Biostatistics, 2024, **25**, 919–932
https://doi.org/10.1093/biostatistics/kxad038
Advance Access publication February 9, 2024

A Bayesian nonparametric approach for multiple mediators with applications in mental health studies

Samrat Roy [1], Michael J. Daniels [*,2], Jason Roy [3]

[1] Operations and Decision Sciences, Indian Institute of Management Ahmedabad, Gujarat, India
[2] Department of Statistics, University of Florida, Gainesville, USA
[3] Department of Biostatistics and Epidemiology, Rutgers University, New Brunswick, USA

[*] Corresponding author: Department of Statistics, University of Florida, 230 Newell Dr, Gainesville, FL 32611, USA.
Email: daniels@ufl.edu

SUMMARY

Mediation analysis with contemporaneously observed multiple mediators is a significant area of causal inference. Recent approaches for multiple mediators are often based on parametric models and thus may suffer from model misspecification. Also, much of the existing literature either only allow estimation of the joint mediation effect or estimate the joint mediation effect just as the sum of individual mediator effects, ignoring the interaction among the mediators. In this article, we propose a novel Bayesian nonparametric method that overcomes the two aforementioned drawbacks. We model the joint distribution of the observed data (outcome, mediators, treatment, and confounders) flexibly using an enriched Dirichlet process mixture with three levels. We use standardization (g-computation) to compute all possible mediation effects, including pairwise and all other possible interaction among the mediators. We thoroughly explore our method via simulations and apply our method to a mental health data from Wisconsin Longitudinal Study, where we estimate how the effect of births from unintended pregnancies on later life mental depression (CES-D) among the mothers is mediated through lack of self-acceptance and autonomy, employment instability, lack of social participation, and increased family stress. Our method identified significant individual mediators, along with some significant pairwise effects.

KEYWORDS: enriched Dirichlet process; multiple mediators.

1. INTRODUCTION

In the social, behavioral, and health sciences, including neuroimaging (Woo et al. 2015; Zhao et al. 2018), weight loss management (Daniels et al. 2012), air pollution (Kim et al. 2019), and metagenomics (Wu et al. 2011; Sohn et al. 2019), researchers are often interested in estimating the part of the effect of intervention on outcome that is routed through the potential mediators. Examples include, the role of Sulfur dioxide and Nitrogen oxides emission as a mediation path between the effects of coal-fired power plants on the increase in the air pollution levels (see Kim et al. 2019). Another example (Daniels et al. 2012) relates to weight management trials, where the adherence to behavioral weight management strategies can act as a mediator between the effect of a weight management program on maintaining weight loss.

Received: May 15, 2023. **Revised:** December 14, 2023. **Accepted:** December 15, 2023

Existing methods on multiple mediators are often based on parametric models and thus may suffer from parametric misspecification. Also, much of the available literature do not allow interaction among the mediators and estimate the joint mediation effect as just the sum of the individual mediation effects, ignoring the interaction. For example, Wang et al. (2013) considered estimation of mediation effects with binary outcome and contemporaneously observed multiple mediators under the potential outcome framework. Their approach is based on a parametric model and they define the joint mediation effect as the sum of individual mediator effects. Imai and Yamamoto (2013), VanderWeele and Vansteelandt (2014), and Daniel et al. (2015) developed methodologies which allowed for interaction effect of either contemporaneously observed, or causally related mediators. Among these, VanderWeele and Vansteelandt (2014) confined themselves to estimation of only the joint mediation effect. On the other hand, Daniel et al. (2015) extended the approach of Imai and Yamamoto (2013) in the context of causally ordered mediators, providing the finest possible decomposition of the total effect into various path-specific effects. However, the proposed methodologies for estimation employ parametric approaches. In more recent times, Sohn et al. (2019), in the context of contemporaneously observed compositional mediators, proposed a parametric linear structural equation model (LSEM) approach, assuming no interaction effects of the mediators and thus estimated the joint mediation effect as the sum of the individual effects. To deal with high-dimensional multivariate mediators, Chén et al. (2018) developed an LSEM-based approach that linearly combined the mediators into a relatively smaller number of orthogonal components, where the components are ranked based on their contribution towards the LSEM likelihood. Wang et al. (2019) employed a Bayesian regularized approach to deal with both multiple exposures and mediators. However, their approach is again based on a parametric framework. Moreover, their approach allows estimation of only the joint mediation effect. Some other notable work under the parametric paradigm includes Derkach et al. (2019), Song et al. (2020), Song et al. (2021), and Zhang et al. (2021). While Song et al. (2020), Song et al. (2021), and Zhang et al. (2021) primarily focus on mediator selection, Derkach et al. (2019) propose an approach which formalizes the mediators as latent factors. Thus, whether the mediators are observed contemporaneously or they are causally ordered, most of the existing literature use a parametric approach, and estimate the joint mediation effect directly without any partition, or as the sum of individual mediator effects, ignoring the interactions.

In this article, we develop a novel Bayesian nonparametric method for contemporaneously observed multiple mediators, that overcomes the two aforementioned drawbacks: (i) it does not suffer from potential parametric misspecification, and (ii) it allows for the interaction among the mediators. Hence, we can estimate pairwise or any other possible interaction effects of the mediators, instead of just the joint effect. The BNP model extends the enriched Dirichlet process mixture (Wade et al. 2011; Roy et al. 2018) to a third level to flexibly model the joint distribution of our observed data (outcome, mediators, treatment, and confounders). The first level characterizes the conditional distribution of the outcome given the mediators, treatment and the confounders. The second level corresponds to the conditional distribution of each of the mediators given the treatment and the confounders, and the third level corresponds to the distribution of the treatment and the confounders. The above three-level enriched Dirichlet process mixture can be used in a wider variety of applications besides causal mediation; in particular, in settings when the conditional distributions of outcomes and secondary outcomes are of interest and there are many covariates. We introduce identifying assumptions and use standardization (g-computation; see Robins 1986; Robins and Hernán 2009) to compute causal mediation effects, including the interaction effects. We proved the posterior consistency of EDP3 in Appendix B of the supplementary material.

Finally, we apply our method to a data from Wisconsin Longitudinal Study, where we estimate how the effect of births from unintended pregnancies on later life mental depression (CES-D score) among the mothers is mediated through lack of self-acceptance and autonomy, employment instability, lack of social participation, and increased family stress. There has been an escalated interest in studying the effect of unintended pregnancies on the later life mental health of the women (Bahk et al. 2015; Herd et al. 2016; Barton et al. 2017). Among them, Herd et al. (2016) used the same data and established a significant effect of births from unintended pregnancies on later life depression among the women. However, to the best of our knowledge, none of the existing literature investigated how this effect is mediated through other potential intermediate variables. Our method identified significant individual mediation effects for lack of autonomy, lack of self-acceptance and employment instability, along with some other significant pairwise interaction effects.

