Heterogeneous Mediation Analysis on Epigenomic PTSD and Traumatic Stress in a Predominantly African American Cohort

Fei Xue[1], Xiwei Tang[2], Grace Kim[3], Karestan C. Koenen[4], Chantel L. Martin[5], Sandro Galea[6], Derek Wildman[7], Monica Uddin[7], and Annie Qu[8]

[1]Purdue University

[2]University of Virginia

[3]University of Illinois College of Medicine

[4]Harvard T.H. Chan School of Public Health

[5]The University of North Carolina at Chapel Hill

[6]Boston University

[7]University of South Florida

[8]University of California Irvine

Abstract

DNA methylation (DNAm) has been suggested to play a critical role in post-traumatic stress disorder (PTSD), through mediating the relationship between trauma and PTSD. However, this underlying mechanism of PTSD for African Americans still remains unknown. To fill this gap, in this paper, we investigate how DNAm mediates the effects of traumatic experiences on PTSD symptoms in the Detroit Neighborhood Health Study (DNHS) (2008–2013)

which involves primarily African Americans adults. To achieve this, we develop a new mediation analysis approach for high-dimensional potential DNAm mediators. A key novelty of our method is that we consider heterogeneity in mediation effects across sub-populations. Specifically, mediators in different sub-populations could have opposite effects on the outcome, and thus could be difficult to identify under a traditional homogeneous model framework. In contrast, the proposed method can estimate heterogeneous mediation effects and identifies sub-populations in which individuals share similar effects. Simulation studies demonstrate that the proposed method outperforms existing methods for both homogeneous and heterogeneous data. We also present our mediation analysis results of a dataset with 125 participants and more than $450,000$ CpG sites from the DNHS study. The proposed method finds three sub-groups of subjects and identifies DNAm mediators corresponding to genes such as *HSP90AA1* and *NFATC1* which have been linked to PTSD symptoms in literature. Our finding could be useful in future finer-grained investigation of PTSD mechanism and in the development of new treatments for PTSD.

Key words: Clustering, difference of convex, DNA methylation, high-dimensional mediators, linear structural equation modeling, variable selection

1 Introduction

Post-traumatic stress disorder (PTSD) is a serious mental health disorder that people may develop after they experience or witness a traumatic event, such as a natural disaster, a serious accident, a war, or sexual violence. People suffering from PTSD have symptoms such as disturbing thoughts, nightmares related to the events, mental or physical distress to trauma-related cues, attempts to avoid trauma-related cues, and negative alterations in thinking and feeling (Morrison et al., 2019). As shown in Roberts et al. (2011) and Himle et al. (2009), the prevalence of PTSD is higher in African Americans (AAs) than whites. This is possibly due to that PTSD is an adversity-related mental disorder and that AAs are more likely to encounter socially adverse experiences of discrimination and isolation which have a profound impact on mental health (Hudson et al., 2013;

Cacioppo et al., 2015). However, research for the course of PTSD among AAs is very limited.

As suggested by Rusiecki et al. (2013) and Morrison et al. (2019), DNA methylation (DNAm), an epigenetic mechanism associated with the regulation of gene expression, plays a critical role in the pathophysiology of the PTSD. Since DNAm is a reversible process (Ramchandani et al., 1999), uncovering the role of DNAm in the pathophysiology of PTSD may facilitate the development of new potential treatments for PTSD (Rusiecki et al., 2013). For example, the epigenetics literature suggests that DNAm could mediate the effect of traumatic events on depression (Gao et al., 2019; Vangeel et al., 2015; Dempster et al., 2014; Tyrka et al., 2016; Januar et al., 2015; Schuster et al., 2017; Zhao et al., 2013; Lei et al., 2015; Turecki and Meaney, 2016; van der Knaap et al., 2015). In particular, Peng et al. (2018) discover that DNAm levels at two cytosine-phosphate-guanine (CpG) probes mediate the effects from childhood trauma to depressive symptoms. Rutten et al. (2018) also observe that DNAm mediates the relationship between combat trauma and PTSD symptoms.

Moreover, according to Dickstein et al. (2010), heterogeneity occurs in the course of PTSD. For example, Kim et al. (2019b) observe gender differences in risk of PTSD, indicating a potential mechanism that yields heterogeneity across subjects in the effects of trauma. This motivates us to develop a heterogeneous model. In addition, Heinzelmann and Gill (2013) suggest that epigenetics may play a key role in the heterogeneous responses to trauma and differential risk of PTSD. Furthermore, Orcutt et al. (2004) and Dickstein et al. (2010) reveal distinct trajectories of PTSD symptoms. This is crucial since finding all the prototypical patterns of adaptation to trauma could help us identify biomarkers and risk factors for PTSD. Nevertheless, to the best of our knowledge, heterogeneity in the mediation effects of the DNAm between traumatic events and PTSD symptoms has not been investigated so far.

In fact, heterogeneous mediation effects could arise frequently due to factors such as demographic and genetic characteristics, medical history, lifestyle, and unobserved attributes of subjects. Mediators in different sub-populations could vary or have different effects on the outcome. For example, as illustrated in Figure 1, high pressure on subjects could mediate the effects from the exposure of competitive environments to the performance of subjects. Yet, high pressure may have

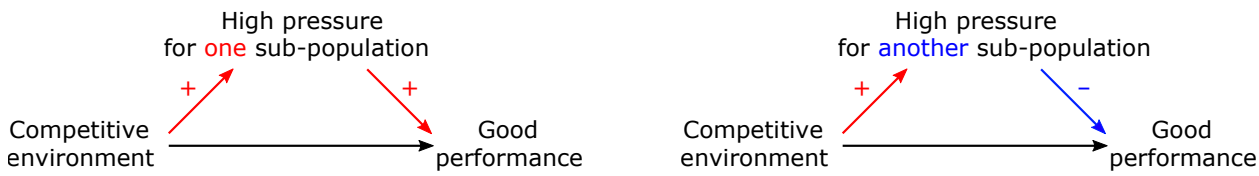

Figure 1: A heterogeneous mediation example.

positive effects on performance of subjects in some sub-population but negative effects for subjects in another sub-population. This heterogeneity among subjects likely comes from stress tolerance and the ability to turn pressure into power, which may vary greatly across different groups of people. In this case, it could be infeasible to identify the true mediator in a homogeneous model, since opposite effects could be canceled out. Recently, several mediation methods have been studied for heterogeneous mediation effects (Qin and Hong, 2017; Dyachenko and Allenby, 2018). However, they either require pre-specified sub-populations or only focus on a single mediator variable.

In this paper, we aim to investigate the high-dimensional heterogeneous mediation effects of DNAm variation on the relationship between trauma exposures and PTSD symptoms using the Detroit Neighborhood Health Study (DNHS) data (Uddin et al., 2010). The DNHS is a representative longitudinal cohort study involving primarily African Americans. In addition, we propose a novel mediator selection method which can identify subgroups of subjects and select mediators in each subgroup simultaneously for high-dimensional potential mediators. We refer to this new method as "the proposed method" in the rest of this paper.

Note that the goal of our study is not to compare differences in PTSD risk by race and thus an interaction between race and trauma is not needed in our model. Rather, AAs' higher prevalence of PTSD is a motivation for us to examine the mediation effects in a sample from this population. Hence the focus of this paper is to investigate the mediation mechanism in a predominantly AA sample (the DNHS data). On the other hand, in fact, there are few non-AA subjects in the DNHS dataset, indicating that it is inappropriate to use this dataset to compare racial differences in PTSD.

Our numerical studies show that the proposed method outperforms existing homogeneous methods in terms of mediation effect estimation and mediator selection. More importantly, the proposed method identifies meaningful DNAm mediators which are not selected by the homo-

geneous mediation methods on the DHNS data. Specifically, the selected DNAm CpG probes correspond to genes including *HSP90AA1*, *SMARCA4*, and *NFATC1* which are indeed associated with PTSD risk (Criado-Marrero et al., 2018; Breen et al., 2019; Kim et al., 2019a; Raabe and Spengler, 2013; Kim et al., 2019a; Kuan et al., 2017). These DNAm mediators can be highly informative in future development of novel interventions for PTSD. In addition, our data analysis shows the potential heterogeneity in mediation effects of DNAm on PTSD risk, which suggests finer-grained comparisons in future PTSD research.

The remainder of this paper is organized as follows. In Section 2, we describe details of the DNHS data which we analyze in this article. In Section 3, we propose the heterogeneous mediation method and illustrate the implementation of the proposed method. Section 4 provides numerical studies through simulations. In Section 5, we apply the proposed method to the DNHS dataset. Finally, we conclude this study with discussion in Section 6.

2 DNHS Data

Our work is motivated by the Detroit Neighborhood Health Study (DNHS) where samples are collected between 2008–2013 from predominantly African American (AA) adults living in Detroit, Michigan. Studies suggest that DNA methylation (DNAm) could play a crucial role as mediators in the underlying relationship between traumatic events and PTSD (Gao et al., 2019; Vangeel et al., 2015; Dempster et al., 2014; Tyrka et al., 2016). Our work was further inspired by initial joint analysis of expression and EWAS data in GRRN genes (Vukojevic et al., 2014; Palma-Gudiel et al., 2015; Kim et al., 2019a; Wani et al., 2021), to find potential functional significance, and the DNHS is one of the few available datasets with both of these data types (Uddin et al., 2010).

In this paper, in order to understand the underlying mechanism of PTSD in AAs, we conduct mediation analysis of DNAm on the relationship between trauma exposures and PTSD symptoms. One significant impact of identifying true mediators is that the occurrence of PTSD can be potentially intervened through the DNAm mediators. This is especially important for PTSD patients

in the DNHS since the independent variables such as the previous trauma experience and adverse events cannot be altered after.

The DNHS is comprised of five survey waves where a total of 2081 participants have completed a 40-minute telephone survey. The survey includes questions on participants' neighborhoods, mental and physical health status, social support, exposure to traumatic events, PTSD symptoms, and various demographics characteristics. All participants have been offered an opportunity to provide a blood specimen for genetic testing of DNA. In particular, the DNHS measures the PTSD symptom severity of participants through the widely-used self-report PTSD Checklist, Civilian Version (PCL-C) (Blanchard et al., 1996; Grubaugh et al., 2007). The PCL-C set contains 17 items corresponding to key symptoms of PTSD. Participants indicate how much they have been bothered by each symptom using a 5-point (1–5) scale in reference to their worst traumatic experience. To access the overall severity, we calculate the average of the 17 items and treat the average PCL-C score as a representative of the PTSD symptom severity for each participant.

The DNHS also records the types of traumatic or stressful events that each participant has experienced. In our survey, we have specific questions such as "Have you experienced exposure to a war zone in the military or as a civilian?" to understand the exact nature of trauma exposures. Thus, a trauma exposure of a subject provides information of a certain type of trauma. We calculate the total number of trauma exposures (i.e., how many different types of traumas a subject experienced) and use it as a trauma feature characterizing overall trauma severity for each subject. Here we use the total number of experienced event types since many studies have used it as a measure of severity of trauma, e.g., Lee and Park (2018); Irish et al. (2013); Harte et al. (2015); Farley et al. (2001); Kessler et al. (2017).

The blood specimens in the DNHS are processed according to Weckle et al. (2015), and the DNA Mini Kit (Qiagen, Germantown, MD) is used to extract genomic DNA from peripheral blood. The extracted DNA samples are bisulfite-converted using the Zymo EZ-96 DNA methylation kit (Zymo Research, Irvine, CA). The converted samples are then profiled through the Illumina Infinium 450 K DNA methylation array (Illumina, San Diego, CA) according to manufacturer pro-

tocols. Specifically, we assess the methylation levels of more than 450,000 CpG sites which cover 99% of reference sequence (RefSeq) genes. More detailed explanation of the processing procedures are provided in Kim et al. (2019b); Ward-Caviness et al. (2020); Wolf et al. (2018); Ratanatharathorn et al. (2017), and Uddin et al. (2018).

There are several existing studies on the DNHS data from various aspects. Uddin et al. (2010) find that Detroit residents have PTSD prevalence more than twice of the corresponding prevalence in the entire US, and that a person's immune-related functions are related to genes with relatively lower levels of methylation. McClure et al. (2018) assess the association between environmental stressors and the Great Recession, while Horesh et al. (2015) and Kim et al. (2019b) reveal gender differences in PTSD. Moreover, Chang et al. (2012); Ratanatharathorn et al. (2017); Uddin et al. (2018) and Nevell et al. (2014) study genetic factors associated with the risk of PTSD in the DNHS. However, none of them have conducted mediation analysis for relationship between traumatic exposures and PTSD symptoms on the DNHS data.

