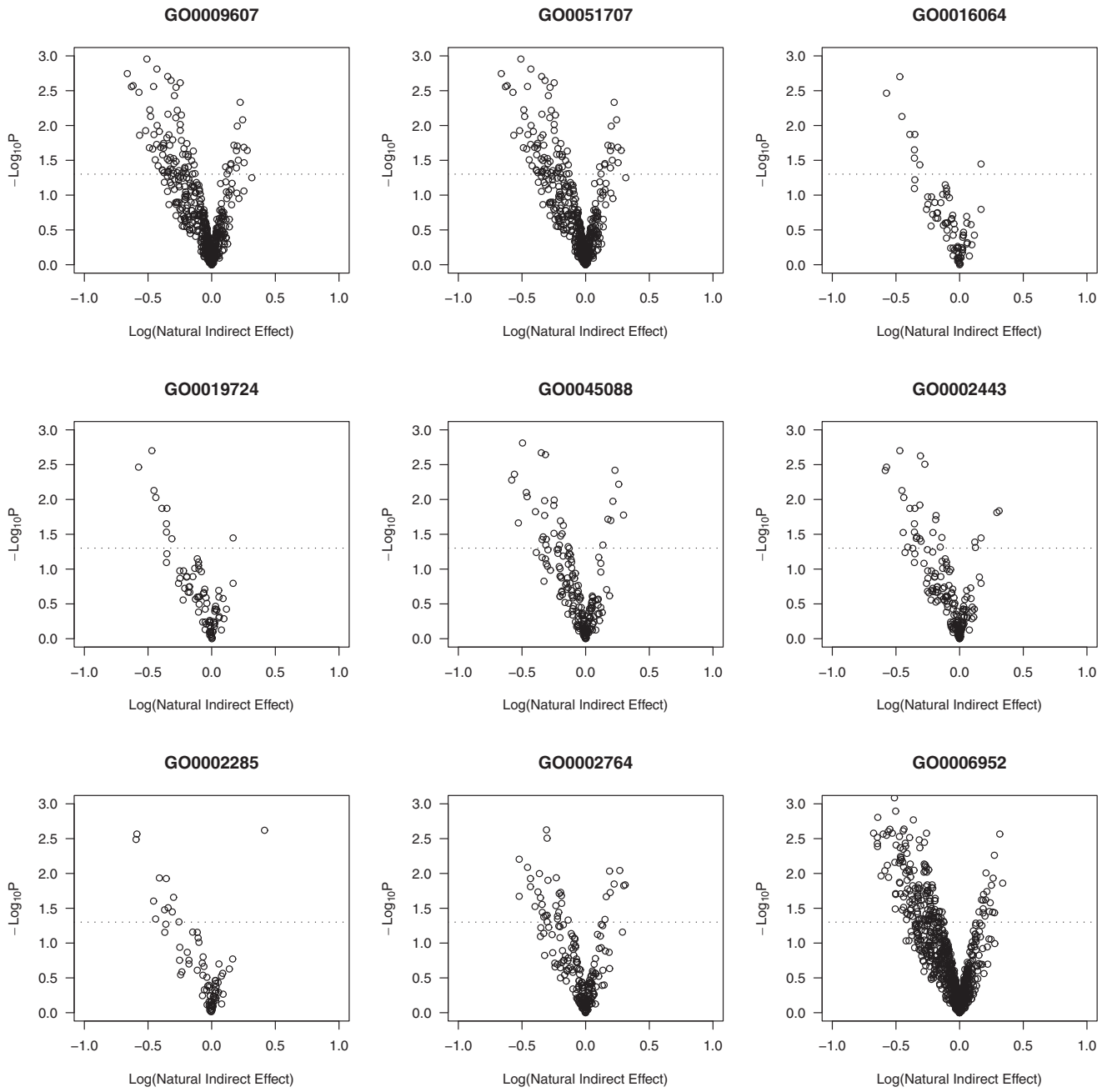


Figure 4. Volcano plots of univariate mediation analyses for genes within the top nine gene sets in Table 2.

and Vansteelandt, 2013), even though they sum up to the marginal mediation effect, $\Delta = \sum_j \delta_j$.

8. Supplementary Materials

Discussion of multivariate extension of assumptions 2) and 4) (Section 2.1), illustration of equivalence of Zhao's method and ours for (11) (Section 4.1), approximation of normal product distribution (Section 4.2), consistency and bias of $\hat{\sigma}^{*2}$ (Section 4.2), additional simulation results including Figures S1–S3 (Section 5, discussion of disjunctive effect including Figures S4 and S5 (Section 5), and a zip file containing R implementation of our methods and analyses is available with this article at the *Biometrics* website on Wiley Online Library.