2. CAUSAL EFFECTS, ASSUMPTIONS, AND IDENTIFIABILITY

In this section, we define the causal mediation effects as introduced in Kim et al. (2019). Suppose we have an outcome Y, a binary exposure A, Q potential mediators, $M_1, M_2, \ldots, M_Q$ that are observed contemporaneously and a vector of p confounders L. For $q = 1, 2, \ldots, Q$, let $a_q \in \{0, 1\}$ be the binary treatment status under which the qth mediator is induced. In other words, following the potential outcome framework introduced in Imai et al. (2010), the vector of potential mediators is denoted by $\{M_1(a_1), M_2(a_2), \ldots, M_Q(a_Q)\}$, which we denote for notational simplicity as $M(a_1, a_2, \ldots, a_Q)$, wherein $M_q(a_q)$, for $q = 1, 2, \ldots, Q$, corresponds to the potential value of the qth mediator under the treatment status $a_q \in \{0, 1\}$. Similarly, the potential outcome is defined as $Y(a; M(a_1, a_2, \ldots, a_Q))$ under the treatment status $a \in \{0, 1\}$ and the mediators $M(a_1, a_2, \ldots, a_Q)$. Thus, the potential outcome notation for Y is the outcome that corresponds to the treatment received (i.e. a) and the mediators, that are set to what they would be under different combinations of treatments $\{a_1, a_2, \ldots, a_Q\}$. Among all possible potential outcomes, only two, namely $Y(1, M(1, 1, \ldots, 1))$ and $Y(0, M(0, 0, \ldots, 0))$ are observable (and only under randomization), while all the remaining ones are (a priori) counterfactual. We define the causal effects as follows:

Total effect (TE): Total effect is the entire effect of the intervention A on the outcome Y and is defined as the expected difference between the two observable potential outcomes, that is, $E[Y(1, M(1, 1, \ldots, 1)) - Y(0, M(0, 0, \ldots, 0))]$. This total effect can be decomposed into two parts: the part that traverses directly from the intervention to the outcome, and the part that is routed through the mediators.

Natural direct effect (NDE): NDE is defined as $E[Y(1, M(0, 0, \ldots, 0)) - Y(0, M(0, 0, \ldots, 0))]$. Thus, this is effect of the intervention on the outcome while the mediators are set to their realizations in the absence of intervention. This effect of A on Y is "direct" in the sense that it is not through the mediators.

Joint natural indirect effect (JNIE): The JNIE of all Q mediators is defined as $TE - NDE = E[Y(1, M(1, 1, \ldots, 1)) - Y(1, M(0, 0, \ldots, 0))]$. Hence, this part of the effect of A on Y passes indirectly through the mediators. Note that, JNIE can be decomposed into natural indirect effects that are attributable to changes in different combinations of the Q mediators. In other words, JNIE comprises various indirect effects, that are mediated through different possible combinations of the Q mediators. For ease of exposition, suppose $Q = 3$. Then, the joint effect of the three mediators JNIE_{123}, is depicted in Fig. H.1 in the supplementary materials. The JNIE can be decomposed into *individual natural indirect effects* (INIE), which are attributable to any one of the three possible mediators (see JNIE_1, JNIE_2, and JNIE_3 in Figure H.1 in the supplementary materials). For example, $\text{JNIE}_1 = E[Y(1, M(1, 1, 1)) - Y(1, M(0, 1, 1))]$. Similarly, one can also define *pairwise natural indirect effects* (PNIE), which are passed through any two of the three mediators (see JNIE_{12}, JNIE_{23}, and JNIE_{13} in Figure H.1 in the supplementary materials). Thus, for example, $\text{JNIE}_{12} = E[Y(1, M(1, 1, 1)) - Y(1, M(0, 0, 1))]$. For more than three mediators, the number of pairwise, three-way, etc. JNIEs gets quite large; for example, for K mediators, the partition of the JNIE will have $\sum \binom{K}{2} + \binom{K}{3} + \ldots + \binom{K}{K-1}$ components. We now summarize the assumptions that are sufficient to identify the causal effects defined above.

Assumption 2.1 $\{Y(a, M(a, a, \ldots, a)), M(0, 0, \ldots, 0), M(1, 1, \ldots, 1)\} \perp\!\!\!\perp A | L = \ell\}$.

The above assumption, known as *ignorable treatment assignment*, says that, conditional on the confounders, the treatment assignment is independent of the observable potential outcomes and the mediators. This assumption is also known as 'no unmeasured confounders'. In the context of our data analysis in Section 6, this assumption implies that the covariates childhood measures, adulthood measures and personality measures (see Table H.13 in the supplementary material) are sufficient to adjust for confounding between the unintended pregnancies (exposure) and the observable potential outcome, mental depression, and mediators, lack of autonomy, lack of self-acceptance, employment instability, lack of social participation, and increased family stress.

Assumption 2.2 For exposure $A = 1$, the conditional distributions of the observable potential outcome $Y(1, M(1, 1, \ldots, 1))$ given values of all potential mediators (and confounders), is the same as that of a priori counterfactual $Y(1, M(a_1, a_2, \ldots, a_Q))$, regardless of whether the mediator values were induced by $A = 1$ or $A = 0$.

Note that, the definition of JNIE involves the observable potential outcome $Y(1, M(1, 1, \ldots, 1))$ and a priori counterfactual $Y(1, M(0, 0, \ldots, 0))$. In terms of the notations, the above assumption implies that, $f_{1,M(0,0,\ldots,0)}(y \mid M(0, 0, \ldots, 0) = m, L = \ell) = f_{1,M(1,1,\ldots,1)}(y \mid M(1, 1, \ldots, 1) = m, L = \ell)$. In other words, when the value of the mediator vector is fixed at 'm' (and the confounder vector is fixed at 'ℓ'), the two conditional distributions stated above are the same, irrespective of the fact that the mediators in the first case are induced in the absence of treatment, while the mediators in the second case are induced under the treatment.

The above assumption also applies to the other counterfactual outcomes, that are present in the decomposition of the JNIE. For example, for JNIE_1 with three mediators (see Figure H.1 in the supplementary materials and relevant discussion in the definition of JNIE), the above assumption takes the following form $f_{1,M(0,1,1)}(y \mid M(0, 1, 1) = m, L = \ell) = f_{1,M(1,1,1)}(y \mid M(1, 1, 1) = m, L = \ell)$. Similarly, for JNIE_{12}, the assumption translates into the following $f_{1,M(0,0,1)}(y \mid M(0, 0, 1) = m, L = \ell) = f_{1,M(1,1,1)}(y \mid M(1, 1, 1) = m, L = \ell)$. In the context of our data in Section 6, the above assumption says that, conditional on the childhood, adulthood, and personality measures (baseline covariates), the chance that a woman will have later-life depression depends only on the values of the mediators (that is, the levels of self-acceptance, autonomy, employment instability, and so on); it does not matter whether those mediator values are induced by the unintended pregnancies (that is, presence of exposure) or not (absence of exposure). Kim et al. (2019) use a similar assumption in the context of their study on power plant emission controls, and they also provide a sensitivity analysis for this assumption.

Assumption 2.3 $M_{j_1}(a) \perp\!\!\!\perp M_{j_2}(a')|L$ for $a \neq a'$ and $j_1, j_2 = 1, 2, \ldots, Q$.