In this paper, we utilize the baseline wave in DNHS for our mediation analysis. In total, there are 125 subjects in the DNHS who have available PTSD measurements, trauma type numbers, and DNAm data. This sample size is relatively small. In fact, many existing studies on PTSD-related mediation analysis have around one hundred participants (Kwon et al., 2021; Ruhlmann et al., 2019; Demir et al., 2020; Kearney et al., 2013; van der Vleugel et al., 2020; Kelly et al., 2019). A relatively small sample size could be a challenge for existing statistical methods, which motivates us to develop more powerful statistical methods to identify mediators. The demographic characteristics of these DNHS participants are summarized in Table 1. The study participants are predominantly (88.8%) African-American (AA). In addition, of the 125 participants, 64.8% are female, and 43.2% are current smoking, referring to any cigarette smoking in the past 30 days. The age of DNHS participants has a wide range from 20 to 89 years with a median of 53 years.

Studies show that the pathophysiology of PTSD is associated with DNAm in glucocorticoid receptor regulatory network (GRRN) genes (Rusiecki et al., 2013). Therefore, we screen DNAm CpG probes via an expression quantitative trait methylation (eQTM) analysis on GRRN-annotated

Table 1: Key demographic characteristics. "AA" represents African American. "EA" represents European American. "Current smoking" refers to any cigarette smoking in the past 30 days.

Gender		Race		
Female	Male	AA	EA	Other
81	44	111	13	1
Median age (Range)		Current smoking		
		No	Yes	Missing
53 (20–89)		70	54	1

DNAm CpG probes and 53 expressed GRRN genes. Here these 53 genes correspond to $1,680$ GRRN-annotated probes. For each probe, we examine whether the probe is significantly correlated with expression levels of the corresponding gene at a significance level of 0.05. If it is significant, we select that probe; otherwise, the probe is not selected. That is, we choose the CpG probes in probe-gene pairs with significant p-values. Through this, we identify 144 CpG probes significantly correlated with the GRRN genes. For mediation analysis, we treat the PCL-C score as a dependent variable, the total number of trauma exposures as an independent variable, and the 144 DNAm CpG probes as potential mediators. Our aim is to select key mediators between the independent and dependent variables from the 144 DNAm CpG probes. We introduce our proposed method in the following section and provide the analysis results of the DNHS data in Section 5.

3 Heterogeneous Mediation Analysis Method

In this section, we propose a heterogeneous mediator selection approach inspired by Tang et al. (2020) to achieve sub-population identification, mediator selection in each sub-population, and mediation effect estimation for heterogeneous data simultaneously. Statistically, to the best of our knowledge, this is the first work which considers heterogeneous mediation effects for high-dimensional potential mediators without pre-specifying subgroups. In addition, we do not assume the subgroup membership depending on observed covariates.

Moreover, to select mediators instead of variables in each sub-population, we propose a new

mediation penalty which jointly penalizes the effect from the independent variable to a mediator (independent-mediator effect) and the effect from the mediator to the outcome (mediator-outcome effect). Essentially, the proposed mediation penalty encourages selection of mediators with large mediation effects. Before introducing the details of the proposed method, we first discuss related existing methods in the following subsection.

3.1 Existing Mediation Analysis methods

Traditional mediation analysis has been conducted via linear regression models (Baron and Kenny, 1986). As an extension, causal mediation analysis imposes "no unmeasured confounding" assumptions and defines direct and indirect effects under a counterfactual framework with potential outcomes (Rubin, 1974; Imai et al., 2010a; Robins and Greenland, 1992; Pearl, 2001).

Methods of multiple mediators have been developed in recent years (Serang et al., 2017; Boca et al., 2013; Imai et al., 2010b; Jirolon et al., 2020). To account for high-dimensional mediators, Zhao and Luo (2016) consider mediation pathway selection for a large number of causally dependent mediators, and present a sparse mediation model using a regularized structural equation model (SEM). In addition, Van Kesteren and Oberski (2018) develop an exploratory coordinate-wise mediation filter approach, and Zhang et al. (2016) propose a high-dimensional mediation analysis (HIMA) approach for DNA methylation. Moreover, Zhou et al. (2020) develop estimation and inference procedures for mediation effects under a high-dimensional linear mediation model. However, these methods are all under the homogeneous model framework.

To investigate heterogeneous mediation effects, Qin and Hong (2017) develop a weighting method to identify and estimate site-specific mediation effects, utilizing an inverse-probability-of-treatment weight (Rosenbaum, 1987) and ratio-of-mediator-probability weighting (Hong et al., 2015). This method assumes the heterogeneity caused by the variation of sites. However, the method is not applicable in general, since the potential mechanism resulting in sub-populations is usually unknown. Dyachenko and Allenby (2018) propose a Bayesian mixture model which combines likelihood functions based on two different outcome models to incorporate heterogene-

ity. Nevertheless, only a single mediator variable is considered and the mixture model requires a pre-specified number of subgroups.

3.2 Notations and Assumptions

In this subsection, we introduce notations and assumptions for the proposed method. Let X_i be an independent variable (e.g., treatment or exposure), $\boldsymbol{Z}_i$ be a $r \times 1$ vector of pre-treatment confounders (e.g., race or gender), $\boldsymbol{M}_i = (M_{i1}, \ldots, M_{ip})^T$ be potential mediators, and Y_i be the outcome for the i-th subject ($1 \leq i \leq N$). Throughout the paper, we make the stable unit treatment value assumption (SUTVA) (Rubin, 1980); that is, the potential outcomes of one subject are unaffected by the assignment of treatments to other subjects, which is a standard assumption for causal inference. Without loss of generality, we also assume that the outcome and all the covariates are centered. We suppose that the entire population can be partitioned into H non-empty subgroups, where mediators and mediation effects within each subgroup are homogeneous. Denote the index set for subjects in the h-th subgroup by $\mathcal{S}(h)$ for $h = 1, \ldots, H$.

3.3 Proposed Subgroup Linear Structural Equations Modeling Framework

We consider the heterogeneous mediation problem under the following subgroup linear structural equations modeling (LSEM) with p potential mediators and r pre-treatment confounders:

$$\boldsymbol{M}_i = \boldsymbol{b}_h X_i + \boldsymbol{\Gamma} \boldsymbol{Z}_i + \boldsymbol{\delta}_i, \tag{1}$$

$$Y_i = \beta_h X_i + \boldsymbol{\theta}_h^T \boldsymbol{M}_i + \boldsymbol{\gamma}^T \boldsymbol{Z}_i + \varepsilon_i, \tag{2}$$

for subject $i \in \mathcal{S}(h)$, where $\boldsymbol{b}_h, \boldsymbol{\delta}_i, \boldsymbol{\theta}_h \in \mathbb{R}^p$, $\boldsymbol{\Gamma} \in \mathbb{R}^{p \times r}$, and $\boldsymbol{\gamma} \in \mathbb{R}^r$. Here, β_h represents the direct effect from the independent variable X_i (e.g., trauma) to the outcome variable Y_i (e.g., PTSD) in the h-th subgroup, $\boldsymbol{\theta}_h^T \boldsymbol{b}_h$ represents the joint mediation (indirect) effect of the treatment, $\boldsymbol{b}_h$ is the parameter relating the treatment to the potential mediators, and $\boldsymbol{\delta}_i \sim N(\boldsymbol{0}, \Sigma)$ and $\varepsilon_i \sim N(0, \sigma^2)$ are random errors, where $\boldsymbol{\delta}_i$ is independent of X_i and $\boldsymbol{Z}_i$, and ε_i is independent of X_i, $\boldsymbol{Z}_i$, and $\boldsymbol{M}_i$. We acknowledge that the LSEM imposes strong assumptions about the linearity and distribution

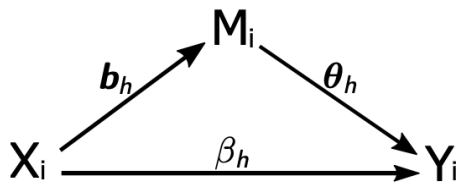

Figure 2: Mediation structure.

of variables, indicating the results based on the proposed subgroup LSEM should be considered exploratory and instructive for generation of potential hypotheses.

We illustrate the relationship of variables in the LSEM for the h-th subgroup without the pre-treatment confounders $\boldsymbol{Z}_i$ in Figure 2. The arrow with β_h represents the direct effects from the independent variable X_i to the response Y_i, while the arrow with $\boldsymbol{\theta}_h$ represents effects from mediators $\boldsymbol{M}_i$ to the response Y_i in model (2). The direct effect β_h and indirect effects $\boldsymbol{\theta}_h^T \boldsymbol{b}_h$ could vary across different subgroups.

In this paper, we assume that the potential mediators are uncausally correlated (Jirolon et al., 2020). That is, the mediators could be conditionally dependent given the independent variable and observed pre-treatment confounders, but are not in any pre-specified causal order. For instance, it is possible that there exist an unmeasured covariate U affecting multiple mediators like two mediators M_1 and M_2 in 3(a), where X denotes the independent variable and Y denotes the outcome. Since M_1 and M_2 are not causally ordered, they are defined as "uncausally correlated". In contrast, M_1 and M_2 in Figure 3(b) are causally ordered, that is, a change in M_2 causes a change in M_1.

In addition, we assume that $P(X_i = x \mid \boldsymbol{Z}_i = \boldsymbol{z}) > 0$ for all x and $\boldsymbol{z}$ if X_i is discrete, and that $f_X(x \mid \boldsymbol{Z}_i = \boldsymbol{z}) > 0$ for all x and $\boldsymbol{z}$ if X_i is continuous, where f_X is the conditional density function of X_i. We also suppose that the effects of pre-treatment confounders $\boldsymbol{Z}_i$ on each mediator and the outcome are homogeneous across all the subgroups in (1) and (2). Moreover, we assume that the mediator models and the outcome model in (1) and (2) are sparse; that is, most true values in $\boldsymbol{b}_h$ and $\boldsymbol{\theta}_h$ are zero for each h.

(a) Uncausally correlated mediators. (b) Causally ordered mediators.

Figure 3: Different situations with two mediators.

Under the proposed model and the assumptions, the sequential ignorability assumption in Jirolon et al. (2020) holds for each subgroup in the parametric model in (1) and (2). Moreover, the natural indirect effects $\delta^j(t)$ on the eighth page of Jirolon et al. (2020) for the j-th mediator is identical to $b_{h,j}\theta_{h,j}$ for subjects from the h-th sub-population, where $b_{h,j}$ and $\theta_{h,j}$ are the j-th element in $\boldsymbol{b}_h$ and $\boldsymbol{\theta}_h$ in equations (1) and (2), respectively. In the following proposition, we have shown that parameters in (1) and (2) are identifiable, which implies that the indirect effects are identifiable under the parametric model. Thus we aim to estimate the causal quantity $b_{h,j}\theta_{h,j}$ under the proposed model.

To state the proposition, we introduce some notations. Let $\boldsymbol{C} = (C_1, \ldots, C_n)^T$ be a vector consisting of all subgroup labels, and $\boldsymbol{\Theta}$ be a vector collecting all elements in $\boldsymbol{\Gamma}, \boldsymbol{\gamma}$ and $\boldsymbol{b}_h, \beta_h, \boldsymbol{\theta}_h$ for $h = 1, \ldots H$. Since the group-specific parameters, $\boldsymbol{b}_h, \beta_h$ and $\boldsymbol{\theta}_h$, are different across sub-groups, we consider the identifiability problem on the set $\mathcal{A} = \{(\boldsymbol{\Theta}, \boldsymbol{C}) : \boldsymbol{b}_h \neq \boldsymbol{b}_{h'}, \beta_h \neq \beta_{h'},$ and $\boldsymbol{b}_h \neq \boldsymbol{b}_{h'}$ for any $h \neq h'\}$.

Proposition 1. *Under the non-degeneracy condition* **??**, *for any* $(\boldsymbol{\Theta}, \boldsymbol{C}), (\boldsymbol{\Theta}', \boldsymbol{C}') \in \mathcal{A}$, *under the proposed model in equations (1) and (2), if the distribution of the observable variables satisfies* $f(X_i, \boldsymbol{M}_i, Y_i, \boldsymbol{Z}_i; \boldsymbol{\Theta}, \boldsymbol{C}) = f(X_i, \boldsymbol{M}_i, Y_i, \boldsymbol{Z}_i; \boldsymbol{\Theta}', \boldsymbol{C}')$ *for* $1 \leq i \leq N$, *then* $(\boldsymbol{\Theta}, \boldsymbol{C})$ *and* $(\boldsymbol{\Theta}', \boldsymbol{C}')$ *are the same up to a permutation of subgroups.*

This proposition states that our model is parametrically identifiable up to a permutation of sub-groups. The proof of the Proposition is provided in Section S.1 of Supplementary Materials. Due to page limit, the non-degeneracy condition 1 is also in Section S.1 of Supplementary Materials.

