Causal Mediation Analysis for Longitudinal Mediators and Survival Outcomes

Shuxi Zeng[1] Elizabeth C.Lange[2] Elizabeth A.Archie[3] Fernando A.Campos [4]
Susan C.Alberts [2,5] Fan Li [1,*]
[1]Department of Statistical Science, Duke University
[2]Department of Biology, Duke University
[3]Department of Biological Sciences, University of Notre Dame
[4]Department of Antropology, University of Texas at San Antonio
[5]Department of Evolutionary Anthropology, Duke University
∗fl35@duke.edu, 014 Old Chemistry Building, Durham, NC, 27708, USA

ABSTRACT

Causal mediation analysis studies how the treatment effect of an exposure on outcomes is mediated through intermediate variables. Although many applications involve longitudinal data, the existing methods are not directly applicable to settings where the mediators are measured on irregular time grids. In this paper, we propose a causal mediation method that accommodates longitudinal mediators on arbitrary time grids and survival outcomes simultaneously. We take a functional data analysis perspective and view longitudinal mediators as realizations of underlying smooth stochastic processes. We define causal estimands of direct and indirect effects accordingly and provide corresponding identification assumptions. We employ a functional principal component analysis approach to estimate the mediator process, and propose a Cox hazard model for the survival outcome that flexibly adjusts the mediator process. We then derive a g-computation formula to express the causal estimands using the model coefficients. The proposed method is applied to a longitudinal data set from the Amboseli Baboon Research Project to investigate the causal relationships between early adversity, adult physiological stress responses, and survival among wild female baboons. We find that adversity experienced in early life has a significant direct effect on females' life expectancy and survival probability, but find little evidence that these effects were mediated by markers of the stress response in adulthood. We further developed a sensitivity analysis method to assess the impact of potential violation to the key assumption of sequential ignorability.

KEY WORDS: Causal Inference, Functional Principal Component Analysis, Mediation, Functional Data

1 Introduction

A common pursuit in biological studies is to understand mediation, that is, the causal relationships between an exposure or treatment Z, an outcome Y, and an intermediate variable (i.e. mediator) M that lies on the causal path between Z and Y. As a motivating example, consider a study where we want to investigate the causal effect of early life adversity on survival outcomes, and how that effect is mediated through hormonal markers of the stress response in wild adult baboons. The classic mediation analysis method is the Baron-Kenny method, which fits two linear structural equation models (SEMs)—one on Y predicted by Z, M and one on M predicted by Z—and interprets specific model coefficients as causal effects (Baron and Kenny 1986; MacKinnon 2012). Recently there is a surge of research in combining the potential outcome framework for causal inference (Neyman 1923; Rubin 1974) and the Baron-Kenny method (Robins and Greenland 1992; Pearl 2001; Sobel 2008; Imai et al. 2010b; Tchetgen Tchetgen and Shpitser 2012; Daniels et al. 2012; VanderWeele 2016). In particular, Imai et al. (2010b) proved that the Baron-Kenny estimator can be interpreted as a causal mediation estimator given a set of structural assumptions under the potential outcome framework. It has since led to many new methodological advancements and applications to disciplines beyond the traditional domains of SEM, including imaging, neuroscience, and environmental health (Lindquist and Sobel 2011; Lindquist 2012; Zigler et al. 2012; Kim et al. 2017, 2019). Advanced Bayesian modeling for mediation analysis has been also been developed (Daniels et al. 2012; Kim et al. 2017, 2018). Comprehensive reviews on causal mediation analysis are given in VanderWeele (2015) and in Nguyen et al. (2020).

Traditionally in mediation analysis the exposure Z, mediator M and outcome Y are all measured at a single time point. Recent studies increasingly involve time-varying data, where at least one of the triplet (Z, M, Y) is measured repeatedly and the data pattern varies in specific applications. For example, in health studies, subjects' clinical information is often measured in multiple scheduled visits. However, the majority of causal mediation research with time-varying data focuses on regularly observed data (van der Laan and Petersen 2008; Roth and MacKinnon 2012; Lin et al. 2017a; VanderWeele and Tchetgen Tchetgen 2017), and the analysis often utilizes marginal structural models (Robins et al. 2000). Another line of research takes a functional data analysis perspective (Ramsay and Silverman 2005) when the observations are made on a dense grid. For example, motivated by applications in neuroimaging, Lindquist (2012) and Zhao et al. (2018) view densely recorded functional magnetic resonance imaging (fMRI) mediators as functional data, and employed functional models as SEMs.

None of the above methods is directly applicable to irregular longitudinal data. To address the gap in methodology, Zeng et al. (2020) adopted the functional mediation perspective and proposed a general method applicable to both regular and irregular longitudinal mediators and outcomes. Specifically, they view the observed time-varying

mediators or outcomes as realizations of underlying stochastic processes, and impose a functional principal component analysis (FPCA) model (Yao et al. 2005; Jiang and Wang 2010, 2011; Han et al. 2018; Kowal and Bourgeois 2020) to impute the entire process. However, Zeng et al. (2020) focused on continuous outcomes, while survival outcomes are common in real applications. It is challenging to handle longitudinal mediators and survival outcomes simultaneously (Lange et al. 2012; VanderWeele 2011). The existing related literature focuses on regular longitudinal mediators (Zheng and van der Laan 2017; Lin et al. 2017b). There are two main complications in analyzing time-varying mediators with survival outcomes: (i) the mediator value for a given subject in the study is not well defined after the subject dies; and (ii) at each step, both the value of the mediator and the survival outcome depend upon prior survival, which renders prior survival as a post-treatment confounder between mediator and outcome and thus standard identification results no longer hold (Didelez 2019a,b; Vansteelandt et al. 2019). Didelez (2019b) tackled these problems by separating the primary treatment into a treatment on mediator and a treatment on outcome, and provided corresponding identification assumptions.

In this paper, we extend the functional data analysis perspective in Zeng et al. (2020) to accommodate longitudinal mediators on an arbitrary grid with a survival outcome. Viewing the longitudinal mediator observations as functional data provides a principled and flexible way to adjust for the correlation between time-varying mediators and directly model its relationship with the survival outcomes, and thus bypasses the two aforementioned complications. We define relevant causal estimands and provide assumptions for nonparametrically identifying these estimands. For estimation, we proceed under the two-SEM mediation framework (Imai et al. 2010b). Similar to Zeng et al. (2020), we specify a Bayesian functional principal component analysis (FPCA) model (Kowal and Bourgeois 2020) to project the mediator trajectories to a low-dimensional representation and impute the underlying mediator process. We also specify a Cox proportional hazard model for the survival outcome, and derive an analytical formula to express the causal estimands by the model coefficients using g-computation Robins (1986). We apply the proposed method to a prospective and longitudinal observational data set from the Amboseli Baboon Research Project (Alberts and Altmann 2012). We further developed a sensitivity analysis method to assess the impact of potential violations to the key assumption of sequential ignorability.

The remainder of this paper proceeds as follows. Section 2 introduces the scientific premise and data of the motivating application. Section 3 presents the causal mediation analysis framework for the setting of irregular longitudinal mediators and survival outcomes and introduces causal estimands and assumptions necessary to nonparametrically identify the estimands. Section 4 proposes a specific modeling and estimation strategy. Section 5 applies the proposed methods to the baboon study. Section 6 develops a sensitivity analysis. Section 7 concludes.

2 Motivating Application: Early Adversity, Physiological Stress, and Survival

2.1 Biological Background

Experiences during early life and adulthood can have profound effects on adult health and survival. For example, negative socioenvironmental conditions during childhood are linked to dysregulation of the stress response and poor adult survival in humans (Berens et al. 2017; Evans et al. 2013; Felitti et al. 1998; Miller et al. 2009, 2011; Petruccelli et al. 2019). In addition, dysregulation of the stress response in adulthood leading to altered glucocorticoid (GC) hormone profiles is hypothesized to reduce lifespan in humans (Adam et al. 2017; Hertzman 1999; Miller et al. 2011; Schoenle et al. 2021) and is known to do so in wild baboons (Campos et al. In Press). Can we identify the major mediators of early life adversity's effects on adult survival? On the one hand, the effects of early life adversity may be concentrated in one or several relatively simple health indices in adulthood, specifically dysregulation of the stress response (Hertzman 1999; Miller et al. 2011). In this case, we would predict that GC hormone profiles are a major mediator of the link between early adversity and survival. On the other hand, the effects of early adversity may be diffuse and multi-factorial, and/or variation in the adult stress response may have multiple causes, leading to very weak mediation by GC hormone profiles in the link between early adversity and survival. No studies to date have been able to unambiguously link real time data on early life adversity, dysregulation of the stress response in adulthood (via assessment of adult GC profiles), and survival in the same individuals. Therefore, the relative importance of early life adversity versus any independent effects of adult physiology in determining survival remains unclear (Boyce and Hertzman 2018; Harris 2019; Warren 2009).

2.2 Data

In this paper, we investigate the causal mediation relationship between early adversity, GC hormone profiles, and survival. We use data from a well-studied population of savannah baboons in the Amboseli ecosystem in Kenya. Founded in 1971, the Amboseli Baboon Research Project has prospective longitudinal data on early life experiences, and fine-grained longitudinal data on adult fecal GC concentrations (Alberts and Altmann 2012).

Our study sample includes 199 female baboons and 11914 observations in total. Survival was assessed for each female baboon starting at age 4 years, but GC hormone concentrations were measured only for females that had reached menarche (average age at menarche = 4.73 $\pm$ 0.56 years). For each subject, we had information on the experience of six sources of early adversity (i.e., exposure) (Tung et al. 2016; Rosenbaum et al.

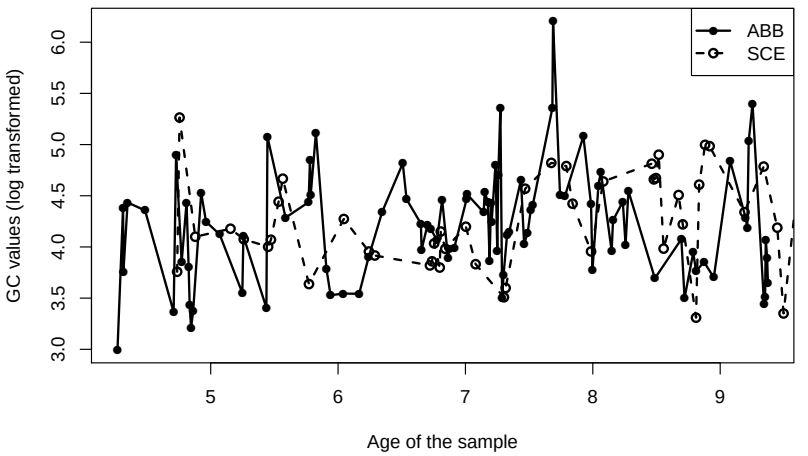


Figure 1: Irregular and sparse GC observations (log transformed) for two randomly selected baboons in the sample.