The above assumption says that any mediator in the absence of treatment is independent of any mediator under the treatment, conditional on the confounders L. This assumption is needed to identify the joint distribution of all the potential mediators. Note that, this assumption does not restrict the dependence of the mediators under the same treatment status. In the context of our data in Section 6, this assumption implies that, any mediator, say employment instability under the presence of unintended pregnancy, is independent of any mediator, say lack of self-acceptance in the absence of unintended pregnancy. However, they can be dependent under the same exposure status (that is, both employment stability and self-acceptance under unintended pregnancy). Like the previous assumptions, this assumption is also untestable; in Appendix D of the supplementary materials, we propose a method for assessing sensitivity to this assumption.

Theorem 2.1 Under Assumptions 2.1, 2.2, and 2.3, the NDE, JNIE (and its decomposition) are identifiable. (see Appendix C.1 in the supplementary materials for the proof).

3. BNP MODEL FOR OBSERVED DATA

We model the joint distribution of the outcome, mediators, treatment, and confounders using an extension of the two-level enriched Dirichlet process mixture (EDPM) model (Wade et al. (2011), Wade et al. (2014)). Denoting $(A, L^T)^T$ by X, we propose a three-level EDP mixture model for the joint distribution of the observed data (Y, M, X):

$$\begin{aligned} Y_i|M_i, X_i; \theta_i &\sim p(y|m, x; \theta_i) \\ M_{iq}|X_i; \omega_i &\sim p(m_q|x; \omega_i) : q = 1, \ldots, Q \text{ (independent over q)} \\ X_{i,r}|\psi_i &\sim p(x_r|\psi_i) : r = 1, \ldots, p+1 \text{ (independent over r)} \\ (\theta_i, \omega_i, \psi_i)|P &\sim P, \quad P \sim EDP3\ (\alpha_\theta, \alpha_\omega, \alpha_\psi, P_0), \end{aligned} \tag{3.1}$$

where, for the ith subject, Y_i, M_i, and X_i represent the outcome, the Q-dimensional mediator vector, and the $(p+1)$ dimensional vector containing the treatment and p confounders, respectively. M_{iq} is the value of the qth mediator for the ith subject and $X_{i,r}$ is the rth element of X_i. The notation, $P \sim EDP3\ (\alpha_\theta, \alpha_\omega, \alpha_\psi, P_0)$ means that $P_\theta \sim DP(\alpha_\theta, P_{0,\theta})$, $P_{\omega|\theta} \sim DP(\alpha_\omega, P_{0,\omega|\theta})$, and $P_{\psi|\theta,\omega} \sim DP(\alpha_\psi, P_{0,\psi|\theta,\omega})$ with base measure $P_0 = P_{0,\theta} \times P_{0,\omega|\theta} \times P_{0,\psi|\theta,\omega}$, where $DP(\alpha, G)$ is a Dirichlet process with base distribution G and concentration parameter α.

The $EDP3$, is discrete and hence, subjects can share the same values of (θ, ω, ψ); such subjects are in the same cluster. By the construction of P, the aforementioned clustering is nested in three levels. The first level clusters correspond to the distinct values of the parameter θ. Given a first level cluster with θ parameter as θ_f, the second level clusters will correspond to those distinct values of the parameter ω for which the θ parameter is fixed at θ_f. Finally, for fixed values of the parameters (θ, ω), say (θ_f, ω_s), the third level clusters are characterized by the distinct values of the parameter ψ (see Section 4 and Figure H.2 in the supplementary materials for more details). The number of clusters at the three levels are controlled by the three concentration parameters, α_θ, α_ω, and α_ψ, respectively, where lower values translate into fewer clusters. Our proposed EDP3 mixture is most suitable when the dimension of Y is small, the dimension of M is moderate, and the dimension of X is large. Our real data example in Section 6 is in line with this setting. To accommodate this setting, one would require small number of y-clusters, moderate number of m-clusters, and large number of x-clusters, which is facilitated by the "enriched" three-level nested clustering of our method.

We assume (local/within cluster) generalized linear models for $Y_i|M_i, X_i; \theta_i$ and $M_{iq}|X_i; \omega_i$ in (3.1). Here, given ψ_i, the $(p+1)$ covariates, $X_{i,1}, X_{i,2}, \ldots, X_{i,(p+1)}$ are assumed to be *"locally"* (that is, intra-cluster) independent. Similarly, given ω_i, and conditional on the covariates X_i, the Q mediators $M_{i1}, M_{i2}, \ldots, M_{iQ}$ are assumed to be locally independent. This notion of local independence is similar to that in latent class models, where given latent class membership, the random variables are assumed to be independent. The assumption of local independence helps in accommodating many mediators and confounders with less computational burden, as the joint distribution is simply the product of the marginals and it does not require complex joint distributions of mediators and counfounders. However, *"globally"* (that is, intercluster) all of the variables are dependent with potentially nonlinear relationships (see various scenarios considered in the Section 5), and this local independence can be weakened; we discuss this in Section 7. In Appendix G of the supplementary material, we present a cube-breaking representation and show that although the local regression models are generalized linear models, the global regression models are computationally tractable, flexible, nonlinear, nonadditive models.

4. COMPUTATIONS

We use a Gibbs Sampler to obtain draws from the posterior distributions by utilizing a further extension of Algorithm 8 in Neal (2000) from Roy et al. (2018). A summary of the Gibbs Sampling is given next, and the details are described in Appendix E of the supplementary materials.

Let $s_i = (s_{i,y}, s_{i,m}, s_{i,x})$ denote the cluster membership for the subject i. Here, $s_{i,m}$ denotes the m-cluster within the y-cluster $s_{i,y}$, to which subject i belongs. Similarly, $s_{i,x}$ characterizes the x-cluster within $(s_{i,y}, s_{i,m})$ (see Figure H.2 in the supplementary materials). We first sample s_i for each subject and then given $s = \{s_i\}_{i=1}^n$, we sample the parameters θ, ω, and ψ (see Section 3) from their conditional distributions given the data and cluster membership. We denote by θ_j^* the θ, that is associated with the jth currently nonempty y-cluster, for $j = 1, 2, \ldots, k$. Similarly, we define $\omega_{l|j}^*$ and $\psi_{u|j,l}^*$ for $l = 1, 2, \ldots k_j$ and $u = 1, 2, \ldots k_{jl}$, where k_j is the number of currently nonempty m-clusters within the j^{th} y-cluster and k_{jl} is the number of currently nonempty x-clusters within the jth y-cluster and lth m-cluster. To demonstrate the postprocessing, it is sufficient to describe the computation of a generic expected potential outcome $E[Y(a, M(a_1, a_2, \ldots, a_Q))]$, since all the causal effects defined in Section 2, involve $E[Y(a, M(a_1, a_2, \ldots, a_Q))]$ for different combinations of $\{a, a_1, a_2, \ldots, a_Q\} \in \{0, 1\}$.