To determine the subgroup membership of each subject, we evaluate how well each subject fits

each sub-population by the loss function

$$L_{i,h}(\boldsymbol{\Theta}_{1h}, \boldsymbol{\Theta}_2) = \text{tr}\{(\boldsymbol{M}_i - \boldsymbol{b}_h X_i - \boldsymbol{\Gamma}\boldsymbol{Z}_i)(\boldsymbol{M}_i - \boldsymbol{b}_h X_i - \boldsymbol{\Gamma}\boldsymbol{Z}_i)^T\} + (Y_i - \beta_h X_i - \boldsymbol{\theta}_h^T \boldsymbol{M}_i - \boldsymbol{\gamma}^T \boldsymbol{Z}_i)^2$$

for the i-th subject and the h-th subpopulation, where $\boldsymbol{\Theta}_{1h} = (\beta_h, \boldsymbol{b}_h^T, \boldsymbol{\theta}_h^T)^T$ and $\boldsymbol{\Theta}_2 = (\boldsymbol{\gamma}, \boldsymbol{\Gamma}^T)$. We propose to group subjects and identify the subgroup label of each subject through finding a smallest $L_{i,h}$ for the i-th subject among all subgroups ($h = 1, \ldots, H$), since a smaller loss function indicates better fitness and greater likelihood of the sample. Let $\boldsymbol{\Theta}_1 = (\boldsymbol{\Theta}_{11}, \ldots, \boldsymbol{\Theta}_{1H})$. Then, the loss function for all the subjects is

$$L_1(\boldsymbol{\Theta}_1, \boldsymbol{\Theta}_2) = \sum_{1 \le i \le N} \min_{1 \le h \le H} \{L_{i,h}(\boldsymbol{\Theta}_{1h}, \boldsymbol{\Theta}_2)\}, \tag{3}$$

which is a sum of all within-cluster loss.

In the proposed method, we do not pre-specify or assume which variables determine the subgroup membership. Instead, these classes are identified in a completely data-driven manner, which is one of the key advantages of our method. We do not impose assumptions on which variables cause or determine the heterogeneity, since the heterogeneity of mediation effects could have complicated reasons. Essentially, analogous to a clustering algorithm, our model aggregates individuals into several classes where the individuals share similar mediation effects within each class but have different effects across classes. In addition, our proposed subgroup identification is different from that in the individualized-multi-directional method (IMDM) (Tang et al., 2021). We provide a comparison of the two methods in Section S.5.1 of the Supplementary Materials.

3.4 Mediation Regularization for Sparsity

In this subsection, we propose a new mediation penalty. As mentioned in Section 3.3, under the proposed subgroup LSEM in (1) and (2), the mediation effect of the j-th mediator in the h-th subgroup is $b_{h,j}\theta_{h,j}$. To identify mediators with large mediation effects in each subgroup, we consider $b_{h,j}$ and $\theta_{h,j}$ jointly for each $1 \le j \le p$, and propose a two-dimensional joint mediation

penalty

$$p_m(\boldsymbol{b}_h, \boldsymbol{\theta}_h) = \sum_{j=1}^{p} \left(1 - \frac{1}{(1 + c_0|b_{h,j}|)(1 + c_0|\theta_{h,j}|)}\right), \tag{4}$$

where c_0 is a constant to adjust the shrinkage. We plot the mediation penalty with $p = 1$ and $c_0 = 0.5$ in Figure 4. As shown in Figure 4, the penalty tends to shrink small values toward zero, and the shrinkage gradually levels off as $|\theta_{h,1}|$ or $|b_{h,1}|$ increases.

The high-dimensional mediation analysis proposed by Zhang et al. (2016) adopts the minimax concave penalty (MCP) (Zhang, 2010) for variable selection in the outcome model. Compared with the traditional Lasso (Tibshirani, 1996), MCP, and smoothly clipped absolute deviation (SCAD) penalties (Fan and Li, 2001), the proposed mediation penalty selects mediators instead of covariates in the sense that the $p_m(\boldsymbol{b}_h, \boldsymbol{\theta}_h)$ penalizes $\boldsymbol{b}_h$ and $\boldsymbol{\theta}_h$ jointly rather than separately. Schaid and Sinnwell (2020) also jointly consider each pair of coefficients in the mediator and outcome model for each mediator using a group Lasso penalty. Compared to the group Lasso penalty, the proposed joint mediation penalty tends to be flat, and the corresponding shrinkage gradually levels off as the coefficients in each pair increase; which can relax the rate of penalization for large coefficients and large mediation effects. Thus, the proposed penalty could reduce the bias due to joint shrinkages.

Intuitively, we need a relatively small shrinkage for $b_{h,j}$ when $\theta_{h,j}$ is large; otherwise it is hard to select the j-th mediator when $b_{h,j}$ is small but $\theta_{h,j}$ and the mediation effect $b_{h,j}\theta_{h,j}$ are large. However, this property does not hold for penalty functions such as $|\theta_{h,j}| + |b_{h,j}|$ and the pathway Lasso penalty in Zhao and Luo (2016) penalizing $\boldsymbol{b}_h$ and $\boldsymbol{\theta}_h$ jointly via $|\theta_{h,j}b_{h,j}|$ ($1 \leq j \leq p$). Specifically, with a penalty of $|\theta_{h,j}| + |b_{h,j}|$, the shrinkage of $b_{h,j}$ does not depend on $\theta_{h,j}$; and with a penalty of $|\theta_{h,j}b_{h,j}|$, the shrinkage on $b_{h,j}$ increases as $\theta_{h,j}$ increases. In contrast, our joint mediation penalty in (4) has the desired property since the shrinkage on $b_{h,j}$ in $p_m(\boldsymbol{b}_h, \boldsymbol{\theta}_h)$ gradually levels off as $\theta_{h,j}$ increases and vice versa, as shown in Figure 4. On the other hand, the proposed mediation penalty will not induce a strong shrinkage for $\theta_{h,j}$ if $b_{h,j} = 0$. In Figure 4, even if $b_{h,1} = 0$, the rate of penalization on $\theta_{h,1}$ will still decrease as $\theta_{h,1}$ itself increases, which is analog to the SCAD penalty imposed on a single coefficient $\theta_{h,1}$.

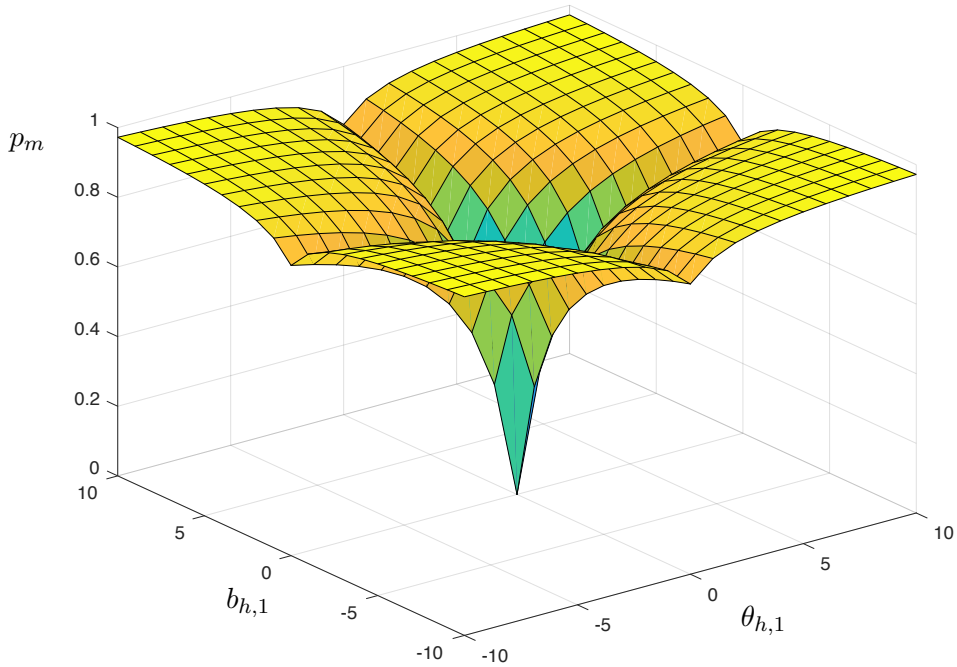

Figure 4: Mediation penalty with $p = 1$ and $c_0 = 0.5$.

3.5 Fusion Penalty for Cross-group Information

In general, it is possible that not every mediator has different mediation effects across subgroups. We use the following between-group fused Lasso penalty (Tibshirani et al., 2005)

$$p_b(\boldsymbol{\Theta}_1) \;=\; \lambda_0 \sum_{1\le h_1,h_2\le H} \left[|\beta_{h_1} - \beta_{h_2}| + \sum_{j=1}^{p} \{|\theta_{h_1,j} - \theta_{h_2,j}| + |b_{h_1,j} - b_{h_2,j}|\} \right] \tag{5}$$

to shrink similar between-group effects together, where λ_0 is a tuning parameter. Specifically, the between-group penalty $p_b(\boldsymbol{\Theta}_1)$ encourages mediators to share the same parameter across different subgroups when the corresponding effects are similar. In this way, we can borrow information across subgroups in estimating the mediation effects.

Consequently, the objective function of the proposed method is

$$f(\boldsymbol{\Theta}_1, \boldsymbol{\Theta}_2) = L_1(\boldsymbol{\Theta}_1, \boldsymbol{\Theta}_2) + N \left\{ \sum_{1\le h\le H} p_w(\boldsymbol{\Theta}_{1h}) + p_b(\boldsymbol{\Theta}_1) + p_c(\boldsymbol{\Theta}_2) \right\}, \tag{6}$$

where

$$p_w(\boldsymbol{\Theta}_{1h}) = p_{SCAD,\lambda_1,a}(\beta_h) + \lambda_1 p_m(\boldsymbol{b}_h, \boldsymbol{\theta}_h) \tag{7}$$

is a within-group penalty, $p_{SCAD,\lambda_1,a}(\cdot)$ is the SCAD penalty with tuning parameters λ_1 and a, $p_c(\mathbf{\Theta}_2) = \lambda_2\{\|\boldsymbol{\gamma}\|_1 + \sum_{j=1}^{p} \|\mathbf{\Gamma}_j\|_1\}$ is a l_1 penalty with tuning parameter λ_2 for pre-treatment confounders to avoid over-fitting, and $\mathbf{\Gamma}_j$ denotes the j-th row of $\mathbf{\Gamma}$. Through minimizing $f(\mathbf{\Theta}_1, \mathbf{\Theta}_2)$ in (6), we not only incorporate the heterogeneity of mediation effects in subjects, but also combine cross-group information for similar effects. Moreover, in the objective function $f(\mathbf{\Theta}_1, \mathbf{\Theta}_2)$, we penalize not only $\boldsymbol{\theta}_h$ but also the direct effect β_h in the outcome model in equation (2) since otherwise the estimation obtained by minimizing the objective function could falsely transfer the effects of mediators $\boldsymbol{M}_i$ to the effect of X_i. In the following subsection, we provide an algorithm to obtain the minimizer of the objective function.

3.6 Implementation

In this subsection, we propose an effective algorithm to minimize the objective function in equation (6). Note that our objective function is non-convex, since the within-group penalty p_w in (7) is not convex and the loss function in (3) involves minimization. To tackle this challenge, we decompose our objective function as a difference of two convex functions and solve the optimization problem based on the difference of convex (DC) algorithm (Le Thi Hoai and Tao, 1997; Shen et al., 2012), which is shown to converge to a stationary point under regularity conditions (Abbaszadehpeivasti et al., 2021). Also, we use a smooth approximation of the between-group fused Lasso penalty.

Specifically, we rewrite the objective function as a difference of two functions

$$f(\boldsymbol{\nu}) = f_1(\boldsymbol{\nu}) - f_2(\boldsymbol{\nu}), \tag{8}$$

where $\boldsymbol{\nu}$ is a vector consisting of all parameters $\boldsymbol{b}_h, \beta_h, \boldsymbol{\theta}_h, \boldsymbol{\gamma}, \mathbf{\Gamma}$ (for $h = 1, \ldots, H$) in models (1) and (2),

$$f_1(\boldsymbol{\nu}) = \sum_{i=1}^{N}\sum_{h=1}^{H} L_{i,h}(\mathbf{\Theta}_{1h}, \mathbf{\Theta}_2) + N\left\{\sum_{h=1}^{H}\{p_d(\mathbf{\Theta}_{1h}) + p_w(\mathbf{\Theta}_{1h})\} + p_b(\mathbf{\Theta}_1) + p_c(\mathbf{\Theta}_2)\right\},$$

$$f_2(\boldsymbol{\nu}) = \sum_{i=1}^{N}\max_{1\le k\le H}\left\{\sum_{1\le h\le H, h\ne k} L_{i,h}(\mathbf{\Theta}_{1h}, \mathbf{\Theta}_2)\right\} + N\sum_{h=1}^{H} p_d(\mathbf{\Theta}_{1h}), \tag{9}$$

and $p_d(\mathbf{\Theta}_{1h}) = \beta_h^2/(a-1) + 2c_0^2\lambda_1(\|\boldsymbol{\theta}_h\|_2^2 + \|\boldsymbol{b}_h\|_2^2)$. It can be shown that $p_d(\mathbf{\Theta}_{1h}) + p_w(\mathbf{\Theta}_{1h})$ is

convex for each $1 \leq h \leq H$, since the corresponding Hessian matrix is block-diagonal and each diagonal block is positive-definite. This implies that $f_1(\boldsymbol{\nu})$ is a convex function. In addition, since the maximum of convex functions is still convex, $f_2(\boldsymbol{\nu})$ is also a convex function.