2020): drought, maternal death, close-in-age younger sibling, high group density, low maternal rank, and maternal social isolation. While only a small proportion of baboons experienced any given source of early adversity, most baboons experienced at least one source of early adversity. In our analysis we also create a binary exposure variable that indicates whether a baboon experienced any source of adversity.

The mediator is each baboon's GC hormone profile across adulthood. These profiles are measured by assessing GC concentrations in fecal samples. For wild baboons, the GC hormone is recorded based on opportunistic collection of fecal samples and is thus measured on an irregular grid. The values of GC range from 7.51 ng/gm to 982.87 ng/gm with mean value at 76.90 ng/gm and standard deviation 39.58 ng/gm. We record the age of the subject at each sample collection as the time index for within-individual observations on GC concentrations. The frequency of observations and time grids of the mediator trajectories vary significantly between baboons: we have on average 59.86 GC observations of each baboon, but the number of observations of a single baboon ranges from 3 to 284. Figure 1 shows the mediator trajectories of two randomly selected baboons (with codenames "ABB" and "SCE") in the sample.

The survival time is measured in years. Figure 2 shows the Kaplan-Meier estimates of the survival function in groups with different number of early adversities. Clearly the baboons experiencing fewer early adversities have better chance of survival. In particular, baboons who experienced two or more early adversities have a sharply decreased survival probability compared with those who had fewer adversities.

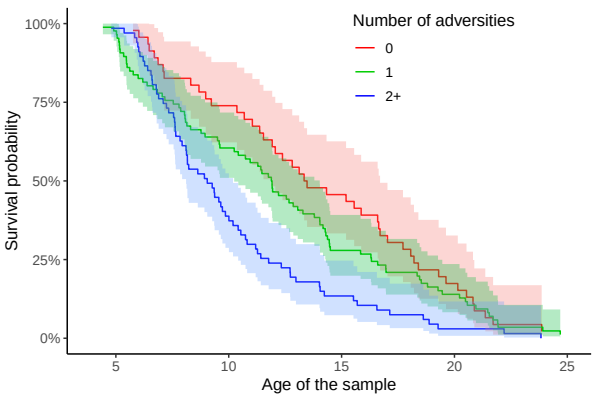


Figure 2: Kaplan-Meier estimates of survival function in groups with different number of early adversities.

The time-varying covariates include reproductive state (i.e. cycling, pregnant, or lactating), density of the social group, group density squared (Markham et al. 2015), max ambient temperature in the 30 days before the fecal sample was collected, whether the sample was collected in wet or dry season, the deviation in rainfall from expected during the three months prior to sample collection, storage time as fecal powder (time between collection of fecal sample and methanol extraction), storage time in methanol (time between methanol extraction and GC measurement), proportional dominance rank, and whether the focal female was top ranked or not. All these covariates are deemed important to wild baboons's physiology and behavior. More information can be found in Rosenbaum et al. (2020) and Levy et al. (2020).

3 Causal Estimands and Identification

3.1 Setup and Causal Estimands

Suppose we have a sample of N subjects; each subject i $(i = 1, 2, \cdots, N)$ is assigned to a treatment $(Z_i = 1)$ or a control $(Z_i = 0)$ group. For each subject i, we make observations at T_i time points $\{t_{ij} \in [0, T], j = 1, 2, \cdots, T_i\}$, and the interval between two consecutive time points can differ within and across subjects. At each time point t_{ij}, we measure a mediator M_{ij}, and a vector of p time-varying covariates $\mathbf{X}_{ij} = (X_{ij,1}, \cdots, X_{ij,p})'$. Let V_i denote the survival time and C_i be the censoring time. The survival time might be right censored when $C_i \leq V_i$ so we observe $\tilde{V}_i = \min(V_i, C_i)$ and the indicator that whether the subject failed within the study period $\delta_i = \mathbf{1}_{V_i \leq C_i}$. In summary, we observe $(Z_i, M_{ij}, \mathbf{X}_{ij}, \tilde{V}_i, \delta_i), j = 1, 2, \cdots T_i$ for each subject i.

We view the observed mediator values drawn from a smooth underlying process $M_i(t)$,

$t \in [0, T]$, with errors drawn from Normal distribution:

$$M_{ij} \quad = \quad M_i(t_{ij}) + \varepsilon_{ij}, \quad \varepsilon_{ij} \sim \mathcal{N}(0, \sigma_m^2). \tag{1}$$

We aim to investigate the relationship between Z_i, the stochastic processes $M_i(t)$, and the survival outcome V_i. In particular, we wish to answer two questions: (a) how big is the causal impact of the treatment on the survival time, and (b) how much of that impact is mediated through the mediator process?

Following the standard notation of potential outcomes in causal inference (Imbens and Rubin 2015), we move the time index of the mediator process to the superscript: $M_i(t) = M_i^t$ from now on. Also, we use bold font to denote a process until time t: $\mathbf{M}_i^t \equiv \{M_i^s, s \leq t\} \in \mathcal{R}^{[0,t]}$. Similarly, we denote covariates between the jth and $j+1$th time point for subject i as $\mathbf{X}_i^t = \{X_{i1}, X_{i2}, \cdots, X_{ij'}\}$ for $t_{ij'} \leq t < t_{ij'+1}$. Further, let $\mathbf{M}_i^t(z) \in \mathcal{R}^{[0,t]}$ for $z = 0, 1$ and $t \in [0, T]$, be the potential values of the unobserved smooth mediator process for subject i until time t under the treatment status z; let $V_i(z, \mathbf{m}) \in \mathcal{R}^{[0,T]}$ be the potential survival time for subject i under the treatment status z and the mediator process taking value of $\mathbf{m} \in \mathcal{R}^{[0,T]}$. The most important potential outcomes are those corresponding to the mediator value being the potential mediator under an intervention z', $\mathbf{V}_i^t(z, \mathbf{M}_i^t(z'))$. For each subject, we can only observe one realization from the potential mediator process and at most one potential survival time if not being censored:

$$\mathbf{M}_i^t = \mathbf{M}_i^t(Z_i) = Z_i \mathbf{M}_i^t(1) + (1 - Z_i)\mathbf{M}_i^t(0), \tag{2}$$

$$V_i = V_i(Z_i, \mathbf{M}_i^T(Z_i)) = Z_i V_i(1, \mathbf{M}_i^T(1)) + (1 - Z_i)V_i(0, \mathbf{M}_i^T(0)). \tag{3}$$

We define the survival function for the potential survival time when a subject's treatment status is z and the mediator process takes the value as if the subject was treated by z', as $S_{z,z'}(t)$,

$$S_{z,z'}(t) = \Pr(V_i(z, \mathbf{M}_i^T(z')) > t), \text{for any } z, z' = 0, 1. \tag{4}$$

When $z \neq z'$, the potential outcome $\mathbf{V}_i^t(z, \mathbf{M}_i^t(z'))$ is called *cross-world counterfactual* (Imai et al. 2010a) because the initial intervention is different from the hypothetical intervention for the mediator. Cross-world counterfactuals are philosophically controversial (Lok 2016; Lok and Bosch 2021) but they are critical in defining causal estimands and have been widely adopted in the causal mediation literature, which we also follow in this paper.

We define the total effect (TE) of the treatment on the expected survival time as:

$$\tau_{\text{TE}}^{h,t} \quad = \quad E[h\{V_i(1, \mathbf{M}_i^T(1)); t\} - h\{V_i(0, \mathbf{M}_i^T(0)); t\}]. \tag{5}$$

where t is a fixed time point, and $h(\cdot; t)$ is a function that transforms the survival outcome and thus defines causal estimands on different scales. For example, when $h(x; t) = x \wedge t$

(i.e. the truncation function), τ^t_{TE} compares the restricted mean survival time. If we let $t \to \infty$, τ^t_{TE} reduces to the standard average treatment effect (ATE) that compares the expected difference. When $h(x;t) = 1_{\{x>t\}}$ (i.e. the at-risk function), τ^t_{TE} becomes the comparison on survival probability. TE can be decomposed into direct and indirect effects (Robins and Greenland 1992; Pearl 2001; Imai et al. 2010a). Specifically, we define the average causal mediation (or indirect) effect (ACME) and the average natural direct effect (ANDE): for $z = 0, 1$

$$\begin{aligned} \tau^{h,t}_{\text{ACME}}(z) &\equiv E[h\{V_i(1, \mathbf{M}_i^T(z))\} - h\{V_i(0, \mathbf{M}_i^T(z))\}], \quad (6) \\ \tau^{h,t}_{\text{ANDE}}(z) &\equiv E[h\{V_i(z, \mathbf{M}_i^T(1))\} - h\{V_i(z, \mathbf{M}_i^T(0))\}]. \quad (7) \end{aligned}$$

ACME and ANDE quantifies the portion in the TE that goes through and bypasses the mediators, respectively. ACME is also referred as the *natural indirect effect* (Pearl 2001), or the *pure indirect effect* for $\tau^{h,t}_{\text{ACME}}(0)$ and *total indirect effect* for $\tau^{h,t}_{\text{ACME}}(1)$ (Robins and Greenland 1992). It is easy to verify that TE is the sum of ACME and ANDE:

$$\tau^{h,t}_{\text{TE}} = \tau^{h,t}_{\text{ACME}}(z) + \tau^{h,t}_{\text{ANDE}}(1-z), \quad z = 0, 1. \quad (8)$$

Therefore, we only need to identify two of the three quantities $\tau^{h,t}_{\text{TE}}$, $\tau^{h,t}_{\text{ACME}}(z)$, $\tau^{h,t}_{\text{ANDE}}(z)$. In this paper, we estimate $\tau^{h,t}_{\text{TE}}$ and $\tau^{h,t}_{\text{ACME}}(z)$, which can be expressed as functions of the survival function, $S_{z,z'}(t)$. Specifically, with the at-risk function $h(x;t) = 1_{\{x>t\}}$, we have

$$\tau^{h,t}_{\text{TE}} = \int_0^t \{S_{1,1}(u) - S_{0,0}(u)\}\mathrm{d}u, \quad \tau^{h,t}_{\text{ACME}}(z) = \int_0^t \{S_{z,1}(u) - S_{z,0}(u)\}\mathrm{d}u,$$

and with the truncation function $h(x;t) = x \wedge t$, we have

$$\tau^{h,t}_{\text{TE}} = S_{1,1}(t) - S_{0,0}(t), \quad \tau^{h,t}_{\text{ACME}}(z) = S_{z,1}(t) - S_{z,0}(t).$$

Further, for simplicity we only consider the estimands with $h = x \wedge t$ and $t = \infty$, which contrasts the expected potential survival time. Alternative estimands such as difference in restricted mean or survival probability (VanderWeele 2011) can be derived in a similar manner within our framework.