Given each posterior sample of cluster-specific parameters and concentration parameters (and the total number of subjects in each cluster), we conduct the following postprocessing steps to compute $E[Y(a, M(a_1, a_2, \ldots, a_Q))]$: (a) Draw the covariates l (see Step a in Appendix A of the supplementary materials), (b) Given the covariates from step (a), draw the mediators m in such a way that the qth mediator is induced under the treatment status a_q, for some fixed set $\{a_1, a_2, \ldots, a_Q\} \in \{0, 1\}$ (see Step b in Appendix A of the supplementary materials), (c) Given the values from (a) and (b), compute $E(Y|A = a, L = l, M = m, \theta^*, \omega^*, \psi^*, s)$ (see Step c in Appendix A of the supplementary materials). θ^*, ω^*, and ψ^* are used to denote $\{\theta_j^*\}_j$, $\{\omega_{l|j}^*\}_{j,l}$, and $\{\psi_{u|j,l}^*\}_{j,l,u}$, respectively from the particular posterior sample and $s = \{s_i\}_{i=1}^n$ denotes the corresponding cluster memberships, (d) Repeat steps (a)–(c) T times and use Monte Carlo Integration to compute $E(Y(a, M(a_1, a_2, \ldots, a_Q)))$ (see Step d in Appendix A of the supplementary materials). Using the above steps, one can compute any of the expected potential outcomes required to construct the causal effects. The computations described above are a postprocessing step, that can be done outside the Gibbs Sampler. These steps correspond to doing MC integration (G-computation) on the following integral to compute the potential outcomes needed to define any of the causal effects of interest:

$$E[Y(a, M(a_1, a_2, \ldots, a_Q))] = \\ \int\int E[Y|A = a, m, L = l]d\{F(M_{vec}^0|A = 0, L = l) \times F(M_{vec}^1|A = 1, L = l)\}dF(l),$$

where the qth mediator in m is induced under the treatment status a_q, for some fixed set $\{a_1, a_2, \ldots, a_Q\} \in \{0, 1\}^Q$; as such, we define m_{vec}^0 (and m_{vec}^1) as the set of those mediators under $A = 0$ and under $A = 1$, respectively. The G-computation can easily be parallelized.

5. SIMULATION STUDIES

In this section, we evaluate the performance of our proposed methodology on synthetic data under different scenarios. The true data generating mechanism and the numerical results for each of the scenarios are discussed next. We report point estimates (posterior means), 95% CI widths and empirical coverage probabilities corresponding to the estimation of NDE, JNIE, TE and compare the results with those obtained using a LSEM approach with bootstrap. We also present results on INIE, and for Scenario 3, we additionally report the results on PNIE for some pairs of mediators. Scenario 2, Scenario 3, and Scenario 4 are discussed next, while Scenario 1, Scenario 5, Scenario 6, and Scenario 7 are deferred to Appendix F of the supplementary material. Also, some of the tables related to this section are presented in Appendix H in the supplementary material.

Scenario 2: *Continuous outcome and mediators, complex functional forms involving nonlinear and interaction terms*

In Scenario 2, the steps of generating the data remain the same as in Scenario 1 in Appendix F of the supplementary material. However, here we introduce more complex functional forms involving nonlinear and interaction terms. To that end, while generating A, the Bernoulli probability now has the form $\text{logit}^{-1}(0.3 \sum_{j=1}^4 L_{p_1+j})$. The means of the normal distributions are considered as $\mu_{m1} = -4 + 2A - 0.5L_{p_1+2} - L_{p_1+3} + 0.5L_{p_1+4}$, $\mu_{m2} = 4 + 0.4A + 0.5L_{p_1+2}^2 - 0.8L_{p_1+3}(L_{p_1+3} > 0)$, $\mu_{y1} = -4 + 2A - 0.5L_{p_1+2} * M_Q - L_{p_1+3} * M_Q + 0.5L_{p_1+4} * M_Q$ and

Table 1. Scenario 2 results for NDE, JNIE, and TE

		Estimate	*CI width*	*Coverage*
True NDE=1.51	BNP ($n = 1,000$)	1.55	1.78	0.96
	BNP ($n = 2,000$)	1.53	1.55	0.97
	LSEM ($n = 1,000$)	0.93	0.81	0.21
	LSEM ($n = 2,000$)	0.94	0.57	0.23
True JNIE=0.41	BNP ($n = 1,000$)	0.47	1.71	0.99
	BNP ($n = 2,000$)	0.42	1.51	0.99
	LSEM ($n = 1,000$)	0.95	0.46	0.46
	LSEM ($n = 2,000$)	0.93	0.32	0.46
True TE=1.92	BNP ($n = 1,000$)	2.01	0.90	0.99
	BNP ($n = 2,000$)	1.96	0.74	0.99
	LSEM ($n = 1,000$)	1.88	0.89	0.88
	LSEM ($n = 2,000$)	1.89	0.63	0.89

$\mu_{y2} = 4 + 0.4A + 0.5L^2_{p_1+2} - 0.8L_{p_1+3}(L_{p_1+3} > 0)$. Finally, δ_m and ζ_y are specified as follows: $\delta_m = \frac{exp\{-2(L_{p_1}+1)^2\}}{exp\{-2(L_{p_1}+1)^2\}+exp\{-2(L_{p_1}-2)^2\}}$ and $\zeta_y = \frac{exp\{-2(M_Q+1)^2\}}{exp\{-2(M_Q+1)^2\}+exp\{-2(M_Q-2)^2\}}$.

The results are summarized in Table 1 (NDE, JNIE, and TE) and Table H.2 in the supplementary materials (INIE). As shown in Table 1, the NDE and JNIE results based on LSEM approach suffer from parametric misspecification and the BNP approach significantly outperforms LSEM, though the credible intervals for BNP are a bit conservative. The INIE results, summarized in Table H.2, are very good for the true mediator (the Qth mediator, $Q = 10$; see Scenario 1 for more details) and moreover, they show improvement for increased sample size (Table H.3 in the supplementary materials).

Scenario 3: *Continuous outcome and mediators, mediators are correlated and outcome model involves interaction terms among the mediators*