Based on this DC decomposition on f, we iteratively construct a sequence of convex approximations for $f(\boldsymbol{\nu})$. We first calculate the subdifferential of $f_2(\boldsymbol{\nu})$ in the following:

$$\partial f_2(\boldsymbol{\nu}) = \sum_{i=1}^{N} \partial \left\{ \max_{1 \leq k \leq H} F_{ik}(\boldsymbol{\nu}) \right\} + N \sum_{h=1}^{H} \partial p_d(\boldsymbol{\Theta}_{1h}) = \sum_{i=1}^{N} \text{co} \left\{ \cup_{k \in J(\boldsymbol{\nu})} \partial F_{ik}(\boldsymbol{\nu}) \right\} + N \sum_{h=1}^{H} \partial p_d(\boldsymbol{\Theta}_{1h}),$$

where $F_{ik}(\boldsymbol{\nu}) = \sum_{1 \leq h \leq H, h \neq k} L_{i,h}(\boldsymbol{\Theta}_{1h}, \boldsymbol{\Theta}_2)$, $J(\boldsymbol{\nu}) = \{1 \leq k \leq H : F_{ik}(\boldsymbol{\nu}) = \max_{1 \leq k \leq H} F_{ik}(\boldsymbol{\nu})\}$, and "co" stands for the convex hull. At the m-th iteration, given the previous estimate $\boldsymbol{\nu}^{(m-1)}$, we replace $f_2(\boldsymbol{\nu})$ in (8) by its affine minorization: $f_2^{(m)}(\boldsymbol{\nu}) = f_2(\boldsymbol{\nu}^{(m-1)}) + \left\langle \boldsymbol{\nu} - \boldsymbol{\nu}^{(m-1)}, \boldsymbol{\mu}^{(m-1)} \right\rangle$, where $\boldsymbol{\mu}^{(m-1)} \in \partial f_2(\boldsymbol{\nu}^{(m-1)})$. Then, $f_1(\boldsymbol{\nu}) - f_2^{(m)}(\boldsymbol{\nu})$ is an upper convex approximating function for $f(\boldsymbol{\nu})$ at the m-th iteration. Through this, we convert the non-convex objective function into a convex relaxation via a tangent approximation of $f_2(\boldsymbol{\nu})$.

In addition, since the between-group fused Lasso $p_b(\boldsymbol{\Theta}_1)$ in $f_1(\boldsymbol{\nu})$ is non-smooth and non-separable, we approximate it by a smooth function. Specifically, we reformulate the fused Lasso as $p_b(\boldsymbol{\Theta}_1) = \lambda_0 \|\boldsymbol{D}\boldsymbol{\nu}\|_1 = \lambda_0 \max_{\|\boldsymbol{\eta}\|_\infty \leq 1} \boldsymbol{\eta}^T \boldsymbol{D}\boldsymbol{\nu}$, where $\boldsymbol{D}$ is a difference operator corresponding to the differences in $p_b(\boldsymbol{\Theta}_1)$. That is, the $p_b(\boldsymbol{\Theta}_1)$ is equivalent to the maximum of the maximization problem for $\boldsymbol{\eta}^T \boldsymbol{D}\boldsymbol{\nu}$. Hence, we let $\tilde{p}_b(\boldsymbol{\nu}; \rho) = \lambda_0 \max_{\|\boldsymbol{\eta}\|_\infty \leq 1} (\boldsymbol{\eta}^T \boldsymbol{D}\boldsymbol{\nu} - \frac{\rho}{2}\|\boldsymbol{\eta}\|_2^2)$, where ρ is a positive smoothing parameter. This function $\tilde{p}_b(\boldsymbol{\nu}; \rho)$ approximates $p_b(\boldsymbol{\nu})$ as $\rho \to 0$ (Nesterov, 2005). Let $\boldsymbol{\eta}^* = S(\boldsymbol{D}\boldsymbol{\nu}/\rho)$, where

$$S(x) = \begin{cases} x, & -1 \leq x \leq 1, \\ 1, & x > 1, \\ -1, & x < -1. \end{cases}$$

Then we have $\tilde{p}_b(\boldsymbol{\nu}; \rho) = \lambda_0 \left\{ (\boldsymbol{\eta}^*)^T \boldsymbol{D}\boldsymbol{\nu} - \frac{\rho}{2}\|\boldsymbol{\eta}^*\|_2^2 \right\}$, which is convex and differentiable in $\boldsymbol{\nu}$ (Chen et al., 2012). In our implementation, we choose $\rho = 10^{-4}$ following Chen et al. (2012).

Consequently, at the m-th iteration, we replace the $f_2(\boldsymbol{\nu})$ and $p_b(\boldsymbol{\nu})$ in (8) by $f_2^{(m)}(\boldsymbol{\nu})$ and

$\tilde{p}_b(\boldsymbol{\nu};\rho)$, respectively, and obtain

$$\boldsymbol{\nu}^{(m)} = \underset{\boldsymbol{\nu}}{\operatorname{argmin}} \tilde{f}^{(m)}(\boldsymbol{\nu}), \tag{10}$$

where

$$\tilde{f}^{(m)}(\boldsymbol{\nu}) = \sum_{i=1}^{N}\sum_{h=1}^{H} L_{i,h}(\boldsymbol{\Theta}_{1h},\boldsymbol{\Theta}_2) + N\sum_{h=1}^{H}\{p_d(\boldsymbol{\Theta}_{1h}) + p_w(\boldsymbol{\Theta}_{1h})\} + N\{\tilde{p}_b(\boldsymbol{\nu};\rho) + p_c(\boldsymbol{\Theta}_2)\} - f_2^{(m)}(\boldsymbol{\nu}).$$

We solve the minimization problem in equation (10) through the gradient descent algorithm (Curry, 1944) with the backtracking line search (Shi, 2004; Stanimirović and Miladinović, 2010) for the step size. The above algorithm can be summarized in Algorithm 1. We also provide an expanded version of this algorithm in Section S.2 of the Supplementary Materials.

Algorithm 1

1. Set the number of subgroups H, tolerance ϵ, and the tuning parameters.
2. Obtain initial subgroup labels $C_i^{(0)}$ for $i = 1,\ldots,N$, via the k-means clustering method (MacQueen, 1967) on all the variables.
3. Calculate $\boldsymbol{\nu}^{(0)}$ based on $C_i^{(0)}$ for each subgroup separately.
4. At the m-th iteration, given $\boldsymbol{\nu}^{(m-1)}$ from the $(m-1)$-th iteration:
   (a) Calculate $\boldsymbol{\mu}^{(m-1)} \in \partial f_2(\boldsymbol{\nu}^{(m-1)})$.
   (b) Solve $\boldsymbol{\nu}^{(m)}$ in (10) through the gradient descent.
   (c) Update subgroup labels $C_i^{(m)}$ based on the loss function in (3) for $i = 1,\ldots,N$.
5. Iterate Step 4 until $\|\boldsymbol{\nu}^{(m)} - \boldsymbol{\nu}^{(m-1)}\|_1 < \epsilon$.

To determine the tuning parameters and the number of subgroups H, we propose the following Bayesian information criterion (BIC) type criterion:

$$\begin{aligned} BIC(\boldsymbol{\lambda},H) &= N\cdot\log\{RSS_M(\boldsymbol{\lambda},H)/N\} + df_M(\boldsymbol{\lambda},H)\cdot\log(N) \\ &\quad + [N\cdot\log\{RSS_Y(\boldsymbol{\lambda},H)/N\} + df_Y(\boldsymbol{\lambda},H)\cdot\log(N)], \end{aligned} \tag{11}$$

where $df_M(\boldsymbol{\lambda},H)$ and $df_Y(\boldsymbol{\lambda},H)$ are numbers of non-zero estimated coefficients in mediator models and the outcome model, respectively, $\boldsymbol{\lambda} = (\lambda_0,\lambda_1,\lambda_2)$, $RSS_Y(\boldsymbol{\lambda},H) = \sum_{1\le h\le H}\sum_{\{i:\ \widehat{C}_{i,\boldsymbol{\lambda}}=h\}}$

$(Y_i - \hat{\beta}_{h,\boldsymbol{\lambda}} X_i - \hat{\boldsymbol{\theta}}_{h,\boldsymbol{\lambda}}^T \boldsymbol{M}_i - \hat{\boldsymbol{\gamma}}_{\boldsymbol{\lambda}}^T \boldsymbol{Z}_i)^2$, and $RSS_M(\boldsymbol{\lambda}, H) = \sum_{1 \le h \le H} \sum_{\{i:\, \widehat{C}_{i,\boldsymbol{\lambda}} = h\}} (\boldsymbol{M}_i - \hat{\boldsymbol{b}}_{h,\boldsymbol{\lambda}} X_i - \widehat{\boldsymbol{\Gamma}}_{\boldsymbol{\lambda}} \boldsymbol{Z}_i)^T (\boldsymbol{M}_i - \hat{\boldsymbol{b}}_{h,\boldsymbol{\lambda}} X_i - \widehat{\boldsymbol{\Gamma}}_{\boldsymbol{\lambda}} \boldsymbol{Z}_i)$. Here, $\hat{\boldsymbol{b}}_{h,\boldsymbol{\lambda}}$, $\widehat{\boldsymbol{\Gamma}}_{\boldsymbol{\lambda}}$, $\hat{\beta}_{h,\boldsymbol{\lambda}}$, $\hat{\boldsymbol{\theta}}_{h,\boldsymbol{\lambda}}$, $\hat{\boldsymbol{\gamma}}_{\boldsymbol{\lambda}}$, and $\widehat{C}_{i,\boldsymbol{\lambda}}$ are the estimates of $\boldsymbol{b}_h$, $\boldsymbol{\Gamma}$, β_h, $\boldsymbol{\theta}_h$, $\boldsymbol{\gamma}$, and C_i, respectively, based on H subgroups and tuning parameters $\boldsymbol{\lambda}$. Recall that the between-group fused Lasso penalty $p_b(\boldsymbol{\Theta}_1)$ in (5) encourages shared parameters for similar effects across different subgroups. In df_M and df_Y, the shared parameters are counted without multiplicity.

We select the optimal tuning parameters and the optimal number of subgroups through minimizing $BIC(\boldsymbol{\lambda}, H)$, incorporating information from both the mediator models and the outcome model. In the implementation for the following sections, we mainly tune λ_0, λ_1, and H using a grid search to minimize the BIC. We do not tune λ_2 since it is for penalization of the pre-treatment confounders, which are not involved in our simulations and real data application.

4 Simulated Data Experiments

In this section, we investigate the performance of the proposed method compared with existing homogeneous mediation methods via simulation studies. We simulate data following models in (1) and (2) with $r = 0$. The proposed method is implemented based on Algorithm 1 with $c_0 = 10$ and $a = 3.7$. We apply the "HIMA" package (`https://cran.r-project.org/web/packages/HIMA/index.html`) in R to implement the high-dimensional mediation analysis (HIMA) method (Zhang et al., 2016) for comparison, which is a homogeneous mediation approach. Our results are summarized based on 100 replications.

To evaluate the performance of each method, we calculate the average of all individuals' mediator false negative rates (FN) and the average of all individuals' mediator false positive rates (FP) for mediator selection as follows:

$$FN = \frac{1}{N} \sum_{i=1}^{N} \frac{\sum_{j=1}^{p} I(\hat{\theta}_{\widehat{C}_i j} \hat{b}_{\widehat{C}_i j} = 0, \theta_{C_i j} b_{C_i j} \neq 0)}{\sum_{j=1}^{p} I(\theta_{C_i j} b_{C_i j} \neq 0)}, \quad FP = \frac{1}{N} \sum_{i=1}^{N} \frac{\sum_{j=1}^{p} I(\hat{\theta}_{\widehat{C}_i j} \hat{b}_{\widehat{C}_i j} \neq 0, \theta_{C_i j} b_{C_i j} = 0)}{\sum_{j=1}^{p} I(\theta_{C_i j} b_{C_i j} = 0)},$$

where $\hat{\theta}_{hj}$, $\hat{b}_{hj}$, and $\widehat{C}_i$ are estimators of θ_{hj}, b_{hj}, and C_i, respectively. Specifically, the FN and FP

represent proportions of unselected true mediators and selected noises, respectively. A method with smaller FN+FP selects more accurate mediators. In practice, the FP and FN may have different costs. For example, in scenarios such as a pregnancy test or a COVID-19 test, FN costs much more than FP. However, in other scenarios such as criminal conviction or identifying spam emails, FP costs much more than FN. Given specific application context and background information, we may assign different weights to FP and FN, respectively. Since simulation studies do not involve any real situations, we treat them equally and just use FN+FP as an evaluation criterion.