3.2 Identification assumptions

Because we only observe a portion of all the potential outcomes, we need additional assumptions to identify causal estimands from the observed data. Below we present a set of assumptions that are sufficient for nonparametrically identifying ACME and ANDE .

The first assumption extends the standard ignorability (or unconfoundedness) assumption and rules out the unmeasured treatment-outcome confounding.

Assumption 1 (Ignorability)**.** *Conditional on the observed covariates, the treatment is unconfounded with respect to the potential mediator process and the potential survival time:*

$$\{V_i(1,\mathbf{m}), V_i(0,\mathbf{m}), \mathbf{M}_i^t(1), \mathbf{M}_i^t(0)\} \perp\!\!\!\perp Z_i \mid \mathbf{X}_i^t,$$

for any t *and* $\mathbf{m} \in \mathcal{R}^{[0,t]}$.

In our application, Assumption 1 indicates that there is no unmeasured confounding, conditioning on the observed covariates, between the early adversity, the process of adult physiological stress response, and survival. Equivalently, early adversity can be viewed as randomized among the baboons with similar covariates values. This assumption is likely to hold in our application because the early adversity events for the wild baboons were largely determined by nature.

The second assumption generalizes the sequential ignorability assumption in (Imai et al. 2010b; Forastiere et al. 2018) to the functional data setting.

Assumption 2 (Sequential Ignorability)**.** *There exists* $\varepsilon > 0$*, such that for any* $0 < \Delta < \varepsilon$*, the increment of the mediator process from time* t *to* $t + \Delta$ *is independent of the potential survival time conditional on the observed treatment status, covariates and the mediator process up to time* t*:*

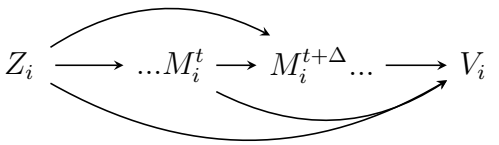
$$V_i(z,\mathbf{m}) \perp\!\!\!\perp (M_i^{t+\Delta} - M_i^t) \mid \{Z_i, \mathbf{X}_i^t, \mathbf{M}_i^t\},$$

for any $z, 0 < \Delta < \varepsilon, t, t+\Delta \in [0,T], \mathbf{m} \in \mathcal{R}^{[0,T]}$.

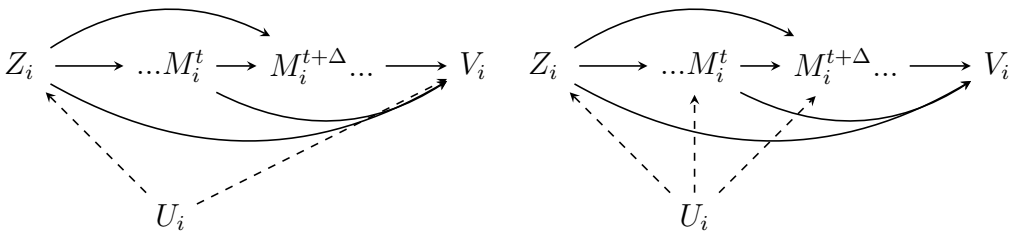
Assumption 2 implies that given the early adversity status, covariates, and the physiological stress history up to a given time point, change in the physiological stress within a sufficiently small time interval Δ is randomized with respect to potential survival time. Namely, we assume there are no unobserved mediator-outcome confounders in a sufficiently small time interval. Though taking a different form, Assumption 2 is essentially the same sequential ignorability assumption we make with the regularly spaced observations, as in Tchetgen (2011); Didelez (2019b); Vansteelandt et al. (2019), but it differs from the sequential ignorability assumption for a time-varying continuous outcome in Zeng et al. (2020). Assumption 2 is fundamental for mediation analysis, yet it is generally untestable even in randomized trials because it involves cross-world counterfactuals. Therefore, it is crucial to conduct sensitivity analysis to assess the impact of potential violations to this assumption, as we did in Section 6.

Assumptions 1 and 2 are demonstrated by the directed acyclic graphs (DAG) in Figure 3a, which implicitly condition on the covariates $\mathbf{X}_i^t$ and a window between sufficiently close time points t and $t+\Delta$. The arrows between Z_i, M_i^t, Y_i^t represent a causal relationship, with solid and dashed lines standing for the measured and unmeasured relationships,

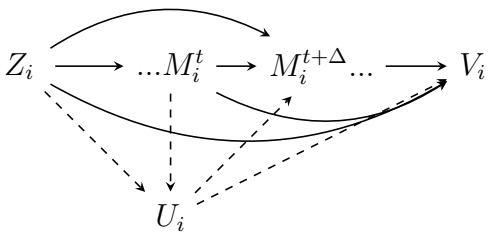
respectively. Figure 3b and 3c depicts possible cases where Assumptions 1 and 2 are violated, respectively, where U_i stands for an unmeasured confounder.



(a) DAG of Assumption 1 and 2



(b) DAG of two examples of violation to Assumption 1 (ignorability)



(c) DAG of examples of violation to Assumption 2 (sequential ignorability)

Figure 3: Directed acyclic graphs (DAG) of Assumptions 1 and 2, and potential scenarios of violation. U_i is an unmeasured confounder. The arrows between variables represent a causal relationship, with solid and dashed lines representing measured and unmeasured relationships, respectively.

The third assumption imposes independent censoring mechanism, which allows us to identify the distribution of survival time from censored data.

Assumption 3 (Independent censoring)**.** *The censoring time is independent of all remaining variables, including covariates, treatment, mediators and outcome:*

$$C_i \perp\!\!\!\perp \{X_i, Z_i, \boldsymbol{M}_i^t(z), V_i(z, \mathbf{m})\},$$

for any $z, t \in [0, T], \mathbf{m} \in \mathcal{R}^{[0,T]}$.

In our application, the time for a wild baboon to exit the study is largely random and therefore this assumption is deemed reasonable. This assumption can be readily extended to covariate-dependent censoring.

Under Assumptions 1,2 and 3, we can identify the causal estimands, which is equivalent to identify the survival function $S_{z,z'}(t)$, nonparametrically from the observed data. This result is summarized in the following theorem.

Theorem 1. *Let $t_1 < t_2 <, \cdots, < t_k <, \cdots, < t_K < \cdots$ be the time grid where we observe event ($\delta_i = 1$) and consider a fixed time point t that $t_K \leq t \leq t_{K+1}$. Under Assumptions 1, 2 and 3, and some regularity conditions (specified in the Supplementary Material), the TE, ACME and ANDE can be identified nonparametrically from the observed data: for $z, z' = 0, 1$, we have*

$$\begin{aligned} S_{z,z'}(t) &= \int\int \Pr(V_i > t | \mathbf{M}_i^t = \mathbf{m}, \mathbf{X}_i^t = \mathbf{x}^t, Z_i = z) \mathrm{dF}_{\mathbf{M}_i^t | Z_i = z', \mathbf{X}_i^t = \mathbf{x}^t}(\mathbf{m}) \mathrm{dF}_{\mathbf{X}_i^t}(\mathbf{x}^t) \\ &= \int\int \{ \prod_{k=1}^{K} \Pr(\tilde{V}_i > t_k | \tilde{V}_i > t_{k-1}, \mathbf{M}_i^{t_k} = \mathbf{m}^{t_k}, \mathbf{X}_i^{t_k} = \mathbf{x}^{t_k}, Z_i = z) \} \times \\ &\quad \mathrm{dF}_{\mathbf{M}_i^t | Z_i = z', \mathbf{X}_i^t = \mathbf{x}^t}(\mathbf{m}) \mathrm{dF}_{\mathbf{X}_i^t}(\mathbf{x}^t), \end{aligned}$$

where $F_W(\cdot)$ and $F_{W|U}(\cdot)$ denotes the cumulative distribution of a random variable or a vector W and the conditional distribution given another random variable or vector U, respectively.

We provide the proof of Theorem 1 in the Supplementary Material. Theorem 1 indicates that estimating the causal effects requires specifying two models: (a) the conditional survival probability given the treatment, covariates, and the observed mediator process, $\Pr(V_i^t > t | Z_i, \mathbf{X}_i^t, \mathbf{M}_i^t)$, and (b) the conditional distribution of the observed mediator process given the treatment and covariates, $\mathrm{F}_{\mathbf{M}_i^t | Z_i, \mathbf{X}_i^t}(\cdot)$. These two models are in parallel to the two linear SEMs in the Baron-Kenny framework. In the next section, we specify these two models and express the TE and ACME in terms of the model parameters.

4 Modeling mediators and survival outcome

4.1 Model for the mediators

For the mediator process, we follow Zeng et al. (2020) to employ a functional principal component analysis (FPCA) approach to impute the entire mediator process from sparse and irregular longitudinal data (Yao et al. 2005; Jiang and Wang 2010, 2011). In particular, we employ a Bayesian FPCA model similar to Kowal and Bourgeois (2020) to account for the uncertainty due to estimating the functional principal components (Goldsmith et al. 2013). The mediator model is the same as that proposed in Zeng et al. (2020)

and thus we only present the main model form and refer the readers to Zeng et al. (2020) for details. Our approach bypasses the conceptual challenges for mediation analysis with survival outcomes and time-varying mediators, without decomposing the treatment into different components or altering the definition of mediation effect, as in Didelez (2019b). First, by treating the mediator process as a whole from the functional data perspective, the mediator is defined on the same domain for all subjects. The mediator value is well-defined even at the time point after the subject fails. Second, we can separate out the survival status and mediator process by viewing the mediator as a function or a process for each subject. By doing so, we ensure that prior survival is no longer a post-treatment confounder that affects both the mediator and the survival in the future.