In Scenario 3, the data-generating steps are mostly in line with that of Scenario 2. However, here we induce correlation among the Q mediators and also include interaction among the mediators in the outcome model. As defined in Scenario 1, each row of the mediator matrix $M \in \mathbb{R}^{n\times Q}$ corresponds to a particular individual or observation. Thus to induce correlation among the mediators, we generate the rows of M, denoted by $\{M^i\}_{i=1}^n$ independently as follows:

$$M^i|A, L \sim \delta^i_m N(\mu^i_1, \Sigma_M) + (1 - \delta^i_m)N(\mu^i_2, \Sigma_M), \text{ for } i = 1, 2, \ldots, n$$

where δ^i_m is the ith element of δ_m and $\mu^i_1 = \{\mu^i_{m1}, \mu^i_{m1}, \ldots, \mu^i_{m1}\}^T \in \mathbb{R}^Q$ and μ^i_{m1} is the ith element of μ_{m1}. μ^i_2 is defined as μ^i_1 and $\Sigma_M \in \mathbb{R}^{Q\times Q}$ represents the covariance structure among the mediators, whose diagonal elements are 1 and the off-diagonal elements are 0.45. Finally, to introduce the interaction terms among the mediators, the means of the outcome model are taken as $\mu_{y1} = -4 + 2A - 0.5L_{p_1+2} * M_Q - M_{Q-1} * M_Q + 0.5M_{Q-2} * M_{Q-1}$ and $\mu_{y2} = 4 + 0.4A + 0.3M_{Q-2} * M_Q - 0.8L_{p_1+3}(L_{p_1+3} > 0)$. This scenario is more difficult than the previous one for the following reasons. First, the outcome model considered here includes the interaction among the mediators. Thus, as opposed to the previous two scenarios, here we have three true mediators, namely, mediators 8, 9, and 10 (that is, M_{Q-2}, M_{Q-1}, and M_Q for $Q = 10$). In addition to that, this scenario allows the mediators to be correlated. Table 2 summarizes the NDE, JNIE, and TE results under both BNP and LSEM approach. As expected, the LSEM results again suffer from parametric misspecification and BNP significantly outperforms LSEM. For example, the estimated NDE for BNP with $n = 1,000$ is 1.43 which is quite close to the true NDE 1.49, as compared to the LSEM estimate 2.40. For JNIE, though the performances are similar in terms of bias (the

Table 2. Scenario 3 results for NDE, JNIE, and TE

		Estimate	*CI width*	*Coverage*
True NDE= 1.49	BNP ($n = 1,000$)	1.43	1.07	0.87
	BNP ($n = 2,000$)	1.50	0.92	0.93
	LSEM ($n = 1,000$)	2.40	0.64	0.00
	LSEM ($n = 2,000$)	2.39	0.44	0.00
True JNIE= 4.61	BNP ($n = 1,000$)	5.40	2.20	0.88
	BNP ($n = 2,000$)	5.21	1.97	0.92
	LSEM ($n = 1,000$)	3.81	1.71	0.62
	LSEM ($n = 2,000$)	3.82	1.21	0.63
True TE= 6.13	BNP ($n = 1,000$)	6.31	2.17	0.87
	BNP ($n = 2,000$)	6.24	1.85	0.93
	LSEM ($n = 1,000$)	6.98	1.84	0.92
	LSEM ($n = 2,000$)	6.75	1.43	0.93

estimates are 5.40 and 3.81 for BNP and LSEM respectively, while the true value is 4.61), the coverage is significantly worse for LSEM. Note that, for this scenario, the BNP coverage values are slightly lower than the earlier scenarios. This is potentially due to the fact that the mediators are correlated here, while BNP assumes local independence (intra-cluster independence, see Section 3) of the mediators. However, even under this difficult scenario, BNP performs considerably well and the coverage approaches the desirable level as we increase the sample size. The INIE results, summarized in Table H.4 in the supplementary materials, are also quite good for the true mediators (mediators 8, 9, and 10) and moreover, as shown in Table H.6 in the supplementary materials, the bias for INIE decreases as we increase the sample size. Since this scenario has more than one true mediator, we report the results for pairwise indirect effects (PNIE). As summarized in Table H.5 of the supplementary materials, the results for PNIE are good in terms of both bias and coverage.

Scenario 4: *Conditional distributions comprising three mixture components*

As in Scenario 1, we generate a covariate matrix $L \in \mathbb{R}^{n\times(p_1+p_2)}$, wherein the first p_1 columns contain the discrete covariates and the remaining p_2 columns are for continuous covariates. Discrete covariates are generated independently from Bernoulli ($p = 0.5$) and the continuous covariates are generated independently of the discrete covariates and are drawn from $N_{p_2}(0, \Sigma_c)$, where $\Sigma_c \in \mathbb{R}^{p_2\times p_2}$ is the covariance matrix of the p_2-variate normal distribution, for which all the off-diagonal elements are 0.3. The vector of binary treatment $A \in \mathbb{R}^n$ is generated from Bernoulli ($p = 0.4$). The mediator matrix is denoted by $M \in \mathbb{R}^{n\times Q}$, for which the columns correspond to the Q mediators and they are generated independently as follows: $M_q|A, L \sim \delta_{m1}\, N(\mu_{m1}, \mathbb{I}_n) + \delta_{m2}\, N(\mu_{m2}, \mathbb{I}_n) + \delta_{m3}\, N(\mu_{m3}, \mathbb{I}_n)$, for $q = 1, 2, \ldots, Q$, where, $\delta_m = [\delta_{m1}, \delta_{m2}, \delta_{m3}]$ is taken as $(0.4, 0.3, 0.3)$. The means of the Normal distributions are considered as $\mu_{m1} = -4 + 2A - 0.5L_{p_1+2} - L_{p_1+3} + 0.5L_{p_1+4}$, $\mu_{m2} = 4 + 0.4A + 0.5L^2_{p_1+2} - 0.8L_{p_1+3}(L_{p_1+3} > 0)$, and $\mu_{m3} = 4 + 0.6A + 0.9L^3_{p_1+2} - 0.7L^2_{p_1+3}$. Finally, the outcome vector $Y \in \mathbb{R}^n$ is generated as follows:

$$Y|M, A, L \sim \zeta_{y1}\, N(\mu_{y1}, \mathbb{I}_n) + \zeta_{y2}\, N(\mu_{y2}, \mathbb{I}_n) + \zeta_{y3}\, N(\mu_{y3}, \mathbb{I}_n) \tag{5.2}$$

where, $\zeta_y = [\zeta_{y1}, \zeta_{y2}, \zeta_{y3}]$ is taken as $(0.4, 0.3, 0.3)$. The means of the normal distributions are considered as $\mu_{y1} = -4 + 2A - 0.5L_{p_1+2} * M_Q - L_{p_1+3} * M_Q + 0.5L_{p_1+4} * M_Q$ and $\mu_{y2} = 4 + 0.4A + 0.5L^2_{p_1+2} - 0.8L_{p_1+3}(L_{p_1+3} > 0)$ and $\mu_{y3} = 4 + 0.6A + 0.9L^3_{p_1+2} * M_Q - 0.7L^2_{p_1+3} * M_Q$. The results are summarized in Table H.11 of the supplementary materials. As in case of other scenarios, our method outperforms the LSEM approach, which demonstrates the desired performance of our method even if the conditional distributions have more than two mixture components.

Scenarios 1, 5, 6, and 7: As mentioned in the beginning of this section, Scenarios 1, 5, 6, and 7 are deferred to Appendix F in the supplementary material. Scenario 1 corresponds to simple functional forms of the conditional distributions. Both Scenarios 5 and 6 correspond to complex functional forms; the former with a skewed error distribution, and the latter with correlated binary covariates. Finally, Scenario 7 is similar to Scenario 1, and that is intended to check the posterior mean of the number of clusters. The results corresponding to these scenarios, like the scenarios described in this section, are also good in terms of both bias and coverage.