We also evaluate each method via the mean-squared-errors (MSE) of mediation effects in an average of all individuals as $\sum_{i=1}^{N}\sum_{j=1}^{p}(\hat{\theta}_{\widehat{C}_{ij}}\hat{b}_{\widehat{C}_{ij}}-\theta_{C_{ij}}b_{C_{ij}})^2/N$. For the proposed method, we also report the proportion of replications where the number of subgroups is correctly selected via the $BIC(\boldsymbol{\lambda}, H)$ in (11), which we refer to as the correct rate of subgroup number selection. Moreover, we compute the misclassification rate of subjects based on the proposed method for evaluation of subgroup identification.

We consider the following three settings, where the non-zero coefficients in $\boldsymbol{b}_h$ and $\boldsymbol{\theta}_h$ share the same signal strengths b_{hs} and θ_{hs}, respectively, for $1 \leq h \leq H$. For the sensitivity of the proposed approach to misspecification, we investigate situations without heterogeneity in the first setting, which involves a homogeneous underlying true model with only one sub-population. In contrast, Settings 2 and 3 assume heterogeneous true models with two sub-populations. In addition, we consider high-dimensional situations in Setting 3.

Setting 1. *Let $N = 200$, $H = 1$, $n_1 = 200$, and $p = 30$. True coefficients in the model are illustrated in Figure 5 with $\beta_1 = 0.5$, $b_{1s} = 1$, and $\theta_{1s} = 0.2, 0.3,$ or 0.4. As shown in Figure 5, we have four true mediators with mediation effects θ_{1s}. In addition, we generate X_i and ε_i from a standard normal distribution and $\boldsymbol{\delta}_i \sim N(\mathbf{0}, \boldsymbol{I}_{p\times p})$ for each $i = 1, \ldots, N$.*

Setting 2. *We proceed similarly as in Setting 1 except that $H = 2$, $n_1 = 50$, and $n_2 = 150$. True coefficients in the model are illustrated in Figure 6 with $\beta_1 = 0.5, b_{1s} = 1$, $\beta_2 = -0.5, b_{2s} = -1$, $\theta_{1s} = 0.5, 1,$ or 4 and $\theta_{2s} = -0.5, -1,$ or -4. As shown in Figure 6, we have three true mediators in each subgroup.*

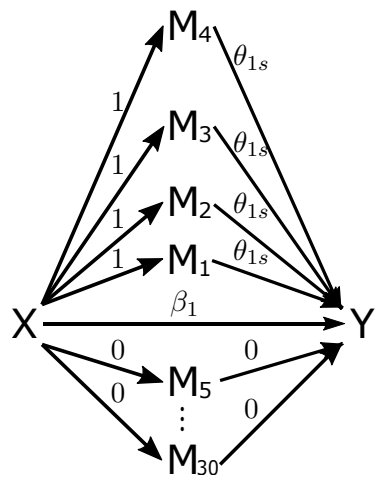

Figure 5: True coefficients for the homogeneous Setting 1. The "X" denotes the independent variable, "M$_i$" denotes the i-th mediator, and "Y" denotes the dependent variable. The value above each arrow represents the true coefficient for the corresponding effect.

Table 2: FN, FP, FN+FP, and MSE under Setting 1. "HIMA" stands for the high-dimensional mediation analysis method. The "θ_{1s}" represents the signal strength in $\boldsymbol{\theta}_1$.

θ_{1s}	Method	FN	FP	FN+FP	MSE
0.2	**Proposed**	0.046	0.002	0.048	0.002
	HIMA	0.570	0.000	0.570	0.003
0.3	**Proposed**	0.000	0.000	0.000	0.002
	HIMA	0.100	0.000	0.100	0.002
0.4	**Proposed**	0.000	0.000	0.000	0.002
	HIMA	0.005	0.000	0.005	0.001

Setting 3. *We investigate a high-dimensional case proceeding similarly as in Setting 2 except that* $N = 100$*,* $n_1 = 30$*,* $n_2 = 70$*,* $p = 150$*,* $\beta_1 = 1$*,* $\beta_2 = -1$*,* $\theta_{1s} = 0.5, 0.8, 1,$ *or* 4 *and* $\theta_{2s} = -0.5, -0.8, -1,$ *or* -4*. The covariance matrix of* $\boldsymbol{\delta}_i$ *has an autoregressive structure of order* 1*, that is, AR(1), with diagonal* 1 *and off-diagonal parameter* ρ*.*

Tables 2–4 provide the results of the proposed method and the HIMA method, and show that the proposed method produces smaller overall FN+FP totals than the HIMA method across all the settings, indicating that the proposed method selects mediators more accurately. Moreover, the proposed method produces smaller MSE of mediation effects, implying that the proposed method is also more effective in estimation of mediation effects.

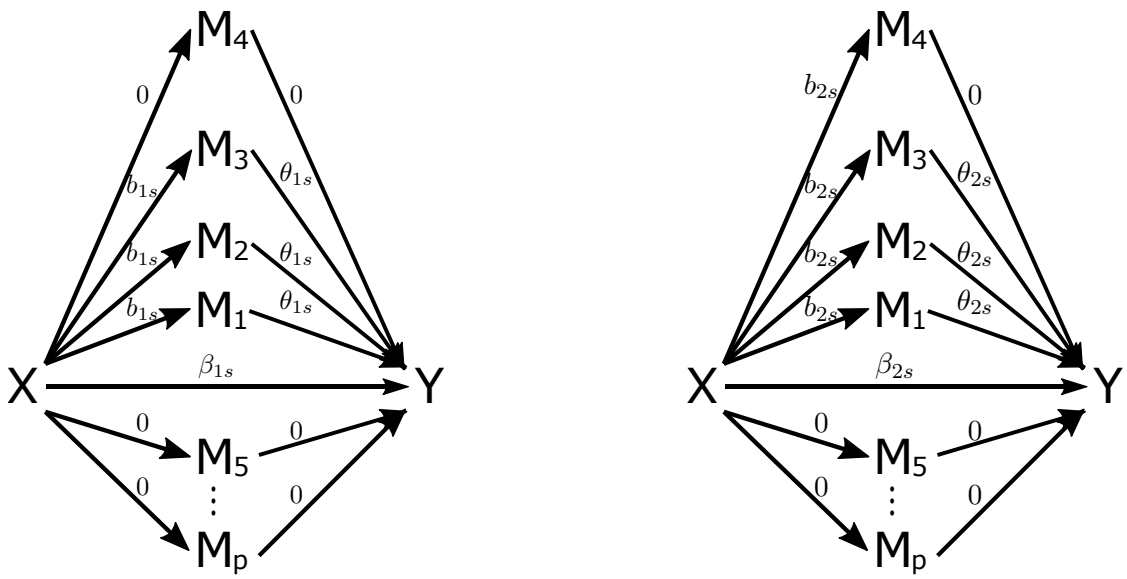

Figure 6: True coefficients for the heterogeneous Settings 2 and 3 with two sub-populations (left and right). The p is 30 and 150 under Settings 2 and 3, respectively. The "X" denotes the independent variable, "M$_i$" denotes the i-th mediator, and "Y" denotes the dependent variable. The value above each arrow represents the true coefficient for the corresponding effect.

Table 3: FN, FP, FN+FP, and MSE under Setting 2. "HIMA" stands for the high-dimensional mediation analysis method. The "θ_{1s}" and "θ_{2s}" represent the signal strength in $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$, respectively.

$(\theta_{1s}, \theta_{2s})$	Method	FN	FP	FN+FP	MSE
(0.5, -0.5)	**Proposed**	0.031	0.007	0.038	0.004
	HIMA	0.963	0.004	0.968	0.025
(1, -1)	**Proposed**	0.002	0.000	0.002	0.006
	HIMA	0.943	0.003	0.946	0.099
(4, -4)	**Proposed**	0.001	0.000	0.001	0.058
	HIMA	0.790	0.002	0.792	1.514

Table 4: FN, FP, FN+FP, and MSE under Setting 3. "HIMA" stands for the high-dimensional mediation analysis method. The "θ_{1s}" and "θ_{2s}" represent the signal strength in $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$, respectively, and "ρ" is a correlation parameter.

	$(\theta_{1s}, \theta_{2s})$	Method	FN	FP	FN+FP	MSE
$\rho = 0$	(0.5, -0.5)	**Proposed**	0.530	0.001	0.531	0.004
		HIMA	0.890	0.001	0.891	0.006
	(0.8, -0.8)	**Proposed**	0.220	0.000	0.220	0.005
		HIMA	0.967	0.001	0.967	0.013
	(1, -1)	**Proposed**	0.143	0.001	0.144	0.007
		HIMA	0.977	0.001	0.977	0.021
	(4, -4)	**Proposed**	0.092	0.001	0.093	0.068
		HIMA	0.963	0.000	0.964	0.314
$\rho = 0.2$	(0.5, -0.5)	**Proposed**	0.554	0.001	0.555	0.003
		HIMA	0.917	0.001	0.917	0.006
	(0.8, -0.8)	**Proposed**	0.269	0.001	0.270	0.006
		HIMA	0.983	0.001	0.984	0.013
	(1, -1)	**Proposed**	0.259	0.001	0.260	0.009
		HIMA	0.977	0.000	0.977	0.021
	(4, -4)	**Proposed**	0.161	0.001	0.162	0.093
		HIMA	0.957	0.000	0.957	0.312

Table 5: Correct rate of subgroup number selection and misclassification rate for different settings. "Correct rate of subgroup number selection" represents the proportion of replications where the number of subgroups is correctly selected via the proposed method. "Misclassification rate" is the average proportion of subjects who are mis-classified by the proposed method.

	$(\theta_{1s}, \theta_{2s})$	Correct rate of subgroup number selection	Misclassification rate
Setting 1	(0.2, –)	0.76	–
	(0.3, –)	1.00	–
	(0.4, –)	1.00	–
Setting 2	(0.5, -0.5)	0.99	0.13
	(1, -1)	1.00	0.11
	(4, -4)	0.92	0.10
Setting 3 $\rho = 0$	(0.5, -0.5)	0.66	0.17
	(0.8, -0.8)	0.68	0.14
	(1, -1)	0.72	0.13
	(4, -4)	0.62	0.12
Setting 3 $\rho = 0.2$	(0.5, -0.5)	0.59	0.17
	(0.8, -0.8)	0.69	0.15
	(1, -1)	0.62	0.15
	(4, -4)	0.66	0.14

In particular, under the homogeneous Setting 1, the proposed method still outperforms the homogeneous HIMA method. For example, when $\theta_{1s} = 0.2$, the FN+FP of the proposed is only 8.8% of that of the HIMA as shown in Table 2. This is likely due to the advantage of the proposed mediation penalty, and that the proposed method correctly identifies the number of subgroups in most situations as shown in Table 5. Moreover, the proposed method falsely un-selects 4.6% of true mediators and falsely selects just 0.2% of noises (variables that are not mediators), even when the mediation effect of each true mediator is as small as 0.2. In addition, under this situation, the MSE of mediation effects is only 0.002, indicating that the proposed method performs consistently well even when there is no heterogeneity and the effect size is small.

In addition, the proposed method also performs much better than the HIMA when there are two subgroups with opposite mediation effects in Settings 2 and 3, that is, with positive mediation effects in one subgroup and negative mediation effects in the other. In this case, the homogeneous HIMA method usually fails to identify true mediators. For instance, the FN of the HIMA is as high as 0.957 when $(\theta_{1s}, \theta_{2s}) = (0.5, -0.5)$ as shown in Table 3, while the corresponding FN of the proposed method is only 0.031.

Moreover, we explore high-dimensional scenarios in Setting 3 to mimic the case of the DNHS data. In this setting, we also consider the situations where the error terms in the mediator models in (1) are correlated, that is, the correlations among mediators may not just come from the independent variable. In all the high-dimensional cases, the proposed method produces smaller FN+FP and smaller MSE than the HIMA method illustrated in Table 4. For example, when $(\theta_{1s}, \theta_{2s}) = (4, -4)$ and $\rho = 0.2$, the FN+FP of the proposed method is only 16.7% of that of the HIMA method, and the MSE is only 29.8% of that of the HIMA.