We assume the potential processes for mediators $\mathbf{M}_i^t(z)$ have the following Karhunen-Loeve decomposition,

$$M_i^t(z) = \mu_M(\mathbf{X}_i^t) + \sum_{r=1}^{\infty} \zeta_{i,z}^r \psi_r(t), \tag{9}$$

where $\mu_M(\cdot)$ are the mean functions of the mediator process $\mathbf{M}_i^t$; $\psi_r(t)$ are the Normal orthogonal eigenfunctions for $\mathbf{M}_i^t$, and $\zeta_{i,z}^r$ are the corresponding principal scores of subject i. The above model assumes that the treatment affects the mediation processes only through the principal scores. We represent the mediator process of each subject with its principal score $\zeta_{i,z}^r$. Given the principal scores, we can transform back to the smooth process with a linear combination.

The underlying process $\mathbf{M}_i^t$ is not observed. We assume the observations M_{ij}'s are randomly sampled from the corresponding underlying processes with errors. For the mediator trajectories, we impose the following model truncating to the first R principal components of the mediator process:

$$M_{ij} = X_{ij}'\beta_M + \sum_{r=1}^{R} \zeta_i^r \psi_r(t_{ij}) + \varepsilon_{ij}, \quad \varepsilon_{ij} \sim \mathcal{N}(0, \sigma_m^2), \tag{10}$$

where $\psi_r(t)$ $(r = 1, ..., R)$ are the orthogonormal principal components, ζ_i^r $(r = 1, ..., R)$ are the corresponding principal scores, and ε_{ij} is the measurement error. We follow the same parametrization and prior distributions used in Kowal and Bourgeois (2020). More details can be found in the Supplementary Material and in Zeng et al. (2020).

We select the minimal truncation term R which renders the fraction of explained variance (FEV), $\sum_{r=1}^{R} \lambda_r^2 / \sum_{r=1}^{\infty} \lambda_r^2$ being greater than 90%. We usually require only 3 or 4 components to explain most of the variation.

4.2 Model for the survival outcome

We posit the following Cox proportional hazards model for the survival time,

$$\lambda(t_{ij}|X_{ij}, Z_i, \mathbf{M}_i^t) = \lambda_0(t_{ij}) \exp\{\alpha Z_i + X_{ij}'\beta_S + f(\mathbf{M}_i^t; \gamma)\}, \tag{11}$$

where $\lambda_0(t_{ij})$ is the baseline hazard rate, and $f(\mathbf{M}_i^t;\gamma)$ is a function of the mediators with parameter γ, which captures the impact of the mediator process on the hazard rate. Specification of $f(\mathbf{M}_i^t;\gamma)$ is crucial, and in this paper we consider two specifications of f:

(i) a *concurrent model* that assumes the hazard rate depends on the instantaneous mediator value, $f(\mathbf{M}_i^t;\gamma) = \gamma M_i(t)$;

(ii) a *cumulative* model that assumes the hazard rate depends the entire mediator process until to time t, $f(\mathbf{M}_i^t;\gamma) = \int_0^t \gamma(s) M_i(s) ds$.

We can express the causal estimands, such as the TE and ACME, as functions of parameters of the mediator model (10) and the survival outcome model (11). First, we express $S_{z,z'}(t)$ via the g-formula as,

$$S_{z,z'}(t) = \exp\{-\Lambda_{z,z'}(t)\},$$
$$\Lambda_{z,z'}(t) = \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{T_i} \lambda_0(t_{ij}) \exp\{\alpha z + X_{ij}'\beta_S + f(X_{ij}'\beta_M + \sum_{r=1}^{R} \chi_{z'}^r \psi_r(s);\gamma)\}(t_{ij} - t_{ij-1}),$$

where $\Lambda_{z,z'}(t)$ denotes the cumulative hazard for $V_i(z, \mathbf{M}_i^T(z'))$. Next, we can calculate τ_{TE} and τ_{ACME} based on $S_{z,z'}(t)$ with the equations in Theorem 1.

We impose a Gamma process prior for the baseline hazard rate $\lambda_0(t)$ (Fahrmeir and Lang 2001; Ibrahim et al. 2014; Wang et al. 2013) and standard normal prior distributions for other coefficients. For the cumulative model, we parameterize the function $\gamma(s)$ as a linear combination of the spline basis $\mathbf{b}(t) = (1, t, b_1(t), \cdots, b_L(t))'$ (Kowal and Bourgeois 2020). Specifically,

$$\gamma(t) = \mathbf{b}(t)'\mathbf{p},$$

where $\mathbf{p}$ is the coefficients with Normal prior, which enables a flexible modeling of how the past mediator history affects the survival outcome.

We perform posterior inference via Gibbs sampling. The credible intervals of the causal effects τ_{TE} and τ_{ACME} can be obtained from the posterior sample of the parameters in the model. We provide the details of the Gibbs sampler in the Supplementary Material.

5 Application to the Amboseli Baboon Research Project

We apply the proposed method to the data described in Section 2.2 to investigate the causal relationship between early adversity, adult stress response, and survival in wild baboons. We perform a separate causal mediation analysis for each source of early

adversity. We posit model (10) for the GC concentrations and Model (11) for the survival outcome. In both models, we added two random effects, one for social group and one for hydrological year. In the mediator model, we use the log transformed GC concentrations instead of the original scale, which allows us to interpret the coefficient as the percent difference in GC concentrations between the adversity and non-adversity groups.

Here we first summarize the results of FPCA of the mediator trajectories, of which the first three functional principal components explain more than 90% of the total variation. Figure 4 shows the first two principal components extracted from the mediator process, which explain 59% and 38% of the total variation, respectively. The first component depicts a relatively stable trend throughout the life span. The second component shows a quick rise until age 6, then steady drop pattern across the lifespan.

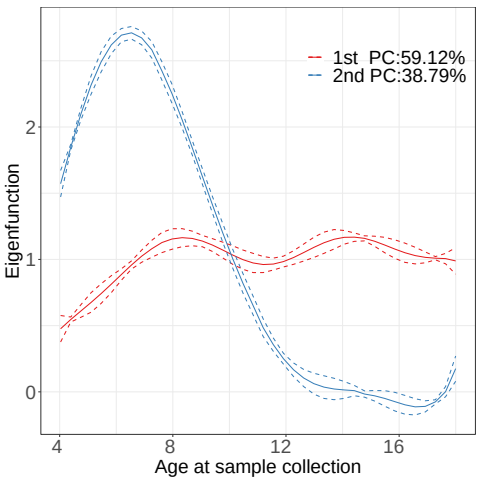


Figure 4: The first two functional principal components of the mediator process, i.e., GC concentrations.

The left panel of Figure 5 displays the observed trajectory of GCs versus the posterior mean of the imputed smooth process of three randomly chosen baboons who experienced zero (EPI), one (ELA), and two (RWA) sources of early adversity, respectively. We can see that the imputed smooth process generally captures the overall time trend of each subject while reducing the noise in the observations. Recall that each subject's observed trajectory is fully captured by its vector of principal scores, and thus the principal scores of the first few dominant principal components adequately represent the whole trajectory. The right panel of Figure 5 shows the principal scores of the first (X-axis) versus second (Y-axis) principal component for the mediator process of all subjects in the sample, color-coded based on the number of early adversities experienced. We can see that significant differences exist in the distributions of the first two principal scores between the group who experienced no early adversity and the group experienced exactly one or the group

with more than one sources of adversity.

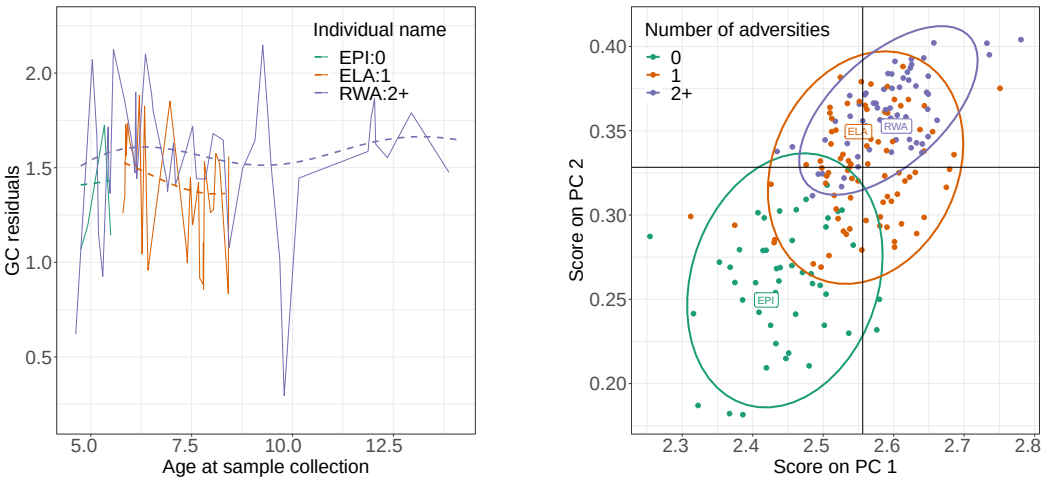


Figure 5: Left panel: Examples of observed trajectory of GCs versus the posterior mean of its imputed smooth process of three baboons who experienced zero (EPI), one (ELA) and two (RWA) sources of early adversity, respectively. Right panel: Principal scores of the first (X-axis) versus second (Y-axis) principal component for the GC process of all subjects in the sample, including the three example subjects that are illustrated in the left panel (with individual names labeling the corresponding points). Color-coding is based on the number of early adversities experienced.

We now summarize the results on the causal estimates. Table 1 presents the posterior mean and 95% credible interval of the total effect (TE), direct effect (ANDE) and indirect effect mediated through social bonds (ACME) of each source of early adversity on life expectancy, as well as the effects of early adversity on the mediator. First, from the first column of Table 1 we can see that experiencing any source of early adversity would increase the GC concentrations in adulthood, which is detrimental to the health of the baboon. The effect is particularly severe for those who experienced drought, high group density, maternal death or low maternal rank in early life. For example, compared with the baboons who did not experience any early adversity, the baboons who experienced drought in the first year of life have 9.7% increase in GC response. Overall, experiencing at least one source of early adversity corresponds to GC concentrations that are 9.4% higher in adulthood.

Second, from the second column of Table 1 we can see a strong negative total effect of early adversity on the life expectancy of female baboons. Baboons who experienced at least one source of early adversity had a life expectancy approximately 1.5 years shorter than their peers who experienced no early adversity. The range of total effect sizes across all individual adversity sources varies from 0.691 to 2.199 years life reduction and

the point estimates are consistently toward a shorter survival time, even for the early adversity sources for which the credible interval includes zero. Among the individual sources of adversity, females who were born during a drought or experienced maternal death experienced a particularly drastic drop in life expectancy, with effect sizes of 1.795 and 2.199 years respectively.