Acknowledgements

The authors are grateful to the editor, the associate editor, and three anonymous referees for their insightful comments that improved the presentation of the article, and would like to thank the prose-editing service of International Biometric Society (Duncan Hedderley). This study is supported by National Institutes of Health grants CA182937 (NCI), AG048825 (NIA).

References

Albert, J. M. and Nelson, S. (2011). Generalized causal mediation analysis. *Biometrics* **67**, 1028–1038.

Aroian, L. A. (1947). The probability function of the product of two normally distributed variables. *Annals of Mathematical Statistics* **18**, 265–271.

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene ontology: Tool for unification of biology. *Nature Genetics* **25**, 25–29.

Avin, C., Shpitser, I., and Pearl, J. (2005). Identifiability of path-specific effects. In *Proceedings of the International Joint Conferences on Artificial Intelligence*, 357–363.

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

Cai, T., Lin, X., and Carroll, R. J. (2012). Identifying genetic marker sets associated with phenotypes via an efficient adaptive score test. *Biostatistics* **13**, 776–790.

Daniel, R. M., De Stavola, B. L., Cousen, S. N., and Vansteelandt, S. (2015). Causal mediation analysis with multiple mediators. *Biometrics* **71**, 1–14.

Harville, D. A. (2001). *Matrix Algebra From a Statistician's Perspective*. Heidelberg: Springer.

Huang, Y. T. (2015). Integrative modeling of multi-platform genomic data under the framework of mediation analysis. *Statistics in Medicine* **34**, 162–178.

Huang, Y. T., Hsu, T., Kelsey, K. T., and Lin, C. L. (2015). Integrative analysis of microRNA, gene expression and survival of glioblastoma multiforme. *Genetic Epidemiology* **39**, 134–143.

Huang, Y. T., VanderWeele, T. J., and Lin, X. (2014). Joint analysis of SNP and gene expression data in genetic association studies of complex diseases. *The Annals of Applied Statistics* **8**, 352–376.

Imai, K., Keele, L., and Yamamoto, T. (2010). Identification, inference and sensitivity analysis for causal mediation effects. *Statistical Science* **25**, 51–71.

Johnson, W. E., Rabinovic, A., and Li, C. (2007). Adjusting batch effects in microarray expression data using empirical bayes methods. *Biostatistics* **8**, 118–127.

Lin, X. (1997). Variance component test in generalised linear models with random effects. *Biometrika* **84**, 309–326.

MacKinnon, D. P. (2008). *Introduction to statistical mediation analysis*. NY: Taylor & Francis.

MacKinnon, D. P., Lockwood, C. M., and William, J. (2004). Confidence limits for the indirect effect: distribution of the product and resampling methods. *Multivariate Behavioral Research* **39**, 99–128.

Pearl, J. (2001). Direct and indirect effects. In *Proceedings of the Seventeenth Conference on Uncertainty and Artificial Intelligence*, 411–420. San Francisco: Morgan Kaufmann.

Preacher, K. J. and Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods* **40**, 879–891.

Robins, J. M. and Greenland, S. (1992). Identifiability and exchangeability for direct and indirect effects. *Epidemiology* **3**, 143–155.

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

Stupp, R., Mason, W. P., van den Bent, M. J., Weller, M., Fisher, B., Taphoorn, M. J., et al. (2005). Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. *New England Journal of Medicine* **352**, 987–996.

VanderWeele, T. J. and Vansteelandt, S. (2010). Odds ratios for mediation analysis for a dichotomous outcome. *American Journal of Epidemiology* **172**, 1339–1348.

VanderWeele, T. J. and Vansteelandt, S. (2013). Mediation analysis with multiple mediators. *Epidemiologic Methods* **2**, 95–115.

Wang, W., Nelson, S., and Albert, J. M. (2013). Estimation of causal mediation effects for a dichotomous outcome in multiple-mediator models using the mediation formula. *Statistics in Medicine* **24**, 4211–4228.

Wu, M., Lee, S., Cai, T., Li, Y., Boehnke, M., and Lin, X. (2011). Rare-variant association testing for sequencing data with the sequence kernel association test. *The American Journal of Human Genetics* **89**, 82–93.

Xue, L. and Zou, H. (2012). Regularized rank-based estimation of high-dimensional nonparanormal graphical models. *Annals of Statistics* **40**, 2541–2571.

Zeger, S. L., Liang, K. Y., and Albert, P. S. (1988). Models for longitudinal data: a generalized estimating equation appraoch. *Biometrics* **44**, 1049–1060.