In Table 5, we provide the correct rates of subgroup number selection and the misclassification rates under various settings. We observe that the proposed method correctly determines the number of subgroups in most situations, especially in Settings 1 and 2 where there are more samples. In addition, the proposed method groups most subjects correctly across different settings due to the low misclassification rates.

In summary, our simulation studies show that the proposed method achieves higher mediator selection accuracy and mediation effect estimation accuracy than the existing homogeneous mediation method across all the settings. One reason is that the proposed method adopts the mediation penalty in (4) which considers effects in mediator models and outcome models jointly, and encourages selection of mediators with large mediation effects. In addition, the proposed method allows heterogeneity among subjects, and thus can identify mediators with heterogeneous mediation effects, which is especially powerful for mediators with opposite effects in different subgroups. We investigate more simulations for larger-scale cases, with different initial values, various coefficients, moderators, and under non-normality of the independent variable in Sections S.3.1, S.3.2, S.3.4–S.3.6 of Supplementary Materials, respectively.

5 DNHS Case Study

In this section, we investigate how DNA methylation mediates the effects of traumatic experiences on development of PTSD based on the DNHS data. Specifically, we apply the proposed method to study the mediation effects of DNAm variation of GRRN genes on PTSD symptom severity. In this study, we use the baseline wave of the DNHS data for our mediation analysis. We treat the total number of trauma exposures as an independent variable, DNAm CpG probes that are significantly correlated with the GRRN genes as potential mediators, and the average PCL-C score as the outcome variable. There are 125 subjects and 144 selected DNAm CpG probes after screening. Our main objective is to identify key DNAm CpG probes from all the potential mediators.

To evaluate the performance of the proposed method compared with existing methods, we randomly split the data into a training set (90% of all samples) and a testing set (10% of all samples) for 100 times. The training sets in the 100 replications are repeated random subsamples (90%) of the complete data. For the proposed method, to identify the subgroup label of the i-th subject in the testing set, we calculate the average prediction error of mediators $(\widehat{\boldsymbol{M}}_i(\widehat{\boldsymbol{b}}_h) - \boldsymbol{M}_i)^T(\widehat{\boldsymbol{M}}_i(\widehat{\boldsymbol{b}}_h) - \boldsymbol{M}_i)/p$, where $\widehat{\boldsymbol{M}}_i(\widehat{\boldsymbol{b}}_h)$ is the predicted value for $\boldsymbol{M}_i$ based on estimated param-

Table 6: The DNHS data results by the proposed method and the high-dimensional mediation analysis (HIMA) method. Here "NS" represents the mean number of selected mediators. The "Mediator" represents average prediction error for mediators, and "Outcome" represents average prediction error for the outcome variable based on 100 replications.

Method	NS	Mediator	Outcome
Proposed	6	0.455	1.235
HIMA	0	4.309	2.774

eters $\hat{\beta}_h$ and $\hat{\boldsymbol{\theta}}_h$ in the h-th subgroup. Then the i-th subject is labeled with a specific subgroup that minimizes $(\widehat{\boldsymbol{M}}_i(\widehat{\boldsymbol{b}}_h) - \boldsymbol{M}_i)^T(\widehat{\boldsymbol{M}}_i(\widehat{\boldsymbol{b}}_h) - \boldsymbol{M}_i)/p$. For each method, we train the model on the training data, and calculate the prediction root-mean-squared errors (PRMSEs) for the mediators and the outcome variable in the testing set; that is, $\sqrt{\sum_{i\in\mathcal{T}}(\widehat{\boldsymbol{M}}_i - \boldsymbol{M}_i)^T(\widehat{\boldsymbol{M}}_i - \boldsymbol{M}_i)/(p|\mathcal{T}|)}$ and $\sqrt{\sum_{i\in\mathcal{T}}(\hat{y}_i - y_i)^2/|\mathcal{T}|}$, respectively, where $\mathcal{T}$ denotes the index set of subjects in the testing set and $|\mathcal{T}|$ is the testing sample size.

We provide mean PRMSE for each method based on 100 replications in Table 6. The proposed method produces much smaller prediction errors for both the mediators and the outcome variable compared to the existing high-dimensional mediation analysis (HIMA) method, indicating that the proposed method is more accurate in terms of prediction. Note that, in both methods, we use all the potential mediators to calculate the prediction errors of mediators, and then take an average of the errors over the mediators.

For the subgroup identification in the DNHS data, we apply the proposed method to all the samples. The proposed method selects four DNAm probes, and identifies three subgroups consisting of 60, 26, 39 subjects, respectively. We provide the estimated coefficients of the four DNAm probes in the three subgroups in Figure 7. Although the size of some estimated coefficients are small, the estimated mediation effect size here is comparable to the findings from other studies for DNA methylation (Tobi et al., 2018), and thus the mediation effects are still non-ignorable. For example, the mediation effect from trauma to PCL score through cg01277438 is $0.02 \times 0.94 \approx 0.02$ for the subgroup in Figure 7(c). Note that the range of the trauma measurement is from 0 to 14 with mean 4.0. If the trauma variable takes its average value 4.0, the influence of this trauma value on

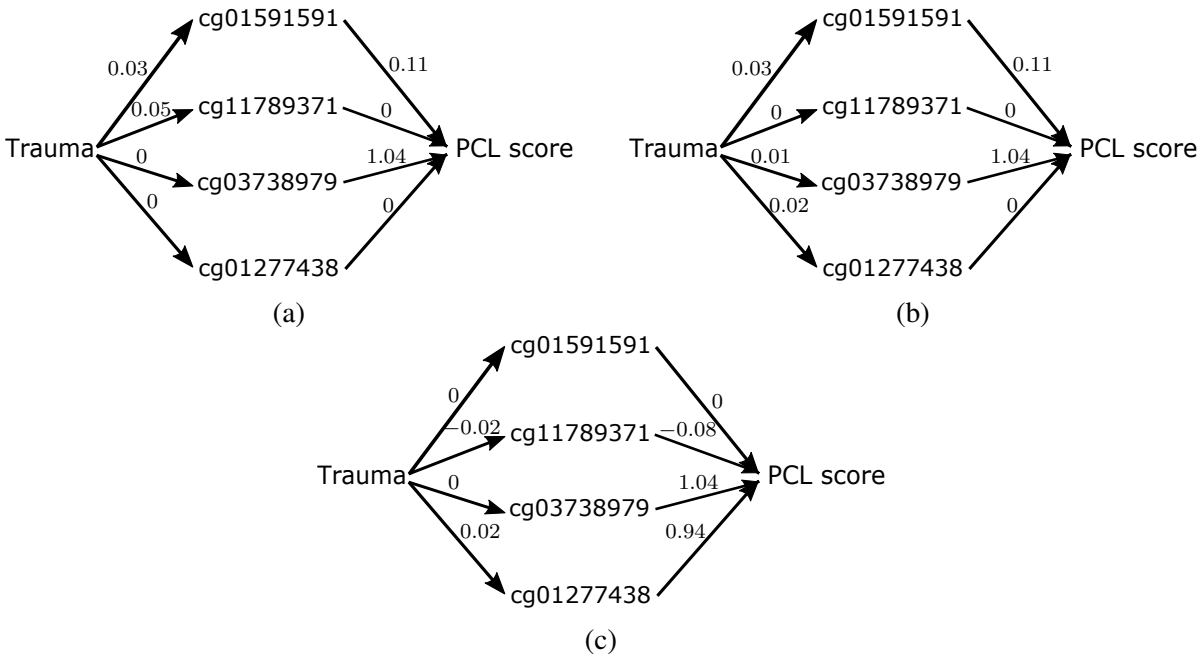

Figure 7: Estimated coefficients for the selected four mediators in the three subgroups identified by the proposed method based on all samples.

PCL score through the mediator cg01277438 is about 0.08. Moreover, if the trauma variable takes its maximum value 14, the influence of this trauma value on the PCL score through the mediator cg01277438 is about 0.28, which is 25% of the standard deviation 1.1 of the PCL score.

Specifically, the four DNAm probes identified by the proposed method correspond to *NFATC1*, *HSP90AA1*, *SMARCA4*, and *CREBBP* genes, among which *HSP90AA1*, *SMARCA4*, and *NFATC1* are indeed related to PTSD based on existing literature (Criado-Marrero et al., 2018; Breen et al., 2019; Kim et al., 2019a; Raabe and Spengler, 2013; Kim et al., 2019a; Kuan et al., 2017). In addition, we apply the HIMA method to each of the three subgroups on the four mediators. At a 0.05 significance level, the HIMA method shows that three ("cg11789371", "cg03738979", "cg01277438") of the four selected mediators are significant, after controlling the false discovery rate (FDR) via the Benjamini–Hochberg procedure (Benjamini and Hochberg, 1995). However, the HIMA method cannot identify any mediator when all potential mediators and samples are used. This confirms that the subgroups and mediators selected by the proposed method are useful in scientific findings.

Furthermore, we find common patterns in subgroup identification across the analyses of the 100 random split datasets and the whole dataset, which are provided in Section S.4.1 of Supplementary Materials. Also, we investigate the racial make-up of the three identified subgroups. The results are provided in Table 13 in Section S.4.2 of Supplementary Materials, showing that the subjects are not grouped based on race since each subgroup contains both AAs and European Americans.

In summary, the proposed method produces smaller prediction errors than the homogeneous mediation method. Moreover, the proposed method identifies important mediators which cannot be detected by existing methods. In addition, our method shows heterogeneous mediation among subjects in the DNHS data.

6 Discussion

Our main contribution is to advance the understanding of the underlying mechanism of PTSD using the DNHS data. Specifically, we conduct a mediation analysis for DNA methylation on the relationship between traumatic events and PTSD symptoms based on the DNHS data. The identification of DNAm mediators presents new statistical challenges due to the heterogeneous nature of PTSD among patients and the high-dimensional structure of the DNAm data. To address these challenges, we develop a heterogeneous mediation model with multiple mediators, incorporating heterogeneity among subjects. Moreover, we propose a novel mediation penalty to incorporate effects in the mediator models and the outcome model jointly for high-dimensional data. Our numerical studies show that the proposed method selects true mediators more accurately than the existing homogeneous high-dimensional mediation analysis method.

In the DNHS case study, we identify meaningful DNAm mediators using the proposed method, which has important impact in practice since it would advance development of new personalized treatments for PTSD. In fact, recent studies have shown that successful PTSD treatments are reflected by significant DNAm changes (Vinkers et al., 2019; Yehuda et al., 2013). Our finding could also suggest future biomedical research on the selected mediators and corresponding genes for

further biological verification. In particular, the selected DNAm CpG probes correspond to genes such as *HSP90AA1*, *SMARCA4*, and *NFATC1* , which have been reported in literature that they are associated with the PTSD (Criado-Marrero et al., 2018; Breen et al., 2019; Kim et al., 2019a; Raabe and Spengler, 2013; Kim et al., 2019a; Kuan et al., 2017).

In addition, the subgroup identification by the proposed method for the DNHS dataset with predominantly African-American subjects indicates potential heterogeneity in the underlying DNAm profiles which mediate risk for PTSD. This is an important discovery, as it could help us to uncover important genetic complexities which have been ignored for treatment and prevention of this debilitating mental disorder.

In this article, we mainly consider the detection of heterogeneous mediators and estimation of corresponding mediation effects. We have not developed statistical inference for the heterogeneous mediation effects, which is a limitation of this paper. It would be of great interest to investigate the statistical inference on these mediation effects in the future; for example, constructing de-biased estimators of sub-population mediation effects for confidence intervals or hypothesis testing. A de-biasing procedure can also reduce the bias for mediation effects incurred from regularization. One possible way is to control the bias of estimators for parameters in both mediator and outcome models, and then use multiplication of each pair of de-biased estimates to estimate the mediation effects of the corresponding mediator.

Moreover, the proposed model in equations (1) and (2) implies that there is no unmeasured confounders. However, to our best knowledge, it is unclear whether there are any common factors impacting both methylation and PTSD symptoms, as the exact biological processes underlying the development of PTSD remain elusive, which is a study limitation given the current state of science. We provide more discussion on the proposed model regarding exposure–mediator interactions, latent subpopulations, and moderated mediation and mediated moderation in Sections S.5.2–S.5.4 of Supplementary Materials, respectively.

Furthermore, our use of cross-sectional rather than longitudinal methylation measurements in the DNHS data application can be considered a study limitation, since the mediation assumption

cannot be verified in cross-sectional data. However, it is consistent with the majority of existing epigenetic work relevant to PTSD published to date. In fact, many methylation studies are actually cross-sectional (Christiansen et al., 2021; Nakatochi et al., 2017; Alghanim et al., 2017; Joehanes et al., 2016; Young et al., 2016; King et al., 2014), and the majority of epigenetic PTSD studies to date have used a similar study design (Sheerin et al., 2021; Young et al., 2021; Qi et al., 2021; Hossack et al., 2020; Kuan et al., 2017). As a future work, we expect the proposed method can be extended to longitudinal data when such data become available.