Table 1: Total, direct and indirect causal effects of individual and cumulative sources of early adversity on life expectancy in adulthood in wild female baboons (measured in year). 95% credible intervals are in the parenthesis.

Source of adversity	effect on mediator (%)	τ_{TE}	τ_{ANDE}	τ_{ACME}
Drought	9.7%	-1.795	-1.596	-0.199
	(1.5%,18.0%)	(-3.300,-0.291)	(-2.852,-0.341)	(-0.594,0.197)
Competing sibling	6.9%	-0.994	-0.886	-0.108
	(1.4%,12.5%)	(-4.038,2.049)	(-2.894,1.121)	(-0.210,-0.006)
High group density	11.9%	-0.691	-0.449	-0.242
	(2.8%,21.0%)	(-3.122,1.740)	(-2.512,1.614)	(-0.460,-0.024)
Maternal death	9.7%	-2.199	-1.972	-0.227
	(1.5%,17.9%)	(-3.856,-0.543)	(-3.527,-0.418)	(-0.466,0.013)
Maternal social isolation	8.0%	-0.692	-0.572	-0.120
	(1.7%,14.4%)	(-3.188,1.805)	(-2.763,1.618)	(-0.255,0.016)
Low maternal rank	11.5%	-1.392	-1.046	-0.346
	(2.6%,20.4%)	(-3.991,1.207)	(-3.201,1.108)	(-0.728,0.036)
At least one	9.4%	-1.494	-1.292	-0.202
	(1.8%,17.0%)	(-2.748,-0.239)	(-2.264,-0.320)	(-0.551,0.147)

Third, while female baboons who experienced harsh conditions in early life have a lower life expectancy, we found no strong evidence that these effects were mediated by GC hormone profiles. Specifically, the mediation effect τ_{ACME} (the fifth column in Table 1) is relatively small; the increase in adult GC concentrations accounted for a reduction in life expectancy of 0.202 years, when comparing the baboons who experienced at least one early adversity to those did not, with a credible interval including zero. In terms of individual early adversity sources, only two out of six individual adversity sources have a negative mediation effect with credible intervals not including zero, and both effects are quite small. On the other hand, the direct effects τ_{ANDE} (the third column in Table 1) are much larger than the mediation effects. When comparing the baboons with or without experiencing any source of early adversity, the direct effect of early adversity on life expectancy was 6.4 times stronger than the mediation effect running through adult physiological stress response. Specifically, for females who experienced at least one source of early adversity, the direct effect accounts of 1.292 years reduction in life expectancy while the mediation effect through GC accounts for only 0.202 years drop in the average survival time.

6 Sensitivity Analysis

The sequential ignorability assumption (Assumption 2), which rules out unmeasured mediator-outcome confounding, is fundamental to our analysis, but it is generally untestable from the observed data. So we develop a sensitivity analysis method to assess the impact of potential violation to sequential ignorability. Given the complex structure of mediation analysis, we adopt a model-based approach with the unmeasured confounders as the augmented variables, along the lines in Imai et al. (2010a); Huang et al. (2020). Specifically, we introduce an unmeasured confounder U_i to characterize the correlation between the mediator process and the survival outcome that is not captured by covariates X_{ij}. Without loss of generality, we posit U_i to be binary. We expanded the mediator model (10) and the outcome model (11) to accommodate U as follows:

$$M_{ij} = M_i(t_{ij}) + \varepsilon_{ij} = \beta_M^T X_{ij} + \sum_{r=1}^{R} \psi_r(t)\{\tau_0^r(1 - Z_i) + \tau_1^r Z_i\} + \zeta_M U_i + \varepsilon_{ij}, \tag{12}$$

$$\lambda(t|X_{ij}, Z_i, \mathbf{M}_i^t) = \lambda_0(t) \exp\{\alpha Z_i + X_{ij}'\beta_S + f(\mathbf{M}_i^t; \gamma) + \zeta_S U_i\}, \tag{13}$$

where ζ_M and ζ_S are the pre-specified sensitivity parameters that measure the correlation between the unmeasured confounder and mediator process and the survival outcome, respectively. When sequential ignorability holds, there is no unmeasured confounder that simultaneously correlates with the mediator process and survival outcome, and thus $\zeta_M\zeta_S = 0$. When both ζ_M and ζ_S are non-zero, sequential ignorability is violated. Therefore, we use (ζ_M, ζ_S) as the sensitivity parameters to measured the degree of violation to Assumption 2.

Our sensitivity analysis consists of the following steps. First, we choose a grid of values of the sensitivity parameters (ζ_M, ζ_S). For example, we choose $(\zeta_M, \zeta_S) \in \{0, 0.1, 0.2, 0.5, 1\} \times \{0, 0.1, 0.5, 1\}$ in our application. Second, with each fixed pair of (ζ_M, ζ_S), we fit the models (12) and (13). Compared with the original models, (10) and (11), here we need to have an additional step of simulating the unmeasured confounder U_i given the observed data $(Z_i, X_{ij}, \delta_i, \tilde{T}_i)$, (ζ_M, ζ_S) and the other model parameters. Next, we estimate the mediation effect τ^{ACME} from the posterior sample following the same procedure in Section 4. We repeat the above steps with the all possible combinations of (ζ_M, ζ_S) on the pre-specified grid and examine how variable the estimates of τ^{ACME} are to the values of (ζ_M, ζ_S), which reflects how sensitive the causal estimates are to the violation of Assumption 2.

Figure 6 summarizes the results of the sensitivity analysis under the aforementioned specified grid of (ζ_M, ζ_S) in our application. First of all, we notice that the point estimate of τ^{ACME} becomes close to zero as ζ_M or ζ_S increases, and the effect size of τ^{ACME} becomes negligible when $\zeta_S \geq 0.5$ and $\zeta_M \geq 0.1$. Also, the credible interval becomes wider when either one of the sensitivity parameters (ζ_M, ζ_S) becomes larger. These patterns indicate

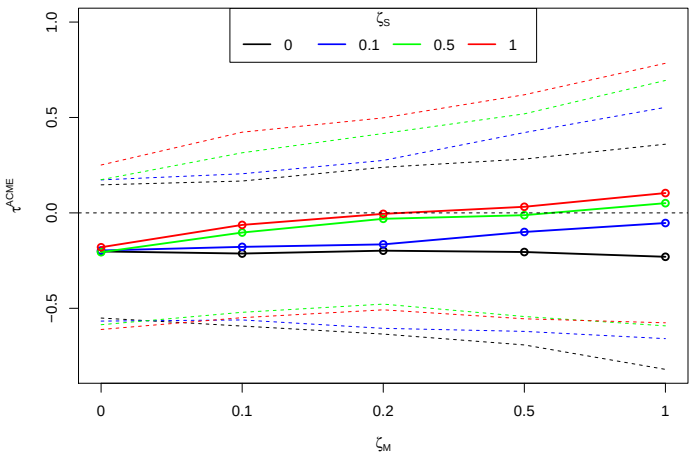


Figure 6: Sensitivity analysis with a grid of (ζ_S, ζ_M). Each value of a fixed ζ_S is coded by a different color; for a given ζ_S, the point estimate and corresponding 95% credible interval of τ^{ACME} as a function of ζ_M is presented by the solid and dashed line, respectively.

our estimation of the mediation effect is sensitive to the sequential ignorability assumption. Recall that the our analysis under Assumption 2 found only a small mediation effect. This sensitivity analysis further suggests that there is no strong evidence supporting that the adult physiological stress response mediates the effect between early adversity and survival.

7 Discussion

We proposed a method for causal mediation analysis with a longitudinal mediator on an arbitrary time grid and a survival outcome. The main idea is to view the time-varying mediator values as realizations from an underlying smooth process and use functional principal component analysis to impute the entire process, which is then used in the structural equation models. This approach naturally bypasses several conceptual and technical challenges in such settings. We defined several causal estimands in such settings and specified structural assumptions to nonparametrically identify these effects. We applied the proposed method to analyze the causal effects of early adversity on adult physiological stress responses and survival in wild female baboons. We found that experiencing adversity early in life significantly increases a baboon's GC response throughout its adulthood and decreases its survival probability. However, we found little evidence that the effect of early adversity on survival is mediated through the chronic elevation

in the GCs, which is linked to poor health and survival in many species (Schoenle et al. 2021). Our results suggest that early adversity and GC in adulthood have independent effects on survival and raise interesting questions in evolutionary biology about alternative causal pathways between early adversity and survival

We developed a model-based method to conduct sensitivity analysis regarding the key assumption of sequential ignorability. Given the complex structure of mediation analysis, related sensitivity analysis usually involves strong and sometimes overly simplified assumptions. For example, our sensitivity analysis depends on the correct specification of the mediator model and the outcome model, while misspecification is common in real applications. Also, to simplify the analysis we assume that the correlation structure between mediator and outcome is constant across time. Nevertheless, even a simplified sensitivity analysis still provides useful insights to causal mediation analysis; in particular it prevents over-interpreting the results and calls for more rigorous investigation of the causal assumptions. We notice that though sensitivity analysis has been standard in causal inference, it has not been routinely performed in causal mediation analysis. We believe more research on interpretable and flexible sensitivity analysis method would help the applied audience to employ causal mediation analysis.

Though motivated by a specific application, the proposed method is readily applicable to other causal mediation studies with similar data structure. For example, comparative effectiveness studies increasingly use electronic health records (EHR) data, where the number of observations usually varies greatly between patients and the time grids are uneven. Moreover, many longitudinal studies in ecology rely on opportunistic sampling of their subjects, resulting in irregularly-spaced observations.

Acknowledgements

The majority of the data represented here was supported by the National Institutes of Health and the National Science Foundation, currently through NIH R01 AG053330 and R01 AG053308, as well as R01 HD088558, P01 AG031719, and NSF IOS 1456832. We also thank Duke University, Princeton University, and the University of Notre Dame for financial and logistical support. For assistance and cooperation in Kenya, we are grateful to the Kenya Wildlife Service (KWS), University of Nairobi, Institute of Primate Research (IPR), National Museums of Kenya, National Environment Management Authority, and National Commission for Science, Technology, and Innovation (NACOSTI). We also thank the members of the Amboseli-Longido pastoralist communities, and the Enduimet Wildlife Management Area for their cooperation and assistance in the field. Particular thanks go to the Amboseli Baboon Project long-term field team (R.S. Mututua, S. Sayialel, J.K. Warutere, Siodi, I.L.), and to T. Wango and V. Oudu for their untiring assistance in Nairobi. The baboon project database, Babase, is expertly managed by N.