6. APPLICATION

Studying the effect of unintended pregnancies on the later life mental health of the women has been an area of significant interest (Bahk et al. 2015; Herd et al. 2016; Barton et al. 2017). We use data from Wisconsin Longitudinal Study (WLS) (Herd et al. 2014). Respondents were women who graduated from a Wisconsin High School in 1957. Survey data on various aspects of those respondents' life course were collected in 1957, 1964, 1975, 1992, 2003, and 2011. Herd et al. (2016) used this data and analyzed the later life mental health consequences of the women who gave births to babies from unintended pregnancies. Unlike the previous work on the effects of unwanted pregnancies, WLS respondents had experienced nearly all their pregnancies before the 1973 *Roe v. Wade* decision. Thus, most, if not all, of these women did not have the opportunity to terminate an unintended pregnancy and hence this data is more reliable for inferring the actual effect of giving birth from an unintended pregnancy. The WLS data consist of a wide range of covariates, including family background, adolescent characteristics, educational and occupational achievement, and aspirations. Finally, the information is collected longitudinally, and this is essential to assess mediation.

6.1. Exposure, outcome, confounders, mediators, and causal assumptions

In this section, we summarize the outcome, exposure, confounders, and mediators for our study. A DAG is given in Fig. 1. We also discuss the validity of the causal assumptions in the context of our data.

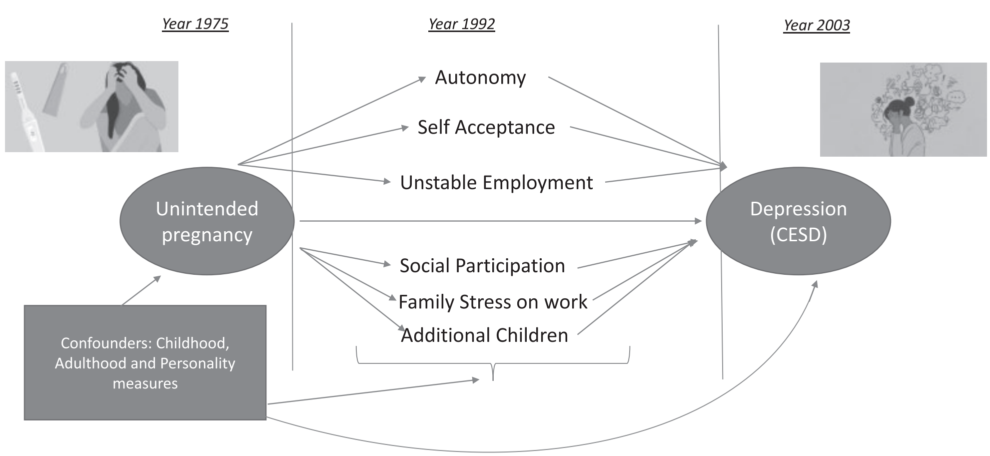

Figure 1. Pictorial illustration of causal pathway: five potential mediators (measured in survey year 1992) between the unintended pregnancies and later life depression.

Exposure: Following Herd et al. (2016), women in the exposed group had given births to at least one unintended pregnancy by the survey year 1975. The control group consists of the women who did not have any births due to any unintended pregnancies by that year. Thus, survey year 1975 is considered as the baseline for our analysis. Our sample consists of 1,701 women in total. 210 women were in the exposed group, and 1,491 women were in the control group (see Table H.13 in the supplementary material).

Outcome: The outcome, used to measure the later-life depression, is the *Center for Epidemiological Studies-Depression* score (CES-D score) as of survey year 2003, when the women were approximately 63 years of age. The CES-D score in WLS ranges over 0 to 140. A higher score implies a higher level of depression. Table H.13 in the supplementary materials summarizes the means and interquartile ranges (IQR) of the CES-D scores for both treated and control groups.

Confounders: We consider 13 baseline confounders, which are categorized into three groups: childhood, adulthood, and personality measures (Herd et al., 2016). Next, we provide a detailed description of the covariates, and Table H.13 in the supplementary materials displays the means and the interquartile ranges of the covariates for both the exposed and control group.

Childhood measures: (i) *High School percentile rank,* (ii) *IQ,* (iii) *Parental socioeconomic status*, and (iv) *Population of town*

Adulthood measures: (i) *Educational attainment:* the number of years of postsecondary schooling. (ii) *Age:* the respondent's age at her first pregnancy, (iii) *Number of children:* this was collected in the survey year 1975. (iv) *Marital status:* marital status of the women as of the survey year 1975. Note that, all these adulthood covariates were measured at the baseline year 1975 (see Herd et al. (2016) for more details).

Personality measures: The five widely used scales of personality are agreeableness, conscientiousness, extroversion, neuroticism, and openness (John et al. 1999), which we use as confounders. These measures were first collected in a 1992 mail survey. However, there are studies (Herd et al. 2016) which argued that the personality remains fairly stable in pregnancy and the transition to parenthood. Thus, even though these measures were recorded post-treatment in 1992, these could be effectively used as pre-treatment covariates (see Herd et al. 2016).

Mediators: We consider six variables as mediators, each measured in survey year 1992, and are potential paths between the effect of unintended pregnancy on later life mental health. (i) *Autonomy:* This is a score based on seven questions related to psychological well-being, and can have values ranging from 1 (lowest autonomy) to 42 (highest autonomy). (ii) *Self-acceptance:* This is also a score related to psychological well-being, where a score of 42 implies the highest level of self-acceptance. (iii) *Unstable employment:* We consider the total number of employment spells of the women as a potential mediator, for which the higher values indicate the higher level of instability. (iv) *Social participation:* This is defined as the number of organizations in which the graduate has very much involvement. Thus, a lower value implies reduced social participation. (v) *Family stress on work:* This is a score based on four questions, where the value 1 implies the lowest level of stress, and the value 20 implies the highest one. (vi) *Additional number of children:* This is the additional number of children born between the baseline year 1975 and 2003.

Assumptions: In the context of our data, Assumption 2.1 implies that the covariates childhood measures, adulthood measures and personality measures (see Table H.13 in the supplementary material) are sufficient to adjust for confounding between the unintended pregnancies (exposure) and the observable potential outcome: mental depression, and mediators. The above list of baseline covariates is in line with the relevant literature on unintended pregnancies (Herd et al. (2016)). Assumption 2.2 says that, conditional on the childhood, adulthood, and personality measures (baseline covariates), the chance that a woman will have later-life depression depends only on the values of the mediators (that is, the levels of self-acceptance, autonomy, employment instability, and so on); it does not matter whether those mediator values are induced under the unintended