Zeng, D. and Lin, D. Y. (2007). Maximum likelihood estimation in semiparametric regression models with censored data. *Journal of the Royal Statistical Society, Series B* **69,** 507–564.

Zhao, S. D., Cai, T. T., and Li, H. (2014). More powerful genetic association testing via a new statistical framework for integrative genomics. *Biometrics* **70**, 881–890.

Received October 2014. *Revised July* 2015.
Accepted August 2015.

Appendix

Appendix: Derivation of the Model $[Y|\mathbf{X}, S, P_j]$

One can derive the expression of the model that relates the continuous outcome to the covariates $\mathbf{X}$, the exposure S and a single transformed mediator P_j:

$$
\begin{aligned}
\mathrm{E}(Y|\mathbf{X}, S, P_j) &= \int y \int \mathrm{P}(y, \mathbf{P}_{(-j)}|\mathbf{X}, S, P_j) d\mathbf{P}_{(-j)} dy \\
&= \int y \int \mathrm{P}(y|\mathbf{X}, S, \mathbf{P}) \mathrm{P}(\mathbf{P}_{(-j)}|\mathbf{X}, S, P_j) d\mathbf{P}_{(-j)} dy \\
&= \int y \int \mathrm{P}(y|\mathbf{X}, S, \mathbf{P}) \mathrm{P}(\mathbf{P}_{(-j)}|\mathbf{X}, S) d\mathbf{P}_{(-j)} dy \\
&= \int \mathrm{E}(Y|\mathbf{X}, S, \mathbf{P}) \mathrm{P}(\mathbf{P}_{(-j)}|\mathbf{X}, S) d\mathbf{P}_{(-j)} \\
&= \mathbf{X}^T \boldsymbol{\beta}_X + S\beta_S + P_j \beta^*_{Gj} + SP_j \beta^*_{Cj} \\
&\quad + \sum_{k \neq j} (\beta^*_{Gk} + S\beta^*_{Ck}) \mathrm{E}(P_k|\mathbf{X}, S) \\
&= \mathbf{X}^T (\boldsymbol{\beta}_X + \sum_{k \neq j} \beta^*_{Gk} \mathbf{A}^*_k) + S(\beta_S + \sum_{k \neq j} \beta^*_{Gk} \alpha^*_{Sk}) \\
&\quad + S\mathbf{X}^T \sum_{k \neq j} \beta^*_{Ck} \mathbf{A}^*_k + S^2 \sum_{k \neq j} \beta^*_{Ck} \alpha^*_{Sk} \\
&\quad + P_j \beta^*_{Gj} + SP_j \beta^*_{Cj}.
\end{aligned}
$$

Under the rare disease assumption where we can approximate logit by log and vice versa, the logistic model that relates the dichotomous Y to $\mathbf{X}$, S, and P_j can be derived as follows:

$$
\begin{aligned}
&\text{logit}[P(Y=1|\mathbf{X}, S, P_j)] \\
&= \text{logit}[\int P(Y=1, \mathbf{P}_{(-j)}|\mathbf{X}, S, P_j) d\mathbf{P}_{(-j)}] \\
&= \text{logit}[\int P(Y=1|\mathbf{X}, S, \mathbf{P}) P(\mathbf{P}_{(-j)}|\mathbf{X}, S, P_j) d\mathbf{P}_{(-j)}] \\
&= \text{logit}[\int P(Y=1|\mathbf{X}, S, \mathbf{P}) P(\mathbf{P}_{(-j)}|\mathbf{X}, S) d\mathbf{P}_{(-j)}] \\
&\approx \text{logit}\Big\{ \exp\left(\mathbf{X}^T \boldsymbol{\beta}_X + S\beta_S + P_j\beta^*_{Gj} + SP_j\beta^*_{Cj}\right) \\
&\quad \times \text{E}[\exp\left(\sum_{k\neq j} P_k(\beta^*_{Gk} + S\beta^*_{Ck})\right)|\mathbf{X}, S]\Big\} \\
&\approx \mathbf{X}^T\left(\boldsymbol{\beta}_X + \sum_{k\neq j}\beta^*_{Gk}\mathbf{A}^*_k\right) + S\left(\beta_S + \sum_{k\neq j}\beta^*_{Gk}\alpha^*_{Sk}\right) \\
&\quad + S\mathbf{X}^T\sum_{k\neq j}\beta^*_{Ck}\mathbf{A}^*_k + S^2\sum_{k\neq j}\beta^*_{Ck}\alpha^*_{Sk} \\
&\quad + P_j\beta^*_{Gj} + SP_j\beta^*_{Cj}.
\end{aligned}
$$