Learn and J. Gordon. Database design and programming are provided by K. Pinc. This research was approved by the IACUC at Duke University, University of Notre Dame, and Princeton University and adhered to all the laws and guidelines of Kenya. For a complete set of acknowledgments of funding sources, logistical assistance, and data collection and management, please visit http://amboselibaboons.nd.edu/acknowledgements/.

Appendix

A.1 Proof for Theorem 1

We first provide the required regularity assumptions. (i) Suppose the potential survival time $V_i(z, \mathbf{m})$ as a function of the mediator process $\mathbf{m}$ is Lipschitz continuous on $[0, T]$ with probability one. Namely, there exists a constant $A < \infty$ such that $|V_i(z, \mathbf{m}) - V_i(z, \mathbf{m}')| \leq A||m - m'||_2$ for any $z, \mathbf{m}, \mathbf{m}'$ almost surely. (ii) Any path of $\mathbf{m}$ we consider is Lipschitz continuous. There exists a constant B, such that $|m(t_1) - m(t_2)| \leq B|t_1 - t_2|$ for any $t_1, t_2 \in [0, T]$.

Fix a time point t and suppose the domain for the covariates is $\mathcal{X}$, with $\mathbf{X}_i^t \in \mathcal{X}$. For any $z, z' \in \{0, 1\}$, we have

$$\begin{aligned}
&\int_{\mathcal{X}} \int_{R^{[0,t]}} \Pr(V_i > t | Z_i = z, \mathbf{X}_i^t = x_i^t, \mathbf{M}_i^t = \mathbf{m}) dF_{\mathbf{M}_i^t | Z_i = z, \mathbf{X}_i^t = x^t}(\mathbf{m}) dF_{\mathbf{X}_i^t}(x^t) \\
= &\int_{\mathcal{X}} \int_{R^{[0,t]}} \Pr(V_i(z, \mathbf{m}) > t | Z_i = z, \mathbf{X}_i^t = x_i^t, \mathbf{M}_i^t = \mathbf{m}) dF_{\mathbf{M}_i^t | Z_i = z, \mathbf{X}_i^t = x^t}(\mathbf{m}) dF_{\mathbf{X}_i^t}(x^t)
\end{aligned}$$

For any path $\mathbf{m}$ on the $[0, t]$, we make equal partitions into H pieces at points $\mathcal{M}_H = \{t_0 = 0, t_1 = t/H, t_2 = 2t/H, \cdots, t_H = t\}$ and corresponding values on path $\mathbf{m}$ are $\{m_0, m_1, \cdots, m_H\}$. Then, we consider using a step function from $[0, t] \to \mathcal{R}$ with jumps at points $\mathcal{M}_H$. Denote the step function as $\mathbf{m}_H$, which is:

$$\mathbf{m}_H(x) = \begin{cases} \mathbf{m}(0) = m_0 & 0 \leq x < t/H, \\ \mathbf{m}(t/H) = m_1 & t/H \leq x < 2t/H, \\ \cdots & \\ \mathbf{m}((H-1)t/H) = m_H & (H-1)t/H \leq x \leq t. \end{cases}$$

We employ this step function $\mathbf{m}_H(x)$ to approximate function $\mathbf{m}$. First, given $\mathbf{m}$ is Lipschitz continuous, there exists $B > 0$ such that $|m(x_1) - m(x_2)| \leq B|x_1 - x_2|$. Therefore, the step function $\mathbf{m}_H$ can approximate the original function $\mathbf{m}$ well as H goes up,

$$||\mathbf{m}_H - \mathbf{m}||_2 \leq \sum_{i=1}^{H} \frac{t}{H} B^2 \frac{t^2}{H^2} \asymp O(H^{-2}).$$

Therefore, we can approximate the survival probability given a continuous mediator process with the mediator values on the jumps, $(m_0, m_1, \cdots, m_H)$. That is,

$$\int_{R^{[0,t]}} \Pr(V_i(z,\mathbf{m}) > t | Z_i = z, \mathbf{X}_i^t = \mathbf{x}^t, \mathbf{M}_i^t = \mathbf{m}) \times \mathrm{d}\{\mathrm{F}_{\mathbf{M}_i^t|Z_i=z', \mathbf{X}_i^t=\mathbf{x}^t}(\mathbf{m})\}$$
$$\asymp \int_{R^{[0,t]}} \Pr(V_i(z,\mathbf{m}_H) > t | Z_i = z, \mathbf{X}_i^t = \mathbf{x}^t, \mathbf{M}_i^t = \mathbf{m}_H) \times \mathrm{d}\{\mathrm{F}_{\mathbf{M}_i^t|Z_i=z', \mathbf{X}_i^t=\mathbf{x}^t}(\mathbf{m}_H)\} + O(H^{-2}).$$

This step applies the regularity condition that the potential survival time $V_i^t(z,\mathbf{m})$ as a function of $\mathbf{m}$ is continuous with the L_2 metrics of $\mathbf{m}$. As the values of steps function $\mathbf{m}_H$ are completely determined by the values on finite jumps, we can further reduce the conditional survival probability to,

$$\asymp \int_{R^H} E(Y_i^t(z,\mathbf{m}_H) | Z_i = z, \mathbf{X}_i^t = \mathbf{x}^t, m_0, m_1, m_2, \cdots m_H)$$
$$\times \mathrm{d}\{\mathrm{F}_{m_0,m_1,\cdots,m_H|Z_i=z', \mathbf{X}_i^t=\mathbf{x}^t}(m_0, m_1, m_2, \cdots m_H)\} + O(H^{-2}).$$

Under Assumption 1, we have,

$$\begin{aligned}
&\mathrm{d}\{\mathrm{F}_{m_0,m_1,\cdots,m_H|Z_i=z', \mathbf{X}_i^t=\mathbf{x}^t}(m_0, m_1, m_2, \cdots m_H)\} \\
&= \mathrm{d}\{\mathrm{F}_{m_0(z'),m_1(z'),\cdots,m_H(z')|\mathbf{X}_i^t=\mathbf{x}^t}(m_0, m_1, m_2, \cdots m_H)\}, \\
&= \mathrm{d}\{\mathrm{F}_{\mathbf{m}_H(z')|\mathbf{X}_i^t=\mathbf{x}^t}(\mathbf{m}_H)\}.
\end{aligned}$$

With a slightly abuse of notations, let $\mathbf{m}_H(z)$ denote the potential step functions induced by the original potential process $\mathbf{M}_i^t(z)$ and $m_i(z)$ to denote potential values of $\mathbf{M}_i^t(z)$ evaluated at point $t_i = it/H$. Under Assumption 2, we can choose a large H such that $t/H \leq \varepsilon$. Then we have the following conditional independence conditions,

$$\begin{aligned}
&V_i(z,\mathbf{m}_H) \perp\!\!\!\perp m_0 | Z_i, \mathbf{X}_i^t, \\
&V_i(z,\mathbf{m}_H) \perp\!\!\!\perp (m_1 - m_0) | Z_i, \mathbf{X}_i^t, \mathbf{m}_H^0, \\
&V_i(z,\mathbf{m}_H) \perp\!\!\!\perp (m_2 - m_1) | Z_i, \mathbf{X}_i^t, \mathbf{m}_H^{t/H}, \\
&\qquad\qquad \cdots \\
&V_i(z,\mathbf{m}_H) \perp\!\!\!\perp (m_H - m_{H-1}) | Z_i, \mathbf{X}_i^t, \mathbf{m}_H^{t(H-1)/H},
\end{aligned}$$

where are equivalent to,

$$\begin{aligned}
&V_i(z,\mathbf{m}_H) \perp\!\!\!\perp m_0 | Z_i, \mathbf{X}_i^t, \\
&V_i(z,\mathbf{m}_H) \perp\!\!\!\perp (m_1 - m_0) | Z_i, \mathbf{X}_i^t, m_0, \\
&V_i(z,\mathbf{m}_H) \perp\!\!\!\perp (m_2 - m_1) | Z_i, \mathbf{X}_i^t, m_0, m_1, \\
&\qquad\qquad \cdots \\
&V_i(z,\mathbf{m}_H) \perp\!\!\!\perp (m_H - m_{H-1}) | Z_i, \mathbf{X}_i^t, m_0, m_1 \cdots, m_{H-1},
\end{aligned}$$

as the step function $m_H^{it/H}, i \leq H$ is completely determined by values at jumps $\{m_0, \cdots, m_i\}$. With the established conditional independence, we have,

$$\Pr(V_i(z, \mathbf{m}_H) > t | Z_i = z, \mathbf{X}_i^t = \mathbf{x}^t, m_0, m_1, m_2, \cdots m_H) = \Pr(V_i(z, \mathbf{m}_H) > t | Z_i = z, \mathbf{X}_i^t = \mathbf{x}^t).$$

With similar arguments, we can show that,

$$\begin{aligned}
&\Pr(V_i(z, \mathbf{m}_H) > t | Z_i = z, \mathbf{X}_i^t = \mathbf{x}^t) = \Pr(V_i(z, \mathbf{m}_H) > t | Z_i = z', \mathbf{X}_i^t = \mathbf{x}^t), \\
&= \Pr(V_i(z, \mathbf{m}_H) > t | Z_i = z, \mathbf{X}_i^t = \mathbf{x}^t, m_0 = m_0(z'), \cdots m_H = m_H(z')), \\
&= \Pr(V_i(z, \mathbf{m}_H) > t | Z_i = z, \mathbf{X}_i^t = \mathbf{x}^t, \mathbf{m}_H(z') = \mathbf{m}_H), \\
&= \Pr(V_i(z, \mathbf{m}_H) > t | \mathbf{X}_i^t = \mathbf{x}^t, \mathbf{m}_H(z') = \mathbf{m}_H).
\end{aligned}$$