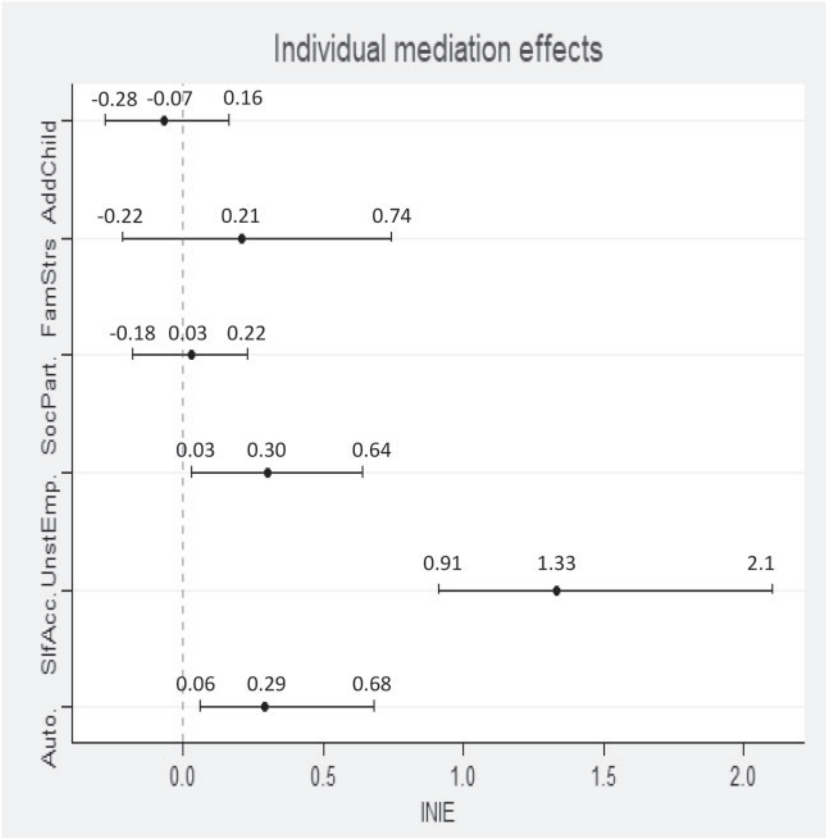

Figure 2. Individual mediation effects (INIE): self-acceptance, autonomy and unstable employment have significant individual mediation effect. Social participation, family stress on work, and additional children do not seem to be significant as individual mediators.

pregnancies (that is, presence of exposure) or not (absence of exposure). Finally, Assumption 2.3 says that, any mediator, say employment instability under the presence of unintended pregnancy, is independent of any mediator, say lack of self-acceptance in the absence of unintended pregnancy. However, they can be dependent under the same exposure status though (that is, both employment stability and self-acceptance under unintended pregnancy). Like the previous assumptions, this assumption is also untestable, however, in Appendix D of the supplementary material, we suggest a method for performing sensitivity analysis for this assumption, and provide numerical results here. As summarized in Tables D.1 and D.2 of the supplementary material, even for the large deviations from Assumption 2.3, the resulting estimates are mostly in accordance with the actual estimates, which implies that our proposed method is not very sensitive to the violation of Assumption 2.3.

6.2. Results

We apply our proposed method (Section 3 and 4) to the WLS data and estimate the posterior distribution of the causal mediation effects. The total effect is 2.21 with 95% credible interval as [0.86, 3.11], and most of that effect is routed through the joint mediation effect (JNIE) which is 2.12 with 95% credible, [1.45, 2.96]. Since the interval does not contain zero, the JNIE is significant. Thus, the effect of giving births to unintended pregnancies on later life mental depression is significantly mediated through the joint effect of the five mediators.

Table 3. Pairwise-NIE for the pairs of mediators: posterior means and 95% credible intervals. Significant estimates (95% CI excludes zero) are marked in bold.

Mediator	INIE estimate	Lower CI	Upper CI
Autonomy, Self-acceptance	**1.58**	0.75	2.21
Autonomy, Unstable employment	0.36	−0.11	0.89
Autonomy, Social participation	0.27	−0.22	0.76
Autonomy, Family stress	0.18	−0.24	0.65
Autonomy, Additional children	0.23	−0.26	0.58
Self-acceptance, Unstable employment	**1.89**	1.23	2.66
Self-acceptance, Social participation	**1.70**	0.87	2.4
Self-acceptance, Family stress	**1.60**	0.65	2.11
Self-acceptance, Additional children	**0.86**	0.1	1.71
Unstable employment, Social participation	0.24	−0.22	0.45
Unstable employment, Family stress	0.05	−0.50	0.35
Unstable employment, Additional children	0.48	−0.26	1.27
Social participation, Family stress	0.04	−0.36	0.38
Social participation, Additional children	0.21	−0.09	0.54
Family stress, Additional children	−0.05	−0.40	0.27

Given the significant joint effect, we investigate how the effects are routed through each individual mediators. As depicted in Fig. 2, there are three mediators, namely, autonomy score, self-acceptance score, and unstable employment, for which the corresponding credible intervals of the INIE do not contain zero, indicating significant individual mediation effects. Among these, the individual mediation effect is the highest (1.33) for self-acceptance score. As mentioned earlier and discussed in the existing literature (Springer and Hauser 2006; Mailick Seltzer et al. 2001), the measure of self-acceptance involves questions on achievements in life, mistakes and ups and downs in the past. Our exploratory data analysis reveals that the women having births from unintended pregnancies tend to show significantly low self-acceptance, which in later life, results in mental depression. The autonomy score, like self-acceptance, is also a psychological well-being score, and it measures the confidence in self opinions, influence of other people on decision making and so on. As displayed in Fig. 2, a small but significant effect (0.29) is mediated through low autonomy. Finally, a similar amount of significant effect (0.30) passes through the mediator unstable employment. As defined earlier, instability of employment is characterized by the higher number of employment spells. Our data analysis showed significantly higher number of employment spells among women with births from unintended pregnancies, which leads to an increase in depression score.

Finally, we look at the pairwise mediation effects (pairwise natural indirect effects) as summarized in Table 3. The pairwise NIE of autonomy and self-acceptance is significant, as is the self-acceptance and unstable employment. In addition, though social participation, family stress, and the additional number of children did not have significant individual NIEs, their pairwise NIEs with self-acceptance were significant.

7. DISCUSSION

In this article, we proposed a methodology which overcomes the two aforementioned drawbacks. Our method is based on a novel Bayesian nonparametric (BNP) approach, which modeled the joint distribution of the observed data (outcome, mediators, treatment, and confounders) flexibly, using an enriched Dirichlet process mixture with three levels. The first level specified the conditional distribution of the outcome given the mediators, treatment, and the confounders, the second level characterized the conditional distribution of each of the mediators given the treatment and the confounders, and the third level characterized the distribution of the treatment and the confounders. Causal effects were identified under some suitable causal assumptions and the proposed method was shown to have desired large sample properties. The efficacy of our proposed method was demonstrated with simulations.

Some future directions of this work are as follows. First, our approach assumes "local" independence among the covariates and among the mediators conditional on the covariates (see Section 3). As an extension, it might be of interest to induce "local" dependence, in particular among the mediators. For binary mediators, the dependence can be induced through the Dirichlet process mixture of the product multinomials (see Dunson and Xing 2009), as the mediators can be considered as unordered categorical variables. Note that this specification makes it difficult to allow the mediators to depend on the covariates locally. Another potential future work could extend our approach to accommodate mediator variable selection. In particular, one might be interested in checking whether a particular mediator (say, the jth mediator) is independent of the outcome, conditional on the exposure, confounders and the other mediators (i.e., $M_j \perp Y|A, L, M_{-j}$) and also whether the mediator is independent of the exposure, conditional on the confounders and the other mediators (i.e. $M_j \perp A|L, M_{-j}$). As opposed to the parametric approach, our nonparametric specification does not have any parameters corresponding to the aforementioned conditional independencies and thus it requires a different way to perform mediator selection (Dhara, Daniels, and Roy 2022, working paper).