In addition, the number of subject in the DNHS study is relatively small, and thus readers should be cautious when directly using the results that the DNHS dataset provides. In the future, with more date collected, we will apply the proposed method to a larger dataset. More discussion on the limitations of the DNHS data is provided in Section S.5.5 of Supplementary Materials.

Acknowledgements

We would like to acknowledge support for this project from the National Institutes of Health grants R01MD011728, R01DA022720, and RC1MH088283; and from the National Science Foundation grants DMS-1821198 and DMS-1952406.

References

Abbaszadehpeivasti, H., de Klerk, E., and Zamani, M. (2021). On the rate of convergence of the difference-of-convex algorithm (DCA). *arXiv preprint arXiv:2109.13566*.

Alghanim, H., Antunes, J., Silva, D. S. B. S., Alho, C. S., Balamurugan, K., and McCord, B. (2017). Detection and evaluation of DNA methylation markers found at SCGN and KLF14 loci to estimate human age. *Forensic Science International: Genetics*, 31:81–88.

Baron, R. M. and Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6):1173.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B*, 57(1):289–300.

Blanchard, E. B., Jones-Alexander, J., Buckley, T. C., and Forneris, C. A. (1996). Psychometric properties of the PTSD Checklist (PCL). *Behaviour Research and Therapy*, 34(8):669–673.

Boca, S. M., Sinha, R., Cross, A. J., Moore, S. C., and Sampson, J. N. (2013). Testing multiple biological mediators simultaneously. *Bioinformatics*, 30(2):214–220.

Breen, M. S., Bierer, L. M., Daskalakis, N. P., Bader, H. N., Makotkine, I., Chattopadhyay, M., Xu, C., Grice, A. B., Tocheva, A. S., Flory, J. D., et al. (2019). Differential transcriptional response following glucocorticoid activation in cultured blood immune cells: A novel approach to PTSD biomarker development. *Translational Psychiatry*, 9(1):1–13.

Cacioppo, J. T., Cacioppo, S., Capitanio, J. P., and Cole, S. W. (2015). The neuroendocrinology of social isolation. *Annual Review of Psychology*, 66:733–767.

Chang, S.-C., Koenen, K. C., Galea, S., Aiello, A. E., Soliven, R., Wildman, D. E., and Uddin, M. (2012). Molecular variation at the SLC6A3 locus predicts lifetime risk of PTSD in the Detroit neighborhood health study. *PloS One*, 7(6):e39184.

Chen, X., Lin, Q., Kim, S., Carbonell, J. G., and Xing, E. P. (2012). Smoothing proximal gradient method for general structured sparse regression. *The Annals of Applied Statistics*, 6(2):719–752.

Christiansen, C., Castillo-Fernandez, J., Domingo-Relloso, A., Zhao, W., Moustafa, J. E.-S., Tsai, P.-C., Maddock, J., Haack, K., Cole, S., Kardia, S., et al. (2021). Novel DNA methylation signatures of tobacco smoking with trans-ethnic effects. *Clinical Epigenetics*, 13(1):1–13.

Criado-Marrero, M., Rein, T., Binder, E. B., Porter, J. T., Koren III, J., and Blair, L. J. (2018). Hsp90 and FKBP51: Complex regulators of psychiatric diseases. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373(1738):20160532.

Curry, H. B. (1944). The method of steepest descent for non-linear minimization problems. *Quarterly of Applied Mathematics*, 2(3):258–261.

Demir, Z., Böge, K., Fan, Y., Hartling, C., Harb, M. R., Hahn, E., Seybold, J., and Bajbouj, M. (2020). The role of emotion regulation as a mediator between early life stress and posttraumatic stress disorder, depression and anxiety in Syrian refugees. *Translational Psychiatry*, 10(1):1–10.

Dempster, E. L., Wong, C. C., Lester, K. J., Burrage, J., Gregory, A. M., Mill, J., and Eley, T. C. (2014). Genome-wide methylomic analysis of monozygotic twins discordant for adolescent depression. *Biological Psychiatry*, 76(12):977–983.

Dickstein, B. D., Suvak, M., Litz, B. T., and Adler, A. B. (2010). Heterogeneity in the course of posttraumatic stress disorder: Trajectories of symptomatology. *Journal of Traumatic Stress*, 23(3):331–339.

Dyachenko, T. L. and Allenby, G. M. (2018). Bayesian analysis of heterogeneous mediation. *Georgetown McDonough School of Business Research Paper*, (2600140).

Fan, J. and Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American Statistical Association*, 96(456):1348–1360.

Farley, M., Minkoff, J. R., and Barkan, H. (2001). Breast cancer screening and trauma history. *Women & Health*, 34(2):15–27.

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.

Grubaugh, A. L., Elhai, J. D., Cusack, K. J., Wells, C., and Frueh, B. C. (2007). Screening for PTSD in public-sector mental health settings: The diagnostic utility of the PTSD checklist. *Depression and Anxiety*, 24(2):124–129.

Harte, C. B., Vujanovic, A. A., and Potter, C. M. (2015). Association between exercise and posttraumatic stress symptoms among trauma-exposed adults. *Evaluation & the Health Professions*, 38(1):42–52.

Heinzelmann, M. and Gill, J. (2013). Epigenetic mechanisms shape the biological response to trauma and risk for PTSD: A critical review. *Nursing Research and Practice*, 2013.

Himle, J. A., Baser, R. E., Taylor, R. J., Campbell, R. D., and Jackson, J. S. (2009). Anxiety disorders among African Americans, blacks of Caribbean descent, and non-Hispanic whites in the United States. *Journal of Anxiety Disorders*, 23(5):578–590.

Hong, G., Deutsch, J., and Hill, H. D. (2015). Ratio-of-mediator-probability weighting for causal mediation analysis in the presence of treatment-by-mediator interaction. *Journal of Educational and Behavioral Statistics*, 40(3):307–340.

Horesh, D., Lowe, S. R., Galea, S., Uddin, M., and Koenen, K. C. (2015). Gender differences in the long-term associations between posttraumatic stress disorder and depression symptoms: Findings from the detroit neighborhood health study. *Depression and Anxiety*, 32(1):38–48.

Hossack, M. R., Reid, M. W., Aden, J. K., Gibbons, T., Noe, J. C., and Willis, A. M. (2020). Adverse childhood experience, genes, and PTSD risk in soldiers: A methylation study. *Military Medicine*, 185(3-4):377–384.

Hudson, D. L., Puterman, E., Bibbins-Domingo, K., Matthews, K. A., and Adler, N. E. (2013). Race, life course socioeconomic position, racial discrimination, depressive symptoms and self-rated health. *Social Science & Medicine*, 97:7–14.

Imai, K., Keele, L., and Tingley, D. (2010a). A general approach to causal mediation analysis. *Psychological Methods*, 15(4):309.

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

Irish, L. A., Gabert-Quillen, C. A., Ciesla, J. A., Pacella, M. L., Sledjeski, E. M., and Delahanty, D. L. (2013). An examination of PTSD symptoms as a mediator of the relationship between trauma history characteristics and physical health following a motor vehicle accident. *Depression and Anxiety*, 30(5):475–482.

Januar, V., Ancelin, M.-L., Ritchie, K., Saffery, R., and Ryan, J. (2015). BDNF promoter methylation and genetic variation in late-life depression. *Translational Psychiatry*, 5(8):e619–e619.

Jirolon, A., Baglietto, L., Birmeli, E., Alarcon, F., and Perduca, V. (2020). Causal mediation analysis in presence of multiple mediators uncausally related. *The International Journal of Biostatistics*.

Joehanes, R., Just, A. C., Marioni, R. E., Pilling, L. C., Reynolds, L. M., Mandaviya, P. R., Guan, W., Xu, T., Elks, C. E., Aslibekyan, S., et al. (2016). Epigenetic signatures of cigarette smoking. *Circulation: Cardiovascular Genetics*, 9(5):436–447.

Kearney, D. J., Malte, C. A., McManus, C., Martinez, M. E., Felleman, B., and Simpson, T. L. (2013). Loving-kindness meditation for posttraumatic stress disorder: A pilot study. *Journal of Traumatic Stress*, 26(4):426–434.

Kelly, M. M., DeBeer, B. B., Meyer, E. C., Kimbrel, N. A., Gulliver, S. B., and Morissette, S. B. (2019). Experiential avoidance as a mediator of the association between posttraumatic stress disorder symptoms and social support: A longitudinal analysis. *Psychological Trauma: Theory, Research, Practice, and Policy*, 11(3):353.

Kessler, R. C., Aguilar-Gaxiola, S., Alonso, J., Benjet, C., Bromet, E. J., Cardoso, G., Degenhardt, L., de Girolamo, G., Dinolova, R. V., Ferry, F., et al. (2017). Trauma and PTSD in the WHO World Mental Health Surveys. *European Journal of Psychotraumatology*, 8(sup5):1353383.

Kim, G., Aiello, A., Koenen, K., Galea, S., Wildman, D., and Uddin, M. (2019a). 27. Differential methylation at glucocorticoid-relevant regulatory regions associated with PTSD in African Americans. *Biological Psychiatry*, 85(10):S11–S12.

Kim, G. S., Smith, A. K., Xue, F., Michopoulos, V., Lori, A., Armstrong, D. L., Aiello, A. E., Koenen, K. C., Galea, S., Wildman, D. E., and Uddin, M. (2019b). Methylomic profiles reveal sex-specific differences in leukocyte composition associated with post-traumatic stress disorder. *Brain, Behavior, and Immunity*, 81:280–291.

King, W. D., Ashbury, J. E., Taylor, S. A., Tse, M. Y., Pang, S. C., Louw, J. A., and Vanner, S. J. (2014). A cross-sectional study of global DNA methylation and risk of colorectal adenoma. *BMC Cancer*, 14(1):1–9.

Kuan, P., Waszczuk, M., Kotov, R., Marsit, C., Guffanti, G., Gonzalez, A., Yang, X., Koenen, K., Bromet, E., and Luft, B. (2017). An epigenome-wide DNA methylation study of PTSD and depression in World Trade Center responders. *Translational Psychiatry*, 7(6):e1158–e1158.

Kwon, A., Lee, H. S., and Lee, S.-H. (2021). The mediation effect of hyperarousal symptoms on the relationship between childhood physical abuse and suicidal ideation of patients with PTSD. *Frontiers in Psychiatry*, 12:376.

Le Thi Hoai, A. and Tao, P. D. (1997). Solving a class of linearly constrained indefinite quadratic problems by DC algorithms. *Journal of Global Optimization*, 11(3):253–285.

Lee, S. Y. and Park, C. L. (2018). Trauma exposure, posttraumatic stress, and preventive health behaviours: A systematic review. *Health Psychology Review*, 12(1):75–109.

Lei, M.-K., Beach, S. R., Simons, R. L., and Philibert, R. A. (2015). Neighborhood crime and depressive symptoms among African American women: Genetic moderation and epigenetic mediation of effects. *Social Science & Medicine*, 146:120–128.

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, volume 1 281–297. Oakland, CA, USA.

McClure, E., Feinstein, L., Ferrando-Martínez, S., Leal, M., Galea, S., and Aiello, A. E. (2018). The Great Recession and immune function. *RSF: The Russell Sage Foundation Journal of the Social Sciences*, 4(4):62–81.

Morrison, F. G., Miller, M. W., Logue, M. W., Assef, M., and Wolf, E. J. (2019). DNA methylation correlates of PTSD: Recent findings and technical challenges. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 90:223–234.

Nakatochi, M., Ichihara, S., Yamamoto, K., Naruse, K., Yokota, S., Asano, H., Matsubara, T., and Yokota, M. (2017). Epigenome-wide association of myocardial infarction with DNA methylation sites at loci related to cardiovascular disease. *Clinical Epigenetics*, 9(1):1–9.

Nesterov, Y. (2005). Smooth minimization of non-smooth functions. *Mathematical Programming*, 103(1):127–152.

Nevell, L., Zhang, K., Aiello, A. E., Koenen, K., Galea, S., Soliven, R., Zhang, C., Wildman, D. E., and Uddin, M. (2014). Elevated systemic expression of ER stress related genes is associated with stress-related mental disorders in the detroit neighborhood health study. *Psychoneuroendocrinology*, 43:62–70.

Orcutt, H. K., Erickson, D. J., and Wolfe, J. (2004). The course of PTSD symptoms among Gulf War veterans: A growth mixture modeling approach. *Journal of Traumatic Stress: Official Publication of The International Society for Traumatic Stress Studies*, 17(3):195–202.