As a conclusion, we have shown that,

$$\begin{aligned}
&\int_{\mathcal{X}} \int_{R^{[0,t]}} \Pr(V_i(z, \mathbf{m}) > t | Z_i = z, \mathbf{X}_i^t = \mathbf{x}^t, \mathbf{M}_i^t = \mathbf{m}) dF_{\mathbf{X}_i^t}(x^t) \times \mathrm{d}\{\mathrm{F}_{\mathbf{M}_i^t | Z_i = z', \mathbf{X}_i^t = \mathbf{x}^t}(\mathbf{m})\}, \\
&\asymp \int_{\mathcal{X}} \int_{R^{[0,t]}} \Pr(V_i(z, \mathbf{m}_H) > t | \mathbf{X}_i^t = \mathbf{x}^t, \mathbf{m}_H(z') = \mathbf{m}_H) \times \mathrm{d}\{\mathrm{F}_{\mathbf{m}_H(z') | \mathbf{X}_i^t = \mathbf{x}^t}(\mathbf{m}_H)\} dF_{\mathbf{X}_i^t}(x^t) + O(H^{-2}), \\
&\asymp \int_{\mathcal{X}} \Pr(V_i(z, \mathbf{m}_H(z')) > t | \mathbf{X}_i^t = \mathbf{x}^t) + O(H^{-2}) \asymp \int_{\mathcal{X}} \Pr(V_i(z, \mathbf{m}(z')) > t | \mathbf{X}_i^t = \mathbf{x}^t) + O(H^{-2}).
\end{aligned}$$

The last equivalence follows from the regularity condition of $V_i(z, \mathbf{m}(z'))$ as a function of $\mathbf{m}(z')$. Let H goes to infinity, we have,

$$\begin{aligned}
&\int_{\mathcal{X}} \int_{R^{[0,t]}} \Pr(V_i > t | Z_i = z, \mathbf{X}_i^t = \mathbf{x}^t, \mathbf{M}_i^t = \mathbf{m}) \mathrm{dF}_{\mathbf{X}_i^t}(\mathbf{x}^t) \times \mathrm{d}\{\mathrm{F}_{\mathbf{M}_i^t | Z_i = z', \mathbf{X}_i^t = \mathbf{x}^t}(\mathbf{m})\} \\
&= \int_{\mathcal{X}} \Pr(V_i(z, \mathbf{m}(z')) > t | \mathbf{X}_i^t = \mathbf{x}^t) \mathrm{dF}_{\mathbf{X}_i^t}(\mathbf{x}^t) = \Pr(V_i(z, \mathbf{m}(z')) > t) = S_{z,z'}(t)
\end{aligned}$$

Under Assumption 3, the conditional survival function can be estimated with a non-parametric Kaplan Meier estimator,

$$\begin{aligned}
\Pr(V_i > t | Z_i = z, \mathbf{X}_i^t = \mathbf{x}^t, \mathbf{M}_i^t = \mathbf{m}) &= \prod_{k=1}^{K} \Pr(\tilde{V}_i > t_k | \tilde{V}_i > t_{k-1}, \mathbf{M}_i^t = m. \mathbf{X}_i^t, Z_i = z) \\
&= \prod_{k=1}^{K} \Pr(\tilde{V}_i > t_k | \tilde{V}_i > t_{k-1}, \mathbf{M}_i = m^{t_k}. \mathbf{X}_i^{t_k}, Z_i = z),
\end{aligned}$$

where $t_1 < t_2 <, \cdots, < t_k <, \cdots, < t_K < \cdots$ is the time grid where we observe failure event ($\delta_i = 1$) and the selected fixed time point t lies between t_K and t_{K+1}. Hence, we complete the proof for Theorem 1.

A.2 Details of Gibbs Sampler

In this section, we provide detailed descriptions on the Gibbs sampler for the model in Section 4. The sampling of mediator process is similar to the one in Kowal and Bourgeois (2020) and Zeng et al. (2020). Therefore, we omit the details for simplicity and refer the reader to Zeng et al. (2021).

Next, we describe the sampling for the survival model. As we have the model,

$$\lambda(t) = \lambda_0(t) \exp(Z_i\alpha + X'_{ij}\beta_S + f\{\mathbf{M}_i^t, \gamma\}).$$

The survival function for a specific subject becomes,

$$S_i(t) = \Pr(V_i > t) = \exp(-H_i(t)) = \exp(-\int_0^t h_i(s)ds)$$

where $H_i(t)$ is the cumulative hazard function, which is a right-continuous increasing function with $H_i(0) = 0$. For a given observation $(\tilde{V}_i, \delta_i, \mathbf{X}_{ij}, M_{ij}, Z_i)$, the likelihood of this observation is,

$$L_{ij} = (1 - \delta_{ij}) \Pr(V_i > t_{ij}) + \delta_{ij} \Pr(V_i = t_{ij}).$$

where δ_{ij} is the indicator for whether the subject is still alive at time point t_{ij}. To derive an explicit formula for the likelihood, we let $\lambda_1, \lambda_2, \cdots, \lambda_K$ be the baseline hazard for a specified time grids $t_1 < t_2, \cdots, < t_K$ that at least one failure happens in each bin $(t_{k-1}, t_k]$. Then the cumulative hazard function becomes,

$$H_i(t) = \sum_{k=1}^{K} \lambda_k U_{i,k}(t, \alpha, \beta_S, \gamma),$$

where

$$U_{i,k}(t, \alpha, \beta_S, \gamma) = (t_k - t_{k-1}) \exp(Z_i\alpha + X_i^{t_k}\beta_S + f\{\mathbf{M}_i^{t_k}, \gamma\}) \quad \text{if } t > t_k$$
$$U_{i,k}(t, \alpha, \beta_S, \gamma) = (t - t_{k-1}) \exp(Z_i\alpha + X_i^{t_{k-1}}\beta_S + f\{\mathbf{M}_i^{t}, \gamma\}) \quad \text{if } t < t_k.$$

As such, we can express the likelihood of the data as a function of parameter $(\{\lambda_k\}_{k=1}^K, \alpha, \beta_S, \gamma)$. First, we describe the prior of baseline hazard rate $\{\lambda_k\}_{k=1}^K$. We specify a Gamma Process prior on λ_k, that is the increments are independent across each other and follow a Gamma distribution, $\lambda_k \sim Gamma(\alpha_k, \beta_k)$. We specify α_k, β_k in the following way that, we let $\alpha_k = A\alpha(t_k)$ and $\beta_k = B$, where $\alpha(t)$ is strictly increasing function that captures the mean of the hazard rate. For example, when $\alpha(t) = t$, then $E\{\lambda_k\} = t_k A/B$. For the hyperparameter A, B, we specify a Gamma prior such that $A, B \sim Gamma(\varepsilon, \varepsilon)$ with $\varepsilon = 0.001$.

Then the conditional posterior distributions of the parameters in the sample are:

- Baseline hazard rate $\lambda_k|-$:

$$p(\lambda_k|-) = Ga(\lambda_k|\alpha_k + n_k, \beta_k + m_k(\alpha, \beta_S, \gamma)),$$

where n_k is the number of failure in $(t_{k-1}, t_k]$ and $m_k(\alpha, \beta_S, \gamma) = \sum_{i=1}^{N} \sum_{j=1}^{n_i} U_{i,k}(t_{ij}, \alpha, \beta_S, \gamma)$.

- The hyperparameter for the Gamma Process A, B—-,

$$p(A|-) \propto A^{\varepsilon-1} \exp(-\varepsilon A) B^{\varepsilon\alpha(t_K)} \prod_{k=1}^{K} \frac{\lambda_k^{A(\alpha(t_k)-\alpha(t_{k-1}))}}{\Gamma(A(\alpha(t_k) - \alpha(t_{k-1}))},$$

$$p(B|-) \propto B^{A\alpha(t)+\varepsilon-1} \exp(-B(\varepsilon + \sum_{k=1}^{K} \lambda_k)$$

This step can be updated using a one step of Metropolis random walk.

- The coefficient for treatment, covariates and mediator process: $\alpha, \beta_S, \gamma|-$

$$p(\alpha, \beta_S, \gamma|-) \propto p(\alpha, \beta_S, \gamma) \exp(\sum_{i=1}^{N} \delta_i (Z_i\alpha + X_{in_i}\beta_S + f\{\mathbf{M}_i^{t_{in_i}}, \gamma) - \sum_{k=1}^{K} \lambda_k m_k(\alpha, \beta_S, \gamma)),$$

This step can be updated efficiently using the adaptive rejection methods in Gilks and Wild (1992) as the density is log-concave in $(\alpha, \beta_S, \gamma)$. The parameterization of the cumulative model $f\{\mathbf{M}_i^t, \gamma)$ is similar in the construction of spline basis in mediator process. We refer the readers to Kowal and Bourgeois (2020) and Zeng et al. (2020) for details.

References

Adam, E. K., Quinn, M. E., Tavernier, R., McQuillan, M. T., Dahlke, K. A., and Gilbert, K. E. (2017), "Diurnal cortisol slopes and mental and physical health outcomes: A systematic review and meta-analysis," *Psychoneuroendocrinology*, 83, 25–41.

Alberts, S. C. and Altmann, J. (2012), "The Amboseli Baboon Research Project: 40 years of continuity and change," in *Long-term Field Studies of Primates*, Springer, 261–287.

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, 1173.

Berens, A. E., Jensen, S. K., and Nelson, C. A. (2017), "Biological embedding of childhood adversity: from physiological mechanisms to clinical implications," *BMC medicine*, 15, 1–12.

Boyce, W. T. and Hertzman, C. (2018), *Early childhood health and the life course: the state of the science and proposed research priorities*, Springer, Cham, 61–93.

Campos, F. A., Archie, E. A., Gesquiere, L. R., Tung, J., Altmann, J., and Alberts, S. C. (In Press), "Glucocorticoids predict survival in female baboons," *Science Advances*.

Daniels, M. J., Roy, J. A., Kim, C., Hogan, J. W., and Perri, M. G. (2012), "Bayesian inference for the causal effect of mediation," *Biometrics*, 68, 1028–1036.

Didelez, V. (2019a), "Defining causal mediation with a longitudinal mediator and a survival outcome," *Lifetime Data Analysis*, 25, 593–610.

— (2019b), "Defining causal mediation with a longitudinal mediator and a survival outcome," *Lifetime Data Analysis*, 25, 593–610.

Evans, G. W., Li, D., and Whipple, S. S. (2013), "Cumulative risk and child development," *Psychological Bulletin*, 139, 1342.

Fahrmeir, L. and Lang, S. (2001), "Bayesian inference for generalized additive mixed models based on Markov random field priors," *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 50, 201–220.

Felitti, V. J., Anda, R. F., Nordenberg, D., Williamson, D. F., Spitz, A. M., Edwards, V., Koss, M. P., and Marks, J. S. (1998), "Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults: The adverse childhood experiences (ACE) study," *American journal of preventive medicine*, 14, 245–258.