8. SUPPLEMENTARY MATERIAL

Supplementary material is available at *Biostatistics Journal* online.

FUNDING

This work was partially supported by National Institutes of Health R01 HL166324, HL158963, and GM112237.

Conflict of interest statement. Authors have no conflict of interest to declare.

REFERENCES

Bahk, J., Yun, S.-C., Kim, Y. and Khang, Y.-H. (2015). Impact of unintended pregnancy on maternal mental health: a causal analysis using follow up data of the Panel Study on Korean Children (PKSC). *BMC Pregnancy Childb* **15**(1), 1–12.

Barton, K, Redshaw, M, Quigley, MA and Carson, C. (2017). Unplanned pregnancy and subsequent psychological distress in partnered women: a cross-sectional study of the role of relationship quality and wider social support. *BMC Pregnancy Childb* **17**(1), 1–9.

Chén, OY, Crainiceanu, C, Ogburn, EL, Caffo, BS, Wager, TD and Lindquist, MA. (2018). High-dimensional multivariate mediation with application to neuroimaging data. *Biostatistics* **19**(2), 121–136.

Daniel, RM, De Stavola, BL, Cousens, SN and Vansteelandt, S. (2015). Causal mediation analysis with multiple mediators. *Biometrics* **71**(1), 1–14.

Daniels, MJ, Roy, JA, Kim, C, Hogan, JW and Perri, MG. (2012). Bayesian inference for the causal effect of mediation. *Biometrics* **68**(4), 1028–1036.

Derkach, A, Pfeiffer, RM, Chen, T-H and Sampson, JN. (2019). High dimensional mediation analysis with latent variables. *Biometrics* **75**(3), 745–756.

Dunson, DB and Xing, C. (2009). Nonparametric Bayes modeling of multivariate categorical data. *J Am Stat Assoc* **104**(487), 1042–1051.

Herd, P, Carr, D and Roan, C. (2014). Cohort profile: Wisconsin Longitudinal Study (WLS). *Int. J. Epidemiol.* **43**(1), 34–41.

Herd, P, Higgins, J, Sicinski, K and Merkurieva, I. (2016). The implications of unintended pregnancies for mental health in later life. *Am J Public Health* **106**(3), 421–429.

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

Imai, K and Yamamoto, T. (2013). Identification and sensitivity analysis for multiple causal mechanisms: revisiting evidence from framing experiments. *Polit Anal*, **21**, 141–171.

John, OP, Srivastava, S et al. The big five trait taxonomy: history, measurement, and theoretical perspectives. In: Handbook of personality: theory and research, University of California, Vol. 2. 1999. p. 102–138.

Kim, C, Daniels, MJ, Hogan, JW, Choirat, C and Zigler, CM. (2019). Bayesian methods for multiple mediators: relating principal stratification and causal mediation in the analysis of power plant emission controls. *Ann Appl Stat* **13**(3), 1927.

Mailick Seltzer, M, Greenberg, JS, Floyd, FJ, Pettee, Y and Hong, J. (2001). Life course impacts of parenting a child with a disability. *Am J Mental Retard* **106**(3), 265–286.

Neal, RM. (2000). Markov chain sampling methods for Dirichlet process mixture models. *J Comput Graph Stat* **9**(2), 249–265.

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. *Math Model* **7**(9-12), 1393–1512.

Robins, JM and Hernán, MA. (2009). Estimation of the causal effects of time-varying exposures. *Longitud Data Anal* **553**, 599.

Roy, J, Lum, KJ, Zeldow, B, Dworkin, JD, Re III, Vo and Daniels, MJ. (2018). Bayesian nonparametric generative models for causal inference with missing at random covariates. *Biometrics* **74**(4), 1193–1202.

Sohn, MB, Li, H et al. (2019). Compositional mediation analysis for microbiome studies. *Ann Appl Stat* **13**(1), 661–681.

Song, Y, Zhou, X, Kang, J, Aung, MT, Zhang, M, Zhao, W, Needham, BL, Kardia, SLR, Liu, Y, Meeker, JD et al. (2021). Bayesian hierarchical models for high-dimensional mediation analysis with coordinated selection of correlated mediators. *Stat Med* **40**(27), 6038–6056.

Song, Y, Zhou, X, Zhang, M, Zhao, W, Liu, Y, Kardia, SLR, Roux, AVD, Needham, BL, Smith, JA and Mukherjee, B. (2020). Bayesian shrinkage estimation of high dimensional causal mediation effects in omics studies. *Biometrics* **76**(3), 700–710.

Springer, KW and Hauser, RM. (2006). An assessment of the construct validity of Ryff's scales of psychological well-being: method, mode, and measurement effects. *Soc Sci Res* **35**(4), 1080–1102.

VanderWeele, T and Vansteelandt, S. (2014). Mediation analysis with multiple mediators. *Epidemiol Methods* **2**(1), 95–115.

Wade, S, Dunson, DB, Petrone, S and Trippa, L. (2014). Improving prediction from Dirichlet process mixtures via enrichment. *J Mach Learn Res* **15**(1), 1041–1071.

Wade, S, Mongelluzzo, S, Petrone, S et al. (2011). An enriched conjugate prior for Bayesian nonparametric inference. *Bayesian Anal* **6**(3), 359–385.

Wang, W, Nelson, S and Albert, JM. (2013). Estimation of causal mediation effects for a dichotomous outcome in multiple-mediator models using the mediation formula. *Stat Med* **32**(24), 4211–4228.

Wang, Y-B, Chen, Z, Goldstein, JM, Buck Louis, GM and Gilman, SE. (2019). A Bayesian regularized mediation analysis with multiple exposures. *Stat. Med.* **38**(5), 828–843.

Woo, C-W, Roy, M, Buhle, JT and Wager, TD. (2015). Distinct brain systems mediate the effects of nociceptive input and self-regulation on pain. *PLoS Biol* **13**(1), e1002036.

Wu, GD, Chen, J, Hoffmann, C, Bittinger, K, Chen, Y-Y, Keilbaugh, SA, Bewtra, M, Knights, D, Walters, WA, Knight, R et al. (2011). Linking long-term dietary patterns with gut microbial enterotypes. *Science* **334**(6052), 105–108.

Zhang, H, Chen, J, Feng, Y, Wang, C, Li, H and Liu, L. (2021). Mediation effect selection in high-dimensional and compositional microbiome data. *Stat Med* **40**(4), 885–896.

Zhao, Y, Luo, X, Lindquist, M and Caffo, B. (2018). Functional mediation analysis with an application to functional magnetic resonance imaging data. arXiv, arXiv:1805.06923, preprint: not peer reviewed.