Palma-Gudiel, H., Córdova-Palomera, A., Leza, J. C., and Fañanás, L. (2015). Glucocorticoid receptor gene (NR3C1) methylation processes as mediators of early adversity in stress-related disorders causality: A critical review. *Neuroscience & Biobehavioral Reviews*, 55:520–535.

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

Peng, H., Zhu, Y., Strachan, E., Fowler, E., Bacus, T., Roy-Byrne, P., Goldberg, J., Vaccarino, V., and Zhao, J. (2018). Childhood trauma, DNA methylation of stress-related genes, and depression: Findings from two monozygotic twin studies. *Psychosomatic Medicine*, 80(7):599.

Qi, R., Luo, Y., Zhang, L., Weng, Y., Surento, W., Xu, Q., Jahanshad, N., Li, L., Cao, Z., Lu, G. M., et al. (2021). Decreased functional connectivity of hippocampal subregions and methylation of the NR3C1 gene in Han Chinese adults who lost their only child. *Psychological Medicine*, 51(8):1310–1319.

Qin, X. and Hong, G. (2017). A weighting method for assessing between-site heterogeneity in causal mediation mechanism. *Journal of Educational and Behavioral Statistics*, 42(3):308–340.

Raabe, F. J. and Spengler, D. (2013). Epigenetic risk factors in PTSD and depression. *Frontiers in Psychiatry*, 4:80.

Ramchandani, S., Bhattacharya, S. K., Cervoni, N., and Szyf, M. (1999). DNA methylation is a reversible biological signal. *Proceedings of the National Academy of Sciences*, 96(11):6107–6112.

Ratanatharathorn, A., Boks, M. P., Maihofer, A. X., Aiello, A. E., Amstadter, A. B., Ashley-Koch, A. E., Baker, D. G., Beckham, J. C., Bromet, E., Dennis, M., et al. (2017). Epigenome-wide association of PTSD from heterogeneous cohorts with a common multi-site analysis pipeline. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 174(6):619–630.

Roberts, A. L., Gilman, S. E., Breslau, J., Breslau, N., and Koenen, K. C. (2011). Race/ethnic differences in exposure to traumatic events, development of post-traumatic stress disorder, and treatment-seeking for post-traumatic stress disorder in the United States. *Psychological Medicine*, 41(1):71.

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

Rosenbaum, P. R. (1987). Model-based direct adjustment. *Journal of the American Statistical Association*, 82(398):387–394.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5):688.

Rubin, D. B. (1980). Randomization analysis of experimental data: The Fisher randomization test comment. *Journal of the American Statistical Association*, 75(371):591–593.

Ruhlmann, L. M., Gallus, K. L., Beck, A. R., Goff, B. S. N., and Durtschi, J. A. (2019). A pilot study exploring PTSD symptom clusters as mediators between trauma exposure and attachment behaviors in married adults. *Journal of Couple & Relationship Therapy*, 18(1):65–84.

Rusiecki, J. A., Byrne, C., Galdzicki, Z., Srikantan, V., Chen, L., Poulin, M., Yan, L., and Baccarelli, A. (2013). PTSD and DNA methylation in select immune function gene promoter regions: A repeated measures case-control study of US military service members. *Frontiers in Psychiatry*, 4:56.

Rutten, B. P., Vermetten, E., Vinkers, C. H., Ursini, G., Daskalakis, N. P., Pishva, E., de Nijs, L., Houtepen, L. C., Eijssen, L., Jaffe, A. E., et al. (2018). Longitudinal analyses of the DNA methylome in deployed military servicemen identify susceptibility loci for post-traumatic stress disorder. *Molecular Psychiatry*, 23(5):1145–1156.

Schaid, D. J. and Sinnwell, J. P. (2020). Penalized models for analysis of multiple mediators. *Genetic epidemiology*, 44(5):408–424.

Schuster, R., Kleimann, A., Rehme, M.-K., Taschner, L., Glahn, A., Groh, A., Frieling, H., Lichtinghagen, R., Hillemacher, T., Bleich, S., and Heberlein, A. (2017). Elevated methylation and decreased serum concentrations of BDNF in patients in levomethadone compared to diamorphine maintenance treatment. *European Archives of Psychiatry and Clinical Neuroscience*, 267(1):33–40.

Serang, S., Jacobucci, R., Brimhall, K. C., and Grimm, K. J. (2017). Exploratory mediation analysis via regularization. *Structural Equation Modeling: A Multidisciplinary Journal*, 24(5):733–744.

Sheerin, C. M., Lancaster, E. E., York, T. P., Walker, J., Danielson, C. K., and Amstadter, A. B. (2021). Epigenome-wide study of posttraumatic stress disorder symptom severity in a treatment-seeking adolescent sample. *Journal of Traumatic Stress*.

Shen, X., Pan, W., and Zhu, Y. (2012). Likelihood-based selection and sharp parameter estimation. *Journal of the American Statistical Association*, 107(497):223–232.

Shi, Z.-J. (2004). Convergence of line search methods for unconstrained optimization. *Applied Mathematics and Computation*, 157(2):393–405.

Stanimirović, P. S. and Miladinović, M. B. (2010). Accelerated gradient descent methods with line search. *Numerical Algorithms*, 54(4):503–520.

Tang, X., Xue, F., and Qu, A. (2020). Individualized multidirectional variable selection. *Journal of the American Statistical Association* 1–17.

Tang, X., Xue, F., and Qu, A. (2021). Individualized multidirectional variable selection. *Journal of the American Statistical Association*, 116(535):1280–1296.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B*, 58(1):267–288.

Tibshirani, R., Saunders, M., Rosset, S., Zhu, J., and Knight, K. (2005). Sparsity and smoothness via the fused lasso. *Journal of the Royal Statistical Society: Series B*, 67(1):91–108.

Tobi, E. W., Slieker, R. C., Luijk, R., Dekkers, K. F., Stein, A. D., Xu, K. M., Biobank-based Integrative Omics Studies Consortium, Slagboom, P. E., van Zwet, E. W., Lumey, L., et al. (2018). DNA methylation as a mediator of the association between prenatal adversity and risk factors for metabolic disease in adulthood. *Science Advances*, 4(1):eaao4364.

Turecki, G. and Meaney, M. J. (2016). Effects of the social environment and stress on glucocorticoid receptor gene methylation: A systematic review. *Biological Psychiatry*, 79(2):87–96.

Tyrka, A., Parade, S., Welch, E., Ridout, K., Price, L., Marsit, C., Philip, N., and Carpenter, L. (2016). Methylation of the leukocyte glucocorticoid receptor gene promoter in adults: Associations with early adversity and depressive, anxiety and substance-use disorders. *Translational Psychiatry*, 6(7):e848–e848.

Uddin, M., Aiello, A. E., Wildman, D. E., Koenen, K. C., Pawelec, G., De Los Santos, R., Goldmann, E., and Galea, S. (2010). Epigenetic and immune function profiles associated with posttraumatic stress disorder. *Proceedings of the National Academy of Sciences*, 107(20):9470–9475.

Uddin, M., Ratanatharathorn, A., Armstrong, D., Kuan, P.-F., Aiello, A. E., Bromet, E. J., Galea, S., Koenen, K. C., Luft, B., Ressler, K. J., et al. (2018). Epigenetic meta-analysis across three civilian cohorts identifies nrg1 and hgs as blood-based biomarkers for post-traumatic stress disorder. *Epigenomics*, 10(12):1585–1601.

van der Knaap, L. J., Oldehinkel, A. J., Verhulst, F. C., van Oort, F. V., and Riese, H. (2015). Glucocorticoid receptor gene methylation and HPA-axis regulation in adolescents. The TRAILS study. *Psychoneuroendocrinology*, 58:46–50.

van der Vleugel, B. M., Libedinsky, I., de Bont, P. A., de Roos, C., van Minnen, A., de Jongh, A., van der Gaag, M., and van den Berg, D. (2020). Changes in posttraumatic cognitions mediate the effects of trauma-focused therapy on paranoia. *Schizophrenia Bulletin Open*, 1(1):sgaa036.

Van Kesteren, E.-J. and Oberski, D. L. (2018). Exploratory mediation analysis with many potential mediators. *arXiv preprint arXiv:1810.06334*.

Vangeel, E., Van Den Eede, F., Hompes, T., Izzi, B., Del Favero, J., Moorkens, G., Lambrechts, D., Freson, K., and Claes, S. (2015). Chronic fatigue syndrome and DNA hypomethylation of the glucocorticoid receptor gene promoter 1F region: Associations with HPA axis hypofunction and childhood trauma. *Psychosomatic Medicine*, 77(8):853–862.

Vinkers, C. H., Geuze, E., van Rooij, S. J., Kennis, M., Schür, R. R., Nispeling, D. M., Smith, A. K., Nievergelt, C. M., Uddin, M., Rutten, B. P., et al. (2019). Successful treatment of post-traumatic stress disorder reverses DNA methylation marks. *Molecular Psychiatry* 1–8.

Vukojevic, V., Kolassa, I.-T., Fastenrath, M., Gschwind, L., Spalek, K., Milnik, A., Heck, A., Vogler, C., Wilker, S., Demougin, P., et al. (2014). Epigenetic modification of the glucocorticoid receptor gene is linked to traumatic memory and post-traumatic stress disorder risk in genocide survivors. *Journal of Neuroscience*, 34(31):10274–10284.

Wani, A. H., Aiello, A. E., Kim, G. S., Xue, F., Martin, C. L., Ratanatharathorn, A., Qu, A., Koenen, K., Galea, S., Wildman, D. E., et al. (2021). The impact of psychopathology, social adversity and stress-relevant DNA methylation on prospective risk for post-traumatic stress: A machine learning approach. *Journal of Affective Disorders*, 282:894–905.

Ward-Caviness, C. K., Pu, S., Martin, C. L., Galea, S., Uddin, M., Wildman, D. E., Koenen, K., and Aiello, A. E. (2020). Epigenetic predictors of all-cause mortality are associated with objective measures of neighborhood disadvantage in an urban population. *Clinical Epigenetics*, 12(1):1–10.

Weckle, A., Aiello, A. E., Uddin, M., Galea, S., Coulborn, R. M., Soliven, R., Meier, H., and Wildman, D. E. (2015). Rapid fractionation and isolation of whole blood components in samples obtained from a community-based setting. *Journal of Visualized Experiments*, (105):e52227.

Wolf, E. J., Maniates, H., Nugent, N., Maihofer, A. X., Armstrong, D., Ratanatharathorn, A., Ashley-Koch, A. E., Garrett, M., Kimbrel, N. A., Lori, A., et al. (2018). Traumatic stress and accelerated DNA methylation age: A meta-analysis. *Psychoneuroendocrinology*, 92:123–134.

Yehuda, R., Daskalakis, N. P., Desarnaud, F., Makotkine, I., Lehrner, A., Koch, E., Flory, J. D., Buxbaum, J. D., Meaney, M. J., and Bierer, L. M. (2013). Epigenetic biomarkers as predictors and correlates of symptom improvement following psychotherapy in combat veterans with PTSD. *Frontiers in Psychiatry*, 4:118.

Young, G. P., Pedersen, S. K., Mansfield, S., Murray, D. H., Baker, R. T., Rabbitt, P., Byrne, S., Bambacas, L., Hollington, P., and Symonds, E. L. (2016). A cross-sectional study comparing a blood test for methylated BCAT 1 and IKZF 1 tumor-derived DNA with CEA for detection of recurrent colorectal cancer. *Cancer Medicine*, 5(10):2763–2772.

Young, R. M., Lawford, B., Mellor, R., Morris, C. P., Voisey, J., McLeay, S., Harvey, W., Romaniuk, M., Crawford, D., Colquhoun, D., et al. (2021). Investigation of C-reactive protein and AIM2 methylation as a marker for PTSD in Australian Vietnam veterans. *Gene*, 803:145898.

Zhang, C.-H. (2010). Nearly unbiased variable selection under minimax concave penalty. *The Annals of Statistics*, 38(2):894–942.

Zhang, H., Zheng, Y., Zhang, Z., Gao, T., Joyce, B., Yoon, G., Zhang, W., Schwartz, J., Just, A., Colicino, E., et al. (2016). Estimating and testing high-dimensional mediation effects in epigenetic studies. *Bioinformatics*, 32(20):3150–3154.

Zhao, J., Goldberg, J., Bremner, J. D., and Vaccarino, V. (2013). Association between promoter methylation of serotonin transporter gene and depressive symptoms: A monozygotic twin study. *Psychosomatic Medicine*, 75(6).

Zhao, Y. and Luo, X. (2016). Pathway Lasso: Estimate and select sparse mediation pathways with high-dimensional mediators. *arXiv preprint arXiv:1603.07749*.

Zhou, R. R., Wang, L., and Zhao, S. D. (2020). Estimation and inference for the indirect effect in high-dimensional linear mediation models. *Biometrika*, 107(3):573–589.