Forastiere, L., Mattei, A., and Ding, P. (2018), "Principal ignorability in mediation analysis: through and beyond sequential ignorability," *Biometrika*, 105, 979–986.

Gilks, W. R. and Wild, P. (1992), "Adaptive rejection sampling for Gibbs sampling," *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 41, 337–348.

Goldsmith, J., Greven, S., and Crainiceanu, C. (2013), "Corrected confidence bands for functional data using principal components," *Biometrics*, 69, 41–51.

Han, K., Hadjipantelis, P. Z., Wang, J.-L., Kramer, M. S., Yang, S., Martin, R. M., and Müller, H.-G. (2018), "Functional principal component analysis for identifying multivariate patterns and archetypes of growth, and their association with long-term cognitive development," *PloS one*, 13, e0207073.

Harris, B. N. (2019), "Stress hypothesis overload: 131 hypotheses exploring the role of stress in tradeoffs, transitions, and health," *General and Comparative Endocrinology*, 113355.

Hertzman, C. (1999), "The biological embedding of early experience and its effects on health in adulthood," *Annals of the New York Academy of Sciences*, 896, 85–95.

Huang, R., Xu, R., and Dulai, P. S. (2020), "Sensitivity analysis of treatment effect to unmeasured confounding in observational studies with survival and competing risks outcomes," *Statistics in Medicine*, 39, 3397–3411.

Ibrahim, J. G., Chen, M.-H., and Sinha, D. (2014), "Bayesian Survival Analysis," *Wiley StatsRef: Statistics Reference Online*.

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

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

Imbens, G. W. and Rubin, D. B. (2015), *Causal inference in statistics, social, and biomedical sciences*, Cambridge University Press.

Jiang, C.-R. and Wang, J.-L. (2010), "Covariate adjusted functional principal components analysis for longitudinal data," *The Annals of Statistics*, 38, 1194–1226.

— (2011), "Functional single index models for longitudinal data," *The Annals of Statistics*, 39, 362–388.

Kim, C., Daniels, M., Li, Y., Milbury, K., and Cohen, L. (2018), "A Bayesian semiparametric latent variable approach to causal mediation," *Statistics in Medicine*, 37, 1149–1161.

Kim, C., Daniels, M. J., Hogan, J. W., Choirat, C., and Zigler, C. M. (2019), "Bayesian methods for multiple mediators: Relating principal stratification and causal mediation in the analysis of power plant emission controls," *The Annals of Applied Statistics*, 13, 1927.

Kim, C., Daniels, M. J., Marcus, B. H., and Roy, J. A. (2017), "A framework for Bayesian nonparametric inference for causal effects of mediation," *Biometrics*, 73, 401–409.

Kowal, D. R. and Bourgeois, D. C. (2020), "Bayesian Function-on-Scalars Regression for High-Dimensional Data," *Journal of Computational and Graphical Statistics*, 29, 629–638.

Lange, T., Vansteelandt, S., and Bekaert, M. (2012), "A simple unified approach for estimating natural direct and indirect effects," *American Journal of Epidemiology*, 176, 190–195.

Levy, E. J., Gesquiere, L. R., McLean, E., Franz, M., Warutere, J. K., Sayialel, S. N., Mututua, R. S., Wango, T. L., Oudu, V. K., and Altmann, J. (2020), "Higher dominance rank is associated with lower glucocorticoids in wild female baboons: A rank metric comparison," *Hormones and Behavior*, 125, 104826.

Lin, S.-H., Young, J., Logan, R., Tchetgen Tchetgen, E. J., and VanderWeele, T. J. (2017a), "Parametric mediational g-formula approach to mediation analysis with time-varying exposures, mediators, and confounders," *Epidemiology (Cambridge, Mass.)*, 28, 266.

Lin, S.-H., Young, J. G., Logan, R., and VanderWeele, T. J. (2017b), "Mediation analysis for a survival outcome with time-varying exposures, mediators, and confounders," *Statistics in Medicine*, 36, 4153–4166.

Lindquist, M. A. (2012), "Functional causal mediation analysis with an application to brain connectivity," *Journal of the American Statistical Association*, 107, 1297–1309.

Lindquist, M. A. and Sobel, M. E. (2011), "Graphical models, potential outcomes and causal inference: Comment on Ramsey, Spirtes and Glymour," *NeuroImage*, 57, 334–336.

Lok, J. J. (2016), "Defining and estimating causal direct and indirect effects when setting the mediator to specific values is not feasible," *Statistics in Medicine*, 35, 4008–4020.

Lok, J. J. and Bosch, R. J. (2021), "Causal Organic Indirect and Direct Effects: Closer to the Original Approach to Mediation Analysis, with a Product Method for Binary Mediators," *Epidemiology*.

MacKinnon, D. (2012), *Introduction to statistical mediation analysis*, Routledge.

Markham, A. C., Gesquiere, L. R., Alberts, S. C., and Altmann, J. (2015), "Optimal group size in a highly social mammal," *Proceedings of the National Academy of Sciences*, 112, 14882–14887.

Miller, G. E., Chen, E., Fok, A. K., Walker, H., Lim, A., Nicholls, E. F., Cole, S., and Kobor, M. S. (2009), "Low early-life social class leaves a biological residue manifested by decreased glucocorticoid and increased proinflammatory signaling," *Proceedings of the National Academy of Sciences*, 106, 14716–14721.

Miller, G. E., Chen, E., and Parker, K. J. (2011), "Psychological stress in childhood and susceptibility to the chronic diseases of aging: moving toward a model of behavioral and biological mechanisms," *Psychological Bulletin*, 137, 959.

Neyman, J. (1923), "On the application of probability theory to agricultural experiments. Essay on principles. Section 9 (translated)." *Statistical Science*, 465–480.

Nguyen, T. Q., Schmid, I., and Stuart, E. A. (2020), "Clarifying causal mediation analysis for the applied researcher: Defining effects based on what we want to learn," *Psychological Methods*, in press.

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

Petruccelli, K., Davis, J., and Berman, T. (2019), "Adverse childhood experiences and associated health outcomes: A systematic review and meta-analysis," *Child abuse & neglect*, 97, 104127.

Ramsay, J. and Silverman, B. (2005), "Functional Data Analysis," .

Robins, J. (1986), "A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect," *Mathematical modelling*, 7, 1393–1512.

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

Robins, J. M., Hernan, M. A., and Brumback, B. (2000), "Marginal Structural Models and Causal Inference in Epidemiology," *Epidemiology*, 11.

Rosenbaum, S., Zeng, S., Campos, F., Gesquiere, L., Altmann, J., Alberts, S., Li, F., and Archie, E. (2020), "Social bonds do not mediate the relationship between early adversity and adult glucocorticoids in wild baboons," *Proceedings of the National Academy of Sciences*, in press.

Roth, D. L. and MacKinnon, D. P. (2012), "Mediation analysis with longitudinal data," *Longitudinal data analysis: A practical guide for researchers in aging, health, and social sciences*, 181–216.

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

Schoenle, L. A., Zimmer, C., Miller, E. T., and Vitousek, M. N. (2021), "Does variation in glucocorticoid concentrations predict fitness? A phylogenetic meta-analysis," *General and Comparative Endocrinology*, 300, 113611.

Sobel, M. E. (2008), "Identification of causal parameters in randomized studies with mediating variables," *Journal of Educational and Behavioral Statistics*, 33, 230–251.

Tchetgen, E. J. T. (2011), "On causal mediation analysis with a survival outcome," *The International Journal of Biostatistics*, 7.

Tchetgen Tchetgen, E. J. and Shpitser, I. (2012), "Semiparametric theory for causal mediation analysis: efficiency bounds, multiple robustness, and sensitivity analysis," *Annals of Statistics*, 40, 1816.

Tung, J., Archie, E. A., Altmann, J., and Alberts, S. C. (2016), "Cumulative early life adversity predicts longevity in wild baboons," *Nature Communications*, 7, 1–7.

van der Laan, M. J. and Petersen, M. L. (2008), "Direct effect models," *The International Journal of Biostatistics*, 4.

VanderWeele, T. (2015), *Explanation in causal inference: methods for mediation and interaction*, Oxford University Press.

VanderWeele, T. J. (2011), "Causal mediation analysis with survival data," *Epidemiology (Cambridge, Mass.)*, 22, 582.

— (2016), "Mediation analysis: a practitioner's guide," *Annual Review of Public Health*, 37, 17–32.

VanderWeele, T. J. and Tchetgen Tchetgen, E. J. (2017), "Mediation analysis with time varying exposures and mediators," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 79, 917–938.

Vansteelandt, S., Linder, M., Vandenberghe, S., Steen, J., and Madsen, J. (2019), "Mediation analysis of time-to-event endpoints accounting for repeatedly measured mediators subject to time-varying confounding," *Statistics in Medicine*, 38, 4828–4840.

Wang, X., Chen, M.-H., and Yan, J. (2013), "Bayesian dynamic regression models for interval censored survival data with application to children dental health," *Lifetime Data Analysis*, 19, 297–316.

Warren, J. R. (2009), "Socioeconomic status and health across the life course: a test of the social causation and health selection hypotheses," *Social Forces*, 87, 2125–2153.

Yao, F., Müller, H.-G., and Wang, J.-L. (2005), "Functional data analysis for sparse longitudinal data," *Journal of the American Statistical Association*, 100, 577–590.

Zeng, S., Rosenbaum, S., Archie, E., Alberts, S., and Li, F. (2020), "Causal mediation analysis for sparse and irregular longitudinal data," *arXiv preprint arXiv:2007.01796*.

Zeng, S., Stacy, R., Susan, A., Elizabeth, A., and Li, F. (2021), "Supplement to "Causal mediation analysis for sparse and irregular longitudinal data."," .

Zhao, Y., Luo, X., Lindquist, M., and Caffo, B. (2018), "Functional Mediation Analysis with an Application to Functional Magnetic Resonance Imaging Data," *arXiv preprint arXiv:1805.06923*.

Zheng, W. and van der Laan, M. (2017), "Longitudinal mediation analysis with time-varying mediators and exposures, with application to survival outcomes," *Journal of Causal Inference*, 5.

Zigler, C. M., Dominici, F., and Wang, Y. (2012), "Estimating causal effects of air quality regulations using principal stratification for spatially correlated multivariate intermediate outcomes," *Biostatistics*, 13, 289